

Daniel Youngho Kim, PhD



daniel.youngho.kim@gmail.com



+82-10-4195-7436



<http://danielykim.me/>



<https://www.linkedin.com/in/danielyounghokim/>

Summary

Full stack data scientist with 14+ years of experience:

- ✓ 4+ years of data science in MakinaRocks, MSD and Samsung SDS
- ✓ 10 years of data science in academia
- ✓ 13 academic papers in total including one highlighted by Nature.com main page
- ✓ 11 awards in academia and business
- ✓ 1 textbook translation and 1 book chapter
- ✓ Advanced Level of English at the British Council

Education

2008-09 – 2015-08	<u>Ph.D. candidate in Physics</u> Korea Advanced Institute of Science and Technology, Korea Supervisor: Prof. Hawoong Jeong
2013-11 – 2014-06	<u>Visiting Research Assistant</u> Institute for New Economic Thinking, University of Oxford, UK
2004-03 – 2008-08	<u>B.S. candidate in Physics</u> Korea Advanced Institute of Science and Technology, Korea Supervisor: Prof. Hawoong Jeong

Professional Experience

2020-09 – Present	<u>Principal Data Scientist</u> MakinaRocks, Seoul, Korea MakinaRocks, Sunnyvale, California, USA (2020-10 – 2020-12)
2019-12 – 2020-09	<u>Lead Data Scientist</u> Merck Sharp & Dohme, Seoul, Korea
2019-06 – 2019-11	<u>Senior Data Scientist</u> Merck Sharp & Dohme, Seoul, Korea
2016-09 – 2019-06	<u>Data Scientist / Senior Consultant</u> Samsung SDS, Seoul, Korea
2016-02 – 2016-05	<u>Academic Visitor</u>

2015-08 – 2016-08 Institute for New Economic Thinking, University of Oxford, UK
Post-Doctoral Researcher
 KAIST Natural Science Research Institute, Daejeon, Korea

2012-03 – 2012-06 Research Assistant
 Samsung Economic Research Institute, Seoul, Korea

2008-01 – 2008-06 Internship
 CYRAM, Seoul, Korea

Project

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

2020-02 – 2020-09 Affiliation: MSD Korea (In-house project)
 ✓ Performed data driven resource reallocation and territory assignment

2020-02 – 2020-09 Affiliation: MSD Korea (In-house project)
 ✓ Designed and implemented data manipulation pipelines

2020-05 – 2020-09 Affiliation: MSD Korea (In-house project)
 ✓ Developed a drug upper-bound price dashboard for strategic bidding

2020-02 – 2020-03 Affiliation: MSD Korea (In-house project)
 ✓ Achieved a 30M KRW cost savings as a result of systematic big data file management

2019-12 – Present Affiliation: MSD Korea
 ✓ Led data science team and managed two juniors
 ✓ Engaged with business unit directors, marketers, and sales managers

2019-11 Affiliation: MSD Korea (In-house project)
 ✓ Accelerated marketed-related tasks by 10 times

2019-10 – 2020-03 Affiliation: MSD Korea (In-house project)
 ✓ Created an AI bot to automatically generate videos with subtitles which saved 0.5B KRW of costs

2019-10 – 2019-11 Affiliation: MSD Korea (In-house project)
 ✓ Developed two self-service dashboards to reduce the failure rate of clinical trials

2019-09 Affiliation: MSD Korea (In-house project)
 ✓ Combined the following Korean public data: house price, population above an age, hospital features, national health insurance, and sales

2019-08 Affiliation: MSD Korea (In-house project)
 ✓ Optimized marketing activities for resource allocation

2019-07 – 2019-08 Affiliation: MSD Korea (In-house project)
 ✓ Performed supply chain optimization
 ✓ Designed and implemented interactive data visualization

2019-07 Affiliation: MSD Korea (In-house project)
 ✓ Accelerated HR tasks by 100 times

2019-06 – Present Affiliation: MSD Korea
 ✓ Led data science projects for a wide range of in-house clients: directors, sales managers and marketers
 ✓ End-to-end responsibility of project execution, understanding the client problems, management, allocation of resources, design and implementation of the solution

2018-02 – 2019-06 Affiliation: Samsung SDS

Project Site: Semiconductor Division, Samsung Electronics

Role

- ✓ Led a data science team for clients in the semi-conductor field
- ✓ Implemented data science projects
- ✓ End-to-end responsibility of understanding the client problems, project execution, management, allocation of resources, design and implementation of the solution.
- ✓ Thought leader on technical topics such as deep learning
- ✓ Achievement: Accelerated legacy distributed anomaly detection algorithms (×1000)

2017-07 – 2017-12

Affiliation: Samsung SDS

Project Site: Semiconductor Laboratory, Semiconductor Division, Samsung Electronics,

Role

- ✓ Implemented data science projects.
- ✓ End-to-end responsibility of understanding the client problems, project execution, management, and implementation of the solution.
- ✓ Thought leader on technical topics such as deep learning
- ✓ Achievement: Improved anomaly image classification accuracy (~97%)

2017-01 – 2017-03

Affiliation: Samsung SDS

Project Site: YES24 (Top e-commerce company in Republic of Korea), Seoul

Role

- ✓ Implemented data science project.
- ✓ Designed and created interactive data visualization applications
- ✓ Achievement: Improved sales forecasting accuracy (~90%)

Technical Expertise

Languages

1. Korean: native speaker
2. English
 - ✓ Advanced Level of English at the British Council
 - ✓ OPIc (Oral Proficiency Interview-Computer): Intermediate High (second highest level of 9 levels)

Specialized Skills

- ✓ Data Analysis with Python, R, C/C++ & Julia.
 1. Data Size: The whole 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).
 2. Data Type: Image, audio, time series, text & graph/network structure
 3. Data Context: Social, economic, artistic, historical, technological, semiconductor manufacturing, and pharmaceutical (commercial & marketing)
- ✓ Deep Learning Frameworks: keras (2017-09 – 2019-05), TensorFlow (2017-10 – 2017-12) & MXNet (2017-04 – 2017-06)
- ✓ Interactive Data Visualization with d3.js (2013 ~ Present) and processing.js (2010 ~ 2016)
- ✓ Numerical Simulation with C/C++, Intel® Math Kernel Library & Intel® Performance Primitives
- ✓ OS: Linux (Cent OS, RHEL & Ubuntu)
- ✓ Programming Languages
 - Python (2010 – Present: Data analysis, ML/DL, & back-end)

- C/C++ (2009 – 2016: Intermediate: numerical simulation)
- R (2015 – 2017: Intermediate: data manipulation & time series forecasting)
- JavaScript (2013 – Present: Intermediate: Interactive visualization & Front-end)
- SQL (2018-02 – Present: Intermediate), Bash (2010 ~ Present: Intermediate)

Publication (Peer-reviewed)

1. Byunghwee Lee, Min Kyung Seo, Daniel Kim, In-seob Shin, Maximilian Schich, Hawoong Jeong, and Seung Kee Han "Dissecting landscape art history with information theory" Proc. Natl. Acad. Sci. U.S.A 117 (43) 26580-26590 (2020).
2. Dong-Kyum Kim, Byunghwee Lee, Daniel Kim, and Hawoong Jeong. "Multi-Label Classification of Historical Documents by Using Hierarchical Attention Networks" J. Korean Phys. Soc. 76, pages368-377 (2020).
3. Yongjoo Baek, Kihong Chung, Meesoon Ha, Hawoong Jeong, and Daniel Kim. "Role of hubs in outbreaks of cooperative contagions" Phys. Rev. E (R) 99, 020301(R) (2019).
4. Francois Lafond and Daniel Kim. "Long-run dynamics of the U.S. patent classification system" Journal of Evolutionary Economics (2019).
5. Byunghwee Lee, Daniel Kim, Hawoong Jeong, Seunghye Sun, Juyong Park. "Heterogeneity in chromatic distance in images and characterization of massive painting data set" PLoS ONE 13(9): e0204430 (2018).
6. 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).
7. Daniel Kim, Daniel Burkhardt Cerigo, Hyejin Youn, Hawoong Jeong. "Technological novelty profile and invention's future impact" EPJ Data Science, 5:8 (2016).
8. Daniel Kim, Seung-Woo Son, Hawoong Jeong. "Large-Scale Quantitative Analysis of Painting Arts" Scientific Reports, 4, 7370 (2014).
9. Kihong Chung, Yongjoo Baek, Daniel Kim, Meesoon Ha, Hawoong Jeong. "Generalized epidemic process on modular networks" Phys. Rev. E, 89, 052811 (2014).
10. 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).
11. Yongjoo Baek, Daniel Kim, Meesoon Ha, Hawoong Jeong. "Fundamental Structural Constraint of Random Scale-Free Networks" Phys. Rev. Lett., 109, 118701 (2012).
12. Daniel Kim, Seung-Woo Son, Hawoong Jeong. "Demographic Studies of Internet Routers" J. Korean Phys. Soc. 60, 585 (2012).
13. 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)

Awards

- ✓ Merck INSPIRE Recognition Program Applause Award (June, 2020)
- ✓ Merck INSPIRE Recognition Program Achieve Award (June, 2020)
- ✓ Merck INSPIRE Recognition Program Applause Award (May, 2020)
- ✓ Merck INSPIRE Recognition Program Applause Award (August, 2019)
- ✓ 3rd grade award, BK21+ graduate student essay contest for overseas visits (July 21, 2015)
- ✓ 3 Best oral presentation awards, Korean Physical Society Meeting (Spring 2015, Fall 2012 & Spring 2011)
- ✓ 2 Best poster awards, Korean Physical Society Meeting (Spring 2012 & Spring 2010)
- ✓ Best team project award, the 3rd KIAS CAC Summer School on Parallel Computing (KIAS, Korea, June 26-28, 2012)
- ✓ Best poster award, the 3rd BK21 Young Physicists Workshop (Seoul National University, Seoul, Korea, January 15-16, 2010)

Translation

Forecasting: Principles and Practice 2nd Ed. English to Korean.

Online Version: <https://otexts.com/fppkr/>

Book Chapter

[How to be a data scientist? \(Korean\)](#) Chapter 2

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.

Honor

2013-10 – 2016-04	Graduate Fellowship, Santa Fe Institute, USA
2008-09 – 2009-08	Graduate Student Research Scholarship, KOSEF, Korea
2006-03 – 2007-11	Department Honor Scholarship, KAIST, Korea

Others

- ✓ Military Service: Republic of Korea Army
(2010-11 – 2014-09)