The Symposium on Spatiotemporal Data Science

Co-Sponsors:

- ☐ Center for Geographic Analysis at Harvard University
- ☐ Future Data Lab
- ☐ China Data Institute
- ☐ GeoComputation Center for Social Sciences at Wuhan University
- ☐ KNIME, Inc.

July 15-16 2023

1730 Cambridge Street, Cambridge MA 02138

Supported by the Spatial Data Lab project
Harvard CGA

Conference Website:

https://projects.iq.harvard.edu/chinadatalab/event/symposium-spatiotemporal-data-science

Conference Registration: https://www.eventbrite.com/e/582353644427

Contact:

Spatialdatalab@lists.fas.harvard.edu

Keynotes Speakers

Daniel Sui
Vice President for Research and Innovation
Virginia Tech

Stepping into the Same River Twice: Whither Reproducibility and Replicability in Spatiotemporal Data Science??

Daniel Sui is an internationally renowned geographer/GIScientist with research interests focusing on GIS-based spatial analysis and modeling for urban, environmental, and public health applications, was appointed Virginia Tech's vice president for research and innovation in the fall of 2020.

Stefano M. Iacus
Director of Data Science and Product Research
Harvard University

Forecasting Asylum-related Migration Flows with Machine Learning and Data at Scale

Stefano M. Iacus started his academic career at the University of Milan, founded the Data Science Lab and served as its director. Since 2019, he has also served at the Joint Research Centre of the European Commission. In 2012, he co-founded a data science firm, Voices from the Blogs, that specialized in social media analysis and big data analytics. In 2017, he co-founded a second startup in the field of quantitative finance.

The Symposium on Spatiotemporal Data Science

Date: July 15 (onsite and online) - July 16 (online only), 2023
Location: 1730 Cambridge Street, Cambridge MA 02138

DAY 1 - Saturday, July 15, 2023, Plenary Sessions, Room CGIS South S020	
8:30 AM	Registration and continental breakfast
9:00 AM	**Welcome Address** Chaowei Phil Yang (George Mason University) Facilitator: Mengxi Zhang (Virginia Tech)
9:20 AM	**Keynote: Stepping into the Same River Twice: Whither Reproducibility and Replicability in Spatiotemporal Data Science?** Daniel Sui (Virginia Tech) Facilitator: Chaowei Phil Yang (George Mason University)
9:55 AM	**Keynote: Forecasting Asylum-related Migration Flows with Machine Learning and Data at Scale** Stefano M. Iacus (Harvard University) Facilitator: Wendy Guan (Harvard University)
10:30 AM	Coffee Break
10:50 AM	**Panel 1: Challenges and Breakthroughs in R&R Spatiotemporal Data Science** Moderator: Joseph Holler (Middlebury College) **Recent Advancements of the Spatial Data Lab Project** Wendy Guan (Harvard Center for Geographic Analysis) **Open Science Practices for Replicating And Extending Spatiotemporal Data Science Research** Peter Kedron (Arizona State University) **Challenges and Breakthroughs in Replicable and Expandable Spatiotemporal Data Science** Joseph Holler (Middlebury College) **KNIME Analytics Platform for Reproducible Data Science** Zachary Durso (KNIME)
12:00 PM	Lunch Break
DAY 1- Saturday, July 15, 2023, Parallel Track I, Room CGIS South S020	
1:00 PM	**Panel 2: Methodologies I - Innovations in Geospatial Analysis** Ting Zhang (University of Baltimore) **Measuring Spatiotemporal Autocorrelation Using Distance Covariance**

	Carlos Baez (University of California, Santa Barbara) **CitySchema.org:A Framework for Cross-Disciplinary, Cross-Temporal Collaboration on City Models** Paul Cote (CitySchema) **Location Factors and Ecosystem Embedding of Sustainability Engaged Blockchain Companies in the US: A Web-based Analysis** Robert Dehghan, Sebastian Schmidt (ISTARI.AI) **Harmonizing multi-sourced multi-scale geospatial data for depression studies: potentials and challenges** Eunhye Yoo (University at Buffalo, SUNY)
2:20 PM	Coffee Break
2:50 PM	**Panel 3: Methodologies II - Advancements in Spatiotemporal Data Science** Moderator: Xintao Liu (Hong Kong Polytech University) **Enabling reproducible spatiotemporal data science: Practice and Challenges** Francisco Rowe (University of Liverpool) **Developing task-driven GeoDataGPT: A powerful paradigm for smart data governance in Hong Kong's Common Spatial Data Infrastructure** Xintao Liu, Zidong Yu (Hong Kong Polytech University) **Geo-social media: A data source for multifaceted spatiotemporal information** Bernd Resch, Sebastian Schmidt (University of Salzburg) **Geospatial grid region name model and experimental system combined with HPC** Daoye Zhu (Harvard University)
DAY 1- Saturday, July 15, 2023, Parallel Track II, Room CGIS South S030	
1:00 PM	**Panel 4: Applications I- Investigating Health Disparities and Accessibility** Moderator: Junghwan Kim (Virginia Tech) **Extreme Weather and Unequal Health Costs: Evidence from Hospital Admission** Yuhang Pan (Peking University) **How do GIS and Public Health Converge in Healthcare Accessibility?** Changzhen Wang (University of Alabama); Mengxi Zhang (Virginia Tech) **Exploring the geographic biases in how ChatGPT introduces environmental justice issues in the United States: A case study of 100 cities** Junghwan Kim (Virginia Tech) **Disparities in Geographic Accessibility of Eating Disorder Treatment** Chuying Huo (Harvard Center for Geographic Analysis)
2:20 PM	Coffee Break

2:50 PM	**Panel 5: Applications II- Urban Inequality and Environmental Impacts** Moderator: Changzhen Wang (University of Alabama) **Heterogeneous effect of extreme heat on employment in Sub-Saharan Africa** Fangyuan PENG (Hong Kong University of Science and Technology) **Forest change and its carbon storage research in China from 1990 to 2021** Zhen Wu (Harvard University) **Poverty Reduction through climate change adaptation in Southeast Asia** Akiyuki Kawasaki (University of Tokyo) **Refining 2SVCA Method for Measuring Telehealth Accessibility of Primary Care Physicians in Baton Rouge** Lingbo Liu(Harvard University)
DAY 1- Saturday, July 15, 2023, Parallel Track III, Room CGIS South S050	
1:00 PM	**Panel 6: Applications III- Urban Dynamics** Moderator: Peter Kedron (Arizona State University) **Dynamic mapping of access to banking in Wales: A path for closing the banking desserts** Sonea Andra (University of Warwick) **A Comparative Study of Methods for Drive Time Estimation on Big Geospatial Data: A Case Study in the U.S** Xiaokang Fu (Harvard University) **Using Remotely Sensed Data in a Machine Learning Framework as a Tool for Forecasting Wildfires** James Zollweg (SUNY at Brockport) **You Can't Hide From Me: Spatiotemporal Data Analysis and Visualization** Hawk Arachy (AForge)
2:20 PM	Coffee Break
2:50 PM	**Panel 7: Applications IV- Social Development** Moderator: Mengxi Zhang (Virginia Tech) **Sophie and the locust curse: Effects of locust plague on human capital accumulations** Pak Hung LAM (Hong Kong Uni of Science and Technology / Yale University) **What contributes to a gender-inclusive park? Understanding female visitor preferences using large-scale social media data** Tianyu Su (Harvard Graduate School of Design) **Living with strangers: housing preferences of professional migrant women in the shared rental market in Nanjing, China** Weiyi Cao (Wageningen University)

	Beyond Suburbanization: Home Location Divergence for Remote and On-Site Workers Ting Zhang (University of Baltimore) Lu Chen (Southeast University,China) Minjia Guo (University of Glasgow)
DAY 2- Sunday, July 16, 2023, Plenary Sessions, Online Only, 9:00AM - 5:00PM	
9:00 AM	**Recap of Day 1 and Overview of Day 2** Mengxi Zhang (Virginia Tech)
9:05 AM	**Keynote:CyberGISX for Reproducible and Scalable Geospatial Data Science** Shaowen Wang (UIUC) Facilitator: Wendy Guan (Harvard Center for Geographic Analysis)
9:35 AM	**Panel 8: Methodologies III - GeoAI** Moderator: Xiao Huang (Emory University) **Development of Reproducible GeoAI Models** Song Gao (Uni of Wisconsin at Madison) **GeoAI in Human Geography** Xiao Huang (Emory University) / Siqin Wang (University of Queensland) **GeoAI Applications in Urban Informatics** Filip Biljecki (National University of Singapore) **Sensing the overlapped nature of mobility network using GeoAI** Di Zhu (University of Minnesota) **On the Opportunities and Challenges of Foundation Models for Geospatial Artificial Intelligence** Gengchen Mai (University of Georgia)
11:00 AM	Break
11:10 AM	**Panel 9: Methodologies IV - GeoAI for Health** Moderator: Mengxi Zhang and Daniel Sui (Virginia Tech) Panelists: Tao Hu (Oklahoma State University) Kai Zhang (University at Albany) Xun Shi (Dartmouth College) Xiao Huang (Emory University)
12:00 PM	Lunch Break
1:00 PM	**Panel 10: Methodologies V - New Frontiers in Urban Modeling** Moderator: Tao Hu (Oklahoma State University) **Digital mental health surveillance using social media data**

	Oliver Gruebner (University of Zurich) **Discrete, Continuous, and Hybrid Methods for Urban Ecological Gradient Analysis** Steven Richter (East Carolina University) **Data modeling with geographically weighted additive models: Cases from forestry and Economics** Slawomir Konopa (University of Agriculture in Krakow) **Utilizing Deep Learning and Google Street View to Analyze Flood Mitigation** Xinyue Ye (Texas A&M University)
1:50 PM	**Panel 11: Applications and Training in Social Sciences** Moderator: Xun Shi (Dartmouth College) **Spatially Integrated Social Sciences** Minxuan Lan (University of Toledo) **Measuring telehealth care accessibility: A case study of the Enhanced 2-Step Virtual Catchment Area (E2SVCA) MethodAF** Yaxiong Shao (Northern Illinois University) **COVID-19 spatiotemporal research with workflow-based data analysis** Srikar Chintala (UCLA) **COVID-19 impact on excess deaths of various causes in the United States** Yogya Kalra (Harvard Center for Geographic Analysis) **Spatial analysis of the influence of pet companionship on human emotions** Ru Wang / Shixuan Li (Wuhan University)
2:50 PM	Break
3:00 PM	**Panel 12: Applications in Human Development** Moderator: Yuhang Pan (Peking University) **Was 1931 Chinese flood the severest? a quantitative study based on model-derived spatiotemporal data changing historical perception** Chang Liu, Akiyuki Kawasaki, Tomoko Shiroyama (University of Tokyo) **Assessing Potable Water Access and its Implications on the Households' Socio-Economic Activities: The Case of Sibi in the Nkwanta North District-Ghana** Kanjin Kingsley (University of Toledo) **Analysis of the influence of built environment on the use modes of household life cycle: a case study of Shanghai** Baoxin Zhai (Northwest University, China)

	Spatial characterization of urban innovation networks constructed based on website hyperlinks: an evidence from China's technology firms Chenxi Liu / Shixuan Li (Wuhan University) **Calculation and Modeling of Historical Block Vitality Based on Behavior Trajectory Data: Taking the Chinese Baroque Historic District in Harbin as an Example** Haixuan Zhu/Zixuan Zhao/Xiaoyu Hou (HIT)
4:00 PM	**Panel 13: Applications in Urban Development** Moderator: Ruishan Chen (Shanghai Jiaotong University) **The Spatio-temporal Impact of Green Space on Life Expectancy at Birth across America Counties** Xiji Jiang (Southwest Jiaotong University) **Spatial-Climatic Finance: Impact of Weather and Airbnb Listings on House Prices** Nakul Gupta (Management Development Institute Gurgaon, India) **Spatio-temporal Visualization of Urban Population Using Spatial Big Data: A Case Study of Kyoto City during the COVID-19 Pandemic** Keiji Yano /Kazumasa Hanaoka / JamesTodd (Ritsumeikan University) **Urban Inequality in The World of Polycrisis: A Case Study of Shanghai and Beyond** Ruishan Chen (Shanghai Jiaotong University) **Computation and Modeling of Linear Historical Space Based on Trajectory Data: Taking Harbin Central Street as an Example** Haixuan Zhu/Cuiling Wu/Xiaoyu Hou (HIT) **Spatial Optimization towards Smart Cities** Kai Cao（East China Normal University）
5:10 PM	**Closing Remark** Wendy Guan and Shuming Bao

Call for Papers: Sustainable Geospatial Analytics and Geoinformatics with Repeatable, Reproducible, and Expandable (RRE) Framework and Design

A Special Issue of The International Journal of Applied Earth Observation and Geoinformation

In recent decades, GIScience has undergone significant development, leading us toward an era that emphasizes sustainability, sharing, repeatability, and reproducibility. Sustainable spatiotemporal data analysis is a critical aspect of understanding social and environmental processes and predicting their long-term outcomes. To achieve this, frameworks and designs that are repeatable, reproducible, and easily re-employed by a wide range of end-users, policymakers, and individuals without technical expertise are required. A sustainable approach to spatiotemporal data analysis involves considering the persistence and sharing of data collection, management, methodologies and workflows. This requires proper training and practice in data collection, including minimizing data errors and biases, using open data standards, and appropriately documenting the data and methods. Additionally, a repeatable and reproducible framework for data analysis ensures that results can be validated and replicated, thus increasing confidence in the findings. In essence, sustainable spatiotemporal data analytics promotes the cutting-edge frontier of human-centered open science, thereby advancing the field and leading us toward a more sustainable and equitable future.

As such, this special issue aims to promote the development of repeatable, reproducible, and expandable (RRE) frameworks, methodologies, and technologies (e.g., use of KNIME workflow) for spatial data analysis, spatial data sharing, and applied research in the fields of spatiotemporal innovation. It welcomes papers with the following topics:

- Review articles on the development of RRE frameworks, methodologies, and technologies
- Technical development of new RRE frameworks and methodologies centered around spatiotemporal data
- Empirical studies using spatiotemporal data as well as RRE frameworks and methodologies in the cross-subdomain of geography and sustainability science, including but not limited to:
- sustainable society
- urban sustainability
- environmental sustainability
- ecological sustainability

All submissions are required to have sharable data, codes, framework and methods, expected to be designated as a RRE workflow e.g., using KNIME, Model Builder, or other workflow software.

Spatial Data Lab Project Partners

Center for Geographic Analysis at Harvard University

The Center for Geographic Analysis (CGA) at Harvard University is a member organization of Harvard's Institute for Quantitative Social Science, and one of the research sites in the Spatiotemporal Innovation Center funded by the US National Science Foundation's Industry–University Cooperative Research Centers Program. Its core mission is to advance and support research and teaching in all disciplines across Harvard University with emerging geospatial technologies.

Future Data Lab

Future Data Lab (FDL) is designed for providing a new generation of data services with cutting-edge methodology and technology. The Future Data Lab is partnering with various academic and business partners on the following missions: spatial data services; development of tools for spatial data analysis; workflow-based spatial data case studies; and training programs for spatial data science.

China Data Institute

China Data Institute is one of the primary data providers for China data services to the world. Its missions are: to support research in the human and natural components of local, regional, and global changes; to promote quantitative research on China; to promote collaborative research in spatial studies; and to promote the use and sharing of China data in teaching and research worldwide.

GeoComputation Center for Social Sciences at Wuhan University

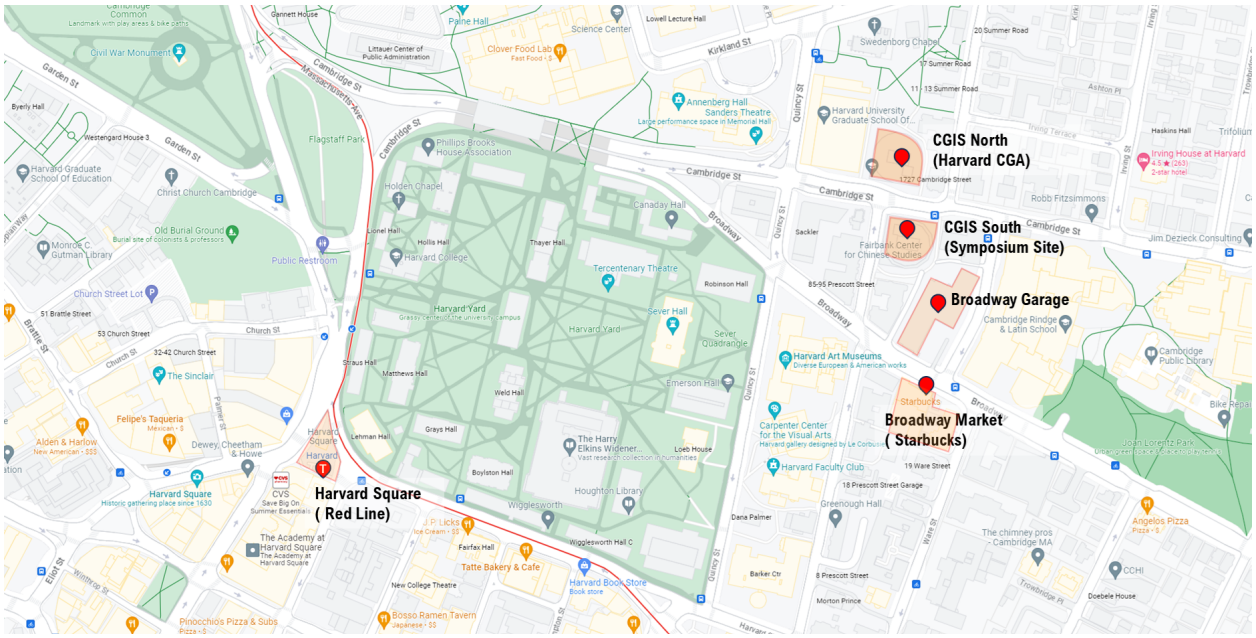
The GeoComputation Center for Social Sciences at Wuhan University is a joint effort by the State Key Lab of Information Engineering in Survey, Mapping and Remote Sensing at Wuhan University and the Center for Spatial Data Sciences at the University of Chicago. It promotes scientific research on the theory and method of spatial data in scientific research, workforce training, international cooperation, and social practice.

KNIME, Inc.

KNIME is an open source platform built to productionize data science from day 1 by a team with a strong scientific background. Today it has a strong user community of 300,000+ across all industries and in over 60 countries.

Campus Map

The location of Harvard CGA, symposium site, Harvard Square metro station (Red line), parking garage, and a nearby market (Broadway Market) are highlighted in red. Harvard Square T Station is the nearest metro station, participants can walk through the beautiful Harvard Yard to the symposium site. If you plan to drive here and need parking in Broadway Garagem, go to https://www.transportation.harvard.edu/parking/visitors for parking permit info.



Outside Access