A. PATRICE SEYED (he/him/his)

EDUCATION

2012 Ph.D. Computer Science, University at Buffalo
2006 MS Computer Science, Boston University
2004 MA Psychology, Boston University
2000 BA Psychology, Minor: CS, North Carolina State University

EXPERIENCE

- Technical product lead / data scientist passionate about applying NLP, computational linguistics, knowledge-based reasoning, and deep learning to improve existing products and establishing a foundation for new products in clinical and other domains.
- Product and project management experience: coordination of key resources and personnel to complete high impact projects on time.
- Strategic Relations: facilitation within and across teams to achieve innovative, impactful goals; develop and manage relationships with high-profile stakeholders and contributors.
- Excellent oral and written communication skills, strong organizational skills, and team-oriented approach to tackling challenges.

2013-Present, 3M Health Informatics Systems

2017-Present NLP Technical Product Owner

- In the position of product owner, led a team of NLP software engineers in developing an NLP pipeline in support of new Unstructured Data Mining project, specifically Implantable Cardioverter Defibrillators (ICD) Registry product
- The team bootstrapped a rules engine in support of this product, and particularly NLP measurement annotations for Ejection Fraction percentage identification
- The team transitioned into supporting the flagship product 360 Encompass System that provides computer-assisted coding for medical reimbursement for ICD10-CM, ICD10-PCS, CPT, and CDI. Activities include improving the context engine, implementing a rules engine for counting procedures for CPT coding
- Developing microservices for a unified Document Structure Analysis engine that is consumed from three major NLP engines at 3M.
- Summer 2019 served as a mentor of Ph.D. Intern project using Active Learning to optimally use the time of subject matter experts for labeling data
- Attended NAACL 2018 and presented to larger team latest deep learning approaches presented there and how they can be leveraged for clinical NLP
- Participated on advisory group for new deep learning project that leveraged convolutional filters for CPT code prediction

NLP Research Engineer / Data Scientist

- Applied research and engineering in Medical Informatics to enhance 3M's clinical natural language processing platform within Scrum Agile Methodology
- Primarily within Clinical NLP engineering environment including UIMA framework and UMLS terminologies to predict diagnosis and procedural codes for ICD10

- Contributed to a multitude of projects involving machine learning, grammar and rule based systems, and hybrid approach for improving performance in the areas of regioning, document typing, and autocoding
- Developed efficient subsumption reasoning using compressed bit vectors
- Leveraged crowdsourcing platforms for generating training data for ICD10 autocoding, through use of corpus-based techniques generating concept-to-code candidate pairs – a project supported by 3M internal funds
- Developed machine learning models for region title and document type disambiguation
- Contributor to ensemble model approach to autocoding for precision improvements
- Contributor to machine learning model for HCA customer pilot

2012-2013, NSF DataONE Project, Tetherless World Constellation of Rensselaer Polytechnic Institute and University of New Mexico

Postdoctoral Fellow / Visiting Research Scientist

- Performed data integration for developing applications for scientific search in the environmental and ecological research domains
- Developed ontologies and vocabulary services leveraging existing ontologies in the geospatial, ecological, environment, and health domains in support of search
- Leveraged DataONE metadata catalogs for developing faceted and free-text search applications with semantic information retrieval techniques
- Led a cross-disciplinary team of undergraduate and graduate students to develop a web-based dataset annotation tool, and also an information retrieval application for the Darrin Freshwater Institute using datasets on water quality and phytoplankton species occurrence

2012, Council for Logistics Research

Ontologist

- Advised Department of Defense customers on development of vocabularies based on business process definition and identifying authoritative data sources
- Managed implementation of data transparency through a standard development and deployment process
- Elicited knowledge from Subject Matter Experts, business process documents, and enterprise architecture products
- Provided ontology development support
- Managed relationships with government and customer points of contact in support of their data strategy implementation

2011, University of Manchester, UK

NSF-Funded Visiting Researcher

- Generated several reformulations of SNOMED's anatomy and Findings axioms with respect to part-whole reasoning in OWL
- Provided analysis of the reformulations based on classification and query times across several Description Logic reasoners, and evaluated query language requirements for various information requests

2010, Stanford Research Institute (SRI) International
Knowledge Engineer Internship

- Within Project Halo, led a team of AP Biology teachers in creating an ontology based on an AP Biology textbook glossary

TECHNICAL SKILLS

- Languages: Python, R, Java, Perl, C++, C, MATLAB, R, Lisp, PHP, Bash shell scripting
- Technology Stack: AWS, Docker, Kubernetes, Gradle, Maven, Jenkins, Git, SVN
- Software: IntelliJ, VSCode, JupyterLab, Eclipse
- DB-Related: MySQL, MongoDB, Neo4j, Redis
- Libraries: Spark, Hadoop, Dask, Pandas, SciPy, Scikit-Learn, TensorFlow, Keras
- Other NLP-Related: Python NLTK, Spacy, UIMA, Weka, Mahout, ClearTk
- Operating Systems: MacOS, Linux, Windows
- Certified Agile Product Owner, Scrum Master

SELECT PUBLICATIONS

- **A. Patrice Seyed**, *Subsumption Preservation as a Comparative Measure for Sense-Directed Embeddings*, Workshop on Evaluating Vector Space Representations for NLP, ACL, Berlin, Germany, August 7-12, 2016.
- **A. Patrice Seyed**, Evan Patton, and Michael Subotin, *Crowdsourcing for ICD10 Code-to-Concept Relationships*, International Semantic Web Conference, Poster Session, Bethlehem, Pennsylvania, October 11-15, 2015.
- **A. Patrice Seyed,** Zachary Fry, and Deborah McGuinness, *Improving Ontology Service-Driven Entity Disambiguation*, Formal Ontology Meets Industry, Formal Ontology in Information Systems, Rio De Janeiro, Brazil, September 22-25, 2014.
- **A. Patrice Seyed**, John Cabral, Daniel Mekonnen, and Michael Corrigan, *Applying an Ontological Theory of Information Assets to Business Process Reengineering*, Fourth International Workshop on Ontology-Driven Information Systems Engineering, July 2012, Graz, Austria.

AWARDS AND HONORS

- Reviewer for: International Journal of Human-Computer Studies, Journal of Applied Ontology, and International Conference on Formal Ontology in Information Systems (2012), OntoContent (2012, 2013, 2014), International Workshop on Definitions in Ontologies (2013), International Conference on Biomedical Ontologies (ICBO) (2013, 2014), International World Wide Web Conference (2013), International Semantic Web Conference Posters and Demos (2014, 2015), Bio-Ontologies Special Interest Group (2012, 2013, 2014, 2015), Semantic Technology For Intelligence, Defense, and Security (STIDS) (2014, 2015, 2016), International World Wide Web Conference 2016.
- Invited speaker at Computational Linguistics at Boston University (CLABU), “Using Perl and regular expressions for lexical token analysis”.
- Guest Lecturer on “Unsupervised Learning” for Ling 467 at Georgetown University, September 2015.