

# Introduction to Geographic Information Systems and Urban Data Science

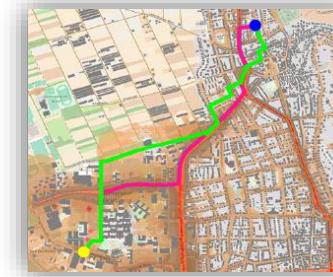
Dr. Tessio Novack  
[\(novack@uni-Heidelberg.de\)](mailto:novack@uni-Heidelberg.de)

# About me

- 2006** Bsc. In Geography (University of Sao Paulo)
- 2009** Msc. in Remote Sensing (National Institute for Space Research)
- 2016** Dr.-Ing. in Remote Sensing (Technical University of Munich)
- since 2016** PosDoc in Geo-informatics (GIScience, Uni Heidelberg)

## Research areas:

- Urban mobility
- VGI data analysis
- Critical GIS



# Schedule for today

<b>10:00 – 11:00</b>	Introduction to GIS and Urban Data Science
<b>11:00 – 11:30</b>	Preliminaries to Ex. 1 and 2 – Basics of GIS
<b>11:30 – 12:30</b>	Ex. 1 and 2 – Basics of GIS
<b>14:00 – 14:45</b>	Preliminaries to Ex. 3 – Spatial Regression with GeoDA
<b>14:45 – 17:00</b>	Ex. 3 – Spatial Regression with GeoDA
<b>17:00 – 17:30</b>	Discussion

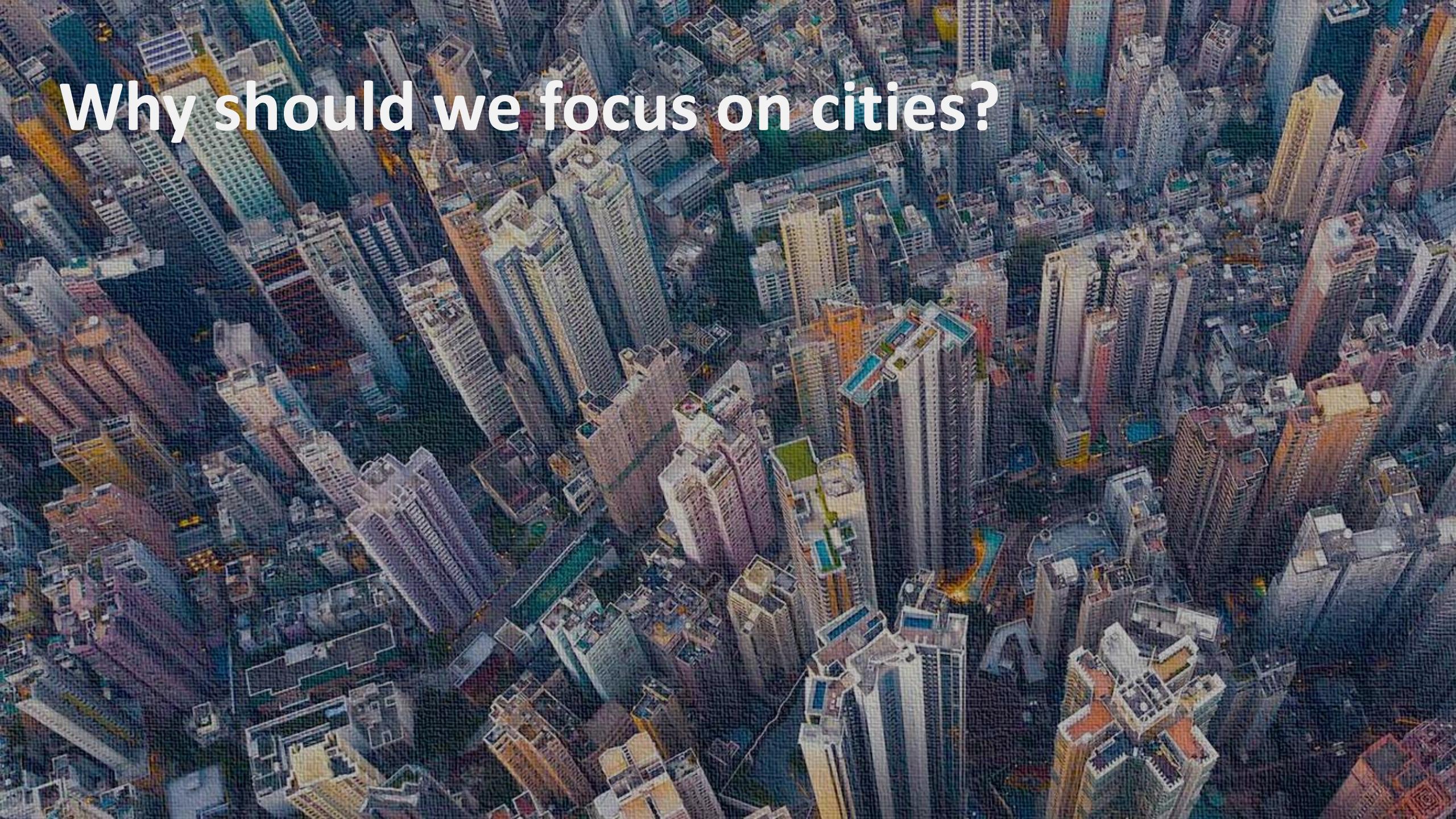
# Outline of this talk

Why cities are important

Why data is important for managing and analysing cities

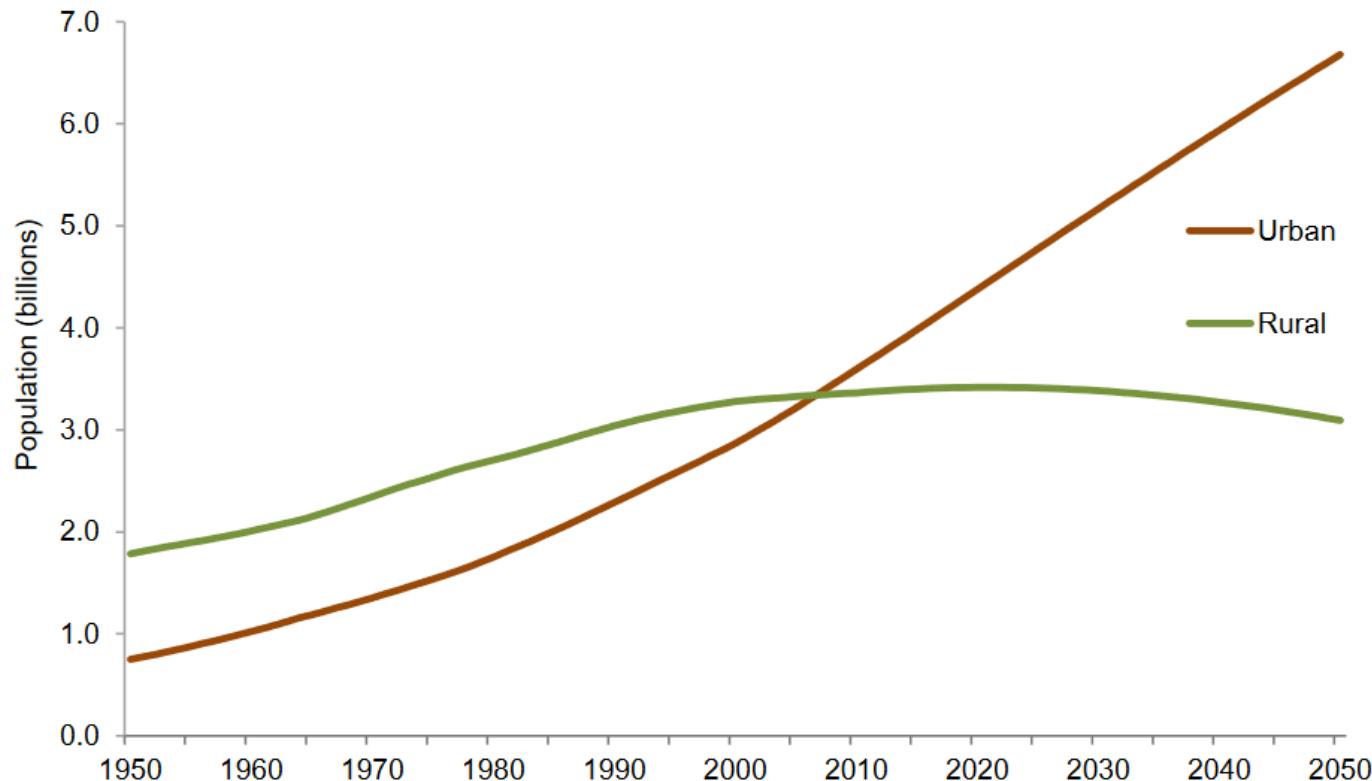
What kinds of data we can rely on

Why GIS is such an useful framework

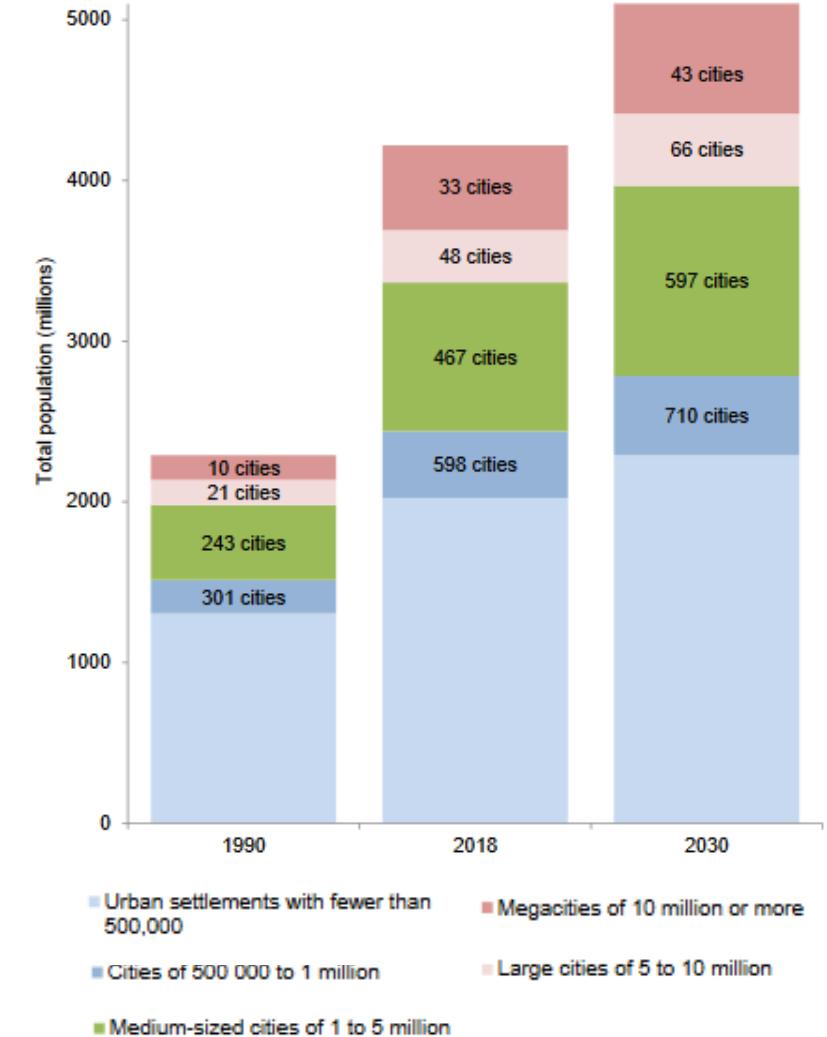
An aerial photograph of a dense urban city, likely Hong Kong, showing a high concentration of skyscrapers, residential blocks, and commercial buildings packed closely together. The image captures the intricate patterns of the city's architecture from a high vantage point.

Why should we focus on cities?

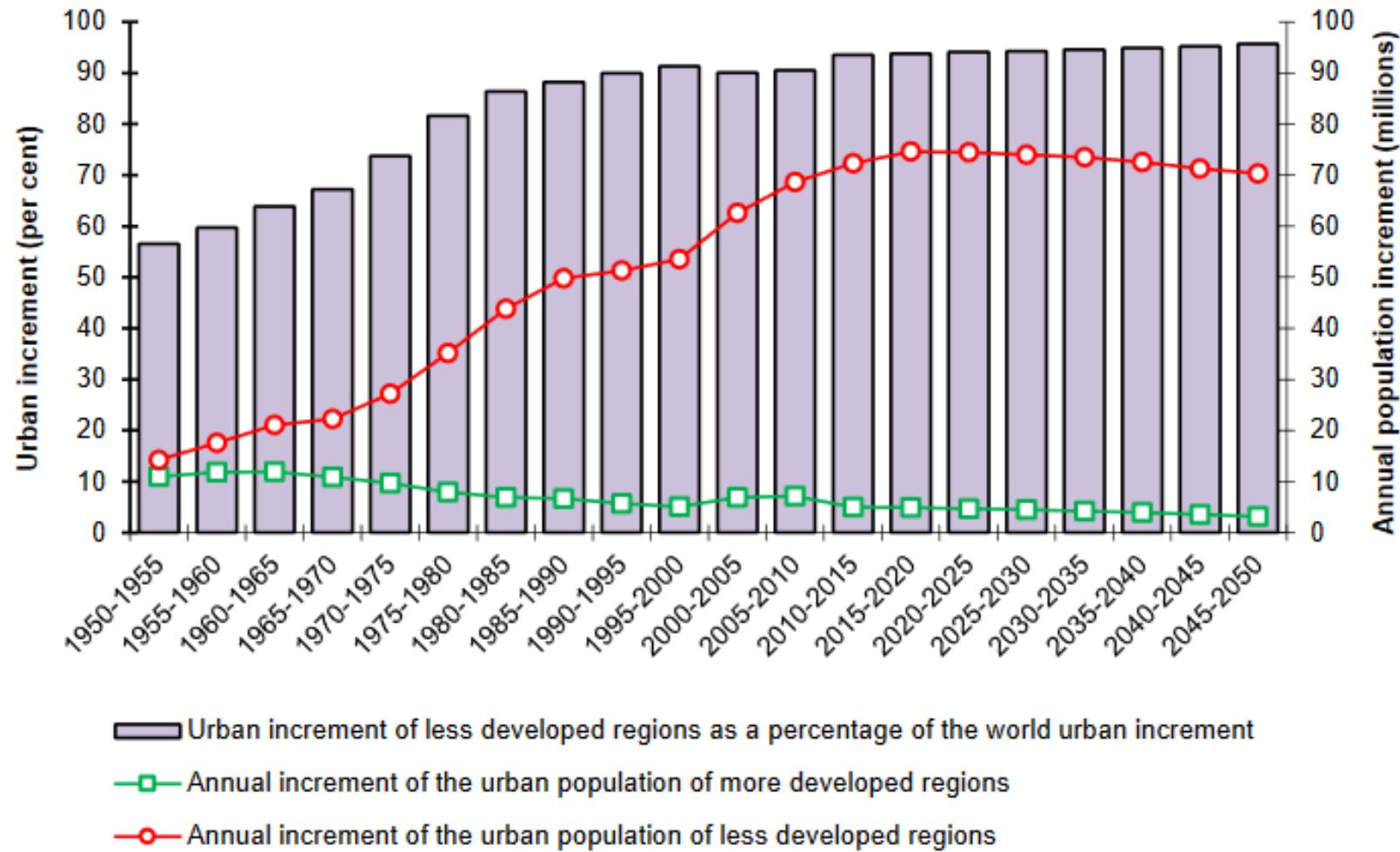
# Worldwide expansion of urban areas



Data source: United Nations, Department of Economic and Social Affairs, Population Division (2018a). *World Urbanization Prospects 2018*.



# Cities are expanding mainly in less developed areas



# Problems associated with unplanned urban growth



Landslides



Water contamination



Noise pollution



Floods



Traffic Jams



Air pollution

# Problems associated with unplanned urban growth

Atlanta and Barcelona have similar populations but very different carbon productivity

Atlanta



Barcelona



Population	Urban area	Transport carbon emissions	Population	Urban area	Transport carbon emissions
<b>2.5</b> million	<b>4,280</b> km <sup>2</sup>	<b>7.5</b> tonnes CO <sub>2</sub> /person (public + private transport)	<b>2.8</b> million	<b>162</b> km <sup>2</sup>	<b>0.7</b> tonnes CO <sub>2</sub> /person (public + private transport)

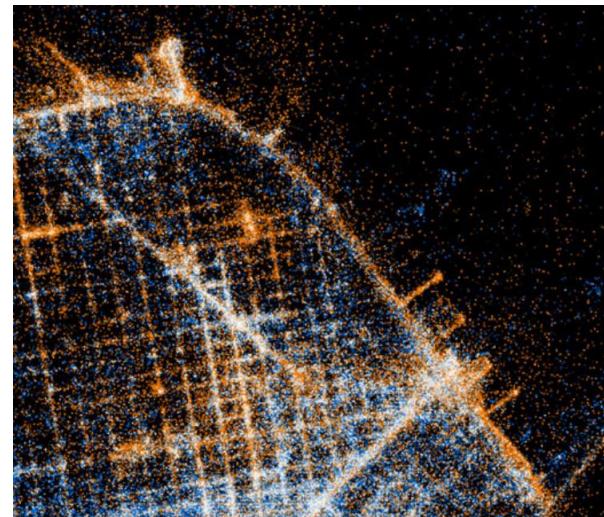
# Planning is everything!

- Aspects of a good urban quality of life
  - Easy access to daily duties
  - Affordable housing
  - Sense of security
  - Adequate environmental conditions
  - Access to public spaces
  - Sense of community
  - Accessibility to services, e.g. health care, children care, cultural facilities, etc.
  - Access to natural areas
  - Access to civic space!



# Urban planning

- Definition
  - Technical and political process concerned with the design of the built environment
- Operates mainly through
  - Policy recommendations
  - Transportation and traffic planning
  - Land-use zoning
- Two major pillars
  - Data and information
  - Public participation



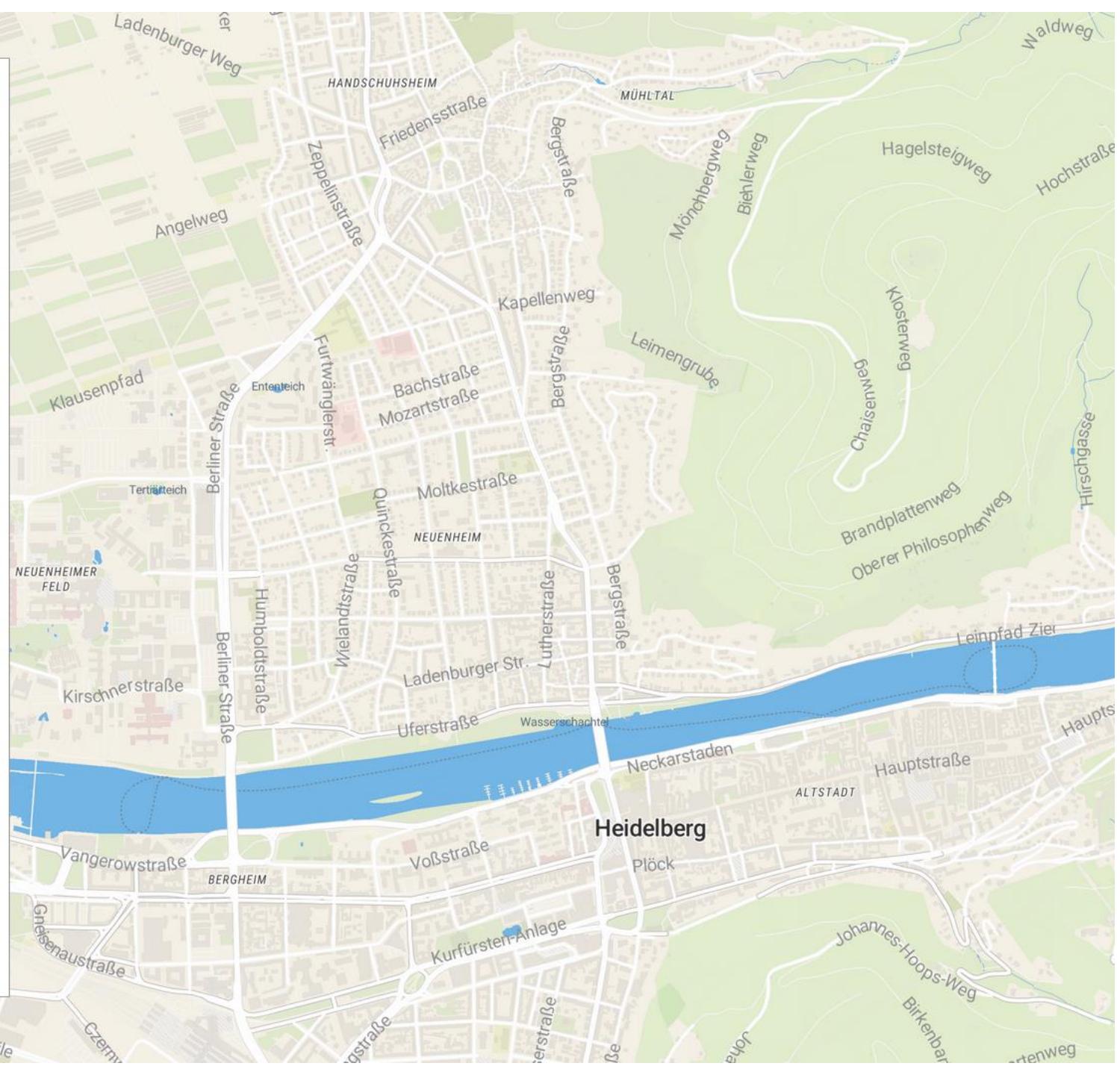
# Healthy Living in Heidelberg

Your opinion is asked on how we can make Heidelberg a healthier and more sociable city concerning possible improvements to green areas and to streets frequently used by pedestrians.

We thank you already for your participation. Choose your preferred language and get started!

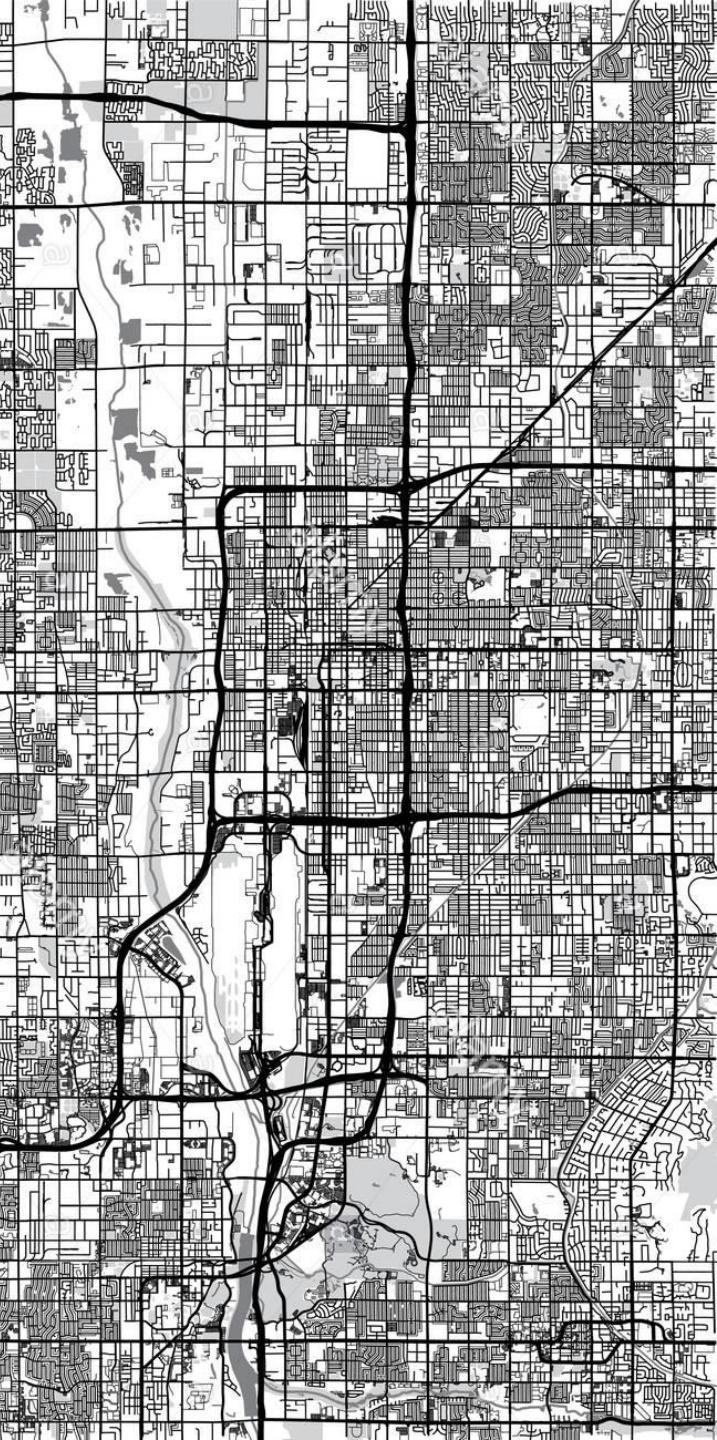
The questionnaire has been closed 7/4/2019

[中文](#) [Português](#) [Deutsch](#) [English](#)



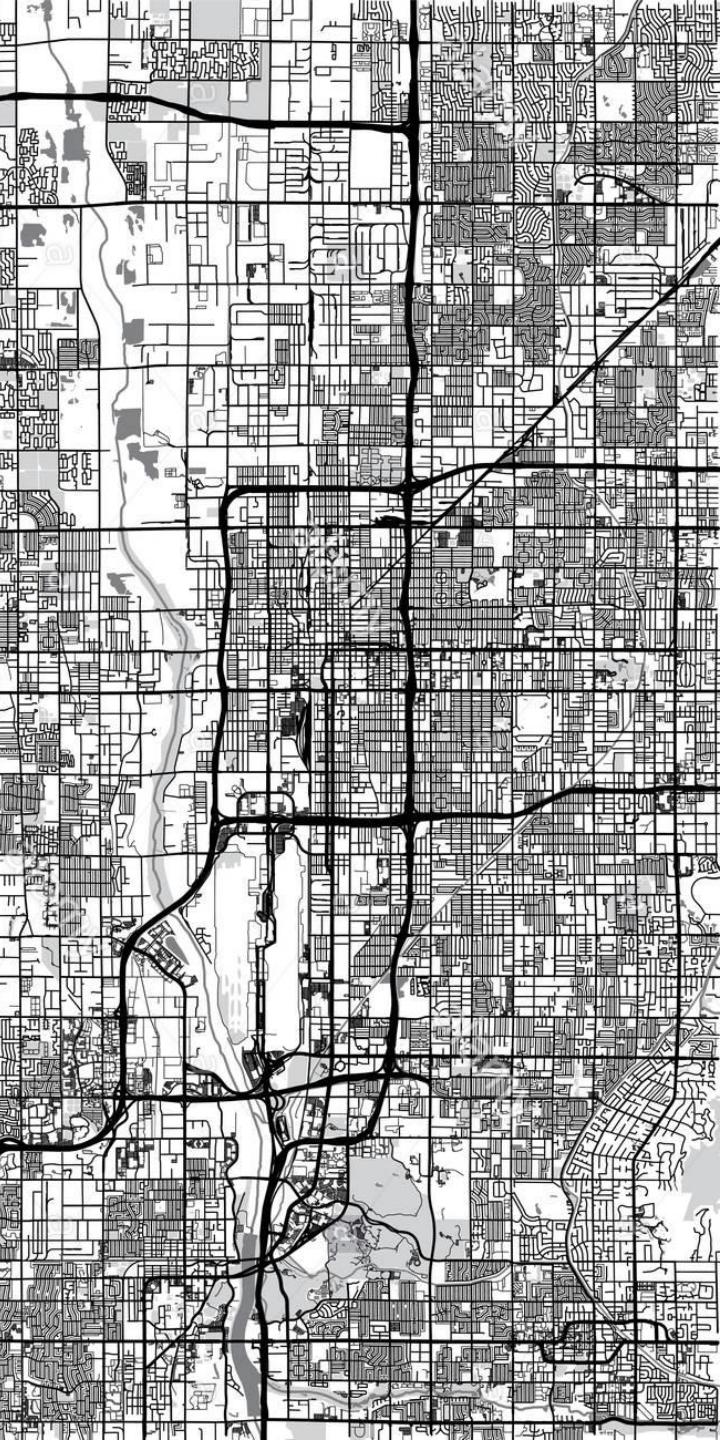
# The city as a system

- Basic spatial components:
  - Locations (e.g., the city hall)
  - Pathways (i.e. streets and subway lines)
  - Regions (e.g., blocks, census tracts, districts)
- Basic flows:
  - People and goods
  - Energy
  - Data and information
- Consumption and production
  - Material goods
  - Solid waste
  - Information



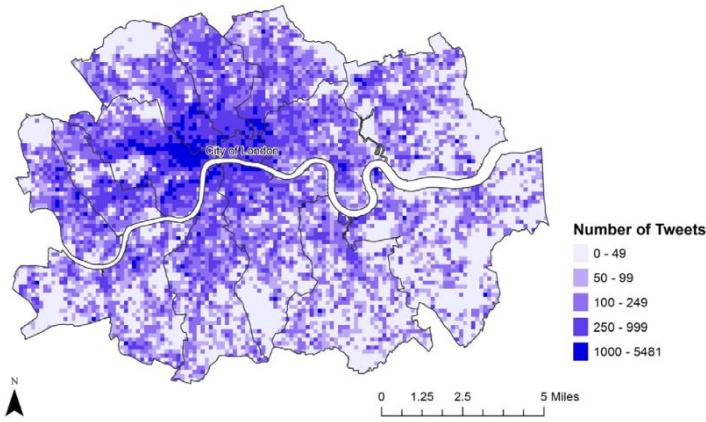
# The city as a system

- Centers of capital reproduction
- Overlapping networks of different extents of materiality
- Spaces emotionally charged with
  - Emotions
  - Memories
  - Feelings
  - Life stories

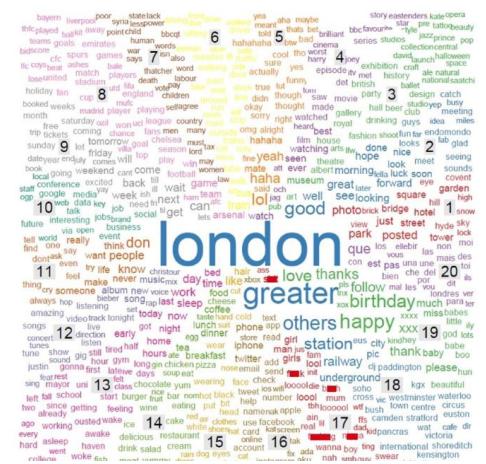


**Urban Data Science** is the multi-disciplinary area of research concerned with using new and emerging forms of data, alongside computational and statistical techniques, to study cities

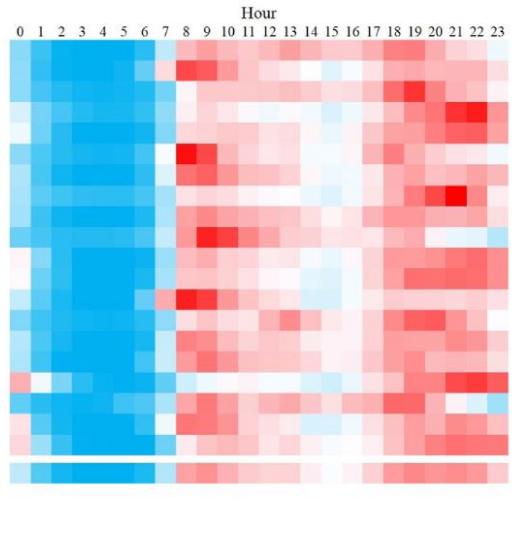
# Case example



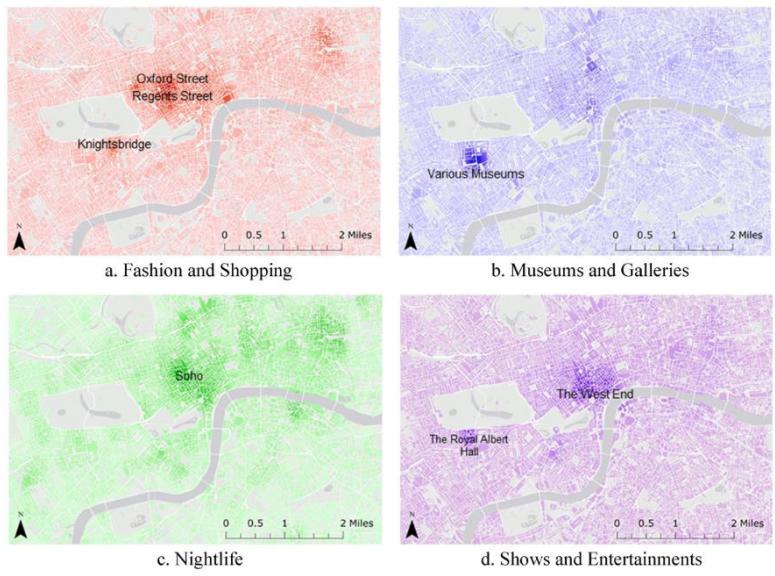
## 1. Twitter data extraction



## 2. Topic modelling



### 3. Spatio-temporal analysis



# Typical applications

- Extracting spatio-temporal patterns of human activity in cities
- Extracting spatio-temporal patterns of urban mobility
- Detecting events
- Extracting emotional patterns tied to urban spaces

# Smart cities



# What are smart cities?

- Cities that rely on technology for their government and management
- Cities that leverage digital data as well as computational, communication and internet technologies for
  - Improving traffic management and other urban operations
  - Increasing security,
  - Saving energy and reducing consumption of resources,
  - Improving governance through
    - Fostering interaction among citizens
    - Empowering citizens with information
    - Facilitating communication between government and communities

Key words: performance, real-time response, interactivity

# Key smart cities technologies

## Sensors



## IoT



## Big urban data analytics



## Cloud computing



## Smart governance



# Smart cities technologies

- Sensors
  - Used for counting pedestrians, bicycles and vehicles in
    - Streets,
    - Train stations,
    - Buses, etc.
  - Surveillance cameras (often harnessed by AI)
    - Government buildings,
    - Fight crime,
  - Monitoring environmental conditions of the city
    - Noise,
    - Humidity,
    - Pollutant quantities, etc.



# Smart cities technologies

- Deep Neural Networks can be trained to detect and distinguish objects, such as
  - Male/female
  - Adult/child
  - Vehicle type
- As well as for
  - Face and plate recognition
  - Color, size, speed, path, direction, etc.
- Video is thus transformed into **structured data**



# Smart cities technologies



**ENVIRONMENTAL SENSOR**

Particles suspended in the air, humidity,  
temperature and battery



**SOUND SENSOR**

Sound level, battery, temperature



**PARKING SENSOR**

Available parking space, occupied space



**HTE MIX**

Ambient humidity, Ambient temperature,  
Soil moisture, Soil temperature, Battery



**AMBIENT HUMIDITY AND TEMPERATURE  
SENSOR**

Environmental temperature, RH, battery



**SOIL MOISTURE SENSOR**

Soil moisture, Soil temperature, Battery



**WASTE SENSOR**

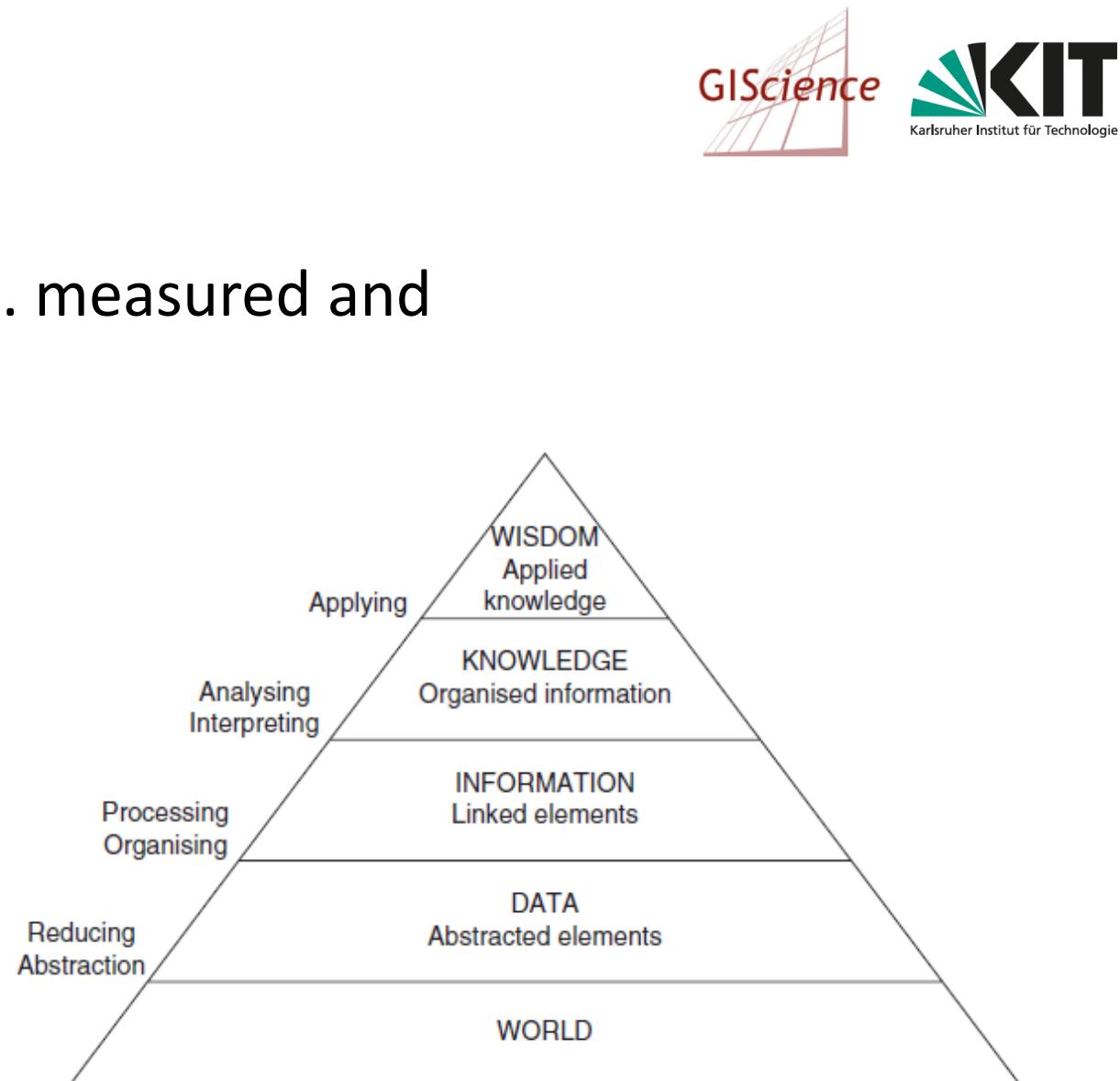
Filling percentage (%) Temperature,  
Battery, Vibration

# Reflecting on data



# What is data?

- Elements that can be abstracted, i.e. measured and recorded, from phenomena
- Selected according to purpose
- Data is an epistemological feature!

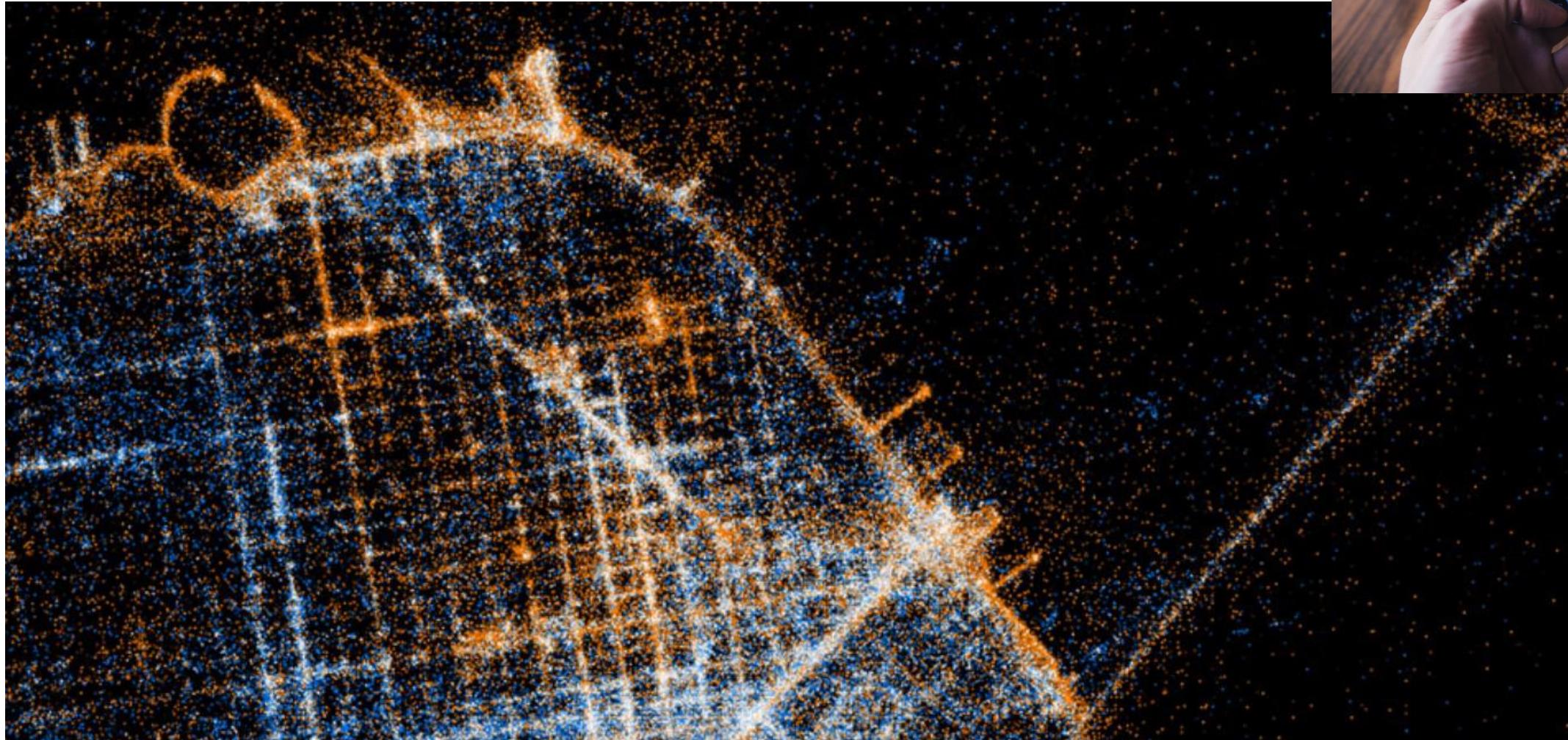


# Open Data

- Freely available data without restrictions of use
- Main sources:
  - Science
  - Governments
  - NGOs
- Important aspect of transparency, social inclusiveness and democratic city management
- Empower people and stimulate research and businesses

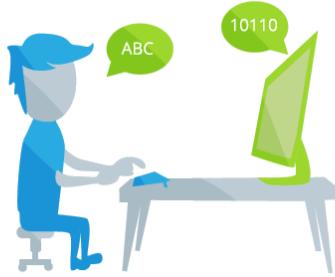


# Geo-social media data



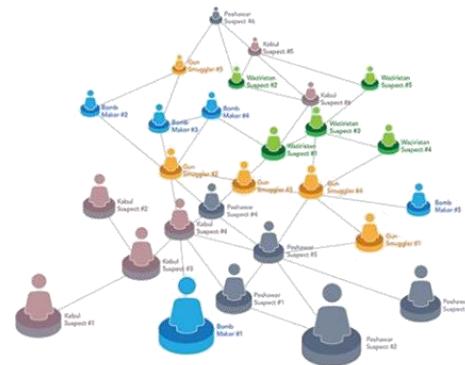
# Geo-social media data

## Natural Language Processing



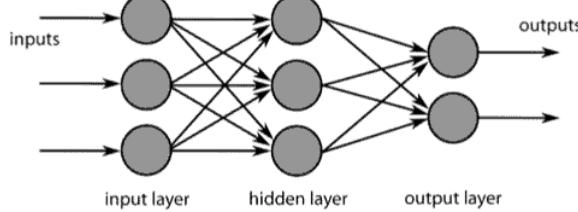
- Word-sense similarity
- String similarity
- Text classification
- Sentiment analysis
- Topic modelling

## Network Analysis

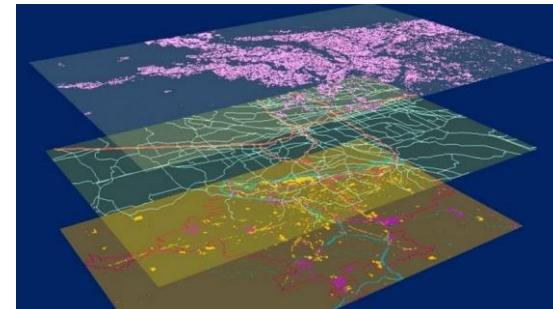


- Nodes' importance
- Networks attributes
- Network segmentation
- Shortest-path calc.

## Image Interpretation – Deep Learning



## GIS and (Spatial) Statistics



- Spatial clustering
- Correlation metrics
- Point to area interpolation

# Volunteered Geographic Information

- Definition
  - Collaborative projects to create a free editable map of the world
  - Users may collect data using manual survey, GPS devices, aerial photography, and other free sources, or use their own local knowledge of the area
- Wikimapia
- OpenStreetMap
  - “OpenStreetMap is a map of the world, created by people like you and free to use under an open license”
  - Definitely the most successful VGI platform to date
  - Geometrically and semantically relatively complete and accurate



# Node: Nando's Southwark Arches (3134099059)

Updated a restaurant

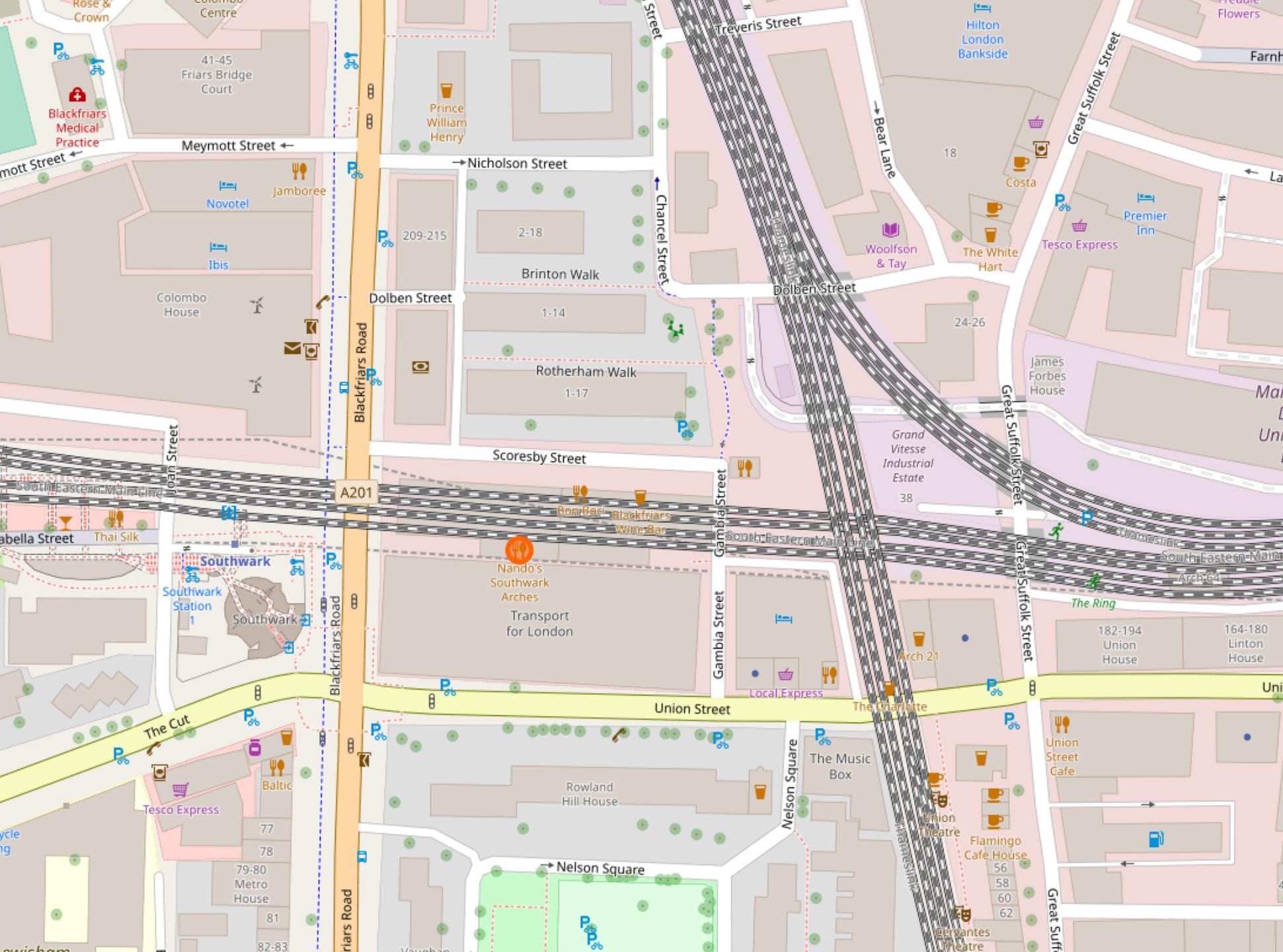
Edited over 2 years ago by [tbt](#)

Version #5 · Changeset #[56442453](#)

Location: 51.5040819, -0.1037565

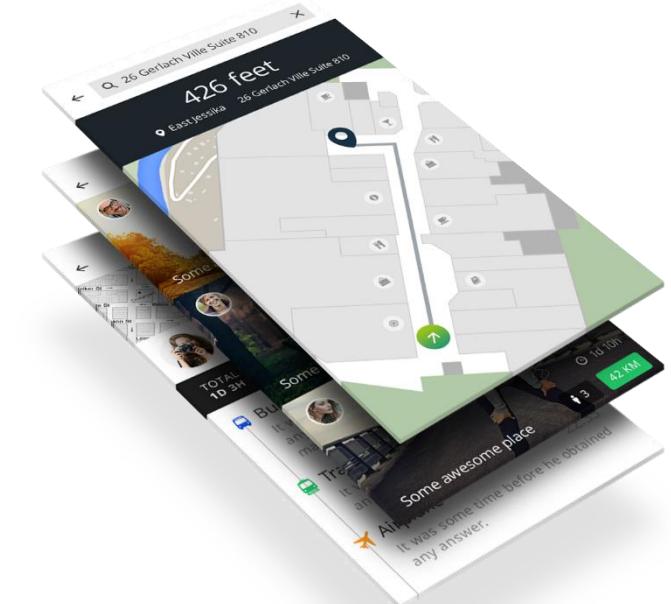
## Tags

addr:postcode	SE1 0XH
addr:street	Blackfriars Road
amenity	restaurant
cuisine	chicken
name	Nando's Southwark Arches
opening_hours	Mo-Th 11:30-22:00; Fr 11:30-23:00; Sa 12:00-23:00; Su 12:00-22:00
operator	Nando's
postal_code	SE1 0XH
website	<a href="https://www.nandos.co.uk/restaurants/southwark-arches-london">https://www.nandos.co.uk/restaurants/southwark-arches-london</a>



# Location-based services

- Smartphone-based applications that provide services and information to users based on their location and geographic data
- Examples:
  - Place recommendation (e.g., Foursquare, Google Places)
  - Routing and navigation
  - Cycling, running and fitness (e.g., MapMyRide, Strava)
  - Social interaction (e.g., Tinder)
  - Augmented reality
  - Mobility (e.g. Uber, Lemon)
- Data is sometimes purchasable in anonymized formats



# Location-based services

<https://www.intelligenttransport.com/transport-news/89550/waze-to-share-traffic-data-with-transport-authorities/>

## e-Scooter



## Navigation

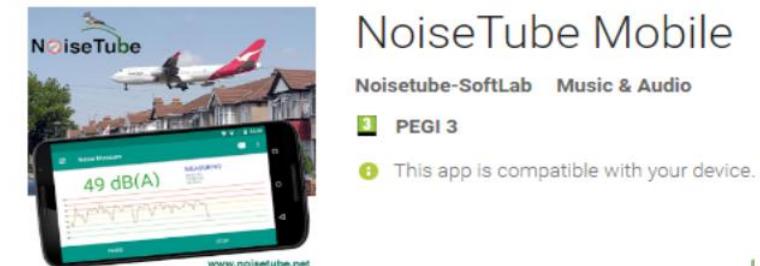


## Ride offer



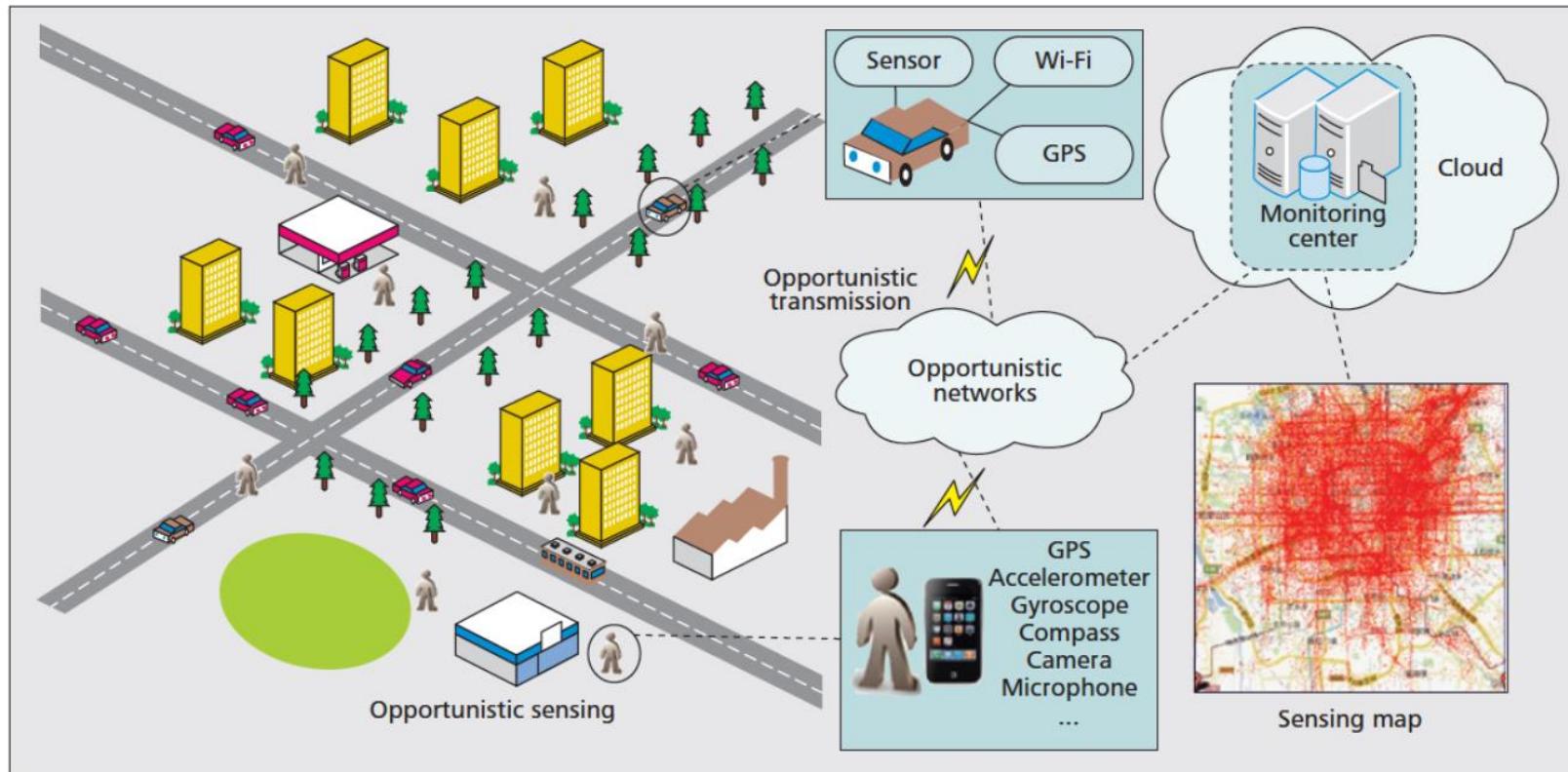
# Location-based services

- Participation in scientific research by citizens through data collection
- Scientists profit from harnessed data collection
- Citizens profit from situation awareness and community empowerment
- Many applications in the environmental sciences
- As well as in the Urban Sciences:
  - Air and noise pollution monitoring
  - Urban fauna monitoring
  - Enhancing people's security



# Crowdsensing

- Crowdsourcing of sensor data from mobile devices
- Can be participatory or opportunistic

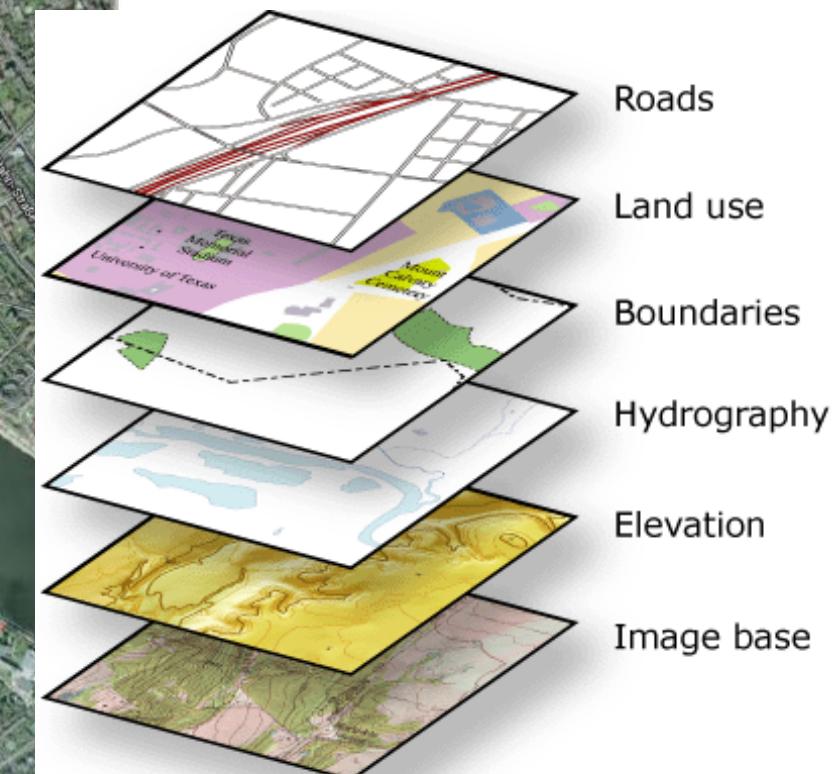
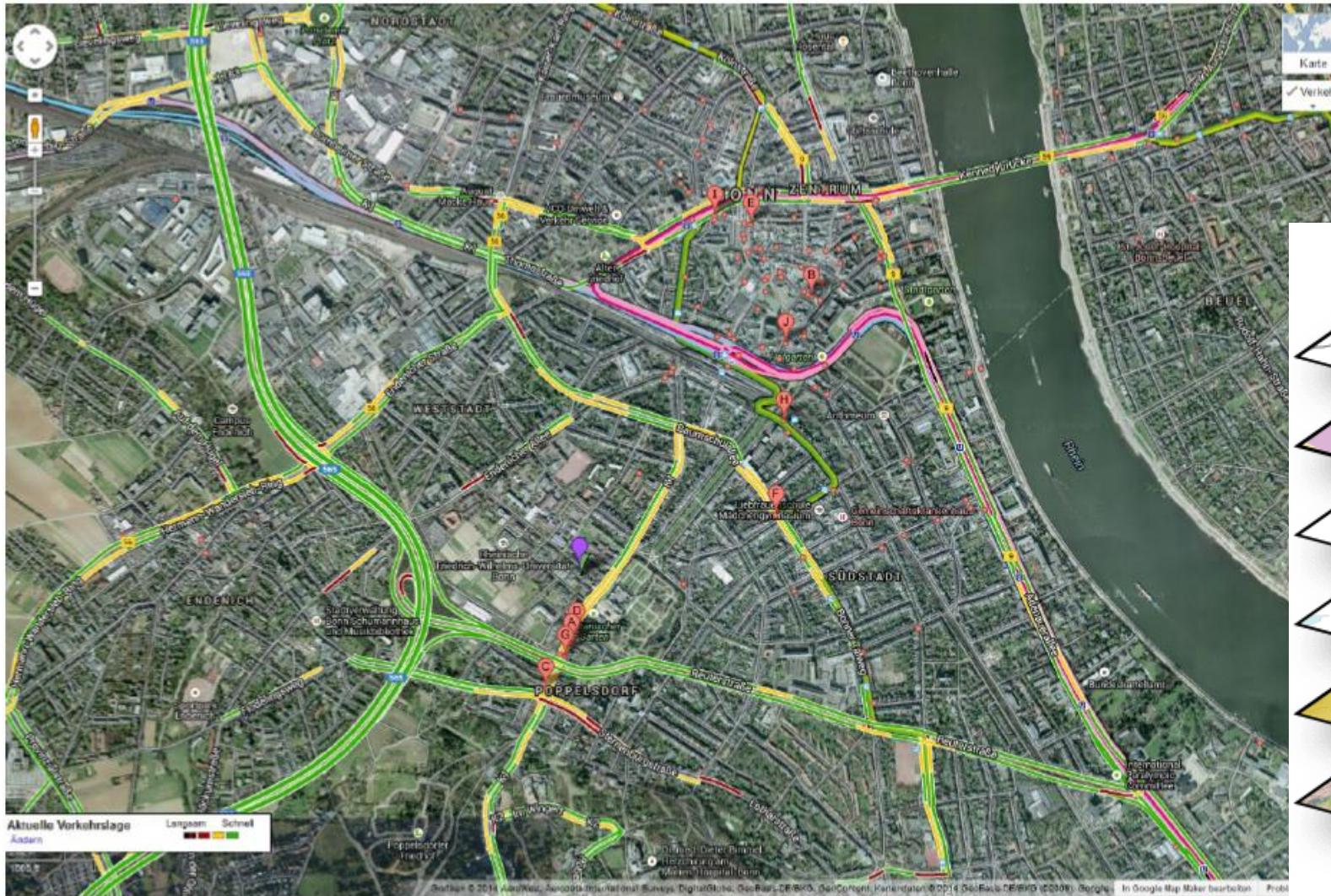


Ma, H., Zhao, D., & Yuan, P. (2014).  
*Opportunities in mobile crowd sensing.*  
*IEEE Communications Magazine*, 52(8), 29–35.

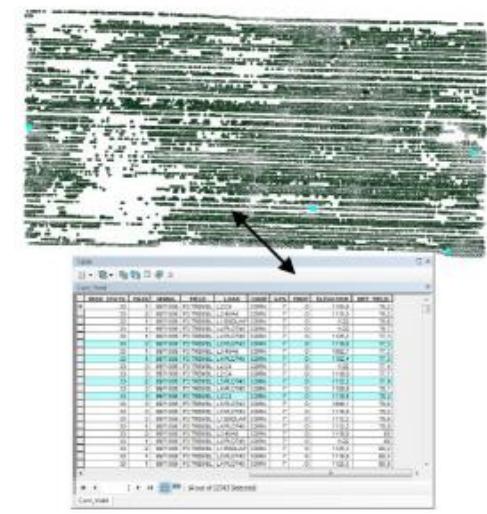
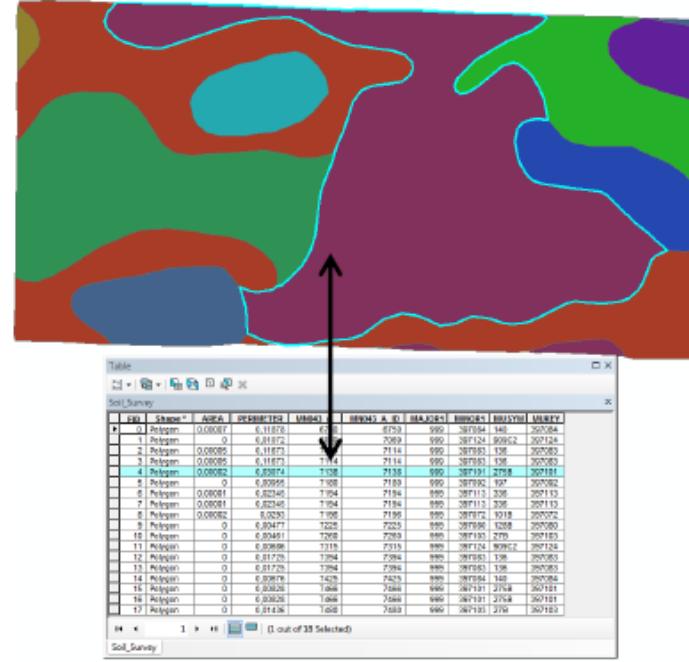
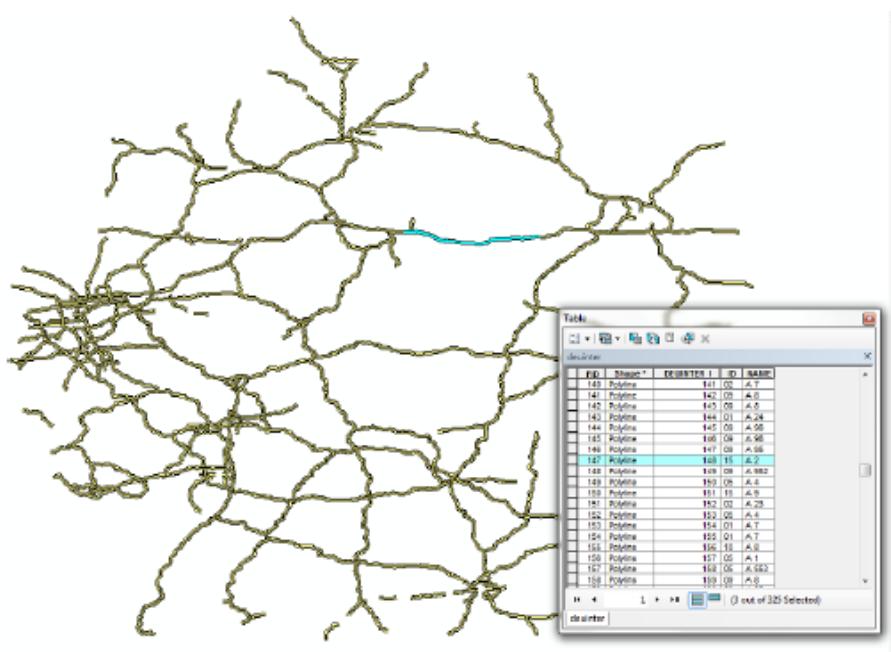
# Geographic Information Systems

Eastern Branch Elizabeth River

# Geographic Information Systems



# Geographic Information Systems



# Geographic Information Systems

- Geographic Information Systems (GIS)
  - Systems designed to encode, manage, analyze and communicate spatial data.

Extract

Manage

Analyse

Visualise

How do I get the  
data in the  
system?

How do I make  
sure the different  
data are free of  
inconsistencies  
and are relatable

How do I derive  
new information  
from the existing  
data?

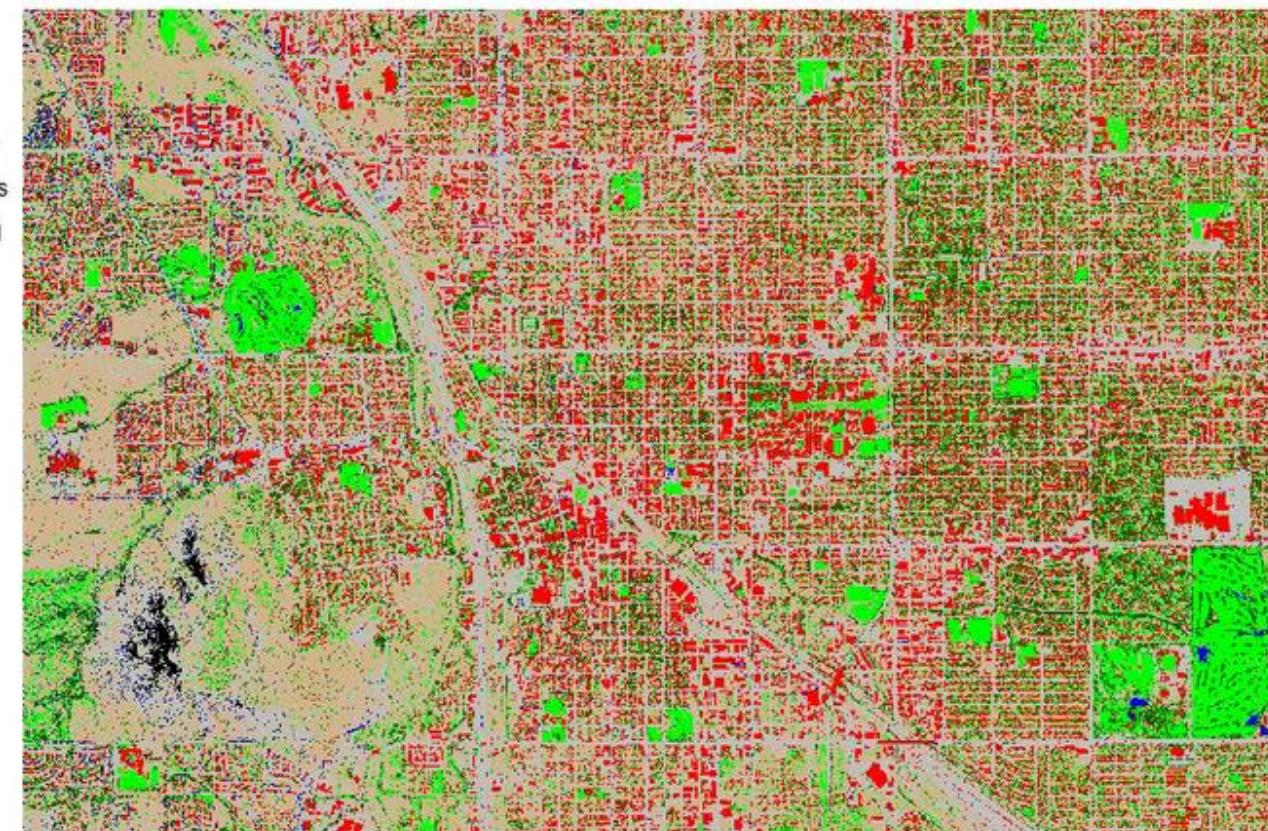
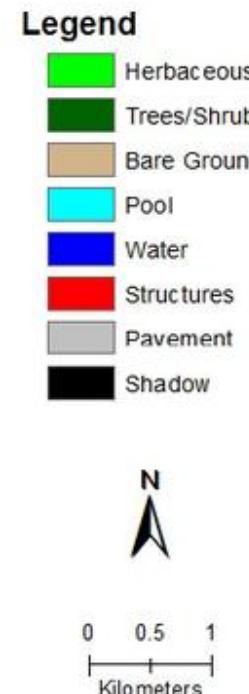
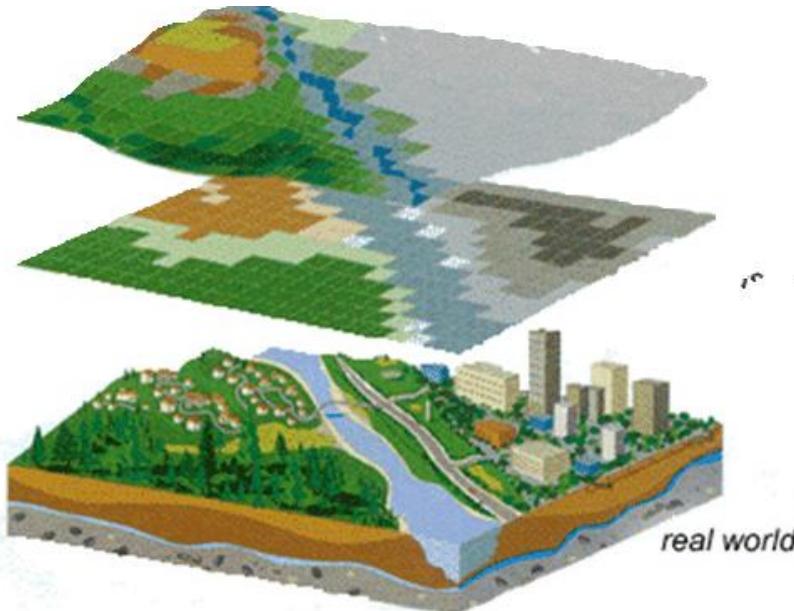
How do I present  
the data?

# GIS – Fundamental concepts

- Geographic Information Systems (GIS)
  - Systems designed to encode, manage, analyze and communicate spatial data.
- GIScience – Multidisciplinary research field dedicated to
  - Development of data structures and computational techniques for representing, analysing and communicating spatial data
  - Studying and understanding geospatial phenomena and dynamics
- Critical GIS
  - Reflects on the social implications and the potential for positive social transformation implicit on GISystems and GIScience's axioms, methods, etc.

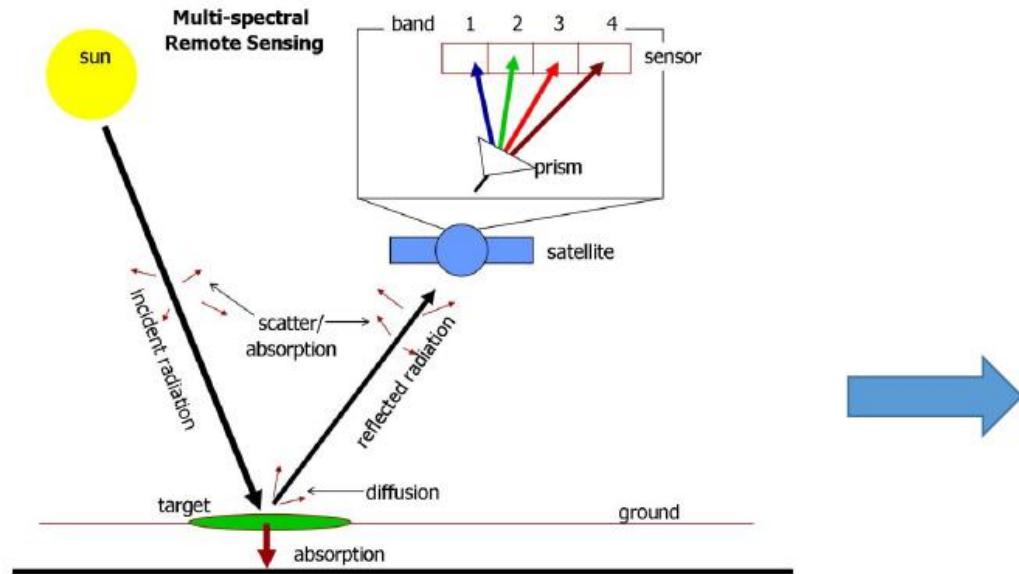
# Geographic data types within a GIS

- Raster data



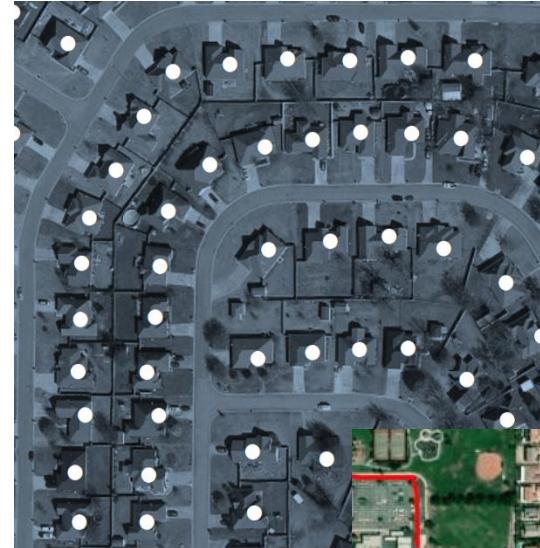
# Geographic data types within a GIS

- Raster data



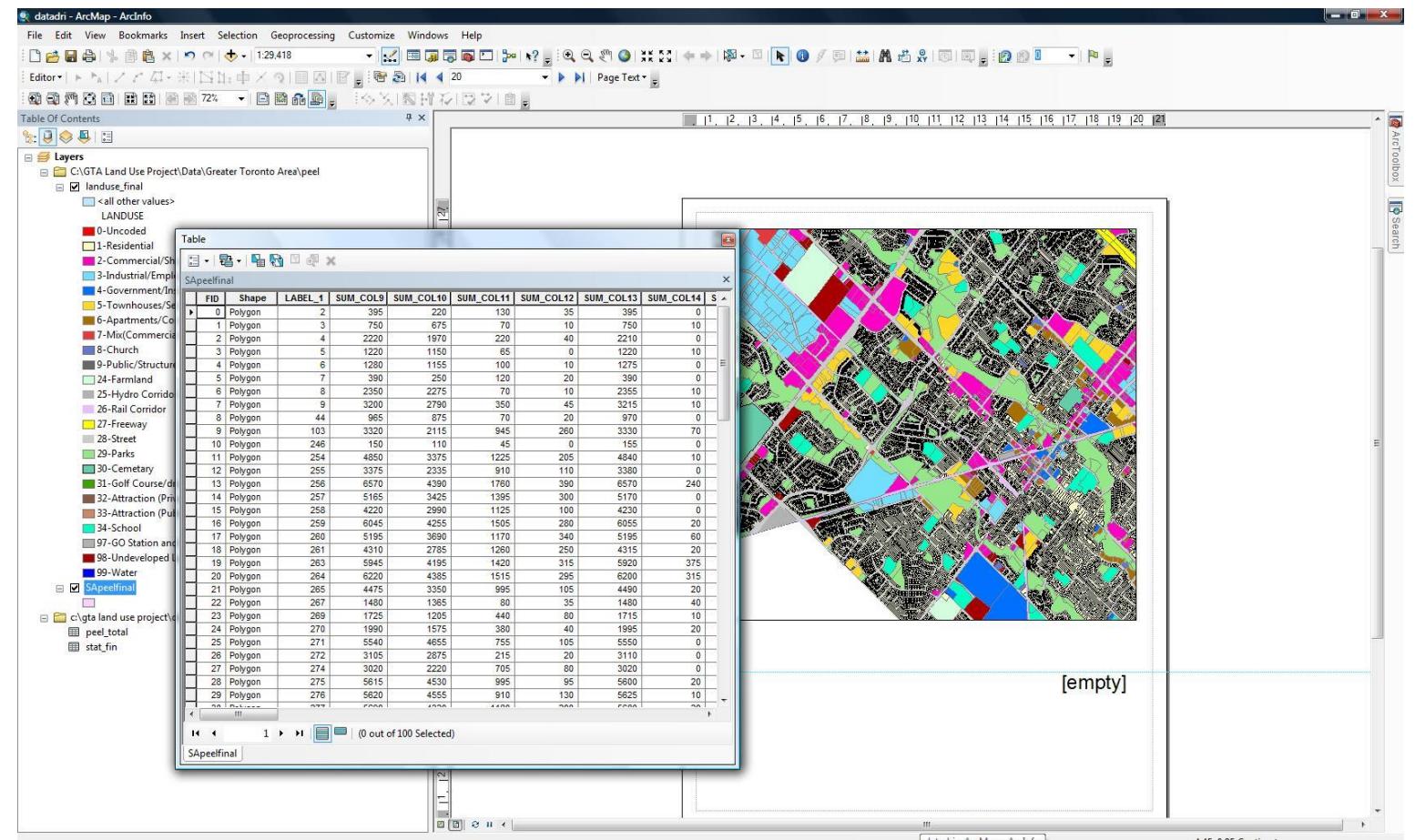
# Geographic data types within a GIS

- Vector data, i.e.
  - Points (bus stations)
  - Lines (streets, subway lines)
  - Polygons (census tracts)
- Alpha-numerical tables
  - Structured data, i.e.
    - Numerical and
    - Nominal variables
  - Unstructured, e.g.
    - Images and text



# Types of GIS – Desktop GIS

- A GIS software operating in one or a group of networked computers
- Types
  - GIS with GUI
  - Spatial databases

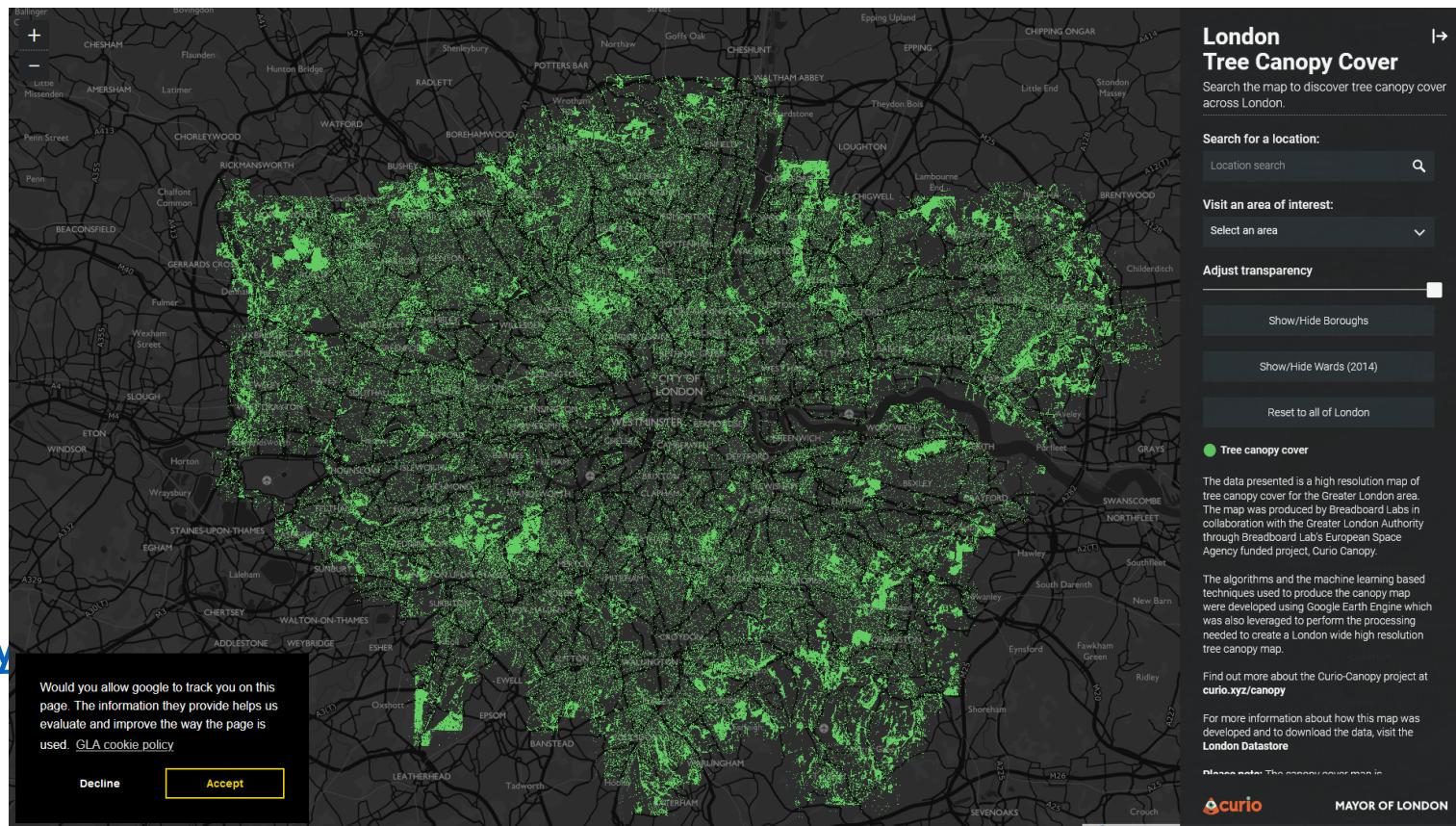


# Types of GIS – WebGIS

- A service by which consumers may choose what the map will show
- Types of WebGIS
  - Analytical web maps
  - Collaborative web maps
  - Online atlases
  - Static web maps

<https://maps.london.gov.uk/ima/>

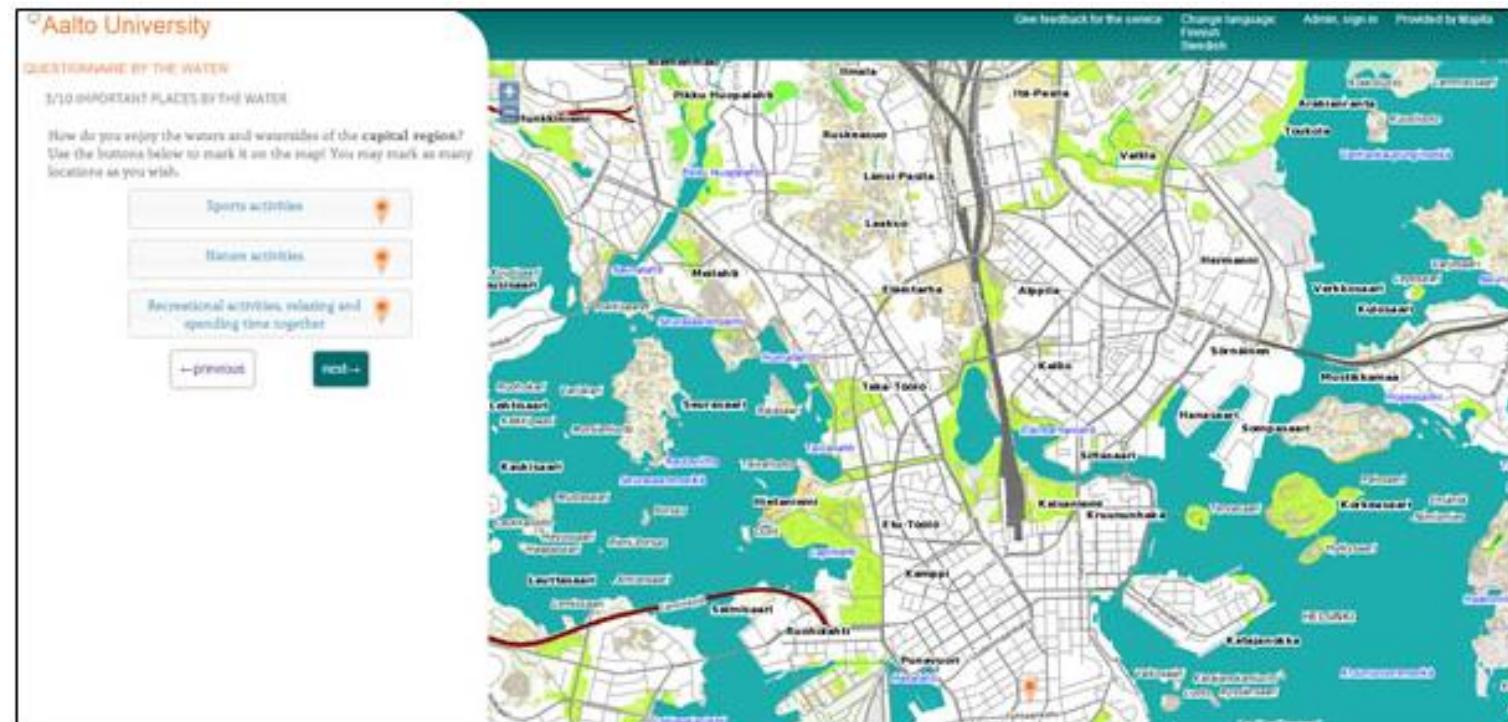
<https://maps.london.gov.uk/canopy>



# Types of GIS – PPGIS

- Public Participation GIS are basically map-based survey interface
- Useful in
  - Participatory planning
  - Participatory research
  - Grassroots movements

<https://maptionnaire.com/>

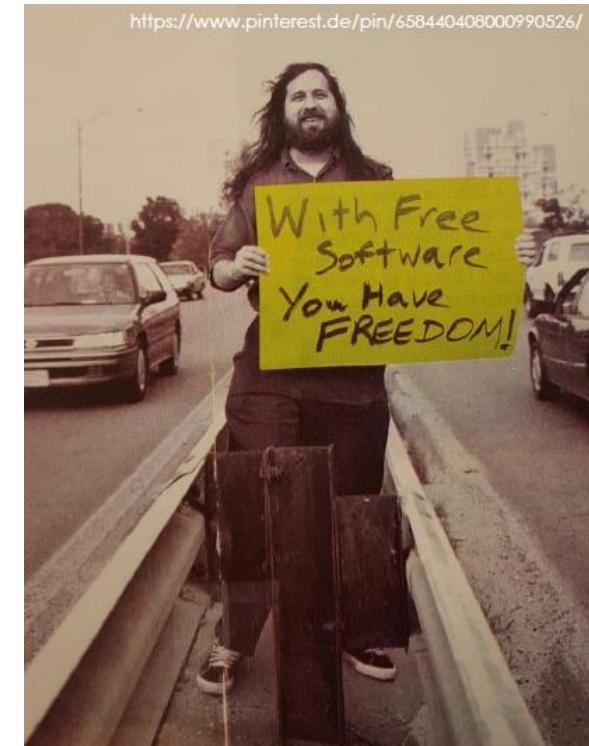


# Types of GIS – PPGIS

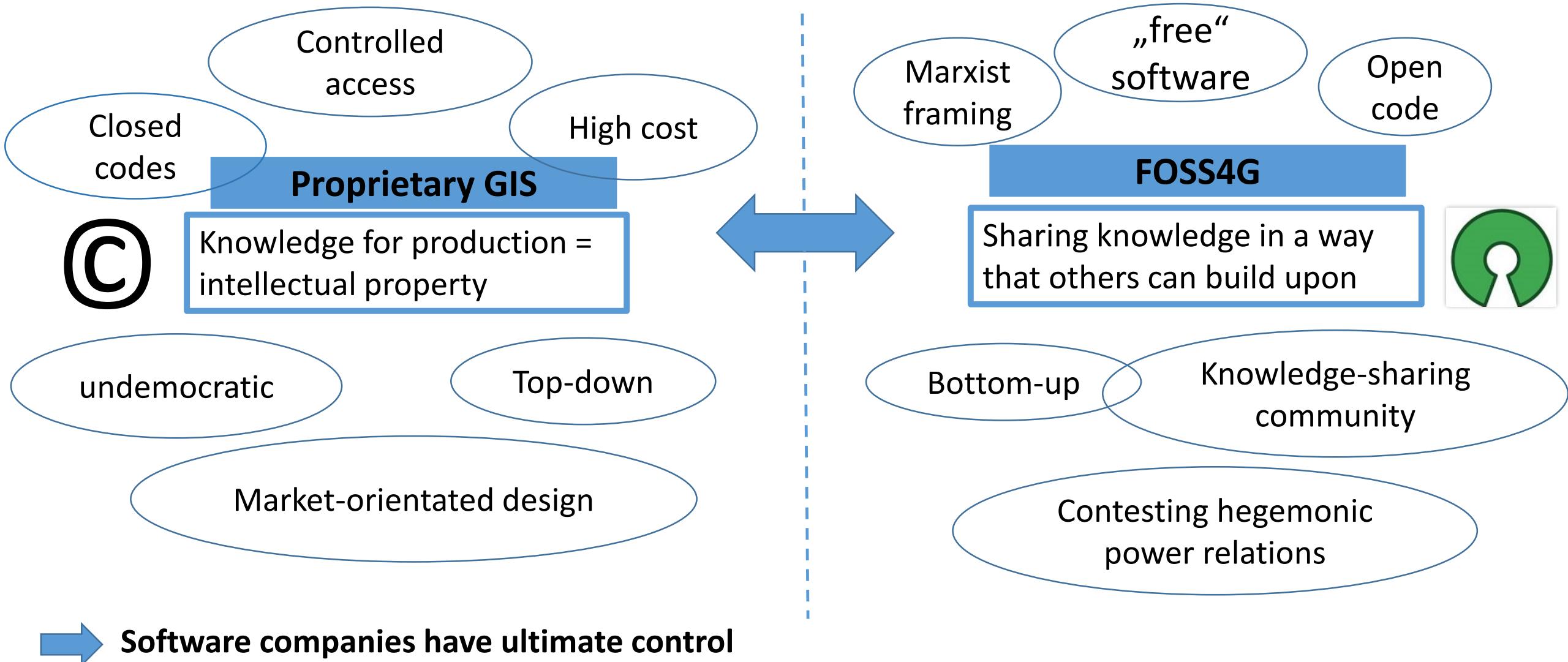
- Central goals of PPGIS
  - Equitable access to spatial data and GIS technologies
  - Incorporation of local knowledge
  - Discouragement of top-down
  - Quantitative & qualitative data
  - Represent complex social processes
  - Leverage geospatial technology to suit the needs of marginalized groups



- Annual conference of practitioners and advocates for FOSS4G (Free and Open Source Software for Geospatial)
- Organized by FOSSGIS – a organization promoting free and open software and data
- Strong exchange with the OSM community



# Proprietary vs. FOSS4G



# Geographic Information Systems (GIS)

A GIS is a **conceptualized framework** for representing, storing, managing, visualizing and analysing spatial and geographic data

Not necessarily a software type, but a **system**, possibly comprised of different independent components

# Critique to GIS

- Representing geographic phenomena using digital objects involves abstraction
- Maps are not objective representations of the Earth!



# Critique to GIS

- Surveillance and privacy
  - Who detains and controls our data?
  - Open data vs. data as a strategic asset
- Power relations
  - Whose agenda is behind digital maps applications and spatial media?
- Simplicity of representation
  - Objects with attributes are caricatures of reality
  - Geography cannot be layered
  - Boolean logic: one feature -> one class
  - Omits personalized views from the ground

