KOR Geospatial Data Science Knowledge Exchange Seminar

Geospatial Data Science for Future Cities and Society

: Focusing on Analysis of Digital Footprint Data

KRIHS Korea Research Institute for Human Settlements | UBAI | 서울시립대학교 UNIVERSITY OF SEOUL | UNIVERSITY OF LIVERPOOL | Geographic Data Science Lab | Consumer Data Research Centre

KOR Geospatial Data Science Knowledge Exchange Seminar

Geospatial Data Science for Future Cities and Society

: Focusing on Analysis of Digital Footprint Data

Date

Thursday 8th ~ Friday 9th June 2023

Venue

8th / Auditorium / KRIHS (Sejong)

9th / Natural Science Building / University of Seoul (Seoul)

KRIHS Korea Research Institute for Human Settlements | UBAI | 서울시립대학교 UNIVERSITY OF SEOUL | UNIVERSITY OF LIVERPOOL | Geographic Data Science Lab | Consumer Data Research Centre

PROGRAM

Day1 Thursday 8th June

Time		Program	Presenter
11:00–11:50	50min	Welcome meeting/KRIHS tour/Photo	Director of GDPC (Global Development Partnership Center) in KRIHS/ Myung-Hwa Hwang
Luncheon (70min) – Lunch Box @ auditorium of KRIHS			
13:00–13:10	10min	Welcoming speech	Director of GDPC in KRIHS
Session 1 : Analysis of Urban spatial structure (moderator: Jae Soen Son)			
13:10–13:40	30min	〈Keynote〉 Urban Grammar – Data-driven understanding of urban form and function	Daniel Arribas-Bel (GDSL)
13:40–14:00	20min	〈Keynote〉 Inferring urban polycentricity from the variability in human mobility patterns	Carmen Cabrera-Arnau (GDSL)
14:.00–14:15	15min	Analysis of centers/subcenters via micro geospatial data: Focusing on Andong in Korea	Youngmin Lee (KRIHS)
14:15–14:30	15min	Mapping cognitive place connectivity using Reddit comments	Cillian Berragan (GDSL)
Coffee Break(30min)			
Session 2 : Analysis of Digital Footprint Data (moderator: Yohan Chang)			
15:00–15:20	20min	〈Keynote〉 A data fusion approach to update survey-based mode-of-transportation estimates using mobile phone data	Francisco Rowe (GDSL)
15:20–15:35	15min	Towards Carbon-Neutral Smart Cities : A Carbon Spatial Mapping Approach	Jae Soen Son (KRIHS)
15:35–15:50	15min	Utilizing relative risk index to diagnose the vulnerability of local markets to COVID-19	GyoungJu Lee (Korea National University of Transportation, KNUT)
15:50–16:05	15min	The economic resilience of a city: the effect of relatedness on the survival of amenity shops during the COVID-19 pandemic	Bogang Jeon (Inha University)
16:05–16:20	15min	Sensing the City: Clustering Pedestrian Behavioural Changes in Melbourne, Post-First and Second Lockdown	Danial Owen (GDSL)
Break Time(10min)			
16:30–17:00	30min	Q&A, Discussion	

Day2 Friday 9th June

Time		Program	Presenter
10:40–11:00	20 min	Registration	seminar room (1st floor) Natural Science Building in UoS
11:00–11:10	10 min	Welcoming speech	Mingyu Kang (UoS) Myung-Hwa Hwang (KRIHS)
PHOTO			
Keynote Speech Session (moderator: Myung-Hwa Hwang)			
11:10–11:40	30 min	Integrating Chat GPT in the Creation of Data Products for Consumer Data Research	Alex Singleton (GDSL)
11:40–12:00	20 min	Shifting patterns of Infectious diseases following social distancing measures for COVID-19 in South Korea	Sangshin Park (UoS)
Luncheon (80min) – Lunch Box			
Session 3 : Analysis of Mobility & Alternative Data (moderator: Mingyu Kang)			
13:20–13:40	20 min	〈Keynote〉 Can we use mobility data to assess the impacts of urban green space interventions?	Mark Green (GDSL)
13:40–13:55	15 min	Examining longitudinal neighborhood change and human mobility patterns with functional data analysis	Paul H. Jung (Inha University)
13:55–14:10	15 min	Sensing population displacement from Ukraine using Facebook data: Identifying potential settlement areas within host countries	Ruth Neville (GDSL)
14:10–14:25	15 min	Analysis of alternative data to monitor migration flow: the case of housing districts in South Korea	Bo Kyeong Lee (KRIHS)
Coffee Break (25min)			
Session 4 : Analysis of Mobility & Sensor Data (moderator: Myung-Hwa Hwang)			
14:50–15:10	20 min	〈Keynote〉 Using Urban Data Sensor Technologies for Enhancing Civic Life in Seoul	MingyuKang (UoS)
15:10–15:25	15 min	What If? Evaluating decarbonization strategies for maritime cargo flows in Liverpool City Region	Patrick Ballantyne (GDSL)
15:25–15:40	15 min	Exploratory Analysis of Smartphone-Based VGI Datasets Collected by Drivers in the Daegu Metropolitan City Area	Jae Hyun Lee (Kyungpook National University, KNU)
15:40–15:55	15 min	Drunk and Disorderly Data: Applying a natural language processing algorithm to classify alcohol-related crimes in police data.	Olivia Horsefield (GDSL)
15:55–16:10	15 min	Distributed computing for traffic-related air pollutant concentration modeling at regional scales	Daejin Kim (Inha University)
Break Time(10min)			
16:20–17:00	40min	Q&A and discussion on further research collaboration between UK and South Korea on data-driven urban sciences	

Session 1 : Analysis of Urban Spatial Structure

Welcoming Speech

Sangkeon Lee

Director of GDPC, Senior Research Fellow
Korea Research Institute for Human Settlements (KRIHS)

Moderator

Jae Soen Son

Associate Research Fellow
Geospatially Enabled Society Research Division
Korea Research Institute for Human Settlements (KRIHS)

Keynote : Urban Grammar - Data-driven understanding of urban form and function

Daniel Arribas-Bel

Professor in Geographic Data Science
Department of Geography and Planning / University of Liverpool (UK)
Deputy Programme Director / Urban Analytics at the Alan Turing Institute
Member of Geographic Data Science Lab (GDSL)

This talk will introduce the Urban Grammar, a project using data science and machine learning to build a detailed, consistent, and scalable characterization of urban form and function in Britain. How we spatially arrange cities matters. Their building blocks include the different elements of the built and natural environments of which cities are composed, but also the purpose they serve. Understanding the former thus requires us to consider urban form, while grasping the latter invites us to examine its function. Urban form and function is relevant in a variety of contexts concerning land use, such as characterising the urban fabric of cities, delineating neighbourhoods, or even understanding urban/rural continuums. Despite its relevance, the measurement of urban form and function has been hampered by the availability of meaningful spatial units, good data, and meaningful metrics, as well as by the complexity of bringing together their multi-dimensional nature. In this talk, we will introduce a new data product -the spatial signatures- that leverages data science to facilitate access to granular descriptions of urban form and function. We will cover its design, implementation, and some of the lessons of relevance for researchers interested in developing similar classifications in different (geographic) contexts. We will conclude with some thoughts on where the project is headed to develop the temporal dimension of the data product.

Keynote : Inferring urban polycentricity from the variability in human mobility patterns

Carmen Cabrera-Arnau

Lecturer in Geographic Data Science
Department of Geography and Planning/ University of Liverpool (UK)
Member of Geographic Data Science Lab (GDSL)

People in cities interact with their environment by developing urban land for different socioeconomic activities. The way in which land use is arranged within a city is normally referred to as urban structure. Through the study of such structures, we can learn more about the spatial behaviours of the societies that, over time, have built them. The polycentric city model of urban structure has gained popularity in spatial planning policy, since it is believed to overcome some of the problems often present in monocentric metropolises, ranging from congestion to difficult accessibility to jobs and services. However, the concept 'polycentric city' has a fuzzy definition and as a result, the extent to which a city is polycentric cannot be easily determined. Here, we leverage the fine spatio-temporal resolution of smart travel card data to infer urban polycentricity by examining how a city departs from a well-defined monocentric model. In particular, we analyse the human movements that arise as a result of sophisticated forms of urban structure by introducing a novel probabilistic approach which captures the complexity of these human movements. We focus on London (UK) and Seoul (South Korea) as our two case studies, and we specifically find evidence that London displays a higher degree of monocentricity than Seoul, suggesting that Seoul is likely to be more polycentric than London.

Analysis of centers/subcenters via micro geospatial data: Focusing on Andong in Korea

Youngmin Lee

Associate Research Fellow
Geospatial Analytics & Monitoring Center, KRIHS DataLab.
Korea Research Institute for Human Settlements (KRIHS)

There is a growing need for a 'compact & network' spatial development strategy to respond to the changes occurring in the national space, such as population decline, aging population, and local extinction crisis. This strategy aims to intensively foster centers/subcenters, which are places where various active populations, including residents and workers, are concentrated, and to maintain and strengthen public transport connections between centers/subcenters and their surroundings. In the land settlement system, centers/subcenters are the places that mediate flows between other nodes and create inflows or outflows to peripheries. In order to support the compact & network spatial development strategy in the era of declining population, we analyzed centers/subcenters in national space. Specifically, we analyzed population and employment concentration using grid data, and based on that, identified 'primary center' and 'secondary center' in Andong. We also identified population agglomerations that are not primary or secondary center but can be considered equivalent and categorized them into separate types: 'periphery-center' and 'periphery-subcenter'. The data used for the population concentration analysis is census population by 1km×1km grid from SGIS(Statistics Korea)(as of 2021), and for the employment concentration analysis, population based on workplace address by 1km×1km grid from Korea Credit Bureau(KCB)(as of 2021). In addition, points-of-interests(POI) data such as administrative buildings, elementary schools, and public health centers were used to analyze and categorize population agglomerations. As a result, one primary center, four secondary centers, 11 periphery-centers, and one periphery-subcenter were identified in Andong.

Session 1 : Analysis of Urban Spatial Structure

Mapping cognitive place connectivity using Reddit comments

Cillian Berragan

PhD Student in Data Analytics and Society
Department of Geography and Planning/ University of Liverpool (UK)
Member of Geographic Data Science Lab (GDSL)

Cognitive place connectivity may be conceptualised as mental models of the relationships between geographic locations. Utilising a large corpus of online discussion data from the social media website Reddit, we experiment on the extraction of such geographic knowledge from unstructured text. First we construct a system to identify place names found in Reddit comments, disambiguating each to a set of coordinates where possible. Following this, we build a collective picture of cognitive place connectivity in the United Kingdom, linking locations that co-occur in user comments and evaluating the effect of distance on the strength of these associations. Exploring these geographies nationally, connections were shown to be typically weaker over greater distances, however this distance decay effect is highly regional, strongest across Scotland, Wales, the South West of England and surrounding London. When comparing major cities across the UK, we observe distinct distance decay patterns, influenced primarily by proximity to cities.

Session 2 : Analysis of Digital Footprint Data

Moderator

Yohan Chang

Associate Research Fellow
Geospatial Analytics & Monitoring Center, KRIHS DataLab.
Korea Research Institute for Human Settlements (KRIHS)

Keynote : A data fusion approach to update survey-based mode-of-transportation estimates using mobile phone data

Francisco Rowe

Professor in Population Data Science
Department of Geography and Planning/ University of Liverpool (UK)
Lead of Geographic Data Science Lab (GDSL)

Up-to-date information on different modes of travel to monitor transport traffic and evaluate rapid urban transport planning interventions is often lacking. Transport systems typically rely on traditional data sources providing outdated mode-of-travel data due to their data latency, infrequent data collection and high cost. To address this issue, we propose a method that leverages mobile phone data as a cost-effective and rich source of geospatial information to capture current human mobility patterns at unprecedented spatiotemporal resolution. Our approach employs mobile phone application usage traces to infer modes of transportation that are challenging to identify (bikes and ride-hailing/taxi services) based on mobile phone location data. Using data fusion and matrix factorisation techniques, we integrate official data sources (household surveys and census data) with mobile phone application usage data. This integration enables us to reconstruct the official data and create an updated dataset that incorporates insights from digital footprint data from application usage. We illustrate our method using a case study focused on Santiago, Chile successfully inferring four modes of transportation: mass-transit (all public transportation), motorised (cars and similar vehicles), active (pedestrian and cycle trips), and taxi (traditional taxi and ride-hailing services). Our analysis revealed significant changes in transportation patterns between 2012 and 2020. We quantify a reduction in mass-transit usage across municipalities in Santiago, except where metro/rail lines have been more recently introduced, highlighting added resilience to the public transport network of these infrastructure enhancements. Additionally, we evidence an overall increase in motorised transport throughout Santiago, revealing persistent challenges in promoting urban sustainable transportation. Findings also point to a rise in the share of taxi usage, and a drop in active mobility, suggesting a modal shift towards less sustainable modes of travel. We validate our findings comparing our updated estimates with official smart card transaction data. The consistency of findings with expert domain knowledge from the literature and historical transport usage trends further support the robustness of our approach.

Session 2 : Analysis of Digital Footprint Data

Towards Carbon-Neutral Smart Cities : A Carbon Spatial Mapping Approach

Jae Soen Son

Associate Research Fellow
Geospatially Enables Society Research Division
Korea Research Institute for Human Settlements (KRIHS)

With extreme weather and disasters becoming more commonplace due to climate change, the need to go carbon neutral is urgent. Making cities carbon neutral, especially those with large populations, is challenging but necessary. To make a city carbon neutral, knowing the exact amount of carbon emissions and carbon absorption in the city is essential. We need information at a fine spatial scale to apply urban planning techniques and understand the effectiveness of carbon reduction policies. For this purpose, a carbon spatial mapping system has been built on a pilot basis. The Carbon Spatial Map System will provide maps and statistics on carbon emissions from buildings and transportation and the absorption capacity of tree species at the grid and administrative district levels. The system is scheduled to be released in the first half of this year, and by 2027, we plan to upgrade the carbon spatial map construction technology through R&D to support the planning of carbon-neutral cities.

Utilizing relative risk index to diagnose the vulnerability of local markets to COVID-19

GyoungJu Lee

Professor
Department of Civil, Environmental, Urban and Transportation Engineering
Korea National University of Transportation (KNUT)

The decline of local markets due to economic downturn during the COVID-19 pandemic has emerged as a major concern. Particularly, industries that rely on face-to-face interactions have experienced significant damages. The persistent closure of numerous stores serves as an indication of severe economic vulnerability within local markets. The objective of this research is to define relative risk index to estimate the level of vulnerability in local markets based on the increased frequency of closures following COVID-19. Using the index, the study aims to analyze spatial patterns that highlight areas with pronounced vulnerability within local markets. Identifying locations within local markets showing high likelihood of severe economic losses provides valuable decision-making information to support targeted interventions, such as the government's focused efforts in disease prevention and control, often referred to as pinset and/or targeted hygiene measures. To that end, empirical analysis was conducted focusing on the food service industry in Seoul, and policy implications were discussed based on the analysis results.

The economic resilience of a city: the effect of relatedness on the survival of amenity shops during the COVID-19 pandemic

Bogang Jun

Professor
Department of Economics
Inha University

Amenity clusters consisting of coffee shops, restaurants, and other small businesses improve urban life and are a source of employment for city dwellers. Although most small business clusters were hit hard by restrictions imposed during the COVID-19 pandemic, some were able to adapt. What determines the economic resilience of amenity clusters? Using store-location data for Seoul from 2016 to 2021, we identify spatial clusters of amenities and from that build an amenity space to examine the effect of relatedness on the resilience of each cluster. We find that businesses are more likely to survive when located in clusters of related amenities.

Sensing the City: Clustering Pedestrian Behavioural Changes in Melbourne, Post-First and Second Lockdown

Danial Owen

PhD Student in Data Analytics and Society
Department of Geography and Planning / University of Liverpool (UK)
Member of Geographic Data Science Lab (GDSL)

The COVID-19 pandemic and subsequent recovery efforts have changed how, when, and where people moved. In Melbourne, Australia, we expect these movement patterns and behaviours to differ, caused by the difference in the stringency and easing of restrictions post-first and second lockdown. This study aims to utilize open-sourced pedestrian counting data across the City of Melbourne to cluster spatio-temporal changes in pedestrian activity between post-first and second lockdown. We aim to investigate the influence of urban morphology, urban setting, and land-use on these changes. Using time-series clustering, our study uncovers three distinct clusters: areas with minimal change in activity encompassing workplace transit stations and residential zones, sustenance areas which experienced an immediate resurgence in activity, and major retail zones which demonstrated a more delayed recovery. Our findings explore the impact of the evolving commuting and lifestyle behaviours on a less explored yet significant mode of transport. The implications of these findings can provide valuable insights for policymakers on changing functionality of certain land-uses and urban settings across Melbourne and support their efforts on their ten-year transport plan to repurpose streets to enhance and encourage pedestrian-friendly environments.

Session 2 : Analysis of Digital Footprint Data

Moderator

Myung-Hwa Hwang

Manager of KRIHS DataLab, Research Fellow
Geospatial Analytics & Monitoring Center
Korea Research Institute for Human Settlements (KRIHS)

Keynote : Integrating Chat GPT in the Creation of Data Products for Consumer Data Research.

Alex Singleton

Professor in Geographic Information Science
Department of Geography and Planning / University of Liverpool (UK)
Deputy Director of the ESRC Consumer Data Research Centre (CDRC)
Director of the ESRC Data Analytics & Society CDT / Member of Geographic Data Science Lab (GDSL)

This talk will introduce the Consumer Research Data Centre (CDRC) which is funded by the Economic and Social Science Research council in the UK to provide a mechanism through which academics can access and use data from commercial organizations for research projects deemed to be in the public good. An overview of our enabling framework will be provided. A key output from the CDRC are a range of data products that convert personally or commercially sensitive "big data" into useable analysis ready forms that can predominantly be disseminated openly. An illustrated example is given for a geodemographic classification of the US, that innovatively uses Chat GPT to develop textual descriptions from numeric summaries of neighbourhood structure.

Keynote : Shifting patterns of Infectious diseases following social distancing measures for COVID-19 in South Korea

Sangshin Park

Vice President of Graduate School of Urban Health
Professor of Urban Health Epidemiology
University of Seoul

In response to the outbreak of COVID-19 in China, South Korea mandated universal use of face masks and recommended modified personal behavior including physical distancing starting in February 2020. As cases surged in South Korea, social distancing was instituted at the national level with varying levels of restrictions from March 2020. The coordinated public health response enforced limits on gatherings in various congregate settings including entertainment venues, places of worship, schools, sporting events, and work environments. Stringent social distancing schemes were enforced between 21 March and 5 May 2020 and were replaced on 6 May 2020 with less restrictive physical distancing in daily life when the number of cases declined. The public health strategy changed to a 3-tier scheme on 28 June 2020 and subsequently to a 5-level model on 17 November 2020 to permit maximum flexibility in response to social and economic upheaval caused by the epidemic.

We hypothesized that nationwide social distancing and other preventive measures for COVID-19 were associated with reduced detection of other respiratory viruses and gastrointestinal pathogens in South Korea.

We analyzed national surveillance data to compare incidence of respiratory viruses during 2016–2019 vs 2020. Results of multiplex reverse-transcription polymerase chain reaction assays for 8 respiratory viruses were included: adenovirus (ADV), parainfluenza virus (PIV), respiratory syncytial virus (RSV), influenza virus (IFV), human coronavirus (HCoV; non–SARS-CoV-2), human rhinovirus (HRV), human bocavirus (HBoV), and human metapneumovirus (HMPV). Results for gastrointestinal pathogens included norovirus, rotavirus, Campylobacter spp., Clostridium perfringens, and Salmonella spp.

During 2016–2019, rates of detection of respiratory viruses were relatively stable: ADV, 3.7%–9.2%; PIV, 1.4%–17.0%; RSV, 0.3%–15.3%; IFV, 0.4%–35.6%; HCoV, 1.5%–8.4%; HRV, 7.0%–25.1%; HBoV, 0.6%–6.3%; and HMPV, 0.7%–14.5%. Following implementation of social distancing in February 2020, rates of detection of enveloped viruses (HCoV, HMPV, IFV, PIV, and RSV) were significantly reduced by up to 100%. However, nonenveloped viruses (ADV, HRV, and HBoV) persisted throughout 2020, and HRV rates in hospitalized patients significantly increased.

We observed substantial declines in the activity of viral enteric pathogens (rotavirus and norovirus, albeit to a smaller extent) following implementation of social distancing measures in 2020. On the other hand, rates of Campylobacter spp., C. perfringens, and Salmonella spp. were either similar or higher in 2020 compared with the previous 5 years.

After implementation of social distancing for SARS-CoV-2 in South Korea, rates of detection of enveloped respiratory viruses decreased significantly, whereas nonenveloped viruses persisted, suggesting that enhanced infection prevention strategies are required to mitigate spread of these viruses. Moreover, social distancing measures have the greatest impact on pathogens that are transmitted persons-to-person, whereas the effect was smaller for foodborne bacteria.

Session 3 : Analysis of Mobility & Alternative Data

Moderator

Mingyu Kang

Head of Big Data Research Center of UBAI
Professor in Department of Urban Administration
University of Seoul

Keynote : Can we use mobility data to assess the impacts of urban green space interventions?

Mark Green

Reader in Health Geography
Department of Geography and Planning / University of Liverpool (UK)
Member of Geographic Data Science Lab (GDSL)

Urban green spaces are important community assets that provide opportunities for people to connect with nature, socialise and exercise. There is considerable policy interest in improving the quantity and quality of urban green spaces to support wellbeing. While local governments and charity groups are commonly changing features of urban green spaces, very little of this work is evaluated and therefore we know little about what works well for encouraging green space utilisation. This talk will present preliminary analyses using mobility data (counts of people visiting places via mobile phone GPS records) on urban green space utilisation to examine whether two interventions in Cheshire and Merseyside were successful: (i) the closing of a golf course, re-wilding and reopening as a public park; and (ii) planting of wildflowers. It will use novel Synthetic Control methods to identify how we can use mobility data to evaluate their impacts.

Examining Longitudinal Neighborhood Change and Human Mobility Patterns with Functional Data Analysis Approach

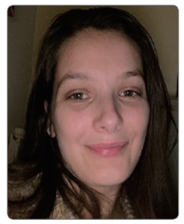
Paul H. Jung

Professor
Division of Asia Pacific Logistics
Inha University & UC Riverside

Recent methodological studies in spatial analysis and modeling have focused on addressing longitudinal spatial dynamics in neighborhood change and spatial interactions. However, incorporating the time dimension in spatial analysis models poses a significant challenge for spatial scientists due to the complexity it entails, including spatial dependency in time and temporal dependence in space. Consequently, representing time dynamics has been limited to merely stacking discrete cross-sectional data over time. For instance, existing neighborhood trajectory models capture neighborhood change by aggregating cross-sectional neighborhood clustering results over multiple years and analyzing the discrete stepwise switching patterns between clusters. Meanwhile, while transportation systems exhibit dynamic characteristics with high temporal variation in weekly usage patterns, accessibility measurement has largely overlooked the fact that accessibility also fluctuates significantly within a day or between days.

In this study, I present two examples showcasing a novel approach known as functional data analysis (FDA) as an alternative method to represent time dynamics in spatial studies: 1) a neighborhood trajectory model and 2) dynamic time accessibility measurement. I propose adopting the FDA framework as a mathematical representation of time dynamics, employing multivariate time-dependent curves (or functions) to identify neighborhood trajectory clusters or changes in accessibility between days. First, the FDA-based neighborhood trajectory model incorporates multivariate functional principal component analysis and k-means clustering. To illustrate its application, I employed this model to analyze neighborhoods in the Charlotte and Detroit metropolitan areas, revealing ongoing racial and socioeconomic segregation patterns and capturing the time dynamics of neighborhood change. Second, I propose a set of dynamic transit accessibility measures to examine the temporal fluctuations of accessibility using public transit modes. By examining Seoul's intermodal integrated transit system, I measure dynamic transit accessibility to capture the day-to-day heterogeneity in accessibility fluctuations across neighborhoods over the course of a week.

While the FDA framework requires further methodological development in various contexts, it demonstrates promising potential in modeling spatiotemporal aspects of neighborhood and transportation phenomena. By considering time-dependent changes, it can provide valuable support for urban policymaking and transportation planning.



Sensing population displacement from Ukraine using Facebook data: Identifying potential settlement areas within host countries

Ruth Neville

PhD Student in Data Analytics and Society
Department of Geography and Planning / University of Liverpool (UK)
Member of Geographic Data Science Lab (GDSL)

The escalation of conflict in Ukraine has triggered the largest refugee crisis in Europe since WWII. As of mid-April 2023, over 8.2 million people have fled Ukraine. Large-scale efforts have been made to identify the major receiving countries. However, less is known about the sub-national areas within host countries where refugees have migrated. Identifying these areas is key for the appropriate allocation of humanitarian aid. By combining digital Facebook API data and traditional data from Eurostat, this paper aims to identify and characterise potential settlement areas of Ukrainians across the main destination countries in Europe. We identify high concentrations of Ukrainians in urban areas with a pre-existing diaspora and tight labour market conditions across southern, northern-west and central Poland and the city of Prague in Czech Republic. We also find potential settlements in key urban agglomerations with a moderate diaspora and high levels of unemployment in Spain. Only in Romania, refugees seem to have settled in rural areas which show a moderate diaspora but low levels of unemployment. Potential settlement areas in Germany, Italy and the United Kingdom are spread across the country. Surprisingly, we do not identify potential settlement areas in bordering regions with Ukraine within neighbouring countries, suggesting that refugees may have used them only as transit points. Our findings point out that different packages of humanitarian assistance may be needed according to the number of refugees and the characteristics of settlement areas.

Analysis of alternative data to monitor migration flows: the case of housing districts in South Korea

Bo Kyeong Lee

Associate Research Fellow
Geospatially Enabled Society Research Division
Korea Research Institute for Human Settlements (KRIHS)

Understanding migration flow and characteristics of population that flows into a place plays important role in establishing the national or regional spatial plan of the central or local governments. Moreover, in the era of extremely low birth rates in Korea, every local government is making various efforts to prevent population outflow and increase inflow, and inter-regional migration due to the housing supply has a direct impact on the increase or decrease in the size of the local population.

The purpose of this study is to develop a big data driven framework to analyze population migration according to housing supply, and to find a policy application plan. We try to establish a model that can answer the question, “When housing was supplied through housing site development, who came to live in the place and where did they come from?”. This study processes private and public big data that showing population migration flows and characteristics of people who move to new houses.

We found that the age composition of the population that flowed into the 281 land development districts was 59.71% in their 30s and 40s out of 1,406,139 people analysed. The proportion of those in their 30s and 40s among the total residents who moved in during the same period within the country (2012-2020) was about 37.6%, indicating that the population that flowed into new houses in land development districts in their 30s and 40s was higher. We also measure the migration distance of about 1.4 million people and it was found that 50% of them moved within a radius of 7km, and most of them (75% of the total movement) moved within about 15 km.

We grouped land development districts into three types according to the base law and the location: group1 – the districts in the metropolitan area, group 2 – Innovation cities/Multifunctional administrative cities in the non-metropolitan area and group 3 – the general districts in the non-metropolitan area. The inflow of population from the group 1 was very likely to have come from a short distance, and even if the distance was far, the probability of moving into new housing in the metropolitan area was above a certain level. Moreover, the case of group 2 shows that the probability of influx from a longer distance was higher than that of other housing districts. On the other hand, the group 3 can be expected to trigger movement within the same city, county, and district, and thus affect the decline of old downtown areas.

When establishing national spatial plans, the suggested framework can be used to preemptively predict population movement. The central and local governments should comprehensively review their housing supply plans using the suggested framework. It is expected that it will be possible to support the establishment of necessary plans in advance by predicting the areas where the population will increase and those where the population will decrease depending on the migration

Session 4 : Analysis of Mobility & Sensor Data

Moderator

Myung-Hwa Hwang

Manager of KRIHS DataLab, Research Fellow
Geospatial Analytics & Monitoring Center, KRIHS DataLab.
Korea Research Institute for Human Settlements (KRIHS)

Keynote : Using Urban Data Sensor Technologies for Enhancing Civic Life in Seoul

Mingyu Kang

Head of Big Data Research Center of UBAI
Professor in Department of Urban Administration
University of Seoul

This study explores the application and impacts of Urban Data Sensor Technologies (UDST) in the dynamic urban landscape of Seoul, South Korea, emphasizing how these technologies can significantly enhance civic life. UDST, which involves the collection, analysis, and application of a range of urban data through sensor technology, has emerged as a vital tool in contemporary urban planning and management.

Seoul, celebrated for its technological advancements and dense population, provides a unique arena for exploring the potentials and challenges of UDST. This research primarily focuses on the use of UDST for monitoring population movements, enhancing the utilization of public spaces, and increasing real-time responsiveness to various urban issues.

The study showcases Seoul's initiatives to track and analyze population movements in real-time using UDST, highlighting the significant role these technologies play. By monitoring and analyzing the movement and behavior of people within the city, policymakers have gleaned significant insights into urban dynamics, informing urban planning decisions and enhancing public services and spaces.

Additionally, the study elaborates on how environmental sensors have been employed not only to manage heatwaves effectively via real-time temperature data but also to monitor air quality and noise levels. The comprehensive, real-time data collected by these sensors has been instrumental in informing more responsive and effective environmental policies, thereby ensuring the safety, comfort, and overall wellbeing of Seoul's residents.

While acknowledging the promising advancements, the study also recognizes potential challenges and ethical considerations related to data privacy and security in the implementation of UDST. It concludes by suggesting future directions for UDST deployment in Seoul, emphasizing its potential in creating a more responsive and sustainable city.

What If? Evaluating decarbonization strategies for maritime cargo flows in Liverpool City Region.

Patrick Ballantyne

Postdoctoral Research Associate in Geographic Data Science
Department of Geography and Planning / University of Liverpool (UK)
Member of Geographic Data Science Lab (GDSL)

Maritime transportation and shipping at the Port of Liverpool represents one of the key sources of economic development within the region, acting as a key hub for goods arriving at the UK. Whilst this has significant economic benefits, the environmental aspects of such economic activities are often overlooked, specifically the CO2 emissions from Heavy Goods Vehicles (HGVs), and their exposure to local, often vulnerable populations. Using HGV origin/destination data provided by Liverpool City Region Combined Authority (LCRCA), we provide a working overview of potential decarbonisation strategies from the regional maritime industry. Firstly, we demonstrate techniques for estimating likely end points for HGVs from the centroids of Great Britain Freight Model (GBFM) zones. Secondly, we estimate likely routes for HGVs to end points and ascribe volumes to the road network, using open-source routing software and route matching techniques. Finally, we discuss how the HGV volumes could be converted into CO2 emissions, used to calculate population exposure, and to support specific interventions which aim to reduce exposure whilst also improving the efficiency of the network. This work, in partnership with LCRCA, has significant implications for transport and maritime planning within the region, drawing attention to the environmental costs of the maritime industry, whilst also providing direct evidence in support of interventions targeting reduced population exposure.

Exploratory Analysis of Smartphone-Based VGI Datasets Collected by Drivers in the Daegu Metropolitan City Area

Jae Hyun Lee

Professor
Department of Geography
Kyungpook National University (KNU)

Urban environments interact with residents daily, exhibiting dynamics that vary within a day and across different timeframes. These dynamics can be categorized into microdynamics, mesodynamics, and macrodynamics. Microdynamics capture daily activity and travel patterns, while mesodynamics encompass habitual behaviors and learning processes. Macrodynamic analysis focuses on long-term travel behavior patterns. Previous research has extensively examined micro- and macro-level travel behavior dynamics using survey and panel datasets spanning multiple years. However, analyzing meso-level travel behavior has been challenging due to data collection difficulties. In this presentation, we present initial findings from an explanatory analysis conducted on a new type of Volunteered Geographic Information (VGI) dataset for meso-level travel behavior analysis in Korea. We discuss the challenges encountered when utilizing this dataset for travel behavior analysis and explore its potential applications. Additionally, we introduce an application of this dataset during the COVID-19 period in the Daegu metropolitan area. One key finding is the significant role of landuse characteristics in understanding travel behavior and a week or month-level data aggregation using location-time-only sensor-based VGI data. Specifically, by comparing landuse changes within a day in a longitudinal manner, we found distinct sequential patterns of travel behavior change. These patterns provide valuable insights into the diverse dynamics of meso-level travel behavior, shedding light on the complexities within urban environments.

Contributors : Soojeong Choi, Jiyoon Shim, Sanghyun Yoon(KNU) Seo Youn Yoon(KRIHS), Myeongho Yeo(WeDrive)

Drunk and Disorderly Data: Applying a natural language processing algorithm to classify alcohol-related crimes in police data.

Olivia Horsefield

PhD Student in Data Analytics and Society
Department of Geography and Planning / University of Liverpool (UK)
Member of Geographic Data Science Lab (GDSL)

Alcohol plays a significant role in crime, and police data are one of the few sources which contains a measure for alcohol-related crime. Researchers utilise the police measure to understand this crime type, however it is not entirely accurate. Our study presents a new measure of alcohol-related crime in police data. We applied a natural language processing algorithm to crime descriptions in police data, to classify crimes as alcohol-related or not. Our algorithm estimated that a higher proportion of crime is linked to alcohol than current police estimates suggest, across crime type and over time. Our approach not only improved estimates of alcohol-related crime, but also our understanding of its nature. We estimated twice as many alcohol-related crimes than the police measure in the most deprived areas. The proportion of alcohol-related crime was however just as high in lesser deprived areas as it was in the most deprived areas, unobserved by current police estimates. Through improved estimates, our approach could progress policing practice and decision-making around alcohol-related crime to reduce its burden.

Distributed computing for traffic-related air pollutant concentration modeling at regional scales

Daejin Kim

Professor
Division of Asia Pacific Logistics
Inha University

Although a variety of modeling tools have been developed to predict potential public exposure to harmful transport emissions at regional and sub-regional scales, computational efficiency remains a critical concern in the design of modeling tools. Microscale dispersion models run at high resolution often require extremely long runtimes for larger roadway networks and high-resolution receptor grids. Motivated by the challenges encountered in the previous modeling efforts, this work develops an advanced modeling framework for region-wide applications of line source dispersion models that integrates a high-performance emission rate lookup system (i.e., MOVES-Matrix), link screening, and innovative receptor site selection routines to further accelerate model implementation within distributed computing modeling framework. The case study in the 20-county metropolitan Atlanta area accounts for an extremely large number of link-receptor pairs (more than 160 thousand transportation links and 1.1 million receptors) demonstrates that the modeling system generates comparable concentration estimates to extremely-high-resolution processes, but with very high computational efficiency. Using AERMOD, the regional analysis required only 10 days to complete the analyses, compared to a total runtime by traditional methods of more than one year. The improvement in computational speeds are attributed to 1) the employment of supervised random forests machine learning model for link screening model, which objectively excludes transportation links that do not significantly affect receptor concentrations, and 2) the use of the supercomputing clustering system, where multiple AERMOD simulation jobs are split and simultaneously processed, thereby significantly reducing the total run-time. The dynamic-grid-receptor model systematically generates the network of receptor sites, based upon road geometry and meteorological conditions, to help minimize model runtime without undermining pollutant concentration predictions. The comprehensive modeling methodology presented in this work will make comparison of air quality impacts across complex project scenarios (and transportation development alternatives over large geographic areas) much more feasible. All these aspects should be of interest to a broad readership engaged in near-road air quality modeling for transportation planning and air quality conformity and for environmental analysis under the National Environmental Policy Act.

MEMO