

2023 
Happy New Year

ANNUAL REPORT 2022

SPATIAL DATA LAB

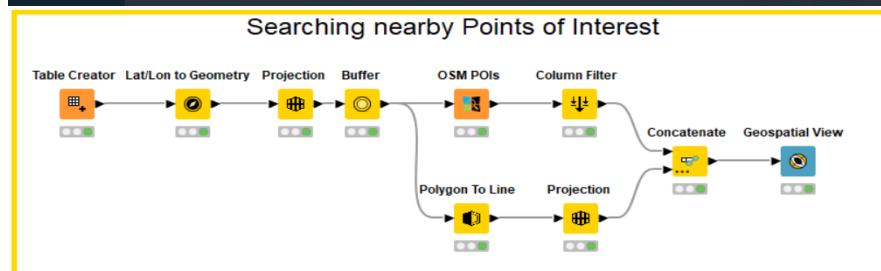
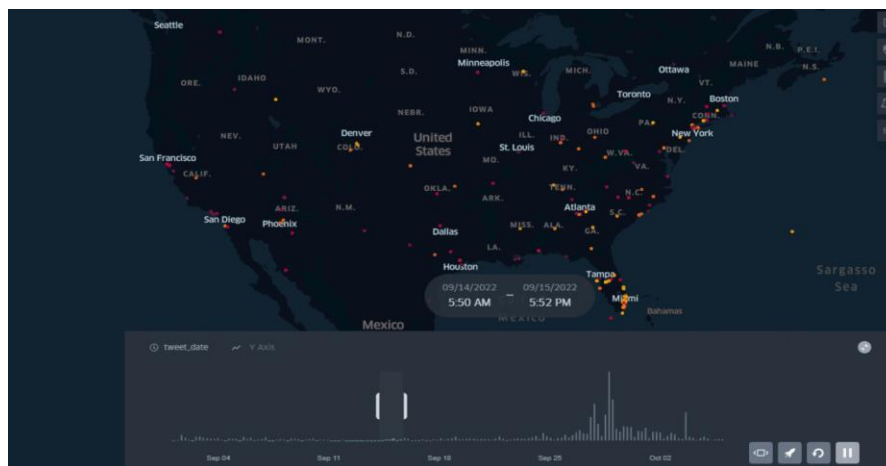
Following the objectives and work plans in our proposal, the Spatial Data Lab (SDL) project, which is an extension to the China Data Lab project, has made some good progress during 2022, including integrating and publishing datasets, developing open GIS tools and case studies, organizing training workshops, and building global collaborations.

DATA COLLECTIONS

The project team updates COVID-19 datasets on [Harvard Dataverse](#) monthly. As of December 20, 2021, global users from 150 countries have downloaded the COVID-19 datasets on [Harvard Dataverse](#) more than 1, 370, 000 times. The team also developed multiple tools based on workflows, facilitating users to access external datasets on Harvard Dataverse, GitHub, China Geo-Explorer database, and KNIME Hub.

Data Link: [China Data Lab Dataverse](#)
[The New Development of China Data Online](#)
[Harvard Dataverse Extension for KNIME](#)(Demo)

TOOL DEVELOPMENT



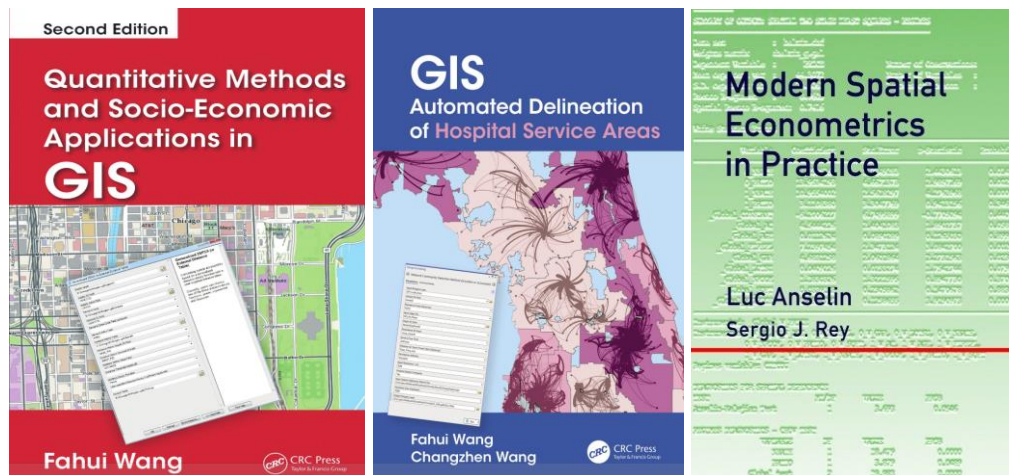
GEOSPATIAL ANALYTICS EXTENSION FOR KNIME

On December 6, 2022, KNIME released the 1st version of [Geospatial Analytics for KNIME](#), which was developed by the Spatial Data Lab as a joint effort by the Center for Geographic Analysis at Harvard, KNIME, and Future Data Lab, This new Geospatial Extension provides the ability to carry out geospatial analysis within KNIME. In its first version, it supports the most common vector data types such as points, lines, polygons, and collections of those. It is a trusted community contribution, and we will continue to improve it further.

In addition to the Geospatial Analytics Extension, we will publish many workflows for the most common geospatial use cases in the [KNIME Community Hub](#) so everyone including non-geospatial experts can hit the ground running.

WORKFLOW-BASED CASE STUDY DEVELOPMENT

The project team built 40+ new workflows based on the three classic GIS textbooks, which strongly support the development of Geospatial Analytics Extension for KNIME. More than 10 workflows have been transplanted onto the KNIME Web portal version. The collaborator can run the model via a web browser. The KNIME workbook for the Third Edition of Quantitative Methods and Socio-Economic Application in GIS will be published next year.

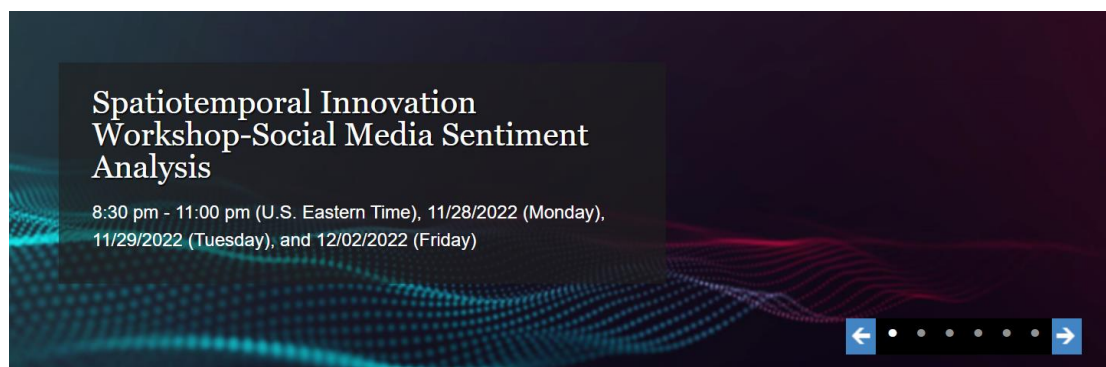


The project team also tested those case studies developed by the previous team, which were built on some peer-reviewed publications to help collaborating researchers easily and quickly conduct analysis, develop methodologies, and deliver education relevant to COVID-19, culture, and society, human mobility, and urban and rural development. More information is available on the [Paper Replication List](#).

WORKSHOPS AND PUBLIC TALKS

SPATIOTEMPORAL INNOVATION TRAINING WORKSHOPS

The team has hosted 1 webinar and 3 online workshops on “Spatiotemporal Innovation Analysis with KNIME” since July 2022. There were more than 270 participants which developed 5 collaborative ongoing research papers.



Spatiotemporal Innovation

[Spatiotemporal Simulation](#) | [Healthcare Accessibility](#) | [Social Media Sentiment Analysis](#)

PUBLIC TALK ON GEOSPATIAL ANALYTICS AT KNIME SUMMIT 2022

On November 14-16, the SDL team with KNIME partners introduced the Geospatial Analysis Extension and the reusable and reproducible blueprints that the Harvard CGA team and KNIME are working on, for users to download and adapt to their own needs. Historically, accessing and analyzing this invaluable data, however, has required niche expertise, demanding deep experience with a wide array of data sources and libraries, as well as coding skills.



[KNIME Summit 2022](#) | [Video Link](#)

NETWORK AND CAPACITY DEVELOPMENT

ESTABLISHMENT OF THE AFFILIATE LABS

To promote academic collaborations on geo-computation for humanity and social data science with the Spatial Data Lab project (SDL), the Future Data Lab (<http://futuredatalab.org>) is accepting applications for affiliate labs/centers. The Future Data Lab, in collaboration with the SDL project team, offers the following support for the affiliate labs, including leadership and technical training, shared case study database, research data support, technical support, promotion of the achievements by the affiliate labs, and recognition of the achievements by the affiliate labs on the Spatial Data Lab project website.

There have been 11 formal affiliate labs with group members all over the world, among which an open community for spatial data labs is growing fast.

[Application Link for Affiliate Lab](#)

RESEARCH CAPACITY DEVELOPMENT

The Spatial Data Lab/China Data Lab team has been collaborating with global scholars from different institutions, including scholars and students from Harvard University, George Mason University, University of Washington, University of South Carolina, Tulane University, the University of Queensland, Colorado State University, Ball State University, Georgia State University, TAMU, CUNY, Tsinghua University, Wuhan University, and many others, on COVID-19 data collection, analysis, and publications. Those teams (mobility, modeling, innovation, and visualization) have published more than 100 peer-reviewed papers. More information is available on the following shared document.

- [Publications](#)

RESEARCH GRANTS SUPPORT

The workflow-based open GIS community development helped grant applications and provided direct support for the following project:

Ongoing Project:

HVD-21-01 *Enabling Replicable Spatiotemporal Research with Virtual Spatial Data Lab.*

HVD-21-07 Developing workbenches for spatial data science.

FUTURE PLANS

- New release of Geospatial Analytics Extension for KNIME 1.01
- New KNIME workbook
- Semi-Annual meetings for Affiliate Labs
- GeoAI extension for KNIME
- GeoImage extension for KNIME based on books such as *Image Analysis, Classification and Change Detection in Remote Sensing, with Algorithms for ENVI/IDL and Python (Third Revised Edition)*, Mort Canty

APPENDIX

[1] Spatial Data Lab project website:

<https://projects.iq.harvard.edu/chinadatalab/home>

[2] Spatial Data Lab project annual reports:

https://projects.iq.harvard.edu/chinadatalab/project-reports?admin_panel=1

[3] Spatial Data Lab Webportal hosted by George Mason University:

<http://knime.futuredatalab.org>

[4] Research data achieving on Harvard Dataverse:

<https://dataverse.harvard.edu/dataverse/2019ncov>

[5] SDL webinar achieving on Harvard Dataverse:

https://dataverse.harvard.edu/dataverse/cdl_training

[6] Spatial Data Lab platform datasets introductions and user guide:

<https://docs.google.com/document/d/1MV0C7ro8n7bFum7q6DYISD0YkKkV1vBMLEMuC-ievmk>

[7] Publications: <https://projects.iq.harvard.edu/chinadatalab/publications-0>

[8] Presentations: <https://projects.iq.harvard.edu/chinadatalab/presentations-0>

[9] The 16th IAB meeting reports:

<https://drive.google.com/drive/folders/1Wgj5ilViBU9QQpaPSq9SqRFNFvWGxAnx>

[10] Spatial Data Lab Webinars:

<https://projects.iq.harvard.edu/chinadatalab/conferences>