

Jaspreet Bhatia

Address: Institute for Software Research, Carnegie Mellon University,
4208 Wean Hall, 5000 Forbes Avenue, Pittsburgh, PA-15213, United States of America
Contact: + 1 412-944-6927 | Email: jbhatia@cs.cmu.edu, jaspreet2709@gmail.com
Website: <http://www.cs.cmu.edu/~jbhatia/>

EDUCATION

8/2014 – Present **Carnegie Mellon University, Pittsburgh, PA, USA**

- Doctor of Philosophy, Software Engineering. Advisor: Prof. Travis Breaux
- Privacy, natural language processing, user studies, statistical data analysis, empirical research, applied deep learning, crowdsourcing, requirements engineering

7/2011 – 8/2013 **Indian Institute of Technology, Delhi, India**

- Master of Science (Research), Computer Science. GPA= 8.71/10, Advisor: Prof. K.K. Biswas
- Ambiguity Detection in Natural Language Requirements Specifications

7/2007 – 5/2011 **Indira Gandhi Institute of Technology, Delhi, India**

- Bachelor of Technology, Computer Science. GPA= 83.28%, Advisor: Prof. Kalpana Yadav
- Open Source Software Reliability, Awarded best undergraduate research project

ACADEMIC EXPERIENCE

Carnegie Mellon University

PA, USA

Graduate Research Assistant – Institute for Software Research, School of Computer Science

8/2014 – Present

- Automated extraction and analysis of privacy requirements from privacy policies using deep learning, crowdsourcing, natural language processing, and machine learning.
- Understanding and measuring perceived privacy risk by conducting user studies.

Indian Institute of Technology – Delhi

Delhi, India

Researcher – Software Engineering Group

7/2013 – 5/2014

- Automated identification of Business Rules in requirements documents using Machine Learning (ML) algorithms.
- Automated conversion of natural language requirements into logical form and UML diagrams – using an intermediate structured representation developed as part of my MS research.

Tata Research Development and Design Centre

Pune, India

Research Intern – Software Engineering Lab

12/2012 – 1/2013

- Manual Analysis of real time industry requirements documents to identify the underlying issues.
- Evaluation of my ambiguity detection approach and requirements structuring approach (developed as part of my MS studies) on documents from the industry.

Sedulity Groups

Delhi, India

Engineering Intern

Summer 2010

- Technical understanding of Network Protocols, Network Attacks, Hacking and Cyber Security.
- Developed hacks for Windows XP and networking games.
- Analyzed low-level cyber-crimes and authored a research based report on Network Security and Ethical hacking.

Computer Maintenance Corporation (a TATA enterprise)

Delhi, India

Engineering Intern

Summer 2009

- Conceptual understanding of Database Management Systems.
- Technical understanding and usage of Oracle Database 10g and Oracle Developer 2000.
- Developed a Database Management System Software (front-end and back end) for an Inventory system.

ONGOING PROJECTS

Identify Semantic Roles in Privacy Policies using Deep Learning

Advisor: Prof. Travis Breaux (Carnegie Mellon University)

- Developing a deep learning framework to identify semantics of data practices in privacy policies.

- Building an annotated corpus of privacy policies, and implementing deep neural network architecture to identify semantic roles associated with data practices.
- Measure the effect of ambiguity in context on privacy risk by conducting user studies.

Empirical Measurement of Perceived Privacy Risk.

Advisor: Prof. Travis Breaux (Carnegie Mellon University)

- Developed an empirically validated framework to measure perceived privacy risk.
- Evaluated the framework for factors which affect the perceived privacy risk experienced by the user, such as – risk likelihood, data types, privacy harms, data purposes, and demographic factors such as age, education level, and ethnicity among others.
- Framework to benefit developers by helping them design systems that take into account user privacy; regulators and privacy policy authors by helping them identify high and moderate risk data practices; and can be used to support the standardization of tools such as Privacy Impact Assessment.

Automated Extraction of Requirements from Privacy Policies.

Advisor: Prof. Travis Breaux (Carnegie Mellon University)

- Automated the extraction of requirements from website privacy policies using crowdsourcing, grounded analysis, natural language processing and machine learning techniques.
- Understanding and measuring the user’s perception about privacy and their interpretation of natural language in privacy policies by conducting empirical studies.

Automated Detection and Measurement of Vagueness and Elasticity in Privacy Policies.

Advisors: Prof. Travis Breaux (Carnegie Mellon University) and Prof. Joel Reidenberg (Fordham University)

- Working on automating the identification and measurement of vagueness and elasticity in website privacy policy text using grounded analysis, user studies, natural language processing and machine learning techniques.

PREVIOUS PROJECTS

Semantically Aware System for Question Answering

Advisor: Prof. Eduard Hovy (Carnegie Mellon University)

- Developed a question answering system that uses lexical, syntactic and semantic features along with machine learning algorithms to answer multiple choice questions from a given story.

Mobile Application for Travelers

Advisor: Prof. Travis Breaux (Carnegie Mellon University)

- Gathered requirements, designed and prototyped a mobile application to benefit travelers in an unknown city.

Thesis: Ambiguity Detection in Natural Language Requirements Specifications

Advisor: Prof. K.K. Biswas (IIT – Delhi)

- Working on automating different activities of Requirements Analysis phase of SDLC.
- Focus on automatic identification of ambiguity (Lexical, Pragmatic, Syntactic and Semantic) in requirements specifications expressed in natural language - using NLP and Machine Learning techniques.
- Automatic structuring of NL requirements in the form of *Frames*, making use of *Grammatical Knowledge Patterns*.

Hybrid Algorithm for Word Sense Disambiguation (WSD)

Advisor: Prof. Niladri Chatterjee (IIT – Delhi)

- Developed a new technique for Word Sense Disambiguation. It is Extended and Weighted Lesk’s Algorithm.
- The approach is “extended” as it makes use of the words semantically related to the bag of words in question by different relationships in the WordNet network.
- A weight is assigned to the words using two metrics, one is the distance between them in the WordNet hierarchy and the second metric is based on the information derived from a sense tagged corpus using the Naïves Bayes classifier.

Flattening Internet Topology

Advisor: Prof. Vinay Ribero (IIT – Delhi)

- The objective of the project was to study if the Flattening Internet Topology theory still holds. And to determine the extent of deployment of the LANs.

- Study showed that the infrastructure of the internet is changing and the big content providers are deploying their own LANs thus bypassing the Tier 1 ISPs.

Discrete Road Event Simulator

Advisor: Prof. S.N. Maheshwari (IIT – Delhi)

- Developed and implemented a road event simulator taking into consideration various road issues such as - uneven traffic; traffic rules; different objects such as traffic lights, pedestrians, cars, heavy vehicles and congestion on the road.
- Conceptual model was based on Object- Oriented Principles, where each entity was modeled as a class with attributes.
- The model implemented was based on the entity classes and the interaction between the classes.

Open Source Software Reliability

Advisor: Prof. Kalpana Yadav (IGIT)

- Analysis of Open Source Software Reliability.
- Comparative study of reliability assessment models based on mathematical and statistical parameters.
- **Awarded best undergraduate research project in the department.**

PUBLICATIONS

Journal Publications

- Jaspreet Bhatia, Travis D. Breaux, "Empirical Framework for Understanding and Measuring Perceived Privacy Risk." Under review: ACM Transactions on Computer-Human Interaction.
- Jaspreet Bhatia, Travis D. Breaux, Florian Schaub. "Privacy Goal Mining through Hybridized Task Re-composition" ACM Transactions on Software Engineering Methodology (TOSEM), 25(3): Article 22, 2016.
- J.R. Reidenberg, J. Bhatia, T.D. Breaux, T.B. Norton, "Automated Comparisons of Ambiguity in Privacy Policies and the Impact of Regulation", Journal of Legal Studies, 45, 2, part 2, Mar 2016. **(Honorable Mention for Privacy Papers for Policymakers)**

Conference Proceedings

- M. C. Evans, J. Bhatia, S. Wadkar, T. D. Breaux, "An Evaluation of Constituency-based Hyponymy Extraction from Privacy Policies." Accepted to: 25th IEEE International Requirements Engineering Conference (RE'17), Lisbon, Portugal, 2017.
- Jaspreet Bhatia, Travis D. Breaux, "A Data Purpose Case Study of Privacy Policies." Accepted to: RE@Next! track at 25th IEEE International Requirements Engineering Conference (RE'17), Lisbon, Portugal, 2017
- J. Bhatia, T.D. Breaux, J.R. Reidenberg, T.B. Norton. "A Theory of Vagueness and Privacy Risk Perception," *Nominated for best paper award*. IEEE 24th International Requirements Engineering Conference (RE'16), Beijing, China, 2016.
- R. Slavin, X. Wang, M.B. Hosseini, W. Hester, R. Krishnan, J. Bhatia, T.D. Breaux, J. Niu. "Toward a Framework for Detecting Privacy Policy Violation in Android Application Code," ACM/IEEE 38th International Software Engineering Conference (ICSE'16), Austin, Texas, 2016.
- Joel Reidenberg, Jaspreet Bhatia, Travis Breaux, "Automated Recognition of Privacy Policy Ambiguity", 8th Annual Privacy Law Scholars Conference, June 4-5, 2015.
- Richa Sharma, Jaspreet Bhatia, KK Biswas, "Automated Identification of Business Rules in Requirements Documents", Fourth International Advanced Computing Conference (IACC 2014), India.

Workshop Proceedings

- J. Bhatia, T. D. Breaux, L. Friedberg, H. Hibshi, D. Smullen, "Privacy Risk in Cybersecurity Information Sharing", 3rd ACM Workshop on Information Sharing and Collaborative Security, 2016.
- Jaspreet Bhatia, Morgan Evans, Sudarshan Wadkar, Travis D. Breaux "Automated Extraction of Regulated Information Types using Hyponymy Relations" IEEE 3rd International Workshop on Artificial Intelligence for Requirements Engineering (AIRE), Beijing, China, Aug. 2016.
- Jaspreet Bhatia, Travis D. Breaux "Towards an Information Type Lexicon for Privacy Policies" IEEE 8th International Workshop on Requirements Engineering and Law (RELAW), Ottawa, Canada, pp. 19-24, Aug. 2015.
- Richa Sharma, Jaspreet Bhatia, and K. K. Biswas. 2014, "Machine Learning for Constituency Test of Coordinating Conjunctions in Requirements Specifications", In Proceedings of the 3rd International Workshop on Realizing Artificial Intelligence Synergies in Software Engineering (RAISE 2014). ACM, New York, NY, USA, 25-31.

- Jaspreet Bhatia, Richa Sharma, KK Biswas, Smita Ghaisas, “Using Grammatical Knowledge Patterns for Structuring Requirement Specifications”, Third International Workshop on Requirements Patterns (RePa 2013) held at 21st International Requirements Engineering Conference (RE’2013), pg 31-34, Brazil.

Others

- Jaspreet Bhatia, Richa Sharma, KK Biswas, “Semantic Ambiguity Detection in Requirements Specifications”, Grace Hopper Conference 2012, Bangalore, India.
- Jaspreet Bhatia, Eshna Jain, Kalpana Yadav, “Comparative Study of Open Source Reliability Assessment Models”, International Conference on Computer Research and Development (ICCRD 2011) Shanghai, China; March 11 – 13, 2011; (ISBN: 978-1-61284-838-9).
- Jaspreet Bhatia, Eshna Jain, Kalpana Yadav, “Open Source Software Reliability-A Review Paper”, Proceedings of 5th National Conference, INDIACOM – 2011, Computing For Nation Development.

TECHNICAL SKILLS

- **Languages:** JAVA, JavaScript (working knowledge)
- **Concepts:** Privacy, Natural Language Processing, User Studies, Statistical Analysis of Qualitative and Quantitative Survey Data, Applied Deep Learning, Software Engineering, Requirements Engineering.
- **Softwares:** Eclipse, WEKA, Stanford Parser, Stanford Tagger and Stanford NLP tools, SPSS, R, Tensorflow.

HONORS & AWARDS

- **Honorable Mention for Privacy Papers for Policymakers 2016** ([Press Release](#))
- **Best Paper Award Nomination (RE’2016)**
- Graduate Research Scholarship, Institute for Software Research, Carnegie Mellon University (2014-Present).
- Awarded Graduate Research Scholarship by Human Resource & Development Ministry, Govt. of India (2011-2013).
- Travel grant for Grace Hopper conference 2012, India.
- 99.5%ile among 0.137 million applicants in Graduate Aptitude Test in Engineering (GATE) 2011.
- Department rank 2 out of 70 students in CSE department during undergraduate studies at IGIT.

POSITIONS OF RESPONSIBILITY (Teaching)

Carnegie Mellon University

Pittsburgh, India

Teaching Assistant, 17-652, 752 – Methods: Deciding What to Design

8/2016 – 12/2016

- This graduate course covers requirements engineering methods, including goal-oriented requirements, use cases and contextual design.
- Responsible for – grading assignments and reading questions; conducting office hours and recitations.
- Guest lecture – Creative requirements engineering.

Indian Institute of Technology – Delhi

Delhi, India

Head Teaching Assistant, CSL 101– Introduction to Computer Science

1/2013 – 5/2013

- Responsible for contributing to exam and assignment questions, preparing marking guidelines, conducting labs, evaluation of assignments, demonstrations, viva, evaluation of answer scripts.
- Coordinator for the group of 400 students and 20 TAs.

Teaching Assistant, CSL 201 – Data Structures

7/2012 – 12/2012

- Responsible for evaluation of assignments, demonstrations, viva, Evaluation of answer scripts and conducting labs.