Big Spatial Data

Data Science and Society

Universiteit Utrecht

18 October 2018

esri Nederland

THE SCIENCE OF WHERE™

Go to **www.menti.com** and use the code **12 62 27**

Mentimeter

Did you create data today? If so how?

- Nope, just garbage
- No
- Streaming music
- Yep, I checked in with my OV Chipcard
- Instagram
- WhatsApp, email, discord, OV
- No
- WhatsApp
- Listened to a podcast
- Ran models
- maybe movement data, but I just got up
- Yes, location.

Pause scroll

34

Go to **www.menti.com** and use the code **12 62 27**

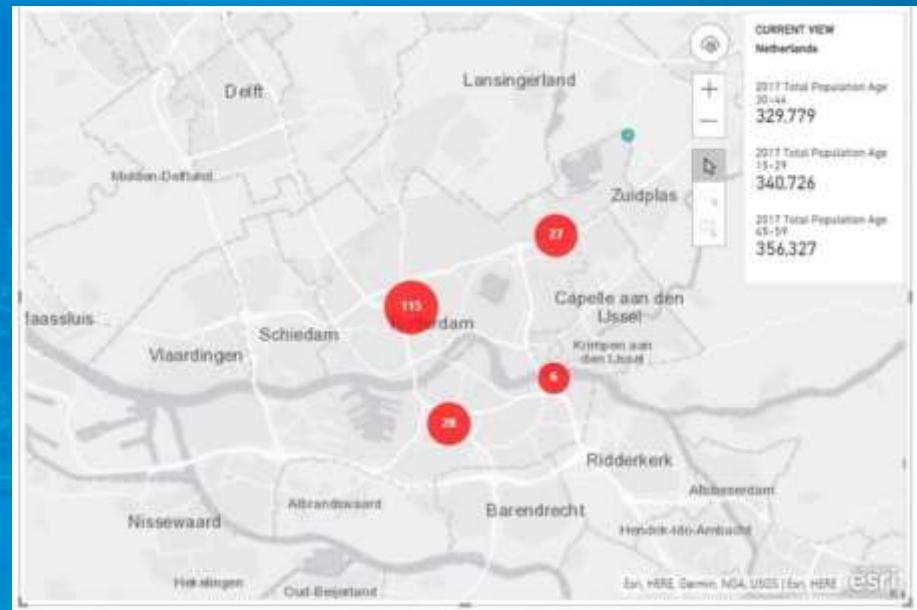
Did you create spatial data today? If so how?

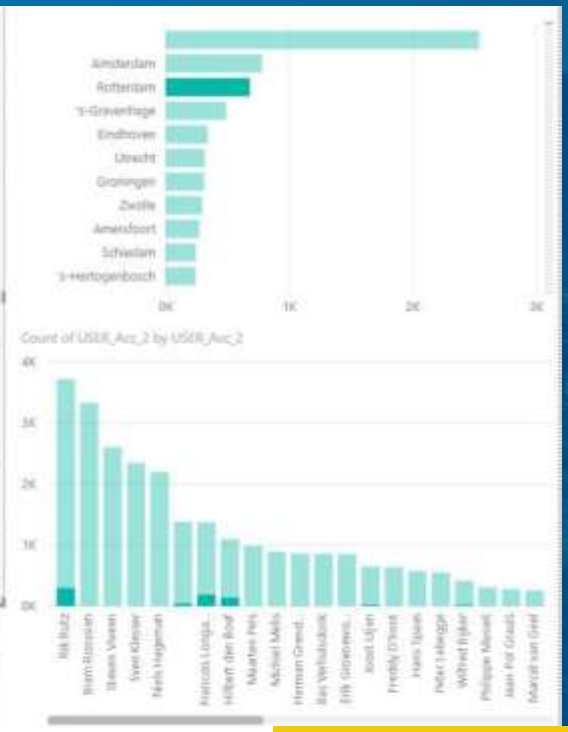
Mentimeter

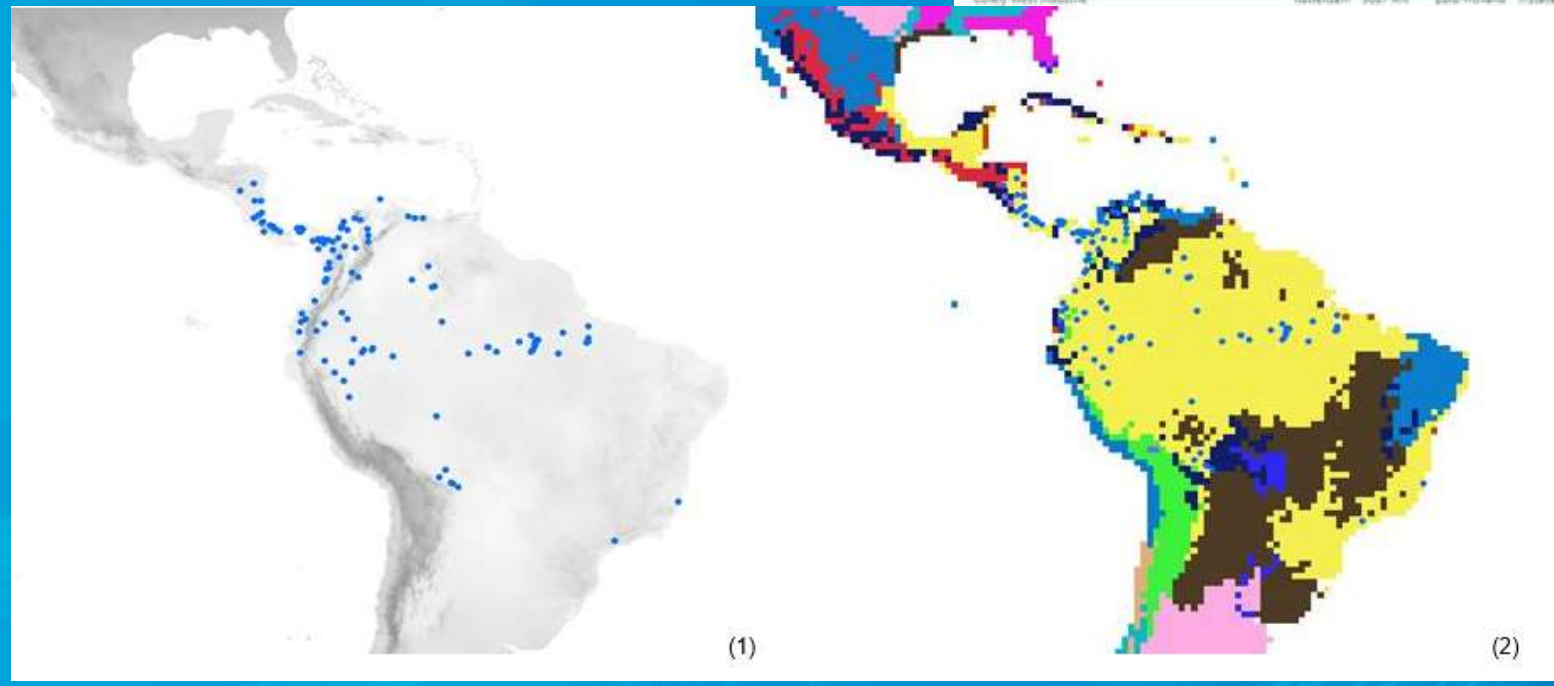
Mapillary	No	Yes, by making a video
no	Find my iphone	No
No	No	Google maps?
Gps tracked probably	Android	Yes, took video of surroundings for mapillary, but uploading was

Pause scroll

22


Delft
Lansingerland
Zuidplas
Capelle aan den IJssel
Schiedam
Vlaardingen
Ridderkerk
Barendrecht
Nissewaard
Albrandswaard
CURRENT VIEW
Netherlands
329.779
340.726
356.327


Amsterdam
Rotterdam
Utrecht
Groningen
Zwolle
Schiedam
Count of USER_Acc_2 by USER_Acc_2



(1)

(2)

Power BI

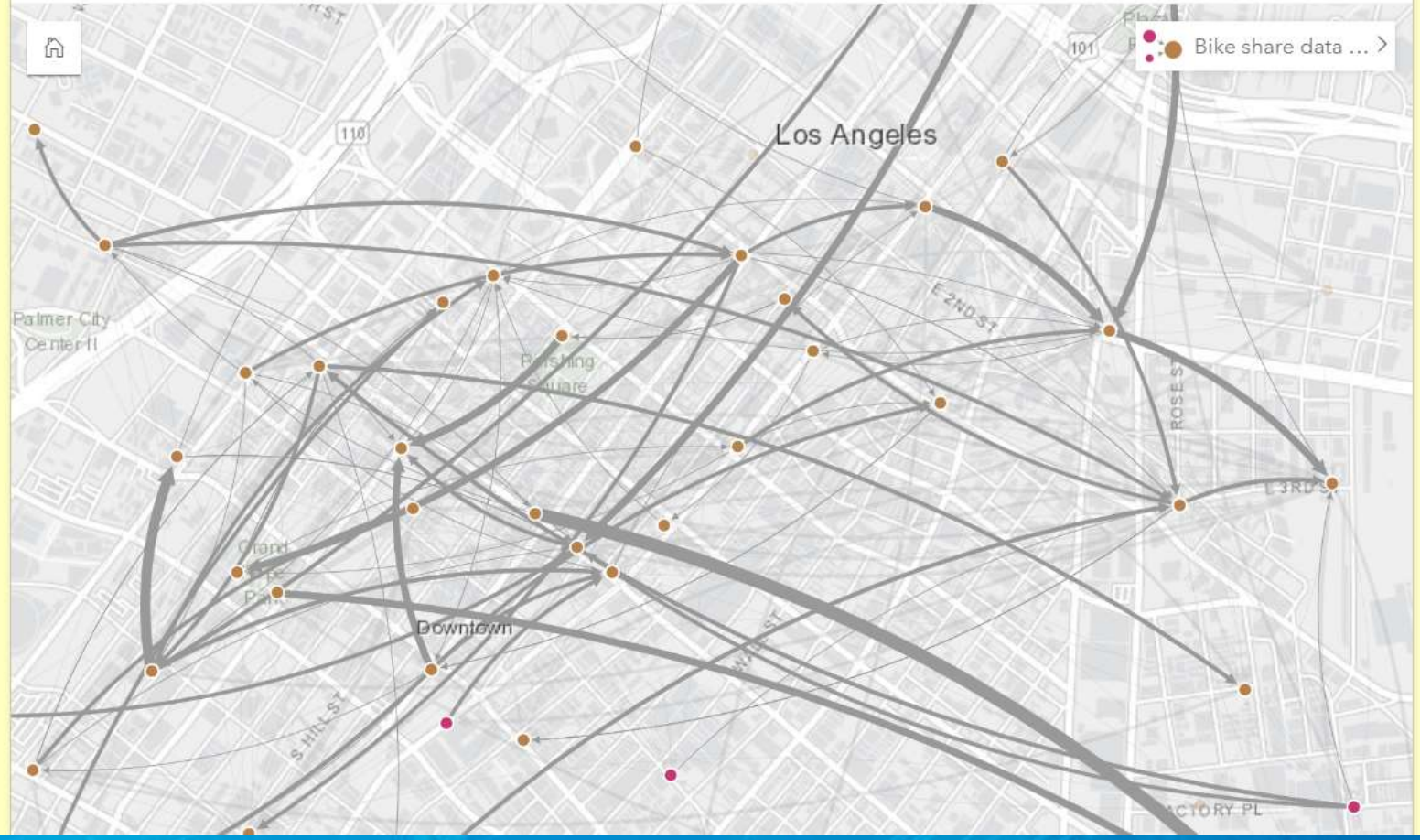
Los Angeles Metro Bike Share Trip Data

Help

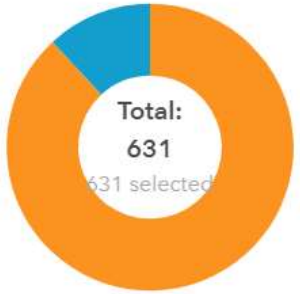
Jan Willem

Explore the data | Network Analysis | Bike sharing in a day

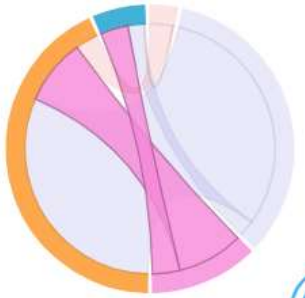
Link charts - weakly geographic. duration per trip



Number of trips per route cate...



Route category per passholde...



Totl duration (min) per passhol...

Learning goals

Understand the role of data science and its societal impact

Recognise the knowledge discovery processes in applied data science

Identify trends and developments in big data technologies

Apply selected big data technologies to solve real-world problems

Learning goals

Understand the role of data science and its societal impact

Recognise the knowledge discovery processes in applied data science

Identify trends and developments in big data technologies

Apply selected big data technologies to solve real-world problems

Assignment

1. Collect data with Mapillary

Capture (at least) 100 images in a stream of images. Start at the UU Campus and focus on a poorly mapped area. You are free the explore other areas and topics of interest as well.

The one student with the most images captured will be rewarded.

2. Read an article on Machine Learning and ArcGIS.

http://www.esri.com/~/media/Files/Pdfs/news/arcuser/0518/Machine%20Learning%20in%20ArcGIS.pdf

Question: Predictive analysis in combination with spatial analysis can assist in finding hot spots for traffic accidents, as illustrated in the article. In this example are traffic accidents predicted based on a number of factors. Can you think of any additional factor that will make more accurate analysis, and one that creates a less accurate analysis?

Assignment

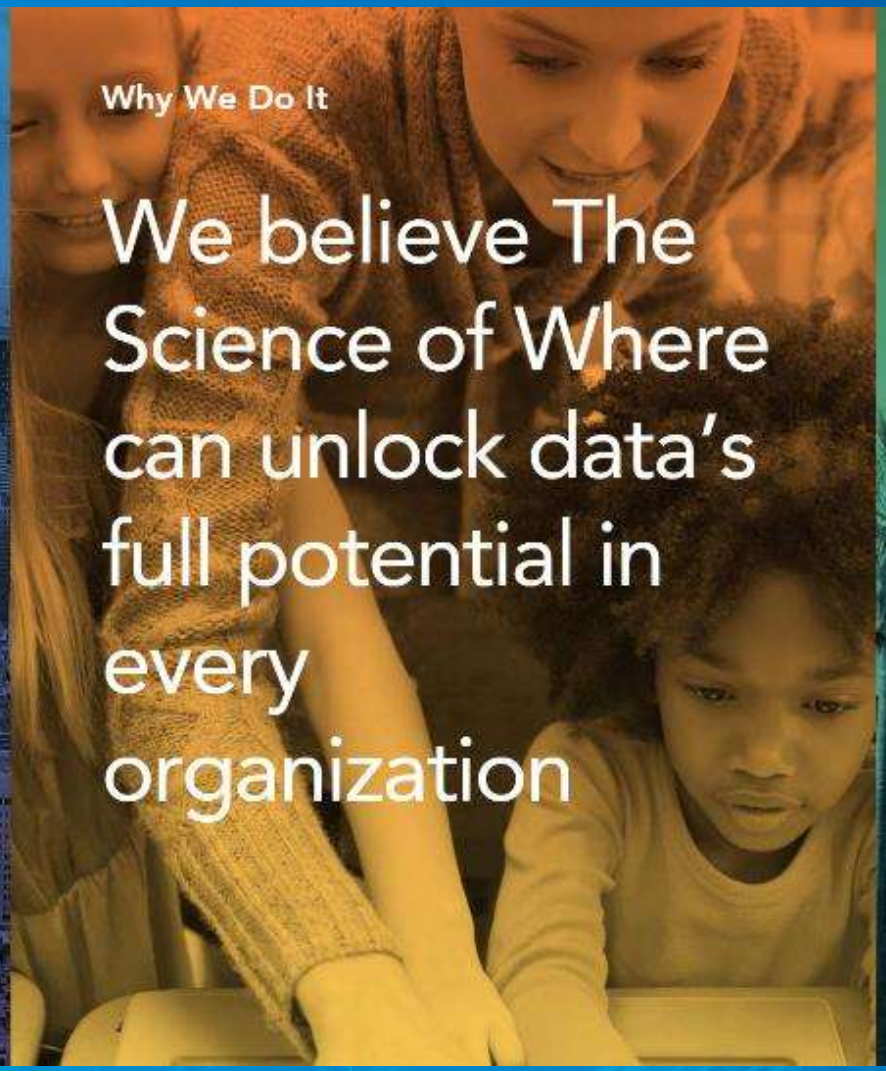
3. Review a storymap on Unlocking Information from Imagery, http://arcg.is/ieHSj

Two questions about this storymap:

Object recognition can be helpful with data acquisition on different zoom levels. Can you think of two examples from different zoom levels were object recognition in combination with neural networks can have meaningful results?

Deep learning can facilitate raster classification, but still can be limited by manual choices such as amount of categories. Can you think of any risks on your analysis that stems from these limitations?


What We Do
We build ArcGIS, the world's most powerful mapping and spatial analytics software


Why We Do It
We believe The Science of Where can unlock data's full potential in every organization


Who We Are
We are the global market leader in GIS and have helped customers improve results since 1969

Programme

Lecture: The spatial perspective on data

Discussion: Spatial applications of big data

Break

Lecture: Big spatial data in society

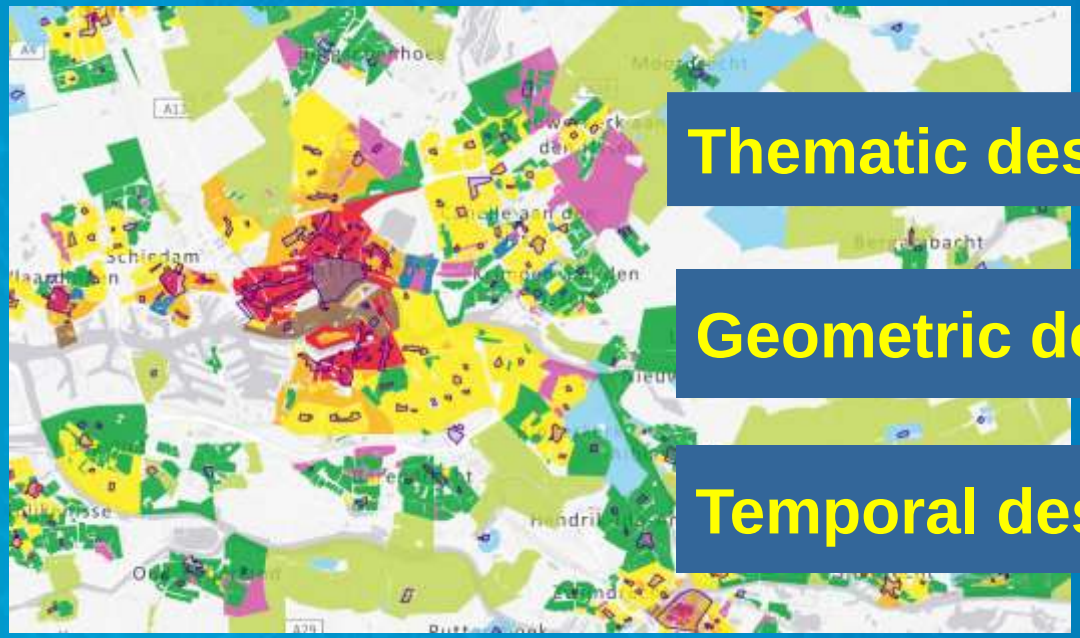
Reflection

The Spatial perspective on data

esri Nederland
THE SCIENCE OF WHERE™

http://radar.zhaw.ch/worldwide.html

Describing reality



Thematic description (what)

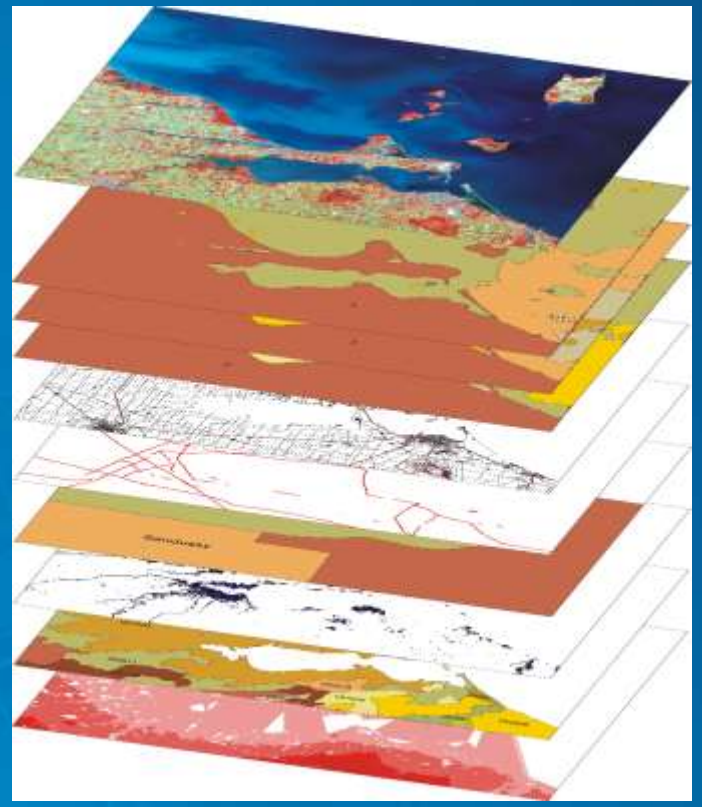
Geometric description (where)

Temporal description (when)

How GIS* works

Data is stored as a collection of thematic layers

….



esri Nederland

Two data-models

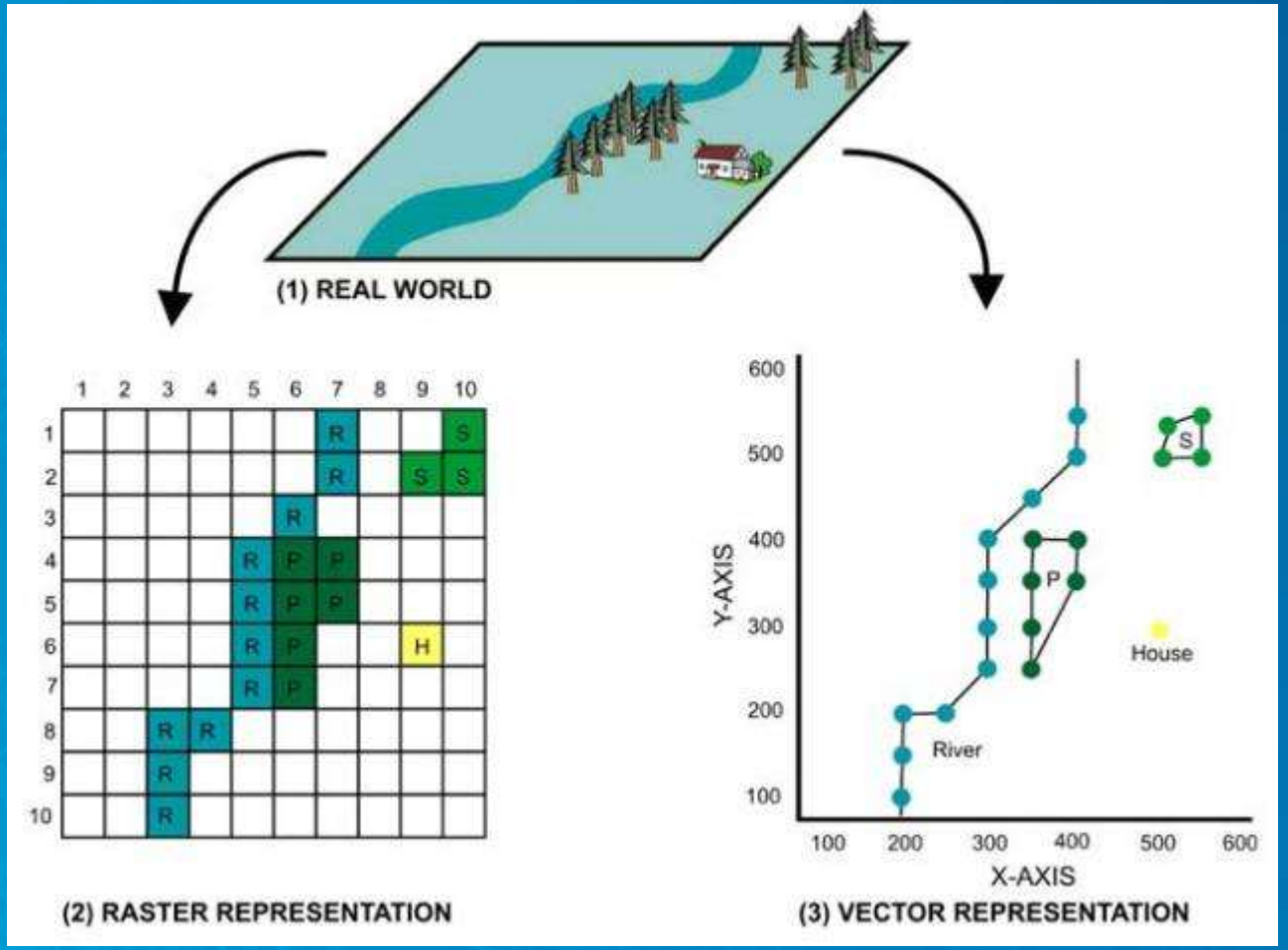
Two components:

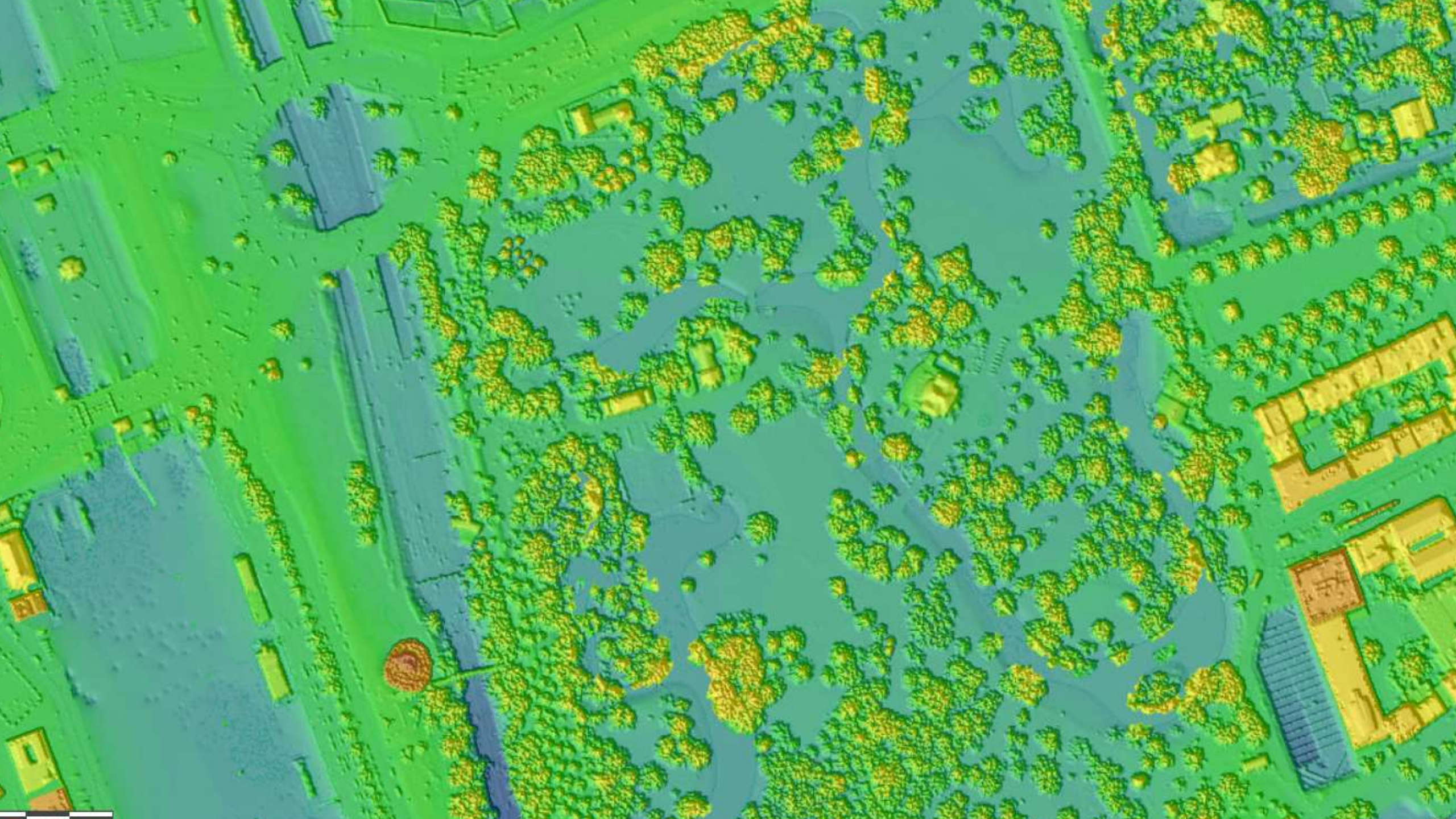
- Geometric data
- Attribute data

Two representation models:

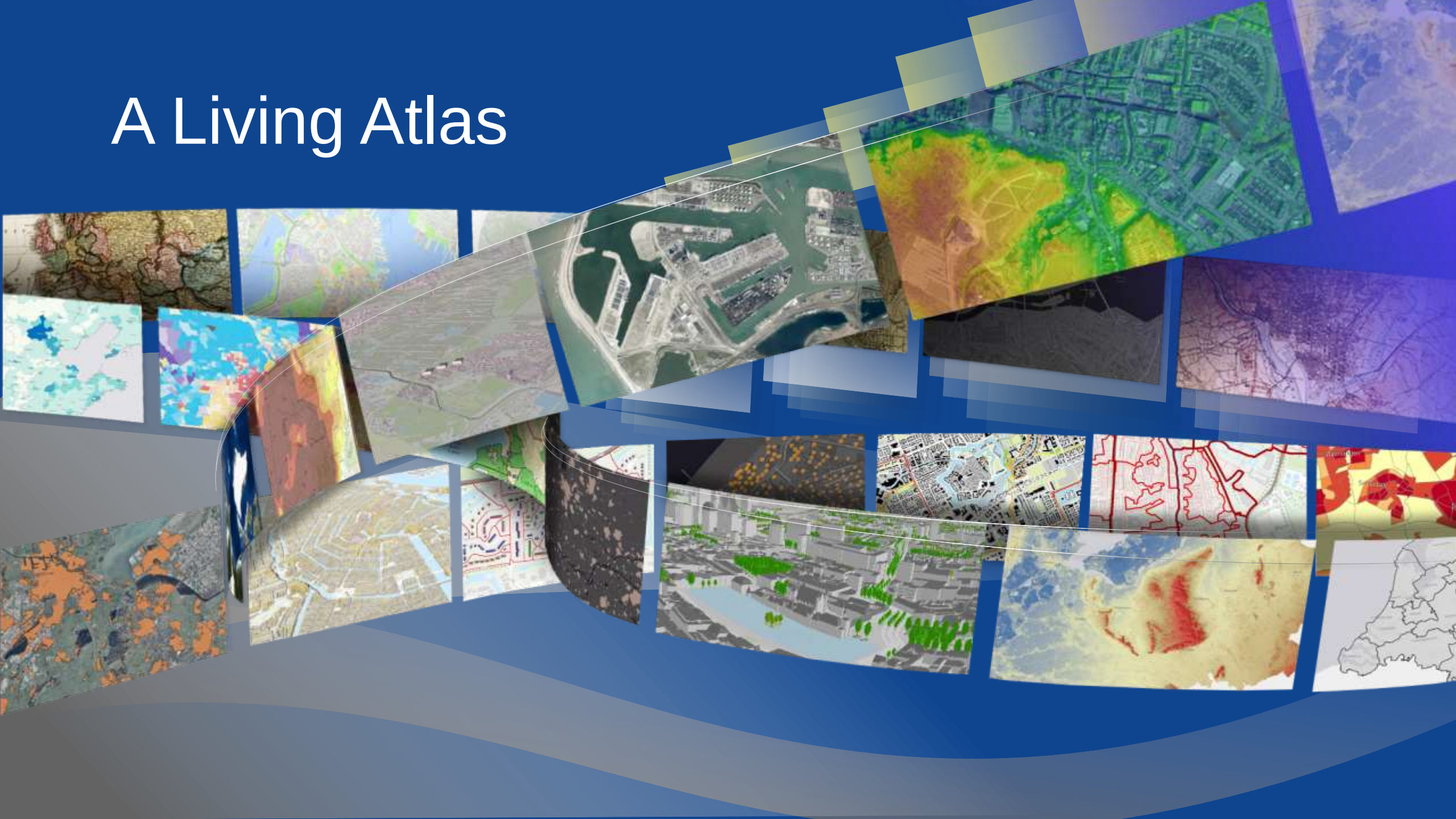
- **Vector data model,** discrete objects (e.g. Clear boundaries, trees, railroad tracks)

- **Raster data model**; Continuous phenomena (e.g. temperature, slope, elevation etc.)


(1) REAL WORLD
1 2 3 4 5 6 7 8 9 10
1
2
3
4
5
6
7
8
9
10
R
S
P
H
(2) RASTER REPRESENTATION
600
500
400
300
200
100
Y-AXIS
100 200 300 400 500 600
X-AXIS
River
House
(3) VECTOR REPRESENTATION



A Living Atlas



ArcGIS Online

Over 1 Million Public Items, ~200 Tb of Data

160 Million Map Requests by 1.6 Million Users ***Per Day***

4-5 Billion Map Tile Requests ***Per Month***

Data management and integration

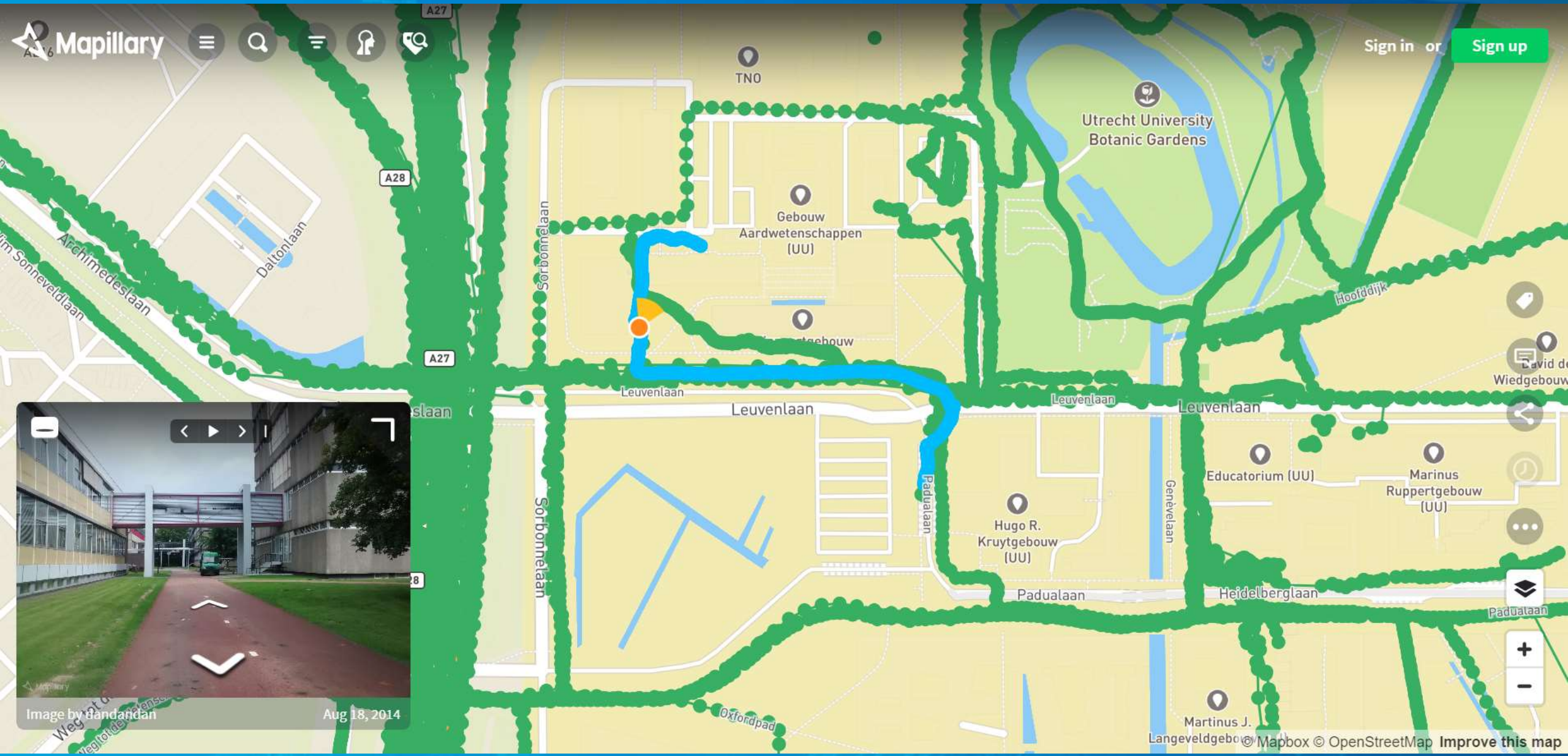
Action

Visualization and mapping

Decision-making

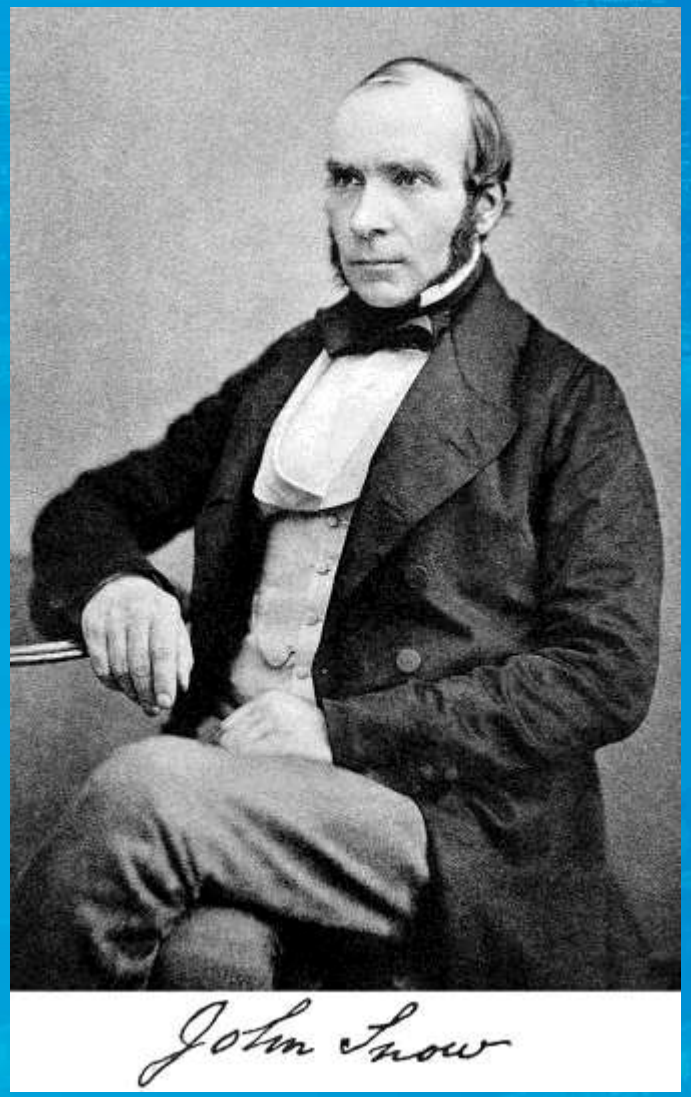
Analysis and modeling

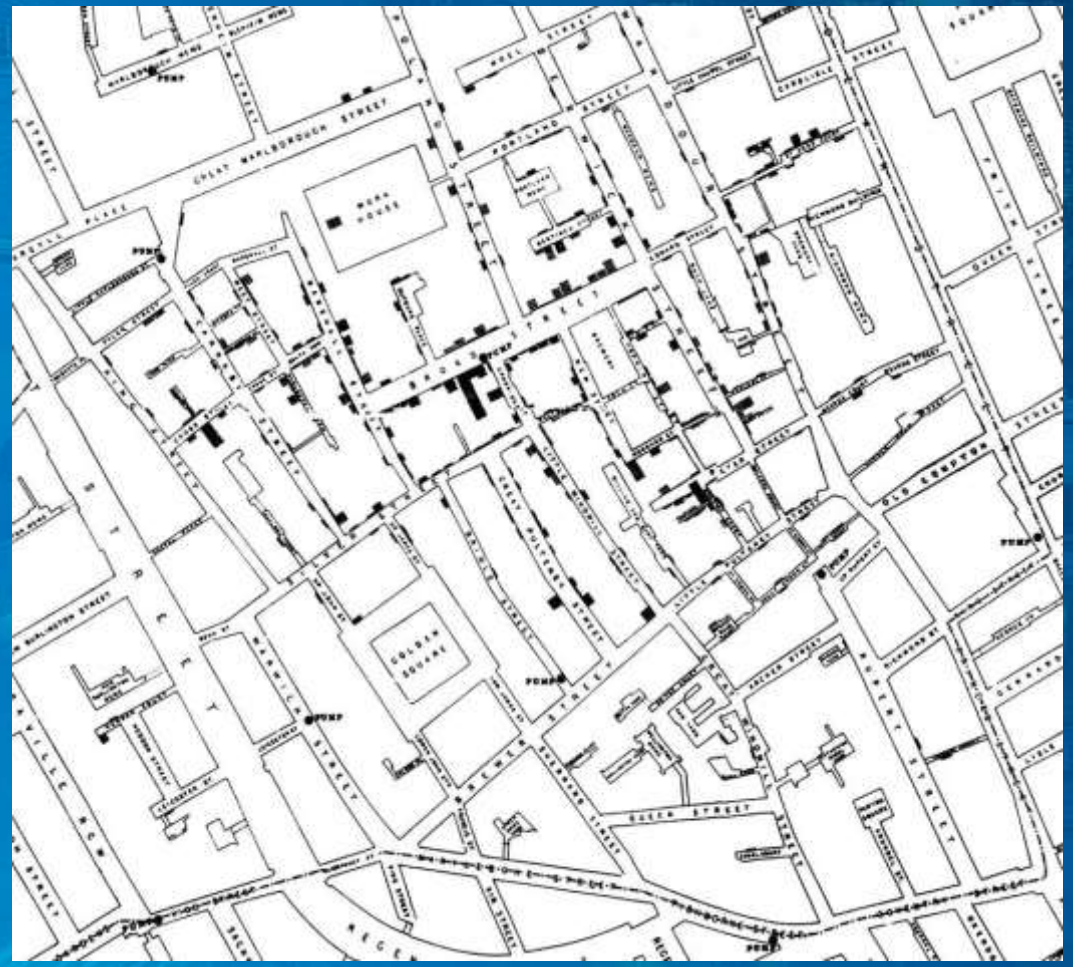
Planning and design


Mapillary
Sign in or Sign up
TNO
Utrecht University Botanic Gardens
Gebouw Aardwetenschappen (UU)
A27
A28
Archimedeslaan
Daltonlaan
Sorbonnelaan
Leuvenlaan
Hoofddijk
Educatorium (UU)
Marinus Ruppertgebouw (UU)
Hugo R. Kruytgebouw (UU)
Padualaan
Genèvelaan
Heidelberglaan
Oxfordpad
Martinus J. Langeveldgebouw
Image by dandandan
Aug 18, 2014
© Mapbox © OpenStreetMap
Improve this map

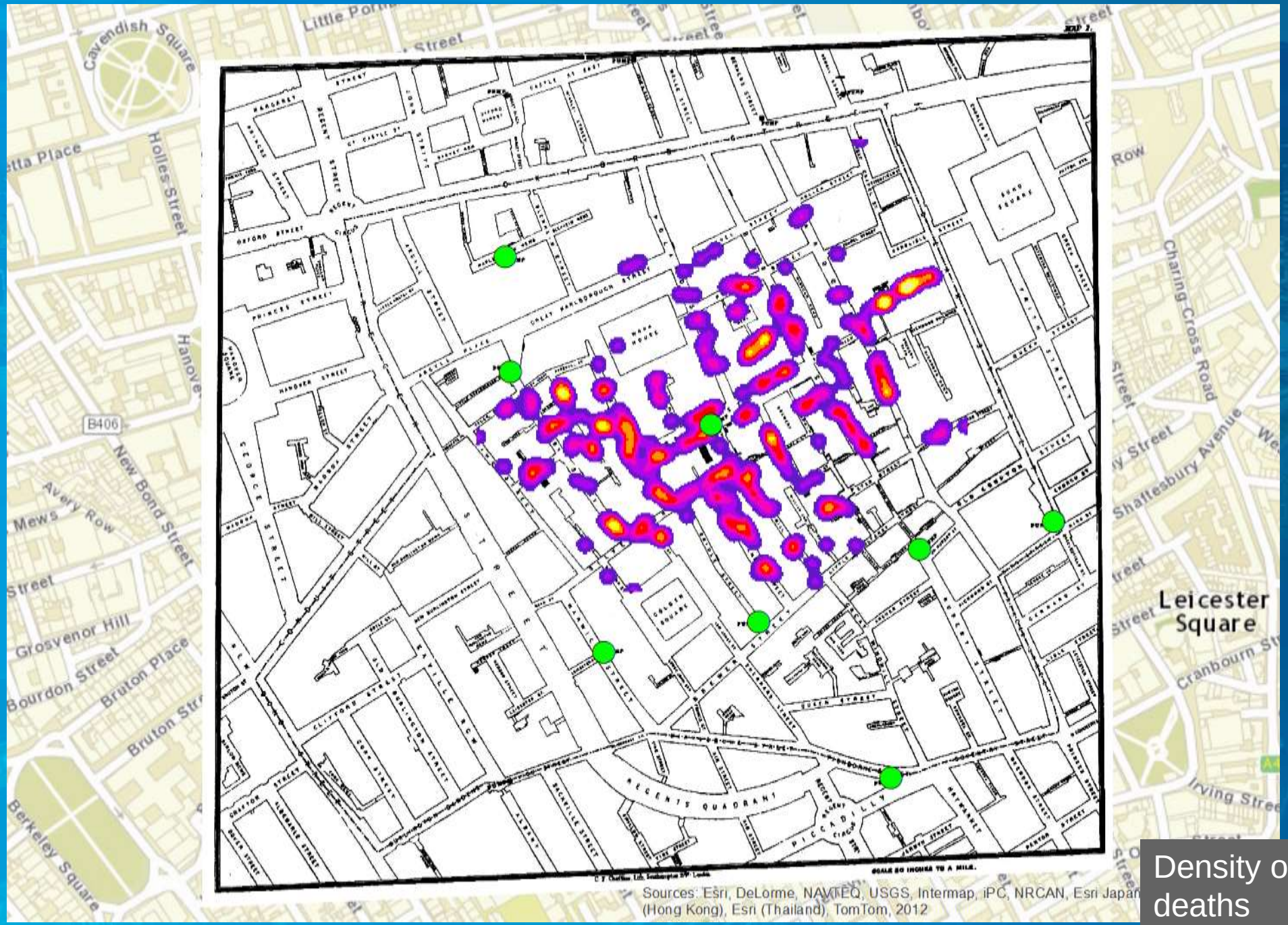

esri Nederland
THE SCIENCE OF WHERE

John Snow

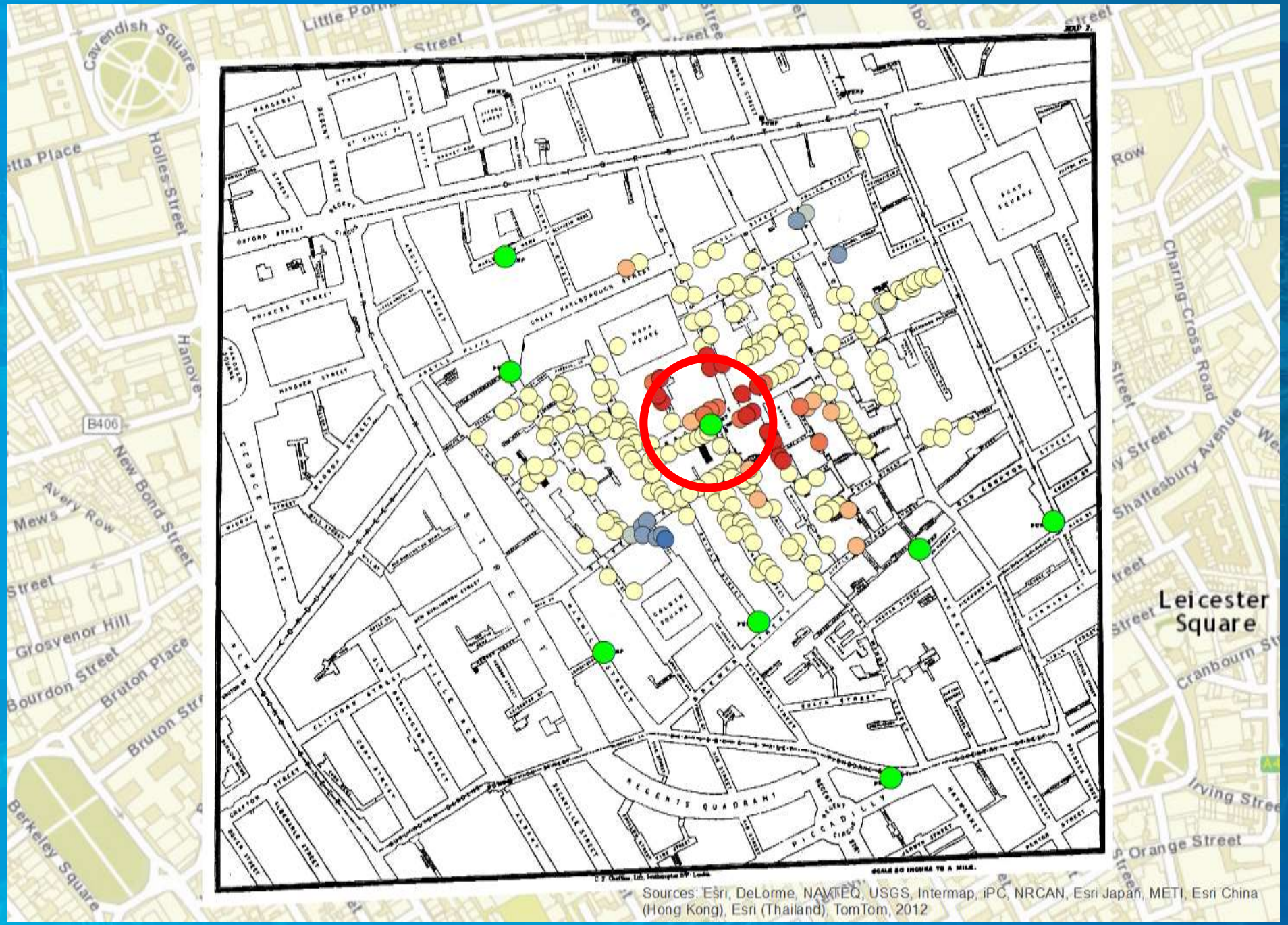




Locations of water pumps and deaths



Density of location of deaths

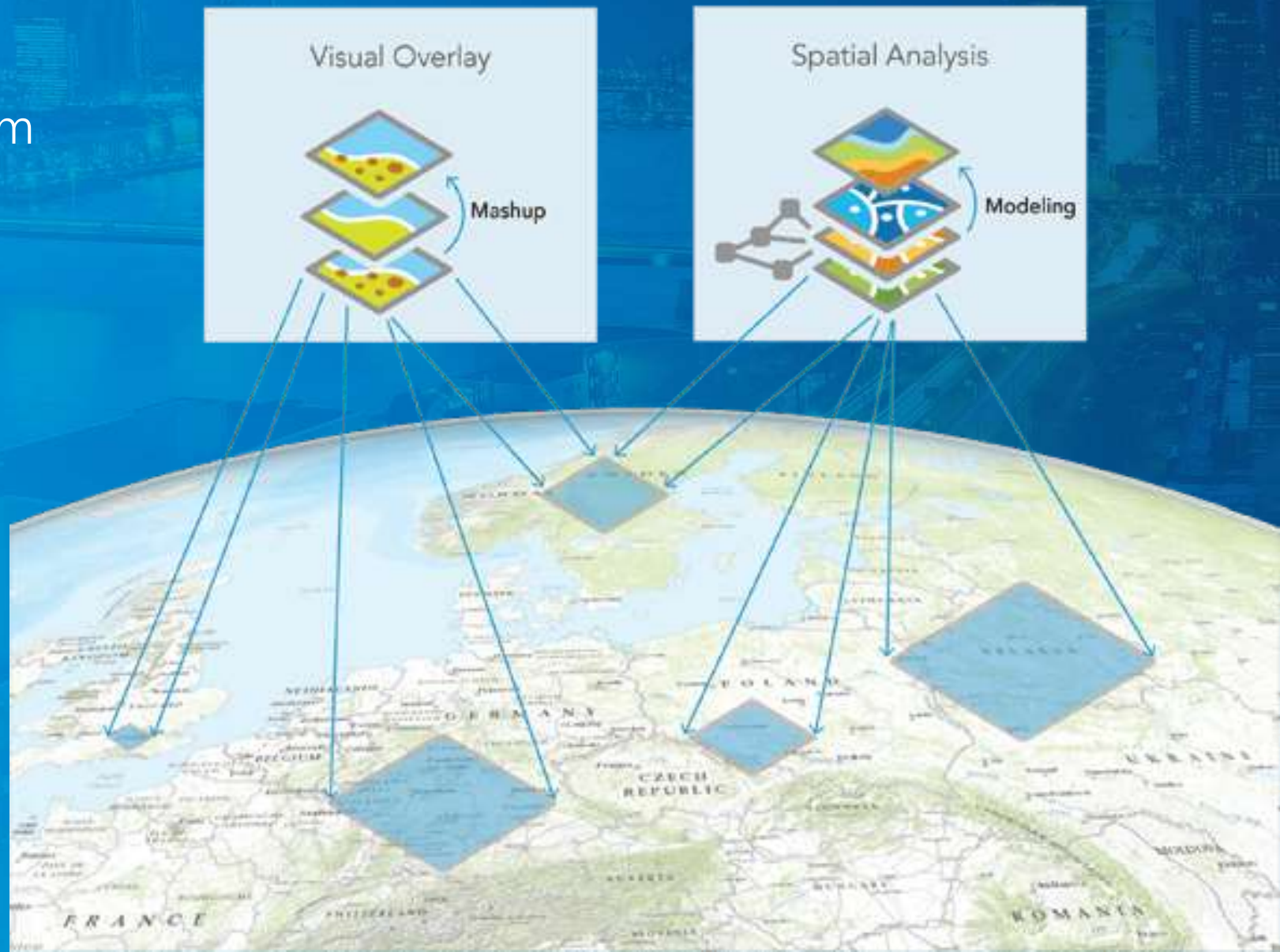

Cavendish Square
Holles Street
B406
New Bond Street
Avery Row
Grosvenor Hill
Bourdon Street
Bruton Place
Berkeley Square
Hanover
Leicester Square
Charing Cross Road
Shaftesbury Avenue
Cranbourn St
Orange Street
Irving Street
Sources: Esri, DeLorme, NAVTEQ, USGS, Intermap, iPC, NRCAN, Esri Japan, METI, Esri China (Hong Kong), Esri (Thailand), TomTom, 2012

What is GIS?

A geographic information system (GIS) lets us:

Visualize
Question
Analyze
Interpret

our data to understand **relationships**, **patterns**, and **trends** to obtain location intelligence.



Statistical Tools in ArcGIS

Classification

- Maximum Likelihood Classification
- Random Trees
- Support Vector Machine

Clustering

- Spatially Constrained Multivariate Clustering
- Density-based Clustering
- Image Segmentation
- Hot Spot Analysis
- Cluster and Outlier Analysis
- Space Time Pattern Mining

Prediction

- Empirical Bayesian Kriging
- Areal Interpolation
- EBK Regression Prediction
- Ordinary Least Squares Regression and Exploratory Regression
- Geographically Weighted Regression

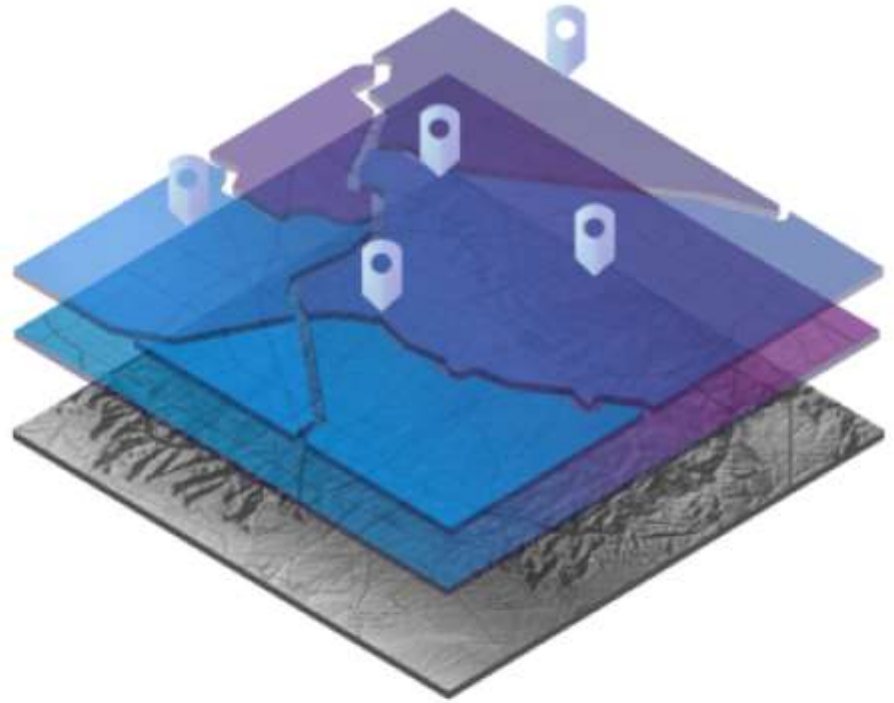
People are saying…

Gartner, 2017:

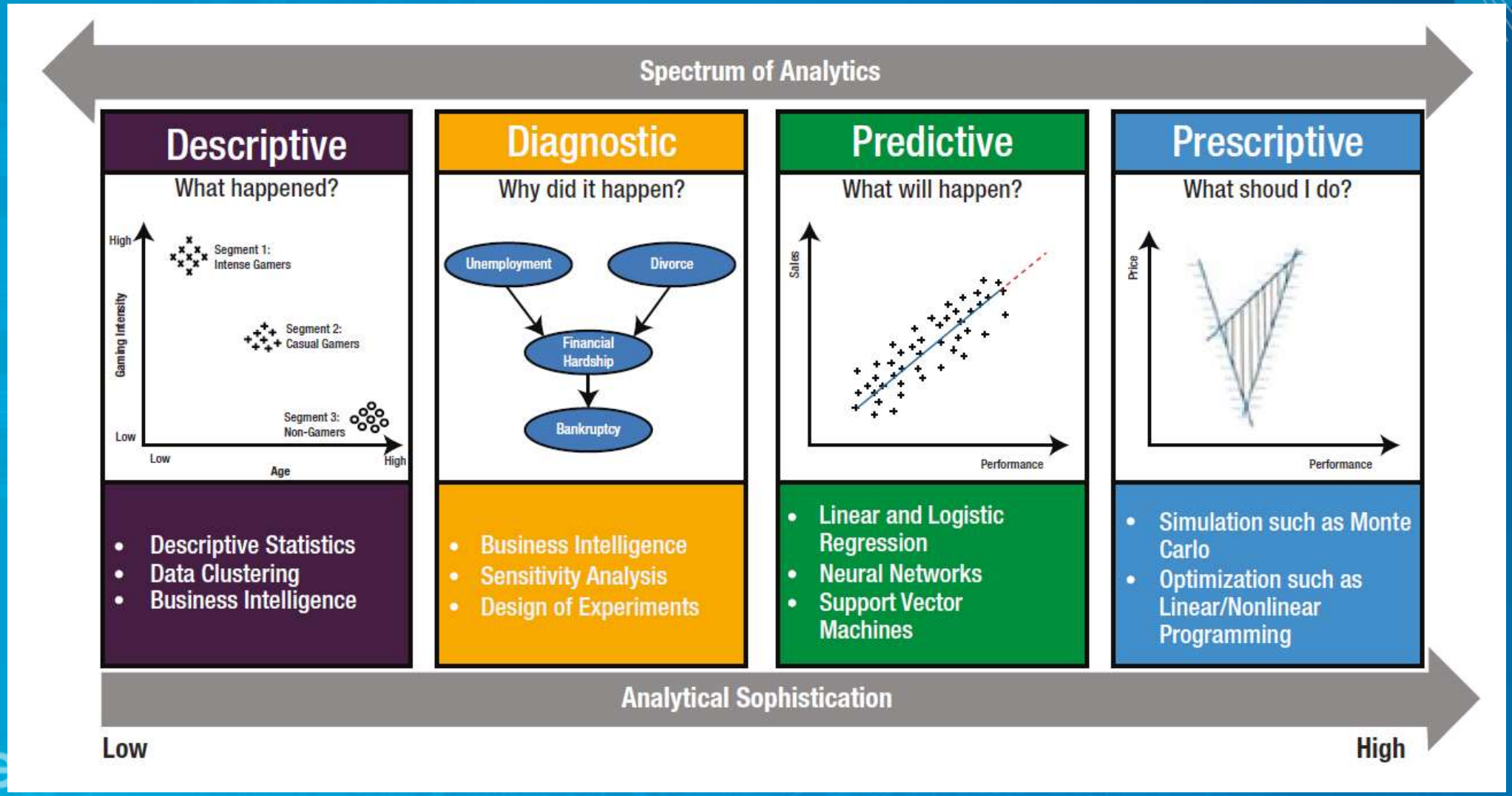
"Properly **analyzing location can provide insights** that support and improve decision making in everything from marketing to supply chain logistics and operations."

"Gaining access to **location-based streams of insight on consumers will be critical** for organizations striving to become digital businesses."

"**New tools and access to data** are now allowing the power of location intelligence to be unleashed across many more business areas and to a much broader base of users."

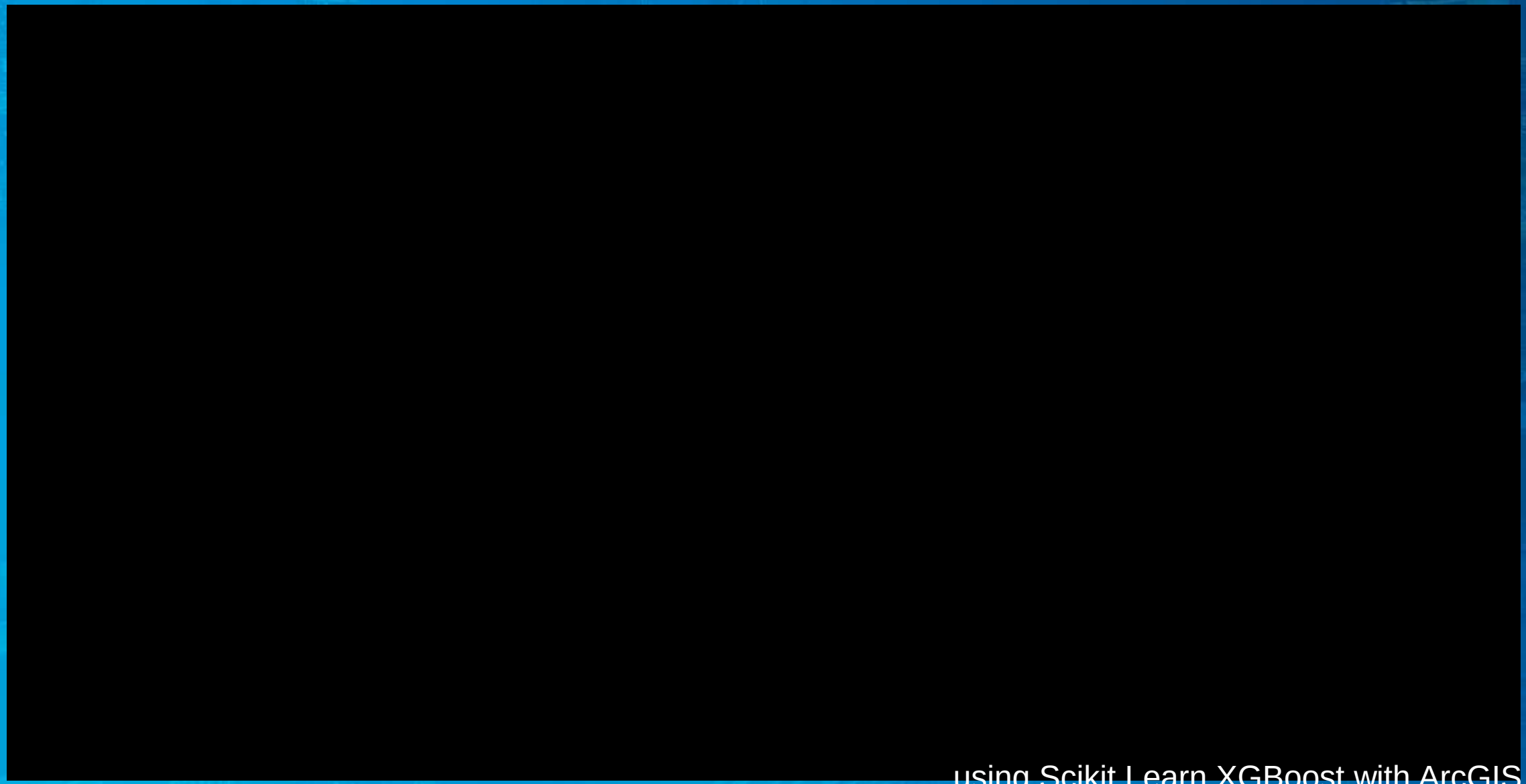


How do we use platforms for Analysis?



(source: courtesy of Brian Hilton)

Accidents Probability Prediction



using Scikit Learn XGBoost with ArcGIS Pro

Machine Learning in ArcGIS

By Lauren Bennett, Esri Spatial Analyst

A2: Can you think of any additional factor that will make more accurate analysis, and one that creates a less accurate analysis?

What would Cause an Accident?

Temperature
Sun, Mon, Fri..

Wind Speed
Fast, Slow..

Visibility
High/Low

Snow Depth
High/Low

Day of the Week
Sun, Mon, Fri..

Time of the Day
12:45, 23:00

Month
Feb, Dec..

Road Width
20-30 M

Road Alignment
Straight / Curved

Proximity to
Intersections

Speed Limit
120 km/h

Sun Direction
East, West

Daily Traffic
AADT

Proximity to
Billboards

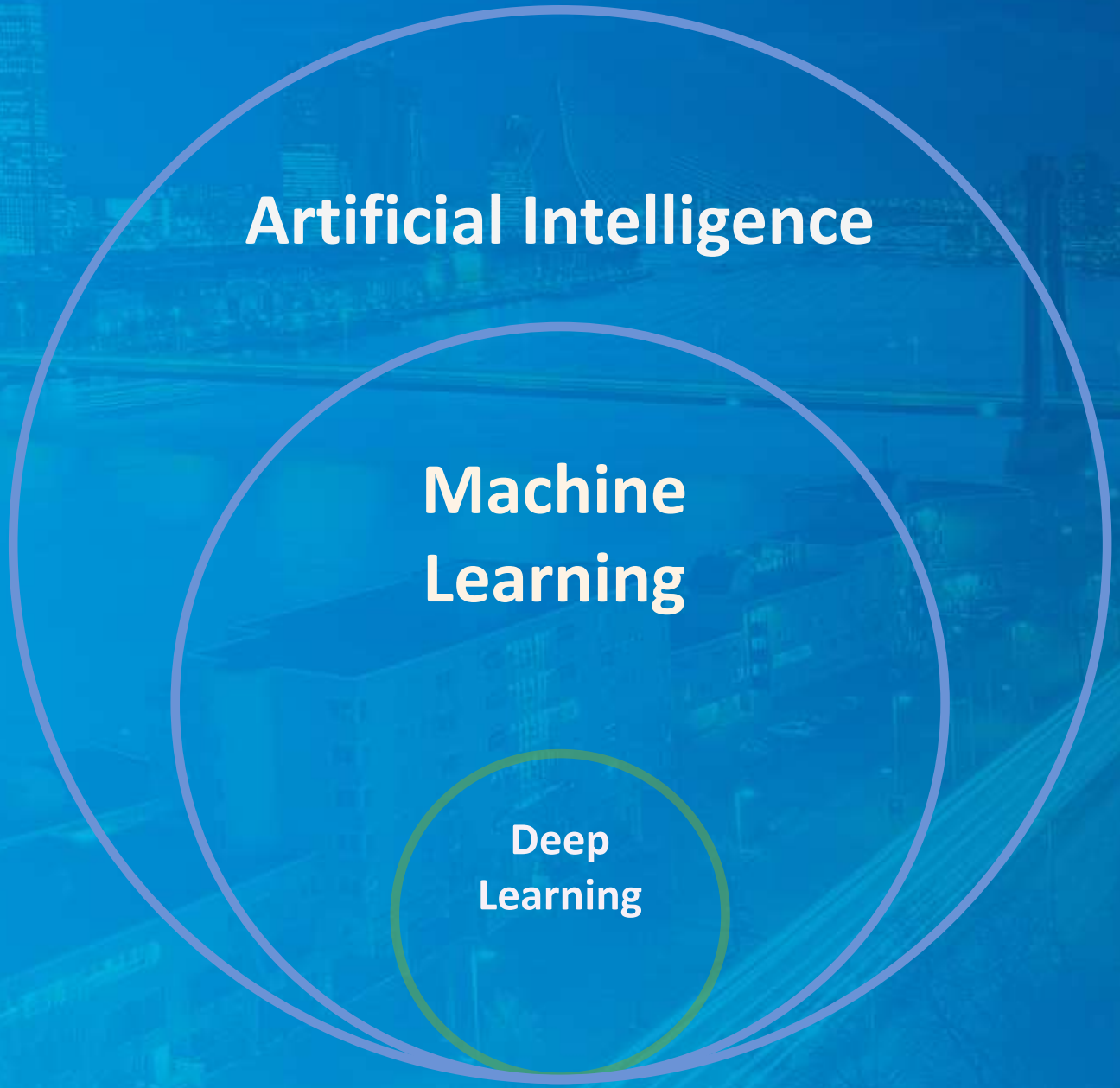
...

Spatial Applications of big data

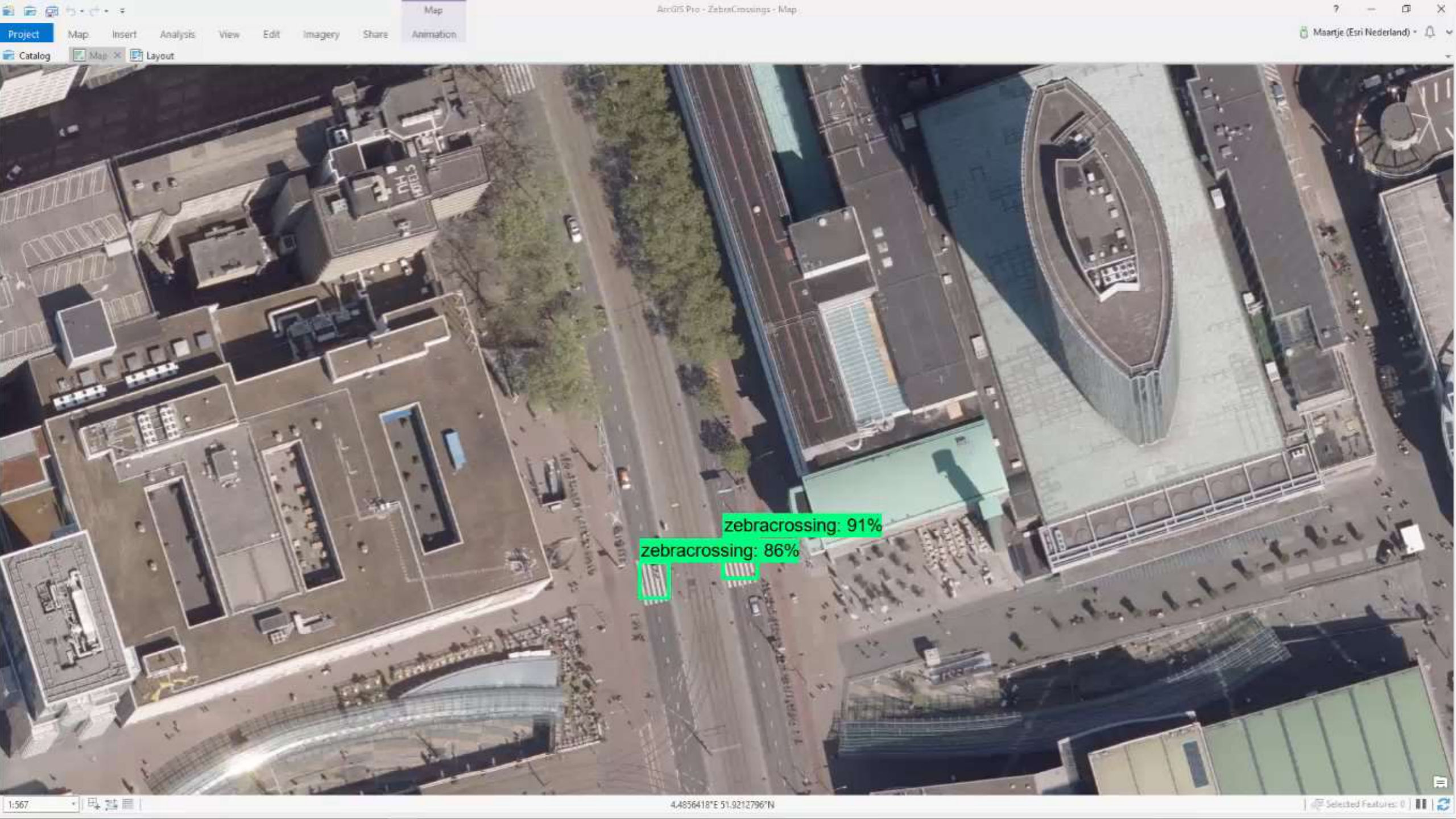
Discussion

esri Nederland

THE SCIENCE OF WHERE™


Artificial Intelligence
Machine Learning
Deep Learning

via Boredpanda.com


zebracrossing: 91%
zebracrossing: 86%

Discussion:

Propose a specific societal challenge, which can best be solved spatially.

esri Nederland
THE SCIENCE OF WHERE™

Break

esri Nederland
THE SCIENCE OF WHERE™

Reflection:

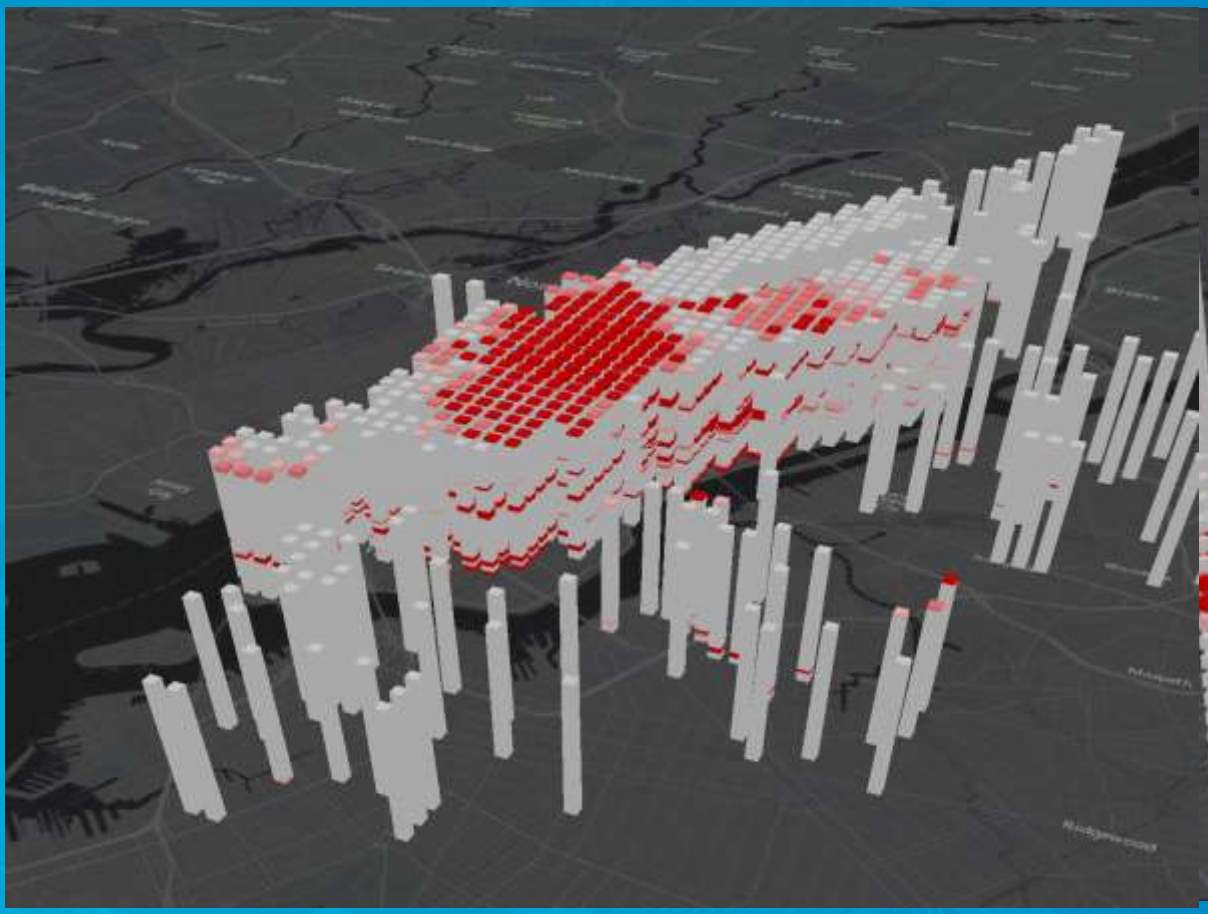
Propose a specific societal challenge, which can best be solved spatially.

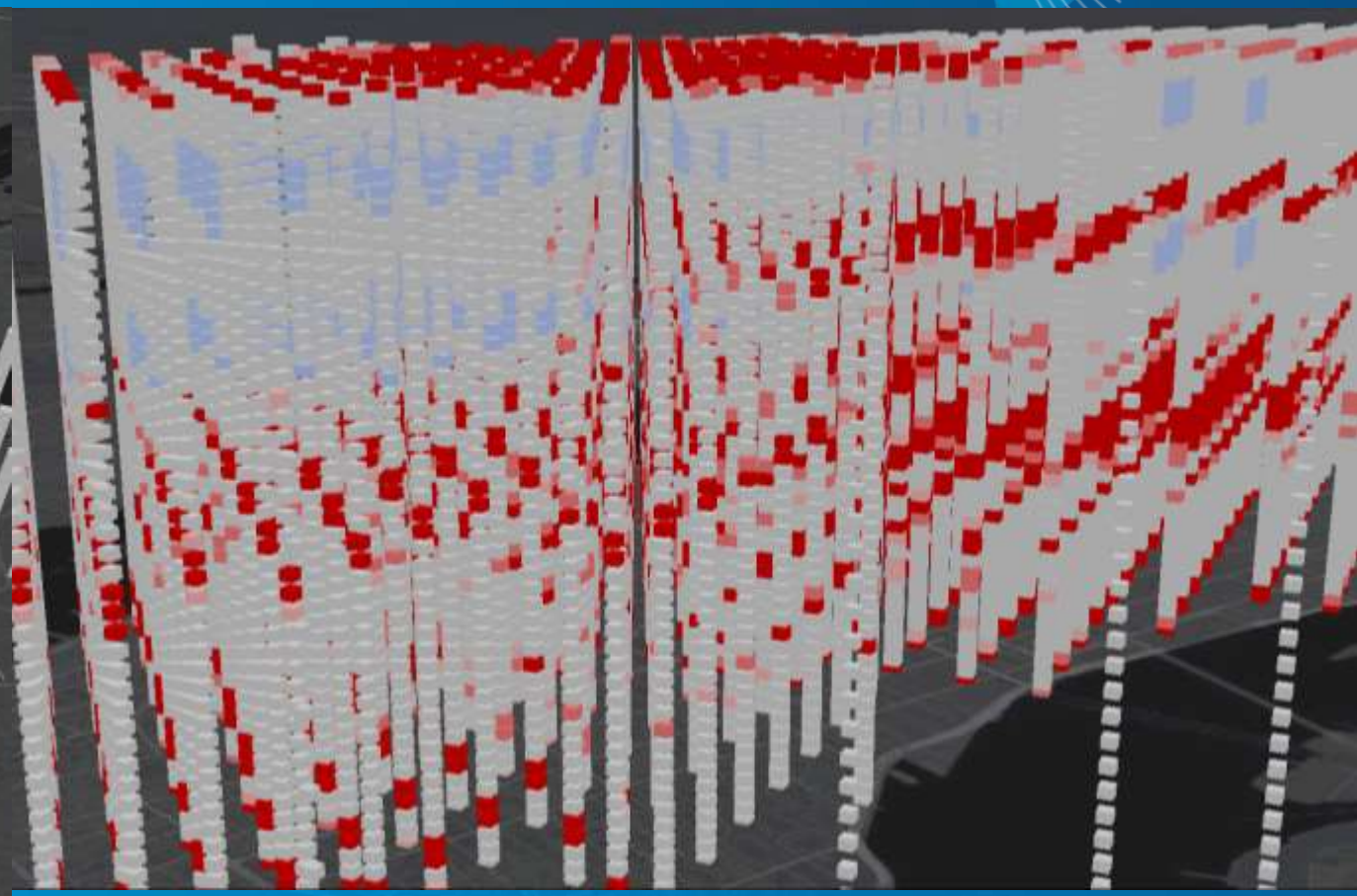
esri Nederland
THE SCIENCE OF WHERE™

Big Spatial Data in society

esri Nederland
THE SCIENCE OF WHERE™

Spatial Big Data and Analytics

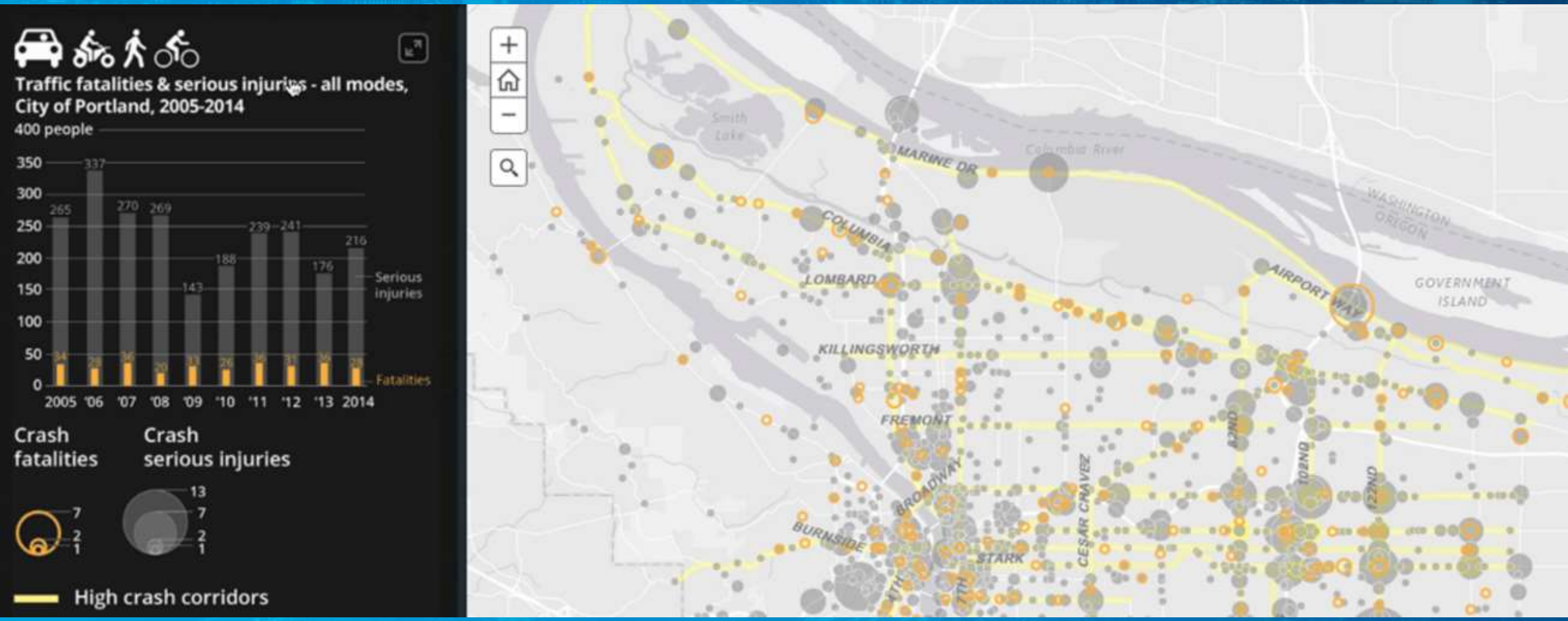




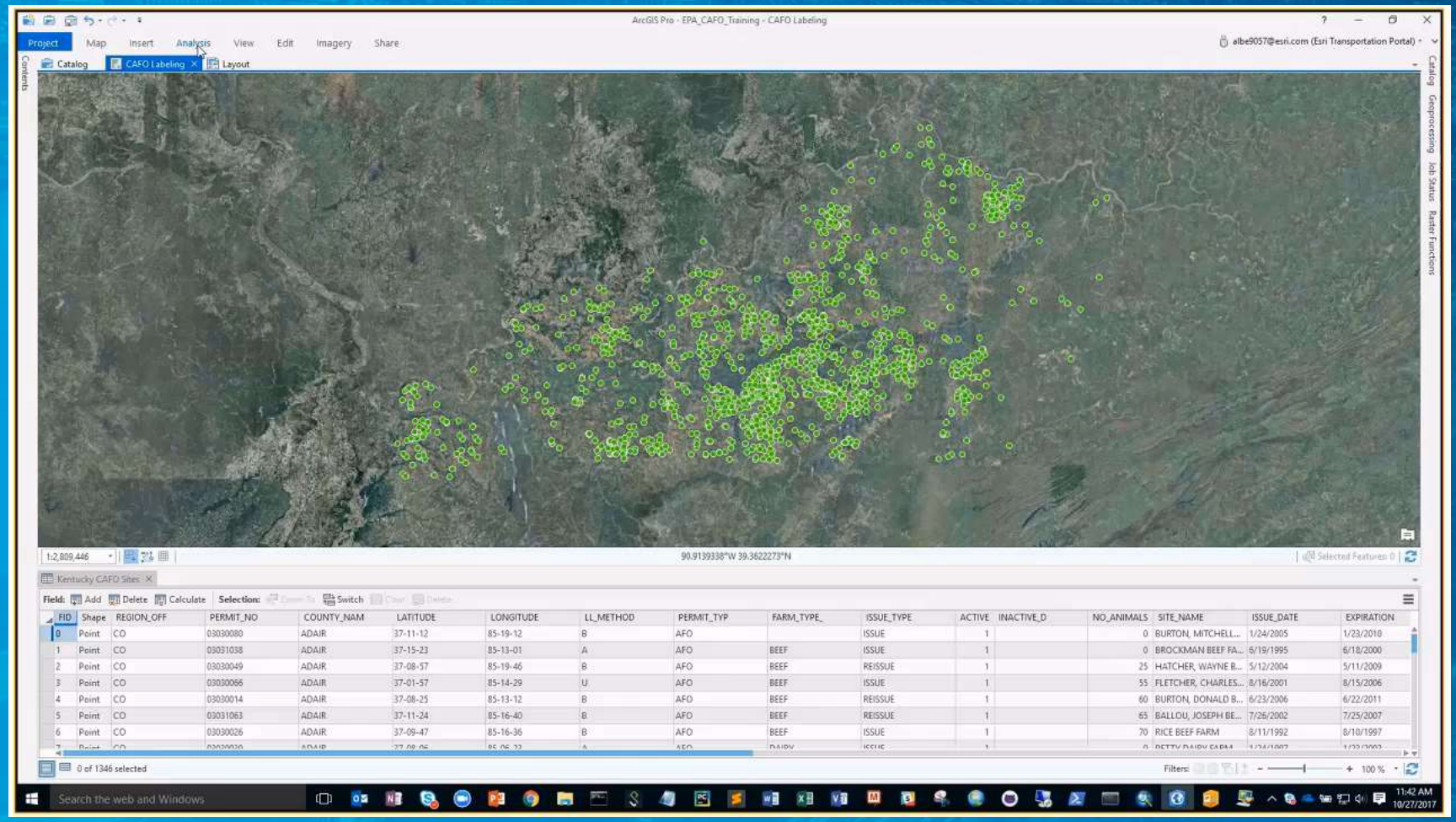
esri Nederland


NYC CRIME TRENDS
Find address or place
500 METERS
1000 METERS
JAN INCIDENTS WITHIN 500 METERS
8
JAN
JUN
DEC

Understand movement patterns

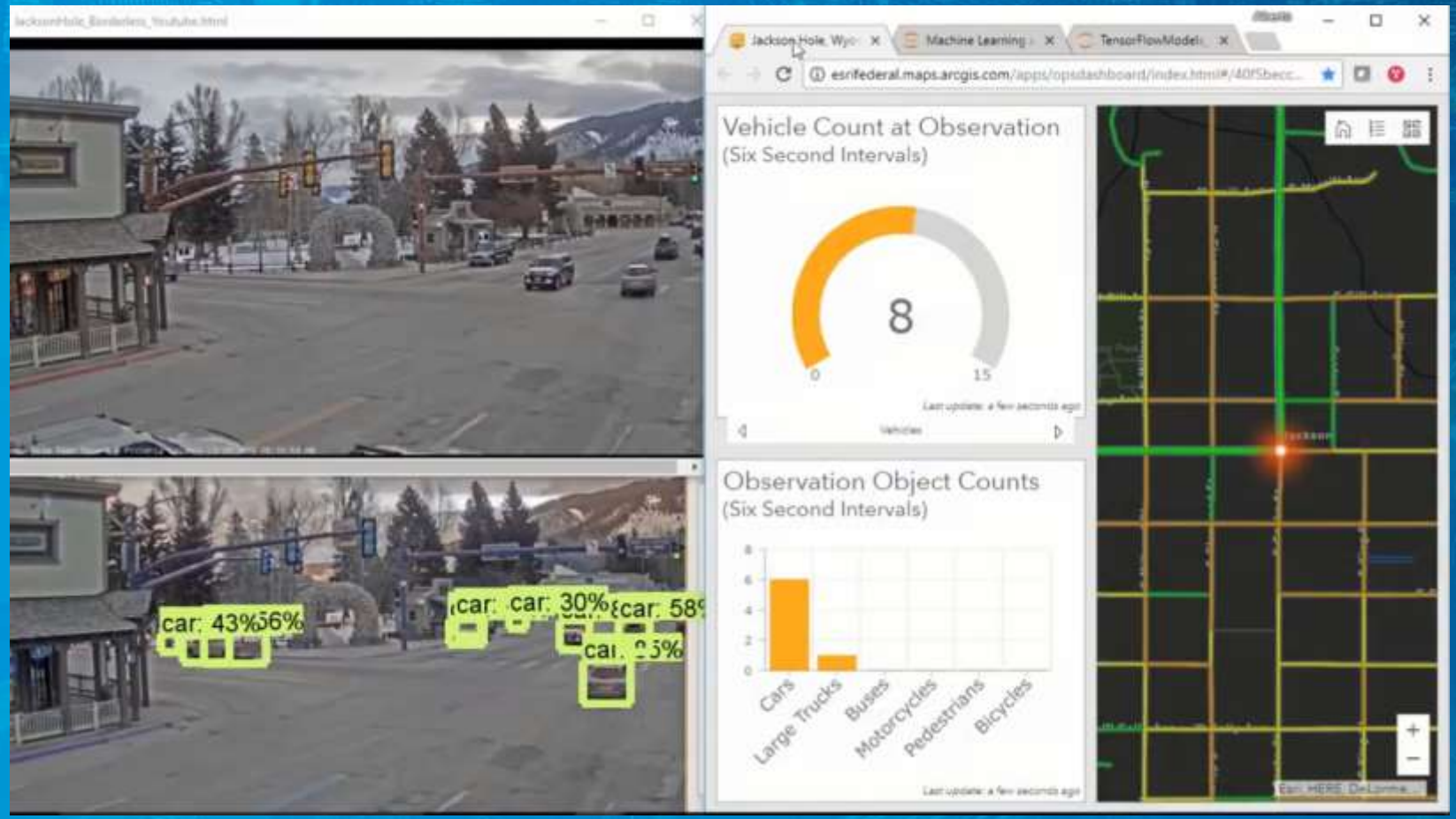


Object Detection from Satellite Imagery



using Deep Learning with ArcGIS Pro

Real-Time Activity Detection



using Deep Learning with ArcGIS API for Python and Ops Dashboard

Programme

Lecture: The Spatial perspective on data

Discussion: Spatial Applications of big data

Break

Lecture: Big Spatial Data in society

Reflection

Learning goals

Understand the role of data science and its societal impact

Recognise the knowledge discovery processes in applied data science

Identify trends and developments in big data technologies

Apply selected big data technologies to solve real-world problems

Space time pattern mining tools, zoals 'Emerging Hotspot Analysis', gebruiken om verborgen patronen te ontdekken in geodata. Krijg jij hier energie van? Dan zijn wij op zoek naar jou!

stage data science met geodata

Als stagiair op de afdeling Product Consultancy hou jij je bezig met de nieuwste tools om ruimtelijke patronen in grote datasets te ontdekken. Deze tools, zoals 'Time Series Clustering' en 'Create Space Time Cube' zijn op hun beurt gebaseerd op recente ontwikkelingen op het gebied van Data Science, zoals machine learning. Jij zet deze tools in om diepgaande analyses uit te voeren op diverse datasets, zoals alle verkeersongevallen van 2003 tot 2017. De patronen die je herkent, worden gebruikt om bijvoorbeeld beleid

Big Spatial Data

Tim Brouwer
tbrouwer@esri.nl

Jan Willem van Eck
jvaneck@esri.nl

esri Nederland
THE SCIENCE OF WHERE