

# Seongtaek Lim

seolim@adobe.com  
sl2838@cornell.edu  
<https://stlim0730.github.io>

## Education

### Cornell University

PhD in Information Science  
Minor in Computer Science  
Aug 22 2016 - Aug 19 2019  
New York, NY

### UC Berkeley

PhD Student  
in Information Science  
Transferred out with my advisor  
Aug 22 2012 - Aug 21 2016  
Berkeley, CA

### Yonsei University

MS in Cognitive Engineering  
Human-Computer Interaction  
Mar 2 2007 - Feb 23 2009  
Seoul, South Korea

### Yonsei University

BS in Computer Science  
Mar 3 2003 - Feb 26 2007  
Seoul, South Korea

## Skills

### Programming / Software

C/C++/C# • Java • Python • HTML  
• JavaScript • CSS • SQL • NoSQL  
• PostgreSQL • Hadoop • Spark •  
MongoDB • Git

### Libraries & Frameworks

TensorFlow • Scikit-learn • Scipy  
• Pandas • Numpy • Django •  
Flask • Tornado • React.js •  
Backbone.js • D3.js • Node.js •  
Express.js • Webpack • Mongoose

### Techniques

Web development • Machine  
learning • Machine learning •  
Natural language processing •  
Data visualization • Statistical  
analysis

### Stat / BI Tools

R • SPSS • Tableau

## Relevant Coursework

- Data Mining and Analytics
- Analyzing Big Data with Twitter
- Cutting-Edge Web Technology
- Information Visualization and Presentation
- Applied Natural Language Processing
- Causal Inferences and Design on Experiments

## Doctoral Research

### Cornell University Software Designs to Support User Collaboration

My research has largely focused on software development and research on the impacts of software designs on users behaviors. I build novel interactive software tools for collaboration, data collection, and data analysis. I'm the creator of GLIDE, MOOCchat, Farmview, CoHeadline, and WordToRI.

## Patent

- H. Ko, Kim, J., H. Park, D. Kang, Park, C., **Lim, S.**, Cha, S. Y., H. Cho, System for Tangible Broadcasting and Control Method using the Same, 10-1076691, Oct. 2011

## Work Experience

### Adobe Software Engineer IV / Data Scientist / Quant UX Researcher

Sep 16 2019 - Present | 345 Park Ave, San Jose, CA 95110

- User analytics engineering for AI-powered products (project Phylo and Camelia)
- Data analysis on product usage data
- Development of web-based data generation software

### ProPublica Research Fellow at Civic News Lab

Jun 20 2016 - Aug 19 2016 | 155 6th Ave 13th Floor, New York, NY 10013

- Development of a data collection software for data-driven journalism
- Development of a browser extension visualizing A/B testing in online journalism
- Work published in ProPublica's third episode of Breaking the Black Box series (<https://www.propublica.org/article/breaking-the-black-box-when-machines-learn-by-experimenting-on-us>)

### Fuji Xerox Palo Alto Laboratory Research Intern

May 26 2014 - Aug 15 2014 | 3174 Porter Dr, Palo Alto, CA 94304

- Topic modeling of dynamic temporal ranges of documents using LDA
- Development of a data visualization tool for exploratory analysis
- Work published in ACM conferences – CSCW15 and IUI16

### KAIST Institute for Information & Electronics Research Military Service

Exempted Research Scientist

Mar 9 2009 - Mar 30 2012 | 335 Gwahangno, Yuseong-gu, Daejeon, South Korea

- Worked on social computing and human-computer interaction research
- Work published in international journals – IJHCS and CHB

## Selected Project Experience (More can be found at <https://stlim0730.github.io/>)

**GLIDE; Git-Learning IDE** Developed and deployed a web-based IDE designed toward systematic collaboration process for novice users' web programming practice. The scaffolding design helped the users' task performance. (Django / Python / React.js / JavaScript / GitHub API / Nginx)

**FarmView** Built an interactive GIS (geographic information system) for farmers in Northern California in trying to alleviate information asymmetry between landlords and farmers.

(Django / Python / JavaScript / Nginx) / ArcGIS) / MapBox)

**DeepProspect** Collected and analyzed 230k+ season records for MiLB data and 96k+ season records for draft data using deep learning and support vector machine to predict promising players among the current young players. Shared the results with Boras Corporation, the largest sports agency in the world. (TensorFlow / Scikit-learn / Pandas / Scipy / BeautifulSoup / Python)

**MOOCchat** Developed and deployed a web-based real-time group formation and discussion tool for online education platforms (MOOCs). The design of the software implemented collaborative learning in an isolated context. (Node.js / Express.js / AJAX / JavaScript / HTML / CSS)

## Awards

### Google

Research Grant

Social Interactions Focused  
Program  
2012

### Cornell University

Research Assistantship  
2016 - 2019

### UC Berkeley

Research Assistantship  
2012 - 2014

### UC Berkeley

Teaching Assistantship  
2014 - 2015

## References

### Tapan Parikh, PhD

Associate Professor  
Information Science  
Cornell University  
tsp53@cornell.edu

### Xin Lu, PhD

Engineering Manager  
Emerging Product Group  
Adobe  
xinl@adobe.com

### Patrick Chiu, PhD

Principal Research Scientist  
Fuji Xerox Palo Alto Laboratory  
chiu@fxpal.com

### Jinwoo Kim, PhD

Professor  
Graduate Program in Cognitive  
Science  
Yonsei University  
jinwoo@yonsei.ac.kr

### Kwangyun Wohn, PhD

Emeritus Professor  
Graduate School of Culture  
Technology  
wohn@kaist.ac.kr  
KAIST

## Leadership Activities

### Social Chair

PhD students organization at  
Cornell Tech (PACT)

### Observation Dept Chair

Yonsei Amateur Astronomical  
Association (YAAA)

**VirtualStadium** Developed a social media where users watch sports games while interacting with each other through gesture-based communication channels.  
(Java / Tomcat / Python / JavaScript / HTML / CSS )

## Publications

- Lim, S. , Varanasi, R., & Parikh, T. (2018). GLIDE (Git-Learning IDE): In-class collaboration in web engineering curriculum for youths. *In Proceedings of the 49th ACM Technical Symposium on Computer Science Education (SIGCSE '18)*. ACM.
- Chen, F., Chiu, P., & Lim, S. . (2016). Topic Modeling of Document Metadata for Visualizing Collaborations over Time. *In Proceedings of the 21st International Conference on Intelligent User Interfaces (IUI '16)*. ACM.
- Lim, S. , & Chiu, P. (2015). Collaboration Map: Visualizing Temporal Dynamics of Small Group Collaboration. *In Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM.
- Coetzee, D., Lim, S. , Fox, A., Hartmann, B., & Hearst, M. A. (2015). Structuring interactions for large-scale synchronous peer learning. *In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. ACM.
- Lim, S. , Coetzee, D., Hartmann, B., Fox, A., & Hearst, M. A. (2014). Initial experiences with small group discussions in MOOCs'. *In Proceedings of the first ACM conference on Learning@ scale conference (L@S '14)*. ACM.
- Lim, S. , Cha, S. Y., Park, C., Lee, I., & Kim, J. (2012). Getting closer and experiencing together: Antecedents and consequences of psychological distance in social media-enhanced real-time streaming video. *Computers in Human Behavior*, 28(4), 1365-1378.
- Lim, S. , Cha, S. Y., Park, C., Lee, I., & Kim, J. (2011). Idioculture in crowd computing: A focus on group interaction in an event-driven social media system. *International Journal of Human-Computer Studies*, 69(10), 632-646.
- Lim, S. (2009). Co-experience: Enhancing the interactive experience of social media users through tangible interaction. *Master's Thesis, Yonsei University*.
- Lim, S. , Park, C., Cha, S. Y., Moon, J., Lee, I., & Kim, J. (2009). The role of event-driven social media and dynamic sports broadcasting system as crowd computing media. *Crowd-Computer Interaction Workshop in the 27th International Conference on Human Factors in Computing Systems (CHI '09)*. ACM.
- Yang, S., Lim, S. , Lee, I, Lee, S., & Kim, J. (2008). A study on interaction factors influencing use intention of interactive video service: Focusing on media synchronicity. *In Proceedings of the Seventh Annual Workshop on HCI Research in MIS (Pre-ICIS '08)*.

## Teaching Experience

### Information Technology High School Guest Lecturer

Oct - Nov 2018 | Queens, NY

- Web front-end programming with HTML, CSS, and JavaScript following Git workflow
- High school seniors

### UC Berkeley Teaching Assistant

Fall 2014 - Spring 2015 | Berkeley, CA

- Data Structure and Algorithms course
- Master of Information and Data Science program in School of Information