

# Ahrim Han, Ph.D.

PRIVACY ENGINEER · MACHINE LEARNING ENGINEER

Irvine, CA (US Permanent Resident / Green Card Holder)

☎ (+1) 949-910-8751 | ✉ ahrimhan@gmail.com | 🏠 ahrimhan.github.io | 📷 ahrimhan | 🌐 ahrimhan

## Summary

I am transitioning from academia to industry to work as a **privacy engineer**. My research area is software engineering in computer science, and I have been worked for software quality-driven development. Over the past 15 years, I develop methods and tools for **assessing and improving software design and code quality** based on **statistics** and **machine learning** techniques. The main focus of my research is on **optimization** and **prediction**. Recently, I have completed the boot camp for the **data science** career track and performed projects specialized in **deep learning**. I am interested in integrating privacy practices into the software development lifecycle. I am particularly interested in researching and participating in open source projects related to differential privacy, federated learning, and potential privacy risk prediction.

## Education

### Korea Advanced Institute of Science and Technology (KAIST)

Ph.D., COMPUTER SCIENCE

Daejeon, South Korea

Feb. 2007 - Aug. 2013

### Korea Advanced Institute of Science and Technology (KAIST)

M.S., COMPUTER SCIENCE

Daejeon, South Korea

Sep. 2004 - Feb. 2007

### Sogang University

B.E., COMPUTER SCIENCE

Seoul, South Korea

Feb. 2000 - Feb. 2004

## Projects

### Data Anonymization for Privacy

Feb. 2020 - May. 2020

- De-identifying sensitive data by implementing anonymization techniques (e.g., generalization for bucketing, suppression, and pseudonymization)
- Implemented the k-anonymity and the information loss for measuring privacy and accuracy of anonymized data
- Observed the accuracy of machine learning models built with data of different levels of privacy (including differential privacy technique)

### Sentimental Analysis of Movie Reviews using a Deep Learning Neural Network (Springboard)

Mar. 2019 - Jul. 2019

- Built the Convolutional Neural Network (CNN) deep learning models with Keras for classifying the sentiment expressed in texts of movie reviews
- Compared various models by observing the parameters affecting the performance in accuracy and discussed the overfitting

### Prediction of Scores for Public Schools in California (Springboard)

Oct. 2018 - Jun. 2019

- Applied data science-related techniques for data wrangling, exploratory data analysis, data visualization, and machine learning modeling
- Provided the prediction models using the regression and classification algorithms for finding the inferior schools that need help

### Verification via Context-Aware Testing (Korea University)

Sep. 2014 - Aug. 2015

- Developed the test case generation algorithm by inferring the resources from the permissions used for running Android applications
- Generated test cases by permuting different resource conditions and prioritized the test cases by weighting the criticality of each resource

### Recommendation for Software Design Quality Improvement (Korea University)

Sep. 2013 - Apr. 2018

- Developed an efficient refactoring recommendation system to help software developers make code changes easier
- Developed a fast refactoring candidate assessment method for measuring maintainability using the matrix computation in order to increase the efficiency of the heavy computation when evaluating a large number of refactoring candidates
- Devised the two-phased assessment approach to improve the efficiency of the refactoring selection process by the search space reduction
- Increased the speed 13.5 (max) times faster compared to the no-reduction approach while still achieving the same amount of improvement

### Bad Smell Detection for Refactoring Candidate Identification (KAIST)

Sep. 2011 - Aug. 2013

- Developed a bad smell (e.g., Feature Envy) detection algorithm to identify refactoring candidates in most frequently used or really in use
- Used the dynamic-profiling technique for extracting run-time information when building the traceability relationships of software entities

### Detecting Implied Scenarios (KAIST)

Mar. 2010 - Aug. 2011

- Developed the algorithm for detecting unexpected behaviors in designs deviated from the contract specifications without state explosion problems
- Used the model checking technique on synthesized scenarios to check the properties extracted from the specifications

### Change Proneness Prediction (KAIST)

Mar. 2006 - Feb. 2010

- Developed a prediction model to help developers to identify the change-prone components and modify them for the better designs
- Devised the new behavioral dependency metrics capturing the dynamic aspects of the program for building a multiple linear regression model
- Increased the model accuracy (R-square) by 8% over the model using only metrics capturing program structures

# Experience

---

## Springboard

DATA SCIENCE CAREER TRACK FELLOW (DEEP LEARNING SPECIALIZATION)

Oct. 2018 - Aug. 2019

- Mastering skills in Python, SQL, data wrangling, data visualization, exploratory data analysis, and machine learning, and deep learning methods
- Performed two data science-related capstone projects:  
“Prediction of Scores for Public Schools in California” and “Sentiment Analysis of Movie Reviews using a Deep Learning Neural Network”

## Korea University

Seoul, South Korea

RESEARCH PROFESSOR

Sep. 2013 - Apr. 2018

- Advised graduate students to develop research topics and conduct the experiments
- Awarded \$158K in grants from the National Research Foundation of Korea (NRF) as a **sole Principal Investigator**
- Led the research projects and published the results in the top tier journals (e.g., IEEE Transactions on Software Engineering)

## Korea Advanced Institute of Science and Technology (KAIST)

Daejeon, South Korea

GRADUATE RESEARCHER

Mar. 2005 - Aug. 2013

- Mentoring and taught computer science courses: “Introduction to JAVA programming” and “Capstone Projects in CS”
- Actively performed the research projects and implemented metric measurement tools using Java and Python

## Peace Corps

Washington, D.C., USA

INTERN

Aug. 2004 - Oct. 2004

- Served in organizing and populating the intranet web pages in the Technical Infrastructure and Support Team

# Skills

---

**Data Analysis, Visualization, and Modeling:** Pandas, Numpy, Scipy, Matplotlib, Pyplot, Seaborn, Scikit-Learn, PySpark, NLTK

**Statistical Methods and Machine Learning:** Classification, Regression, Clustering, Hypothesis Testing, Keras, Tensorflow, Deep Learning

**Programming Languages and Tools:** Python, SQL, Java, R, UML, Fortran, Assembly, Markdown, Latex, SPSS, Jupyter Notebook, Visual Studio Code

**Research Areas:** Software Engineering, Software Quality Driven Development, Change Impact Analysis, Metrics, Software Quality Assessment, Software Design Improvement, Refactoring, Design Patterns, Prediction, Optimization, Natural Language Processing

# Publications

---

[1] **A. Han** and S. Cha, “Two-phase Assessment Approach to Improve the Efficiency of Refactoring Identification,” **IEEE Transactions on Software Engineering**, Vol. 44, No. 10, pp. 1001 - 1023, Oct. 2018.

[2] **A. Han**, D. Bae, and S. Cha, “An efficient approach to identify multiple and independent Move Method refactoring candidates,” *Information and Software Technology*, Vol. 59, pp. 53-66, Mar. 2015.

[3] K. Song, **A. Han**, S. Jeong, and S. Cha, “Generating various contexts from permissions for testing Android applications,” *Proceedings of the 27th International Conference on Software Engineering and Knowledge Engineering*, pp. 87-92, Jul. 2015.

[4] **A. Han** and D. Bae, “An efficient method for assessing the impact of refactoring candidates on maintainability based on matrix computation,” *Proceedings of the 21st Asia-Pacific Software Engineering Conference*, pp. 453-460, Dec. 2014.

[5] **A. Han** and D. Bae, “Dynamic profiling-based approach to identifying cost-effective refactorings,” *Information and Software Technology*, Vol. 55, No. 6, pp. 966-985, Jun. 2013.

[6] I. Song, S. Jeon, **A. Han**, and D. Bae, “An approach to identifying causes of implied scenarios using unenforceable orders,” *Information and Software Technology*, Vol. 53, No. 6, pp. 666-681, Jun. 2011.

[7] **A. Han**, S. Jeon, D. Bae, and J. Hong, “Measuring behavioral dependency for improving change-proneness prediction in UML-based design models,” *Journal of Systems and Software*, Vol. 83, No. 2, pp. 222-234, Feb. 2010.

[8] **A. Han**, S. Jeon, D. Bae, and J. Hong, “Behavioral Dependency Measurement for Change-proneness Prediction in UML 2.0 Design Models,” *Proceedings of 32nd Annual IEEE International Conference on Computer Software and Applications*, pp. 76-83, Jul. 2008.

# Honors and Awards

---

2014 - 2017 **Individual Basic Science Research Grant**, National Research Foundation of Korea (NRF), **Principal Investigator**, **\$125,000**

2016 **Best Paper Award**, Software Engineering Society of Korean Institute of Information Scientists and Engineers, **Prize: \$1,000**

2013 - 2014 **Post-Doctoral Fellowship Grant**, National Research Foundation of Korea (NRF), **Principal Investigator**, **\$33,000**

2011 - 2012 **SAMSUNG Graduate Fellowship**, SAMSUNG Electronics by Video Display Division