GIScience 2021

The 1st International Workshop on Methods, Models, and Resources for Geospatial Knowledge Graphs and GeoAI

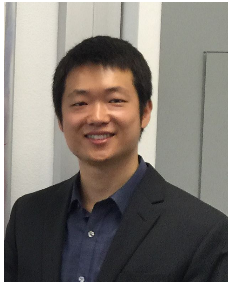
# Geospatial Knowledge Graph and Spatially Explicit AI

web page: https://ling-cai.github.io/GIScience-GeoKG/

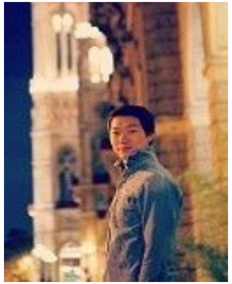

**Gengchen Mai**

Postdoc at Stanford AI Lab, Department of Computer Science, Stanford University
https://gengchenmai.github.io/


Gengchen Mai



Yingjie Hu



Song Gao

Ling Cai

Bruno Martins

Johannes Scholz

Jing Gao

# GeoKG & GeoAI 2021 PC Members

## PC members

- Weiming Huang, GIS Centre, Lund University
- Fei Du, Apple Map
- Cogan Shimizu, Kansas State University
- Bo Yan, Google
- Ni Lao, SayMosaic
- Xi Liu, Google Inc.
- Yao-Yi Chiang, University of Southern California
- Rui Zhu, University of California, Santa Barbara
- Wenwen Li, Arizona State University
- Carsten Keßler, Department of Planning, Aalborg University Copenhagen
- Dalia Varanka, U.S. Geological Survey
- Martin Raubal, ETH Zurich
- Morteza Karimzadeh, University of Colorado Boulder
- Fan Zhang, MIT
- Hongxu Ma, Google X
- Xinyue Ye, Texas A&M University
- Yuhao Kang, University of Wisconsin-Madison

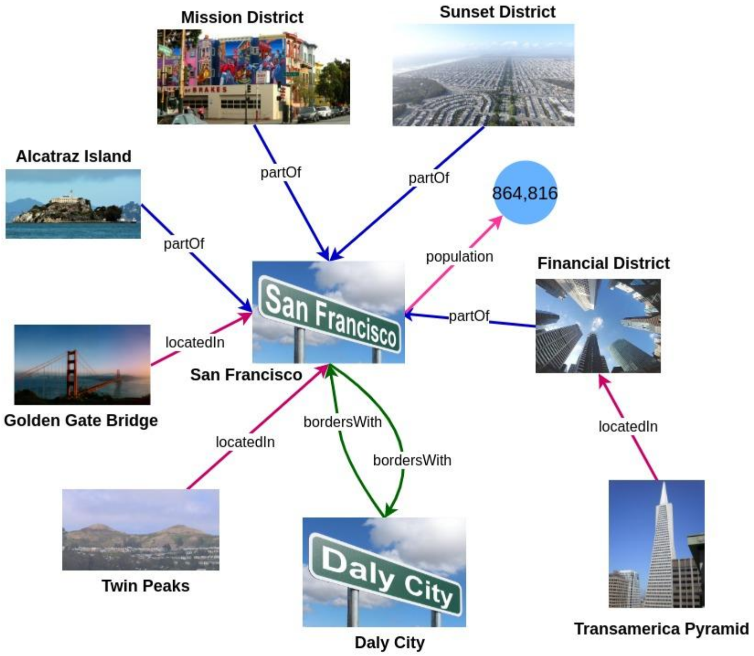
# GeoKG & GeoAI 2021 Program

| Session | Time | Speaker | Title |
|---|---|---|---|
| Opening Introduction | 15:30 - 16:00 | Gengchen Mai, Ling Cai | Geospatial Knowledge Graph and Spatially-Explicit AI |
| Keynote | 16:00 - 17:00 | Pascal Hitzler | The KnowWhereGraph |
| Break | 17:00 - 17:10 | | |
| **Session 1: Geospatial Semantics and GeoKG** | 17:10 - 17:30 | Shirly Stephen, Wenwen Li and Torsten Hahmann | Geo-Situation for Modeling Causality of Geo-Events in Knowledge Graphs |
| | 17:30 - 17:50 | Marvin Mc Cutchan and Ioannis Giannopoulos | Geospatial Semantics and Geographic Aware ANN |
| | 17:50 - 18:10 | Yuanyuan Tian and Wenwen Li | GeoAI for Knowledge Graph Construction: Identifying Causality Between Cascading Events to Support Environmental Resilience Research |
| Break | 18:10 - 18:20 | | |
| **Session 2: GeoAI** | 18:20 - 18:40 | Peng Yue, Boyi Shangguan, Lei Hu, Chenxiao Zhang, Liangcun Jiang and Zhe Fang | Quality Considerations for AI Training Data in Remote Sensing |
| | 18:40 - 19:00 | Jin Xing and Renee Sieber | Challenges of Using XAI for Geographic Data Analytics |
| | 19:00 - 19:20 | Cláudia Rodrigues, Ana Alves, Marco Veloso and Carlos Bento | Identification of the User's Geographic Map |
| | 19:20 - 19:40 | Haojian Liang and Shaohua Wang | A New Approach Based on Graph Neural Network for Solving p-center Problems |
| Discussion & Closing Remark | 19:40 - 20:00 | PC Chair | |

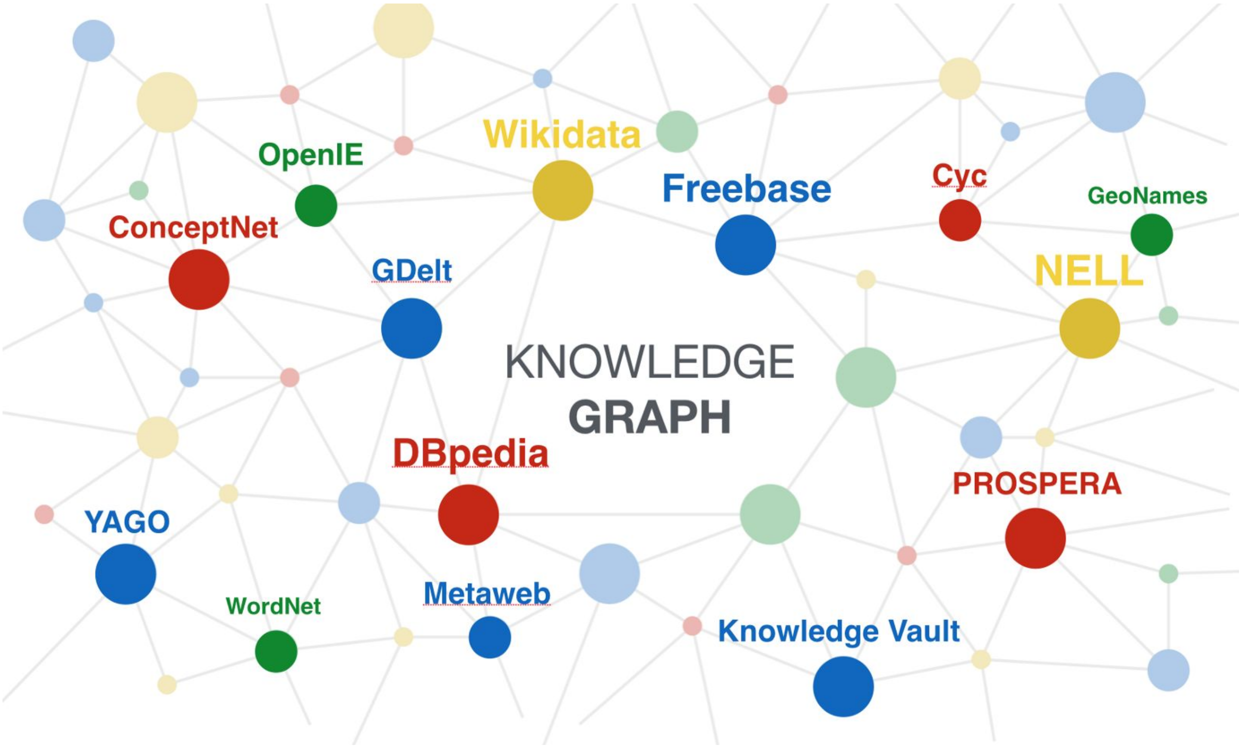
# Knowledge Graph

A **knowledge graph (KG)** is a data repository that stores real-world knowledge under some schema, e.g., an ontology.

- **Directed multi-graphs**
  - Nodes: entities
  - Edges: relationships between entities with relation types as labels
  - Statement: <subject, predict, object>



# Knowledge Graph (Linked Data)




Integration across different datasets

# Spatiotemporal Data in Knowledge graphs

## Geographic Information

- **Geographic Information of Entities**
  - Coordinate information
    - Santa Barbara -> coordinateLocation -> (34°25’33"N, 119°42’51"W)
  - Topological relations
    - (Santa Barbara -> partOf -> California)
- **Other Geospatial-Related Statements**
  - (France -> memberOf -> European Union);
  - (Washington, D.C. -> hasPopulation -> 672,228);
  - (Los Angeles -> twinnedAdministrativeBody -> Berlin);

# Spatiotemporal Data in Knowledge graphs

## Temporal Information

- **Temporal Scope of a Statement**
  - (Poland –> memberOf –> Warsaw Pact, [1955, 1991]);
  - (Washington, D.C. -> hasPopulation -> 672,228, 2015); ...
- **Time as Literals**
  - (Barack Obama –> dateOfBirth –> 4 August 1961);
  - (Santa Barbara –> inception –> 1847); ...
- **Transaction Time**
  - (Fernando Torres–> playFor–>Chelsea, [2011,2015), [09/02/2017])

# Applications

- Geographic question answering
- Geographic knowledge graph summarization
- Qualitative spatio-temporal reasoning
- Geographic information retrieval and Geo-text Analysis
- Geo-ontology engineering
- Geospatial knowledge graph construction
- Geospatial knowledge graph completion

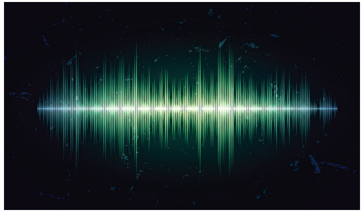
# Why Spatial is Special?

Spatial data have more **irregular structures**

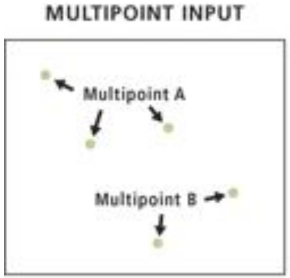
Image

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called "showers".

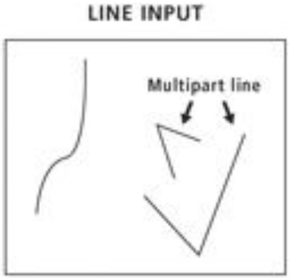
Text



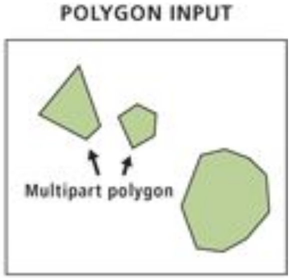
Audio




Point




Line




Polygon

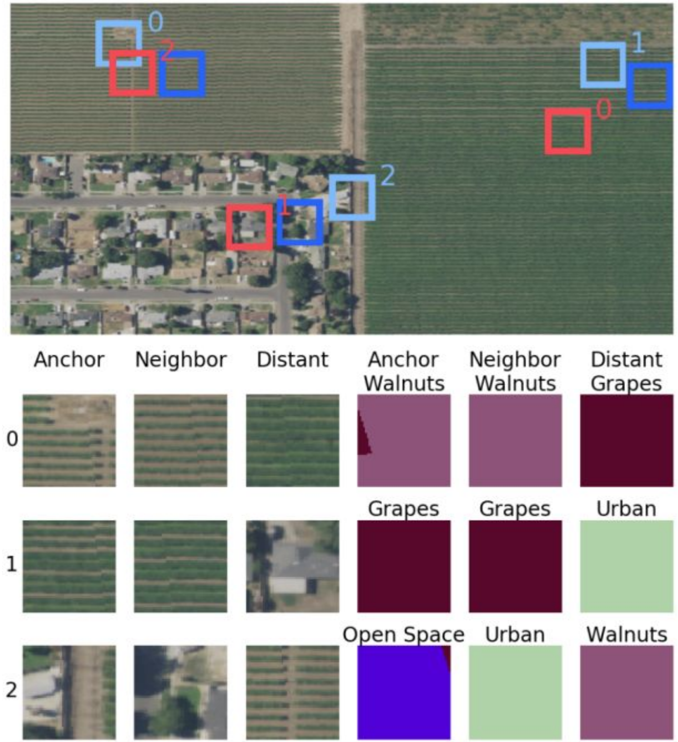
# Why Spatial is Special?

**Spatial Autocorrelation:**

**Tobler's First Law of Geography:**

*"Everything is related to everything else, but near things are more related than distant things."*

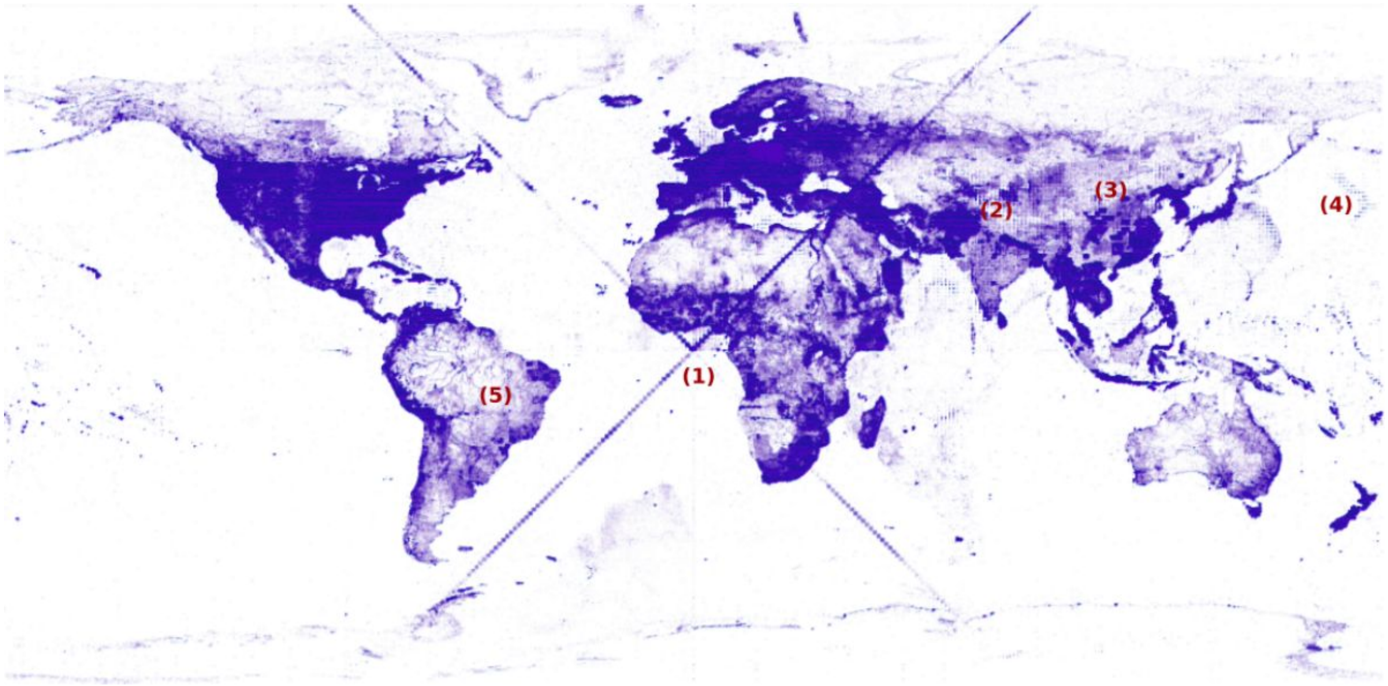
*-- Waldo R. Tobler (Tobler 1970)*




Jean, N., et al.. **Tile2vec: Unsupervised representation learning for spatially distributed data**. AAAI 2019.

# Why Spatial is Special?

**Geographic bias and errors**




Janowicz, K., et al.. **Moon landing or safari? a study of systematic errors and their causes in geographic linked data**. GIScience 2016.

# Spatially Explicit Model

A model is said to be **spatially explicit** when it differentiates behaviors and predictions according to spatial location

- **The invariance test**
  - spatially explicit models are not invariant under relocation
- **The representation test**
  - spatially explicit models include spatial representations in their implementations
- **The formulation test**
  - spatially explicit models include spatial concepts in their formulations
- **The outcome test**
  - spatial structures of inputs and outcomes are different

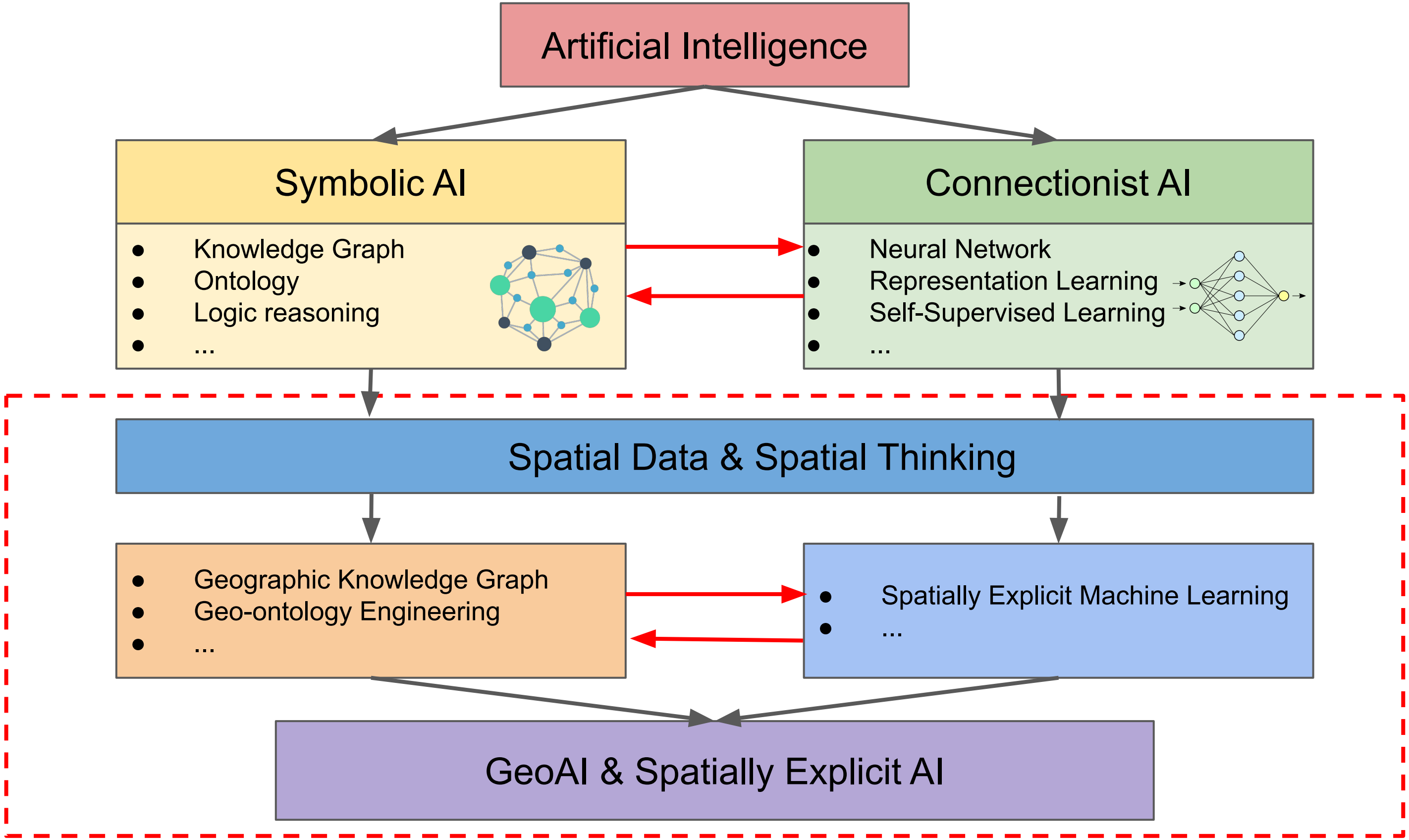
*-- Prof. Michael F. Goodchild*

# Spatially Explicit Machine Learning Model

**Spatially Explicit Machine Learning Model**: Improve the performance of current state-of-the-art machine learning models by using **spatial thinking** and **spatial inductive bias** such as:

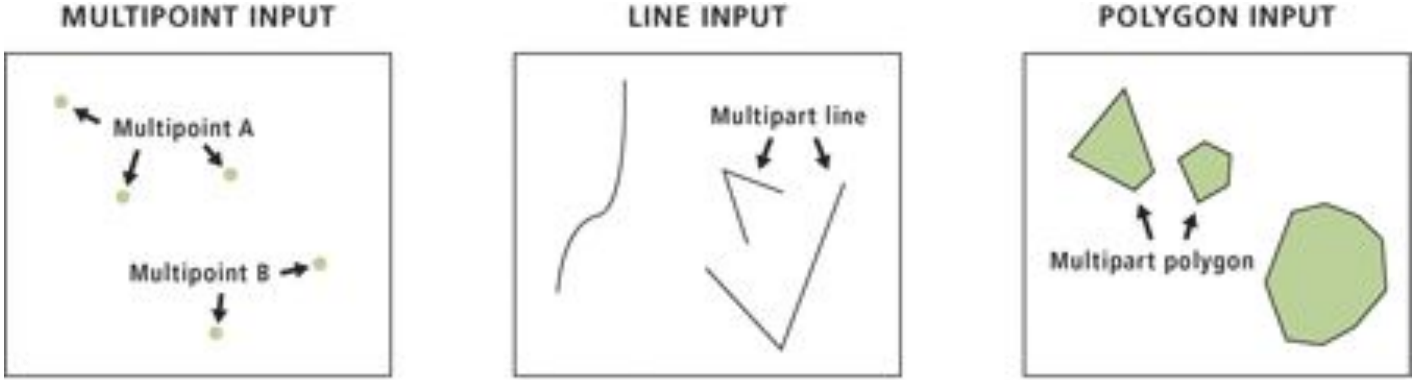
- spatial heterogeneity
- distance decay effect
- map projection

# Spatially Explicit Artificial Intelligence



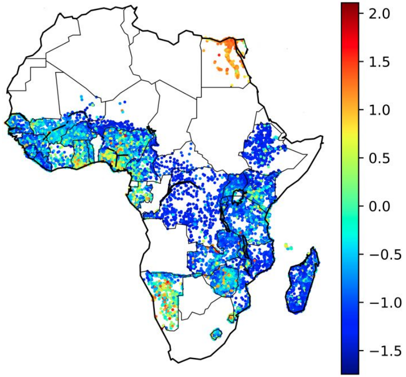
# Represent Spatial Data into the Embedding Space

How to represent different types of spatial data into an embedding space?



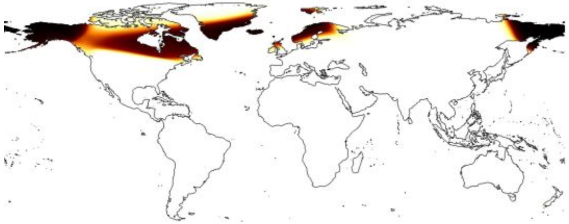
# Various Geospatial Tasks

Economy:
Wealth Index Prediction
(Sheehan et al., 2019)

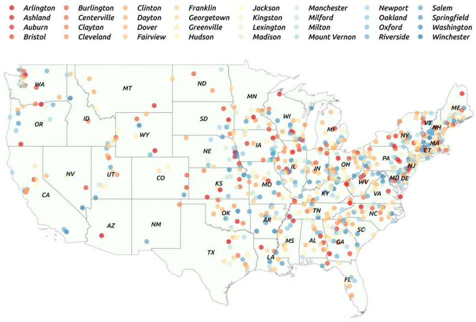



Climate Science:
Precipitation Prediction

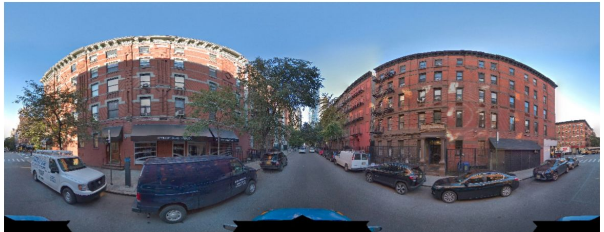

Ecology:
Species Distribution Modeling
(Mac Aodha et al., 2019; Mai et al., 2020)



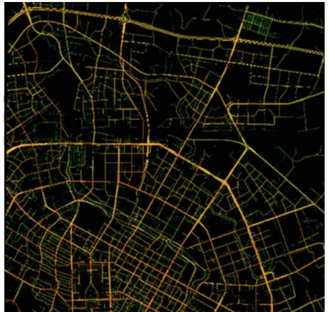
Geospatial Semantics:
Place Name Disambiguation
(Ju et al., 2016)



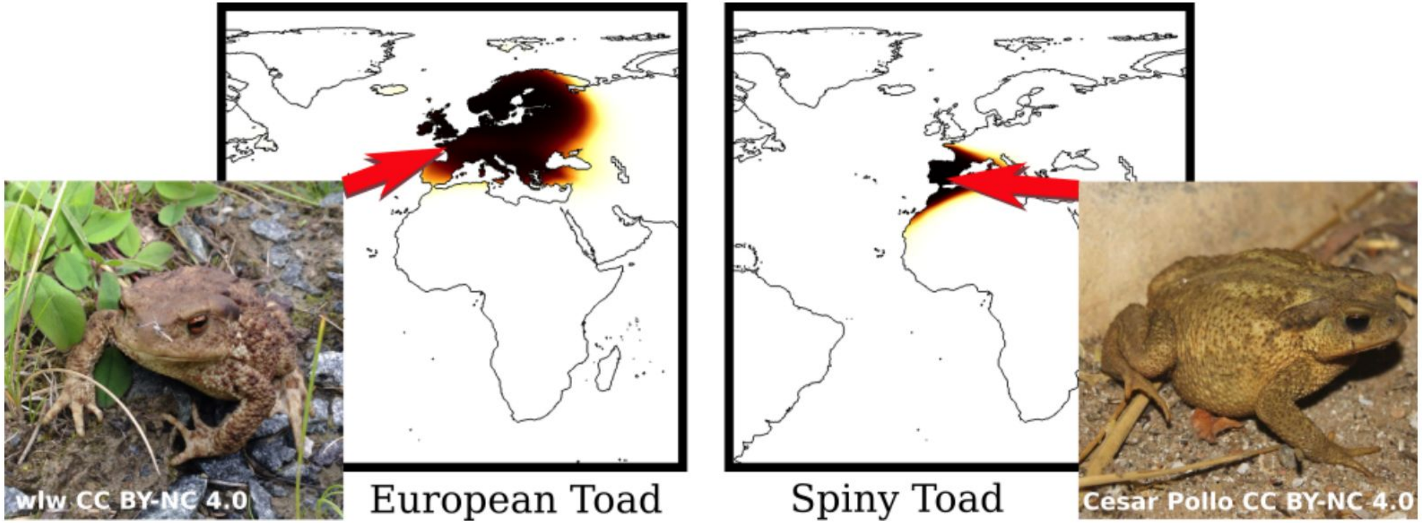

Smart City:
Indoor/outdoor Navigation
(Mehta et al., 2020)



Human Mobility:
Trajectory Prediction
(Gao et al., 2021)

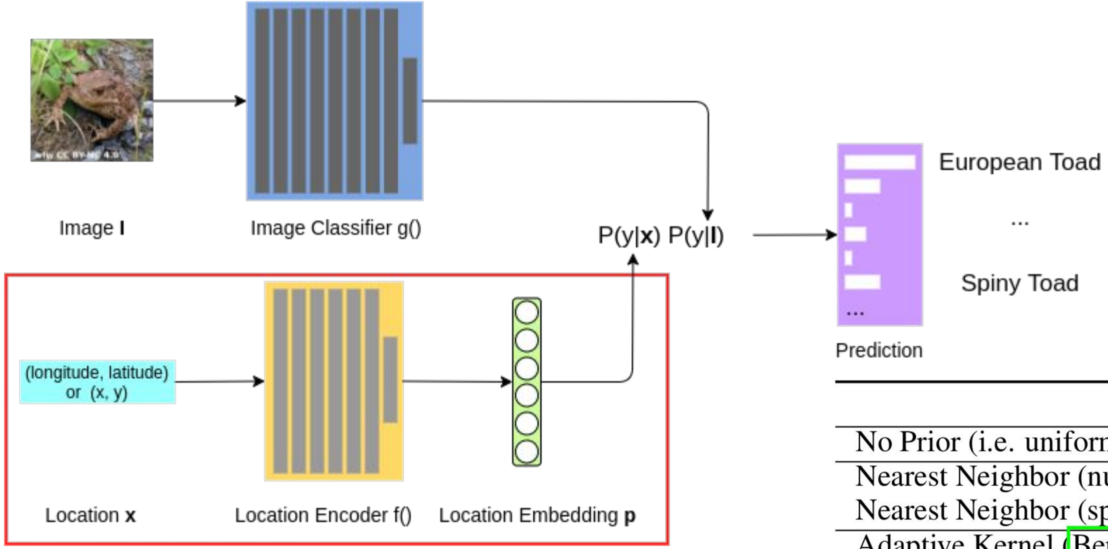


# Spatial Distribution: Geo-Aware Image Classification




Species with similar appearance information may have distinct geographic prior distributions. (Figure from Mac Aodha et al., 2019)

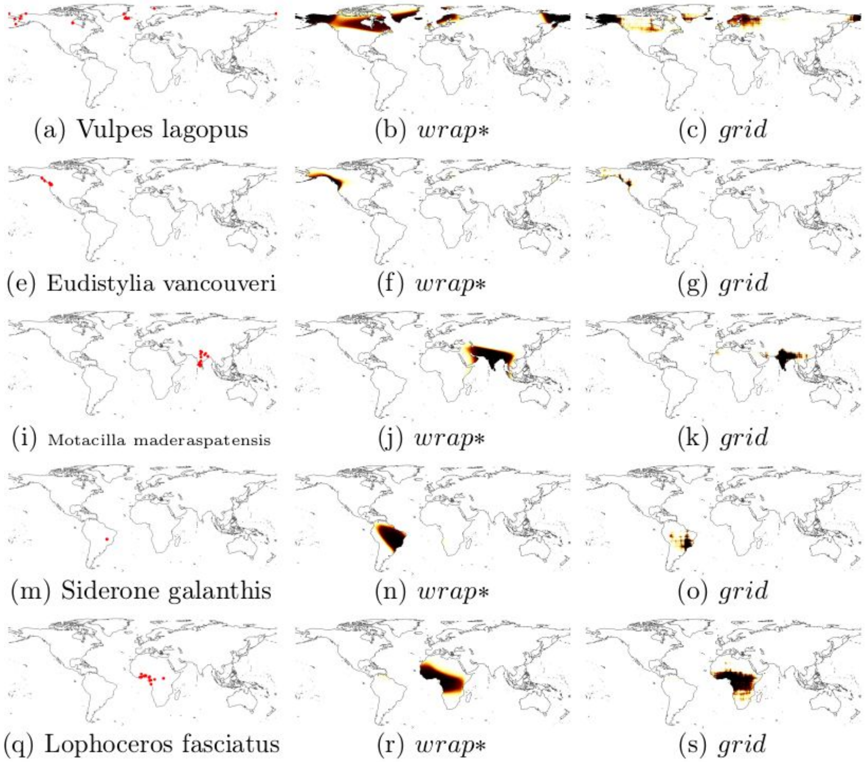
# Geo-Aware Image Classification




| | BirdSnap† | NABirds† |
|---|---|---|
| No Prior (i.e. uniform) | 70.07 | 76.08 |
| Nearest Neighbor (num) | 77.76 | 79.99 |
| Nearest Neighbor (spatial) | 77.98 | 80.79 |
| Adaptive Kernel (Berg et al., 2014) | 78.65 | 81.11 |
| *tile* (Tang et al., 2015) (location only) | 77.19 | 79.58 |
| *wrap* (Mac Aodha et al., 2019) (location only) | 78.65 | 81.15 |
| *rbf* ($\sigma$=1k) | 78.56 | 81.13 |
| *grid* ($\lambda_{min}$=0.0001, $\lambda_{max}$=360, $S$ = 64) | **79.44** | 81.28 |
| *theory* ($\lambda_{min}$=0.0001, $\lambda_{max}$=360, $S$ = 64) | 79.35 | **81.59** |

Mai, G., et al, 2020. **Multi-Scale Representation Learning for Spatial Feature Distributions using Grid Cells**. In International Conference on Learning Representations.

# Species Distribution Prediction




Mai, G., et al, 2020. **Multi-Scale Representation Learning for Spatial Feature Distributions using Grid Cells**. In International Conference on Learning Representations.

# Capturing Distance Decay Effect:

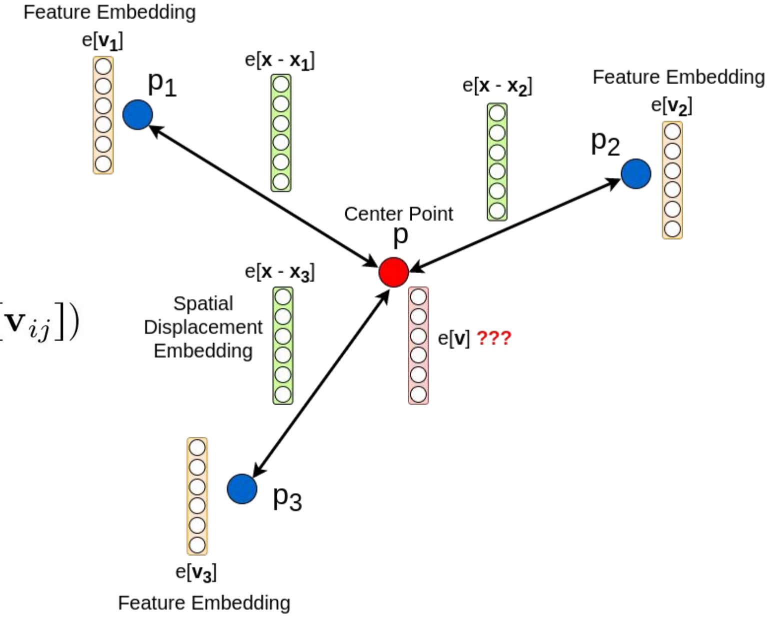
Tasks: **POI type prediction based on spatial context**

**Space-Aware Graph Attention Network Model**:

$$\mathbf{e}[\mathbf{v}_i]' = Dec_c(\mathbf{x}_i, \{\mathbf{e}_{i1}, ..., \mathbf{e}_{ij}, ..., \mathbf{e}_{in}\}; \theta_{\text{dec}_c}) = g(\frac{1}{K}\sum_{k=1}^{K}\sum_{j=1}^{n}\alpha_{ijk}\mathbf{e}[\mathbf{v}_{ij}])$$

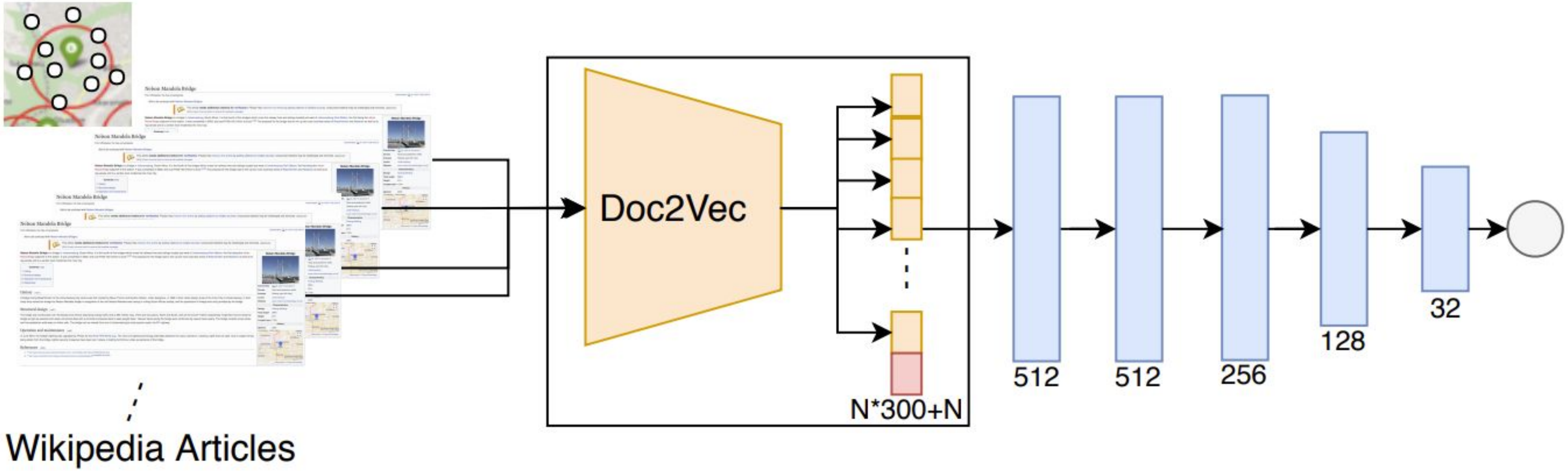
$$\alpha_{ijk} = \frac{exp(\sigma_{ijk})}{\sum_{o=1}^{n} exp(\sigma_{iok})}$$

$$\sigma_{ijk} = LeakyReLU(\mathbf{a}_k^T[\mathbf{e}[\mathbf{v}_i]_{init}; \mathbf{e}[\mathbf{v}_{ij}]; \boxed{\mathbf{e}[\mathbf{x}_i - \mathbf{x}_{ij}]}])$$



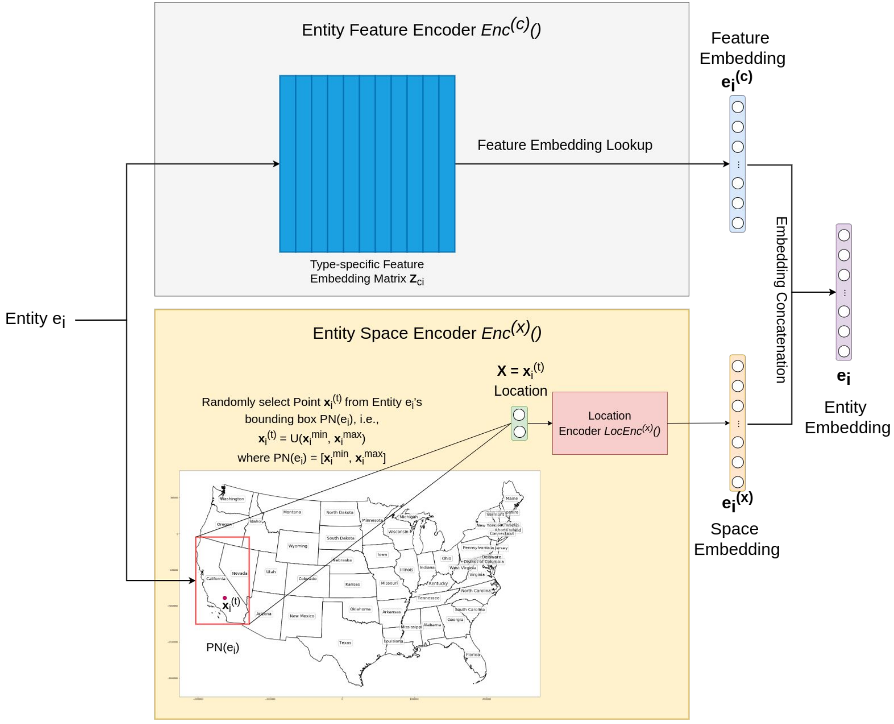

Mai, G., et al, 2020. **Multi-Scale Representation Learning for Spatial Feature Distributions using Grid Cells**. In International Conference on Learning Representations.

# Space-aware Graph Attention Network

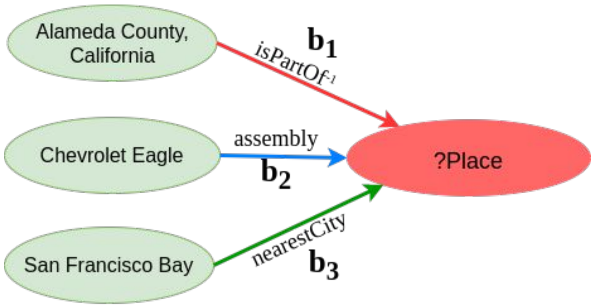



Sheehan, E.,et al.. **Predicting economic development using geolocated wikipedia articles**. In ACM SIGKDD 2019.

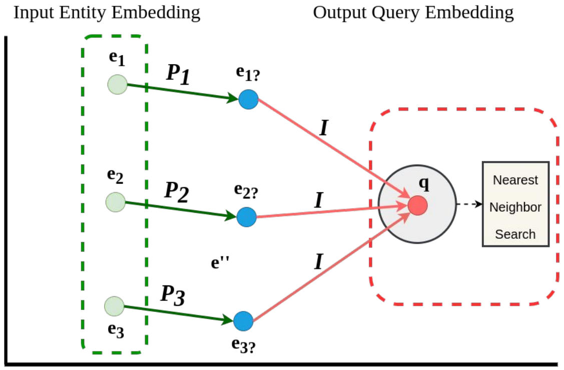
# Geographic Knowledge Graph Embedding




1) Encode Spatial Footprints into Entity Embedding Space




2) Given a GeoKG sub-graph




3) KG Embedding Training

Mai, G., et al., 2020. **SE-KGE: A location-aware Knowledge Graph Embedding model for Geographic Question Answering and Spatial Semantic Lifting**. Transactions in GIS, 24(3), pp.623-655.

# From GeoAI to Spatially Explicit AI

### GeoAI

1. **Remote Sensing:** Apply normal CV models on RS Data
2. **GeoKG:** Apply normal KG embedding on GeoKGs
3. **GeoQA:** Apply normal QA models on GeoQA dataset

### Spatially Explicit AI

1. **Remote Sensing:** Consider distance decay effects in RS image embedding space
2. **GeoKG:** Representing spatial footprints in KG Embedding space
3. **GeoQA:** Let NN learn to do spatial reasoning for GeoQA

### From GeoAI to Spatially Explicit AI

- What are the benefits?
- What are the trade-off?

# What will happen in the future?

| | |
|---|---|
| 5 Year Vision: | A set of basic and **composable spatially explicit AI models** as a tool set |
| | An **interlinked open geospatial knowledge graph** for geospatial tasks |

| | |
|---|---|
| 20+ Year Vision: | An **GIS analytic AI** to answer geographic questions |

# Transactions In GIS Special Issue

SI CFP: https://onlinelibrary.wiley.com/page/journal/14679671/homepage/featured-collections#1

Transactions in GIS

## Call for Papers: Special Issue on Methods, Models and Resources for Geospatial Knowledge and GeoAI

Nowadays, one of the most prominent topics in Artificial Intelligence (AI) is the combination of representation learning techniques (Connectionist Artificial Intelligence) with symbolic representation and reasoning associated with knowledge graphs (Symbolic Artificial Intelligence), in order to develop scalable and interpretable AI models. From a geospatial point-of-view, GeoAI, as an interdisciplinary field of GIScience and AI, advocates the idea of developing and utilizing AI techniques in geography and earth science. Geospatial knowledge graphs, as symbolic representations of geospatial knowledge, go to the core of GeoAI and facilitate many intelligent applications such as geospatial data integration and knowledge discovery.
This special issue seeks new methods, models, and resources for advancing research related to Geospatial Knowledge Graphs and GeoAI.

Full details of the call are available **here**.

Deadline for Submission **February 15th, 2022**

**Submission Guideline**
Submission deadline: Feb 15, 2022
Expected final decision: July 01, 2022

**Guest Editors:**

- Dr. Gengchen Mai, Stanford AI Lab, Stanford University
- Prof. Yingjie Hu, Department of Geography, University at Buffalo
- Prof. Song Gao, Department of Geography, University of Wisconsin-Madison
- Ling Cai, Department of Geography, University of California, Santa Barbara
- Prof. Bruno Martins, Instituto Superior Técnico, University of Lisbon
- Prof. Johannes Scholz, Institute of Geodesy, Graz University of Technology
- Prof. Jing Gao, Department of Geography and Spatial Sciences, University of Delaware
- Prof. Krzysztof Janowicz, Department of Geography, University of California, Santa Barbara
- Dr. Rui Zhu, Department of Geography, University of California, Santa Barbara