

Daniel Youngho KIM, Ph.D.

Personal Information

- Date of Birth: April 17th, 1986
- Address: 1004, Jayang-ro 138, Gwangjin-gu, Seoul 05039, Republic of Korea
- Contact
 - E-Mail: daniel.youngho.kim@gmail.com
 - Mobile: +82-10-4195-7436

Education

- KAIST, Republic of Korea, Ph.D. degree in Physics (September 2008 — August 2015)
 - Supervisor: Prof. Hawoong Jeong (KAIST)
- Institute for New Economic Thinking, University of Oxford, UK (November 2013 — June 2014)
 - Visiting Research Assistant
- KAIST, Republic of Korea, Bachelor's degree in Physics (March 2004 — August 2008)

Summary

- Data Analysis Experience (10 years) including Deep neural network related one
- Research ability: 10 academic papers (data science + physics. One of them used to be posted on Nature.com main page)
- Multidisciplinary collaboration experience with art historians, computer scientists, economists, historians, sociologists, statisticians, and so on. Optimist, open minded.
- 8 awards and 11 data visualizations

Work Experience

Samsung SDS, Republic of Korea (September 2016 — Present)

- Position: Data Scientist
- Affiliation: Data Analytics Cell / IT Innovation Business Division

Institute for New Economic Thinking, University of Oxford, UK (February 2016 — May 2016)

- Position: Academic Visitor

KAIST Natural Science Research Institute, Republic of Korea (August 2015 — August 2016)

- Position: Post-Doctoral Researcher

Samsung Economic Research Institute, Republic of Korea (March 2012 — June 2012)

- Position: Research Assistant

CYRAM, Republic of Korea (January 2008 — June 2008)

- Position: Internship

Project

Major projects in which I played main role are as described below.

Quantifying Anomalous Sensor Time Series Pattern

- **Affiliation:** Samsung SDS
- **Project Site:** Semiconductor Laboratory (Hwasung), Semiconductor Division, Samsung Electronics
- **Period:** February 2018 — Present
- **Role**
 - Responsible for improving anomaly scores
 - Responsible for developing image search system and interactive visualization which shows sensor time series and anomalous score
- **Environment**
 - Hardware: VMware IaaS
 - Software: RHEL 7, GitLab, Greenplum Database 4.3.17, Python 2.6/3.6 (numpy, scipy, pandas, matplotlib, sanic, sanic-jinja2, jupyter, networkx, biopython), jQuery, Bootstrap 4, plotly.js

Anomalous Wafer Classifier with Convolutional Neural Network (CNN)

- **Affiliation:** Samsung SDS
- **Project Site:** Semiconductor Laboratory (Hwasung), Semiconductor Division, Samsung Electronics
- **Period:** July 2017 — December 2017
- **Role**
 - Responsible for training/testing notable CNNs: VGG, ResNet, Inception-ResNet-v2, Xception, NASNet
 - Responsible for developing wafer image search system
 - Responsible for Hadoop ecosystem management and ETL
- **Environment**
 - Hardware: On premise Dell/EMC GPU Cluster (280 Cores + NVIDIA P100 × 20)
 - Software: RHEL 7, GitLab, Ambari, Hive, Elastic Search, Python 2.7/3.6 (numpy, scipy, pandas, matplotlib, opencv-contrib-python, keras, TensorFlowOnSpark, sanic, sanic-jinja2, jupyter)

Custom R Package Maintenance

- **Affiliation:** Samsung SDS
- **Project Site:** Semiconductor Laboratory (Hwasung), Semiconductor Division, Samsung Electronics
- **Period:** June 2017 — July 2017
- **Role**
 - Responsible for maintaining a custom R package
- **Environment**
 - Software: R 3.3, R Studio

Forecasting Book Sales Volume

- **Affiliation:** Samsung SDS
- **Project Site:** YES24 (Top e-commerce company in Republic of Korea), Seoul
- **Period:** January 2017 — March 2017
- **Role**
 - Responsible for forecasting daily book sales volume (January 2012 — January 2017)
 - Responsible for developing interactive data visualization applications
- **Environment**
 - Software: Slack, R 3.3 (dplyr, forecast, lubridate, shiny, shinydashboard, DT, sparkline, plotly), R Studio, d3.js

Skill & Expertise

Language: English

- OPIc (Oral Proficiency Interview-Computer)
 - Intermediate High (The second highest level)
- British Council Language Center
 - Upper Intermediate Level

Specialized Area

- Data Analysis with C/C++, Python and R
 - Data Size Experienced: Entire Korean web graph (including 0.1 billion web pages and 13 billion links), the U.S. patent records (spanning over 2 centuries and including 10 million patents), and so on. (See the publication list below).
 - Data Type Experienced: image, time series, text, graph/network structure
 - Data Context Experienced: social, economic, artistic, historical, technological
- Data Modeling with Deep Neural Networks: keras, TensorFlow, PyTorch, MXNet
- Interactive Data Visualization with d3.js and processing.js
- Numerical Simulation with C/C++, Intel® Math Kernel Library, and Intel® Performance Primitives
- OS: Linux
- Programming Language: Python (Advanced), R (upper intermediate), C/C++ (upper intermediate), JavaScript (intermediate), SQL (intermediate), Bash (intermediate)

Publication

1. Francois Lafond and **Daniel Kim**. "Long-Run Dynamics of the U.S. Patent Classification System" Submitted (2017). Available at SSRN: <https://ssrn.com/abstract=2924387>
 - **Role: data analysis and interactive visualization**
2. Byunghwee Lee, **Daniel Kim**, Hawoong Jeong, Seunghye Sun, Juyong Park. "Understanding the Historic Emergence of Diversity in Painting via Color Contrast" Submitted (2017). Available at ArXiv: <https://arxiv.org/abs/1701.07164>
 - **Role: designing research, code optimization (C++, Intel Performance Primitives), data analysis**
3. Byunghwee Lee, **Daniel Kim**, Dongwoo Kim, Hawoong Jeong. "N-gram Web Service and Stylometric Analysis on Korean Historical Documents" *New Physics: Sae Mulli* **66**, 4 (2016).
 - **Role: designing research, web-crawling, data analysis, interpretation, providing implications**
4. **Daniel Kim**, Daniel Burkhardt Cerigo, Hyejin Youn, Hawoong Jeong. "Technological novelty profile and invention' s future impact" *EPJ Data Science* **5**:8 (2016).
 - **Role: designing research, analyze data, interpretation**
5. **Daniel Kim**, Seung-Woo Son, Hawoong Jeong. "Large-Scale Quantitative Analysis of Painting Arts" *Scientific Reports* **4**, 7370 (2014).
 - **Role: designing research, data analysis, interpretation**
6. Kihong Chung, Yongjoo Baek, **Daniel Kim**, Meesoon Ha, Hawoong Jeong. "Generalized epidemic process on modular networks" *Phys. Rev. E* **89**, 052811 (2014).
 - **Role: code optimization, numerical simulation**
7. Jungjoo Seo, Jinil Kim, Eunsang Kim, **Daniel Kim**, Hawoong Jeong, Sung-Ryul Kim, Kunsoo Park. "Link Structure Analysis of Korean Web Graph" *Journal of KIISE : Computing Practices and Letters* **19**, 7 (2013).
 - **Role: web graph structure analysis**
8. Yongjoo Baek, **Daniel Kim**, Meesoon Ha, Hawoong Jeong. "Fundamental Structural Constraint of Random Scale-Free Networks" *Phys. Rev. Lett.* **109**, 118701 (2012).
 - **Role: numerical simulation, code optimization**
9. **Daniel Kim**, Seung-Woo Son, Hawoong Jeong. "Demographic Studies of Internet Routers" *J. Korean Phys. Soc.* **60**, 585 (2012).
 - **Role: design research, data analysis**
10. Sue Moon, Jinyoung You, Haewoon Kwak, **Daniel Kim** and Hawoong Jeong. "Understanding Topological Mesoscale Features in Community Mining" *COMSNETS 2010* Bangalore, India, January 2010 (2010). (Peer-reviewed)
 - **Role: data analysis**

Media Coverage

1. The paper D. Kim, D. B. Cerigo, H. Jeong, H. Youn, "Technological novelty profile and invention' s future impact" [EPJ Data Science 5:8 (2016)] was introduced in [SpringerOpen blog](#) and [Santa Fe Institute news](#).
2. The paper D. Kim, S.-W. Son, H. Jeong, "Large-scale quantitative analysis of painting arts" , [Sci. Rep. 4, 7370 (2014)] was posted on [Nature.com main page](#), being selected as a Nature Research Highlight (also held the first rank of the most viewed research highlight on Nature.com from 14 December 2014 to 14 January 2015). Subsequently, it was introduced and covered in [PNAS](#), [Daily Mail Online](#), [Business Insider Australia](#), 5 Korean TV news stations (Yeonhap news TV, YTN, KBS, MBC, and TJB) and 18 Korean major newspapers. [Full List](#)
3. In 2012, EBS, a Korean educational TV channel, conducted a questionnaire survey at a Korean middle school in Seoul, and my PhD supervisor and I were involved in analysis of its survey data. The result of my analysis was on TV on 24th April 2012.

Award

1. 3rd grade award, BK21+ graduate student essay contest for overseas visits (July 21, 2015)
2. Best oral presentation award, Korean Physical Society Meeting (Daejeon Convention Center, Daejeon, Korea, April 22-24, 2015)
3. Best oral presentation award, Korean Physical Society Meeting (Phoenix Park, Pyeongchang, Gangwon-do, Korea, October 24-26, 2012)
4. Best team project award, the 3rd KIAS CAC Summer School on Parallel Computing (KIAS, Korea, June 26-28, 2012)
5. Best poster award, Korean Physical Society Meeting (Daejeon Convention Center, Daejeon, Korea, April 24-26, 2012)
6. Best oral presentation award, Korean Physical Society Meeting (Daejeon Convention Center, Daejeon, Korea, April 13-15, 2011)
7. Best poster award, Korean Physical Society Meeting (Daejeon Convention Center, Daejeon, Korea, April 21-23, 2010)
8. Best poster award, the 3rd BK21 Young Physicists Workshop (Seoul National University, Seoul, Korea, January 15-16, 2010)

Honor

- Graduate Fellowship, Santa Fe Institute, USA (October 2013 — April 2016)
- Graduate Student Research Scholarship, KOSEF, Republic of Korea (September 2008 — August 2009)
- Department Honor Scholarship, KAIST, Republic of Korea (March 2006 — November 2007)

Others

- Data Visualizations: [click here and move to my homepage](#)
- Military Service: Republic of Korea Army (November 2010 — September 2014)

Reference

Hawoong Jeong, Ph.D. (KAIST Chair Professor, Department of Physics — Ph.D. supervisor)

- E-mail: hjeong@kaist.edu
- Mobile: 042-350-2543