

# Sundong Kim

---

## CONTACT INFORMATION

Data Science Group, Institute for Basic Science (IBS)  
55 Expo-ro, Yuseong-gu, Daejeon 34126, Republic of Korea  
✉ [sdkim0211@gmail.com](mailto:sdkim0211@gmail.com) ☎ +82-10-2999-1962 🏠 <http://seondong.github.io/>  
<https://scholar.google.com/citations?user=3yQIQfAAAAAJ>

## INTERESTS

Human-like AI (e.g., In these days, I delve into abstraction and reasoning and representation learning)  
Applied Data Mining (e.g., Most of my works are on predictive analytics, user modeling, active learning)  
Democratizing AI (e.g., 2-yr experience on supporting AI transformation of customs administrations)

## POSITIONS

**Young Scientist Fellow, Institute for Basic Science** **Sep 2021 – Present**

- Principal Investigator, Develop computational methods for tabular learning and reasoning.
- Grant: KRW 150 million (132,000 USD) per year, for 3 years

Senior Researcher – Institute for Basic Science Sep 2019 – Aug 2021

- Worked on fraud detection, predictive analytics, representation learning, and online learning.
- Led collaborations between IBS, World Customs Organization, and Korea Customs Service.

Research Intern – Microsoft Research Asia Fall 2018

- Worked on user embedding model using search queries and behavior data.

Coordinating TA – Viterbi iPodia Program 2013 – 2017

- Managed an international class between seven leading universities, worked five semesters.

Research Assistant – KAIST 2013 – 2016

- Deployed entity typing algorithms of the Exobrain, Korean knowledge base.

Consulting Intern – Deloitte Fall 2012

- Helped to renew pricing structure for IT security company.

## EDUCATION

Ph.D. in Knowledge Service Engineering, KAIST (*Advisor: Jae-Gil Lee*) 2015 – 2019

- Thesis: Customer Revisit Prediction Using In-Store Sensor Data

M.S. in Industrial & Systems Engineering, KAIST (*Advisor: Kyoung-Kuk Kim*) 2013 – 2015

- Thesis: Maximizing Influence over a Target User through Friend Recommendation

B.S. in Industrial & Systems Engineering, KAIST 2008 – 2013

- Exchange Student in TU Berlin (Spring 2012)
- Exchange Student in National University of Singapore (Fall 2010)

## PUBLICATION SUMMARY

Recommendation / Fraud Detection

- Customs Fraud Detection: KDD 2020 [8], TKDE 2022 [14], ICDMW 2021 [12], AAAI 2022 [13]
- Social Networks: ICDEW 2015 [1]

User Modeling

- Offline Behavior: ICDM 2018 [2], KAIS 2020 [3], PAKDD 2020 [4], KSC 2019 [O2]
- Online Behavior: ECML-PKDD 2020 [5]
- COVID-19: Frontiers 2021 [10], arXiv:2106.00192 [P1]

Representation Learning

- Spatio-Temporal: IC2S2 2021 [O3]
- Image: ECCV 2020 [9], CVPR 2021 [11]
- Product: KCC 2021 [O2], KAIA 2021 [P2]
- Federated Learning: ECCV 2022 [16]

Active Learning

- Image: MLJ 2020 [6], CIKM 2020 [7]
- Time Series: ICLR 2022 [15]

**PREPRINTS**

(\*: undergraduate interns, †: correspondance)

- [P2] “Explainable Product Classification for Customs”.  
**Korean Artificial Intelligence Association, 2021.**  
<https://arxiv.org/abs/2111.01663>  
*Extended version: Under review at ACM TIST.*
- [P1] “Response to COVID-19 with Probabilistic Programming”.  
**arXiv:2106.00192. IC2S2, 2021.**  
<https://arxiv.org/abs/2106.00192>  
*Extended version: Under revision at ACM TSAS.*

**PUBLICATIONS**

- [16] Sungwon Han, Sungwon Park, Fangzhao Wu, Sundong Kim, Chuhan Wu, Xing Xie, Meeyoung Cha, “FedX: Unsupervised Federated Learning with Cross Knowledge Distillation”.  
**ECCV, 2022, (Acceptance rate: 28%).**
- [15] Yooju Shin, Susik Yoon, Sundong Kim, Hwanjun Song, Jae-Gil Lee, Byung Suk Lee, “Coherence-based Label Propagation over Time series for Accelerated Active Learning”.  
**ICLR, 2022, (Acceptance rate: 32%).**  
<https://openreview.net/forum?id=gjNcH0hj0LM>
- [14] Sundong Kim, Tung-Duong Mai\*, Thi Nguyen Duc Khanh\*, Sungwon Han, Sungwon Park, Karandeep Singh, Meeyoung Cha, “Active Learning for Human-in-the-Loop Customs Inspection”.  
**IEEE Transactions on Knowledge and Data Engineering, 2022, (SCI, IF=6.977)**  
<https://doi.org/10.1109/TKDE.2022.3144299>
- [13] Sungwon Park, Sundong Kim†, Meeyoung Cha†, “Knowledge Sharing via Domain Adaptation in Customs Fraud Detection”.  
**AAAI, 2022 (Special Track on AI for Social Impact, Acceptance rate: 15%).**  
<https://arxiv.org/abs/2201.06759>
- [12] Tung-Duong Mai\*, Kien Hoang\*, Aitolkyn Baigutanova\*, Gaukhartas Alina\*, Sundong Kim†, “Customs Fraud Detection in the Presence of Concept Drift”.  
**ICDM, 2021, (IncrLearn Workshop).**  
<https://doi.org/10.1109/ICDMW53433.2021.00052>
- [11] Sungwon Park, Sungwon Han, Sundong Kim, Danu Kim, Sungkyu Park, Seunghoon Hong, Meeyoung Cha, “Improving Unsupervised Image Clustering With Robust Learning”.  
**CVPR, 2021, (Acceptance rate: 22%).**  
<https://doi.org/10.1109/CVPR46437.2021.01210>
- [10] Yuan Yuan\*, Muzhi Guan, Zhilun Zhou, Sundong Kim†, Meeyoung Cha, Depeng Jin, Yong Li†, “Disruption in the Chinese E-Commerce During COVID-19”.  
**Frontiers in Computer Science - Digital Public Health, 2021.**  
<https://doi.org/10.3389/fcomp.2021.668711>
- [9] Sungwon Han, Sungwon Park, Sungkyu Park, Sundong Kim, Meeyoung Cha, “Mitigating Embedding and Class Assignment Mismatch in Unsupervised Image Classification”.  
**ECCV, 2020 (Acceptance rate: 27%).**  
[https://doi.org/10.1007/978-3-030-58586-0\\_45](https://doi.org/10.1007/978-3-030-58586-0_45)
- [8] Sundong Kim, Yu-Che Tsai, Karandeep Singh, Yeonsoo Choi, Etim Ibok, Cheng-Te Li, Meeyoung Cha, “DATE: Dual Attentive Tree-aware Embedding for Customs Fraud Detection”.  
**KDD, 2020 (Applied Data Science, Acceptance rate: 16.0%).**  
<https://doi.org/10.1145/3394486.3403339>
- [7] Hwanjun Song, Minseok Kim, Sundong Kim, Jae-Gil Lee, “Carpe Diem, Seize the Samples Uncertain at the Moment for Adaptive Batch Selection”.  
**CIKM, 2020 (Acceptance rate: 21%).**  
<https://doi.org/10.1145/3340531.3411898>
- [6] Hwanjun Song, Sundong Kim, Minseok Kim, Jae-Gil Lee, “Ada-Boundary: Accelerating the DNN Training via Adaptive Boundary Batch Selection”.  
**Machine Learning, 109, 1837–1853, 2020 (SCI, IF=2.730, ECML-PKDD Journal track).**  
<https://doi.org/10.1007/s10994-020-05903-6>

- [5] Mingxiao An, Sundong Kim<sup>†</sup>, “Neural User Embedding from Browsing Events”. **ECML-PKDD, 2020** (*Applied Data Science, Acceptance rate: 27%*).  
[https://doi.org/10.1007/978-3-030-67667-4\\_11](https://doi.org/10.1007/978-3-030-67667-4_11)
- [4] Sundong Kim, Hwanjun Song, Sejin Kim, Beomyoung Kim, Jae-Gil Lee, “Revisit Prediction by Deep Survival Analysis”. **PAKDD, 2020** (*Acceptance rate: 21%*).  
[https://doi.org/10.1007/978-3-030-47436-2\\_39](https://doi.org/10.1007/978-3-030-47436-2_39)
- [3] Sundong Kim, Jae-Gil Lee, “A Systematic Framework of Predicting Customer Revisit with In-Store Sensors”. **Knowledge and Information Systems, 62(3), 1005–1035, 2020** (SCIE, IF=2.397)  
<https://doi.org/10.1007/s10115-019-01373-y>
- [2] Sundong Kim, Jae-Gil Lee, “Utilizing In-Store Sensors for Revisit Prediction”. **ICDM, 2018** (*Acceptance rate: 9%*), **Selected as One of the Best Papers**.  
<https://doi.org/10.1109/ICDM.2018.00037>
- [1] Sundong Kim, “Friend Recommendation with a Target User in Social Networking Services”. **ICDE, 2015** (*Ph.D. Symposium*).  
<https://doi.org/10.1109/ICDEW.2015.7129582>

## OTHERS

- [O3] Sundong Kim, “Embedding Heterogeneous Hierarchical Structures”. **IC2S2, 2021** (*Extended Abstract*).  
<https://github.com/seondong/locemb>
- [O2] Sihyeon Kim\*, Eunji Lee, Sungwon Park, Meeyoung Cha, Sundong Kim<sup>†</sup>, “Hyperbolic Interaction Model for Category Classification in E-Commerce”. **Korea Computer Congress & BIEN, 2021**.  
[http://seondong.github.io/assets/papers/2021\\_KSC\\_HyperIM.pdf](http://seondong.github.io/assets/papers/2021_KSC_HyperIM.pdf)
- [O1] Beomyoung Kim\*, Sundong Kim, Sejin Kim, Jae-Gil Lee, “Customer Revisit Prediction Using Macroscale Mobility Information”. **Korea Software Congress, 2019**.  
[https://seondong.github.io/assets/papers/2019\\_KSC\\_Macroscale.pdf](https://seondong.github.io/assets/papers/2019_KSC_Macroscale.pdf)

## SERVICES

### Program Committee

- The Web Conference (WWW), 2020–2022
- International Conference on Machine Learning (ICML) 2020–2022
- Association for the Advancement of Artificial Intelligence (AAAI) 2021–2022
- International Conference on Learning Representations (ICLR) 2021–2022
- ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD) 2022
- ACM International Web Search and Data Mining Conference (WSDM) 2022
- The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD) 2020
- Annual Meeting of the Association for Computational Linguistics (ACL) 2021
- Conference on Neural Information Processing Systems (NeurIPS) 2021

### External Reviewer

- IEEE Transactions on Knowledge and Data Engineering (TKDE)
- IEEE Transactions on Computational Social Systems (TCSS)
- Pervasive and Mobile Computing (PMC)
- ACM Transactions on Intelligent Systems and Technology (TIST)
- Data & Knowledge Engineering (DKE)
- IEEE Access
- The Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)
- The International AAAI Conference on Web and Social Media (ICWSM)

### Organizations:

- World Customs Organization & Korea Customs Service

- Contributed talks: 16th Annual PICARD Conference, 8th Working Group on Revenue Compliance, BACUDA Experts Meeting, Knowledge Academy for Customs and Trade, AMS Regional Seminar on Data Analytics, SACU Trade Facilitation Webinar series

Events:

- IBS-NCKU Joint Seminar on Predictive Analytics
- IBS Data Science Talk Series

**PATENTS**

- [6] Sundong Kim, Eunji Lee, Sihyeon Kim, Meeyoung Cha, “Explainable Classification Using Text Descriptions”. Korean Patent Application No: 10-2022-0034291, Mar 18, 2022.
- [5] Sungwon Park, Sundong Kim, Meeyoung Cha, “Fraud Detection Apparatus and Method Thereof”. Korean Patent Application No: 10-2022-0002535, Jan 7, 2022.
- [4] Sundong Kim, Tung-Duong Mai, Sungwon Han, Sungwon Park, Meeyoung Cha, “Apparatus and Method for Selecting Items for Customs Inspection”. Korean Patent Application No: 10-2021-0075070, Jun 9, 2021.
- [3] Sungwon Han, Sundong Kim, Sungkyu Park, Sungwon Park, Meeyoung Cha, “Image Analysis System and Operating Method of the Same”. Korean Patent Application No: 10-2020-0088732, Jul 17, 2020, PCT No: 17/375650, Jul 14, 2021.
- [2] Jae-Gil Lee, Howon Moon, Hwanjun Song, Minseok Kim, Sundong Kim, “System and Method for Accelerating DNNs Training via Adaptive Batch Selection”. Korean Patent Application No: 10-2020-0044159, Apr 10, 2020.
- [1] Jae-Gil Lee, Sundong Kim, “Type Inference Algorithm for Knowledge Base”. Korean Patent Application No: 10-2017-0004663, Jan 12, 2017.

**AWARDS & HONORS**

- Young Scientist Fellowship, Institute for Basic Science 2021 – Present
  - Project Name: Lifelong Tabular Learning
  - KRW 150 million (132,000 USD) per year, for 3 years.
- Outstanding Research Award, Institute for Basic Science 2021
- One of the Best Papers, IEEE ICDM 2018 2018
- Outstanding Research Award, Dept. of KSE/IsysE, KAIST 2018
- Star of Tomorrow Award, Microsoft Research Asia 2018
- National Scholarship for Graduate Studies, Korea Student Aid Foundation 2013 – 2019
- National Scholarship for Undergraduate Studies, Korea Student Aid Foundation 2008 – 2012

**LANGUAGES**

- Korean: ILR Level 5 – Native
- English: ILR Level 4 – Full Professional Proficiency
- Chinese: ILR Level 1 – Elementary Proficiency

**REFERENCES**

- Prof. Jae-Gil Lee** (*Doctoral Advisor*)  
Associate Professor, KAIST (jaegil@kaist.ac.kr)
- Prof. Meeyoung Cha** (*Postdoctoral Advisor*)  
Chief Investigator, Institute for Basic Science (mcha@ibs.re.kr)
- Dr. Hwanjun Song** (*Collaborator*)  
Naver AI Lab (hwanjun.song@navercorp.com)
- Mr. Yeonsoo Choi** (*Collaborator*)  
Head of ICT and Data Planning Division, Korea Customs Service (yschoi0817@korea.kr)
- Dr. Xing Xie** (*Internship Advisor*)  
Senior Principal Research Manager, Microsoft Research Asia (xingx@microsoft.com)
- Prof. Kyoung-Kuk Kim** (*M.S. Advisor*)  
Associate Professor, KAIST (catenoid@kaist.ac.kr)

[Last updated on 2022-07-09 ]