

The 2024 UK-USA Smart Cities Forum

# Generative A.I. & Applications

Generative AI and Foundation Models for Smart  
Cities

Dr. Andrei Lopatenko

PhD in CS, Vice-President of AI and Engineering

# Unleashing Urban Innovation: The New Frontier of Generative AI in Smart Cities

**Unlocking Potential:** Recent years have seen a dramatic leap in Generative AI and Foundation models, empowering not just big tech but also startups, smaller companies, and government organizations to pioneer smart city projects.

**Broad Spectrum Applications:** These AI advancements span across voice, audio, text, geospatial, traffic, sensor data, satellite imagery, and more, opening up unprecedented opportunities in urban innovation.

**Accessibility & Affordability:** The democratization of these technologies means transformative smart city projects are now achievable without the need for billion-dollar investments, leveling the playing field for all.

**Real-World Impact:** From enhancing public safety and urban planning to improving environmental sustainability and transportation, the applications are as diverse as they are significant, marking a new era in urban development.

# AI for Cities, Examples

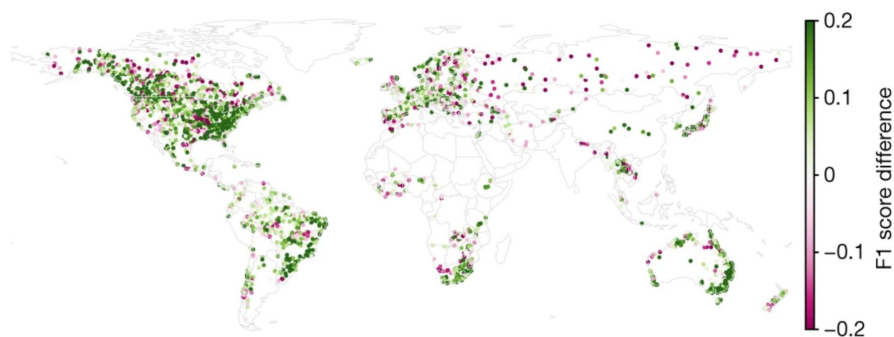
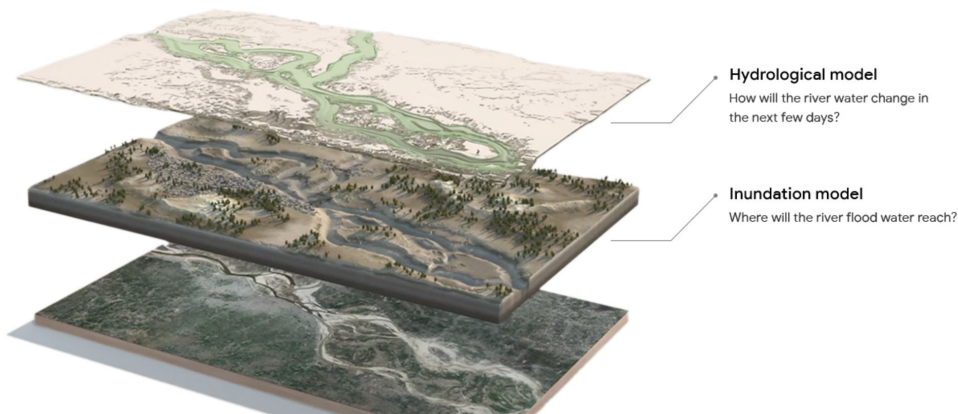


AI Generates 3D City Maps From Single Radar Images

In the hours after a major disaster, these maps could save lives

SAR2Height system from the University of the Bundeswehr Munich

# AI for Cities, Examples



AI model to accurately predict riverine flooding and help protect livelihoods in over 80 countries up to 7 days in advance, including in data scarce and vulnerable regions.

Forecasting up to 7 days in advance of a flood to 460 million people.

Google, European Centre for Medium-Range Weather Forecasts, Institute for Machine Learning, Johannes Kepler University, RAND Corporation,

# About the author

With two decades dedicated to developing AI applications for consumer interfaces for billions of customer around the world at organizations like Google, Apple, Zillow, Walmart, and eBay, and over a decade focused on creating AI, ML, and data platforms for business applications, my career has been deeply intertwined with the practical deployment of artificial intelligence.

This journey is marked by a specialized engagement with geospatial AI applications (14 years), leveraging location and traffic data for urban and beyond urban solutions.

PhD in Computer Science, the University of Manchester, United Kingdom

# AI Foundation models for Smart Cities

In our discussion today, we'll explore the wide array of Foundation Models and the vast opportunities they present for Smart Cities. While Natural Language Processing stands out as the most recognized among these models, there exists a multitude of other AI Foundation Models that offer significant benefits to smart cities. These models unlock new use cases and products by interpreting geospatial information, time series data, images, and videos, in addition to text.

My aim is to highlight the importance of various modalities in Smart Cities and provide an overview of the foundational technologies and the opportunities they herald.

# What are Foundation Models? And Why?

Foundation models are a class of deep neural networks designed to learn from vast datasets. Unlike traditional machine learning approaches that often require a separate model for each task, foundation models excel in handling multiple tasks within a given modality using a single model. This represents a significant shift from the conventional task-specific training approach. Foundation models undergo extensive training on large-scale datasets, enabling them to generate high-quality inferences across a range of tasks. These models can then be further refined or 'fine-tuned' for enhanced performance on specific applications.

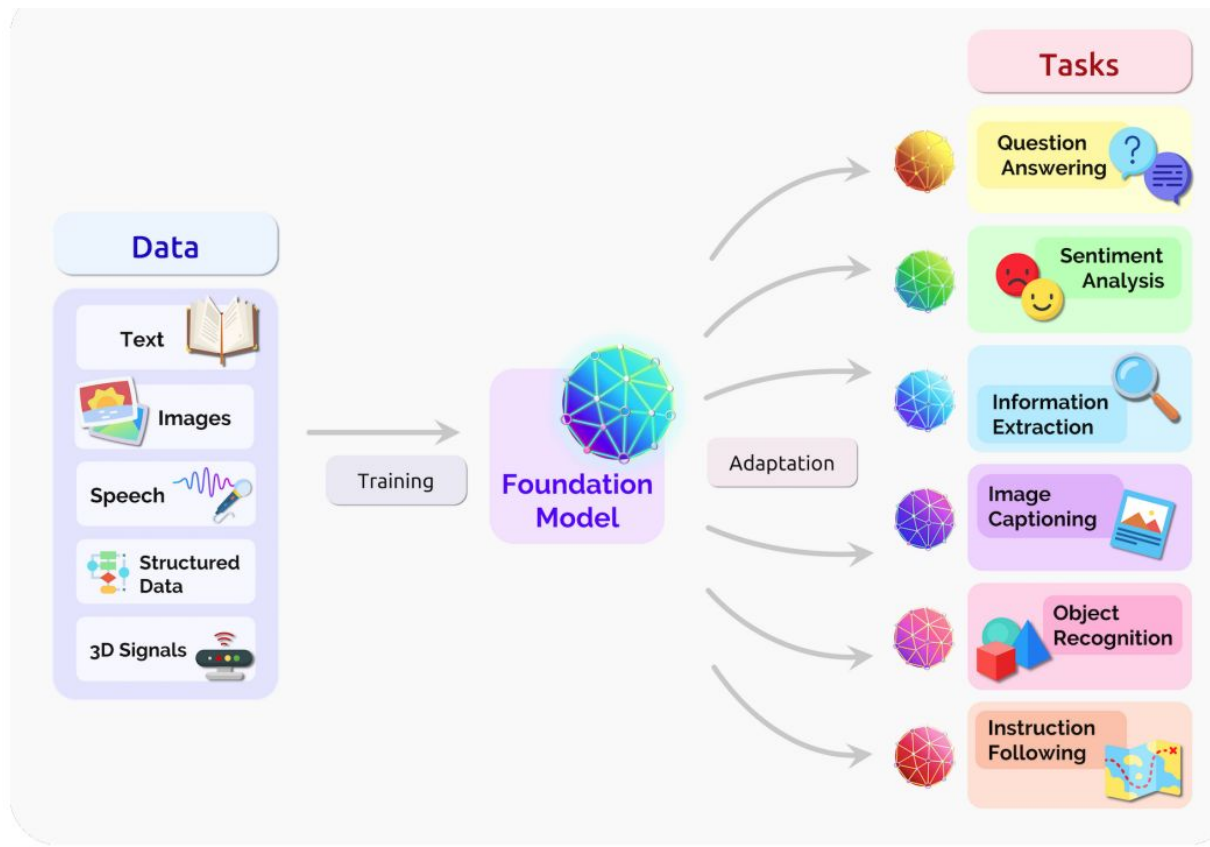
Modern foundation models are notable for their size, with some models boasting 7 billion, 14 billion, 50 billion, or even 70 billion and more parameters. This immense complexity means they necessitate specialized software for both training and inference phases to manage their computational demands effectively.

# Why

Smart cities encompass a broad range of challenges that necessitate profound intelligence and AI, applied across extensive and diverse data sets. Developing models to address smart city issues requires access to essential data (such as traffic data, urban imagery, weather conditions, electrical grids, and sensor outputs) and expertise in crafting intelligence for this kind of data—a capability not universally available to all companies and public institutions. Foundation models tailored to smart cities could level the playing field, enhancing AI accessibility for numerous organizations and enabling them to develop innovative applications and products.



# Foundation Models



From “On the Opportunities and Risks of the Foundation Models”, 2022

# Computer Vision, Images and Video Foundation Models

Images: OpenAI Clip, Microsoft Kosmos-2, DeepMind MagicLens (very promising results, yet to be released), GRIP(Grounded Language-Image Pre-training), InstructBLIP, LLaVa, MoonDream (tiny model, edge deployments) and other fine models to be deployed on the edge

Prithvi, GeoSpatial Foundation Model, trained on satellite images, recognizing geospatial data (NASA and IBM)

Beyond image understanding (object detection, classification, embedding representation) many other CV Gen AI model types can find application in the Smart cities: NeRF, Diffusion models, VAE many other model families depending on the CV tasks

# Computer Vision, Images and Video Foundation Models

- Deep Understanding of locations using satellite and airplane imaging
- Detection of meteorological phenomena, meteorological disasters
- Planning of maintenance work, energy efficiency work
- Computer vision models for vehicles , smart traffic, traffic management, driving safety
- CV models on traffic and other cameras, deep understanding of current state of transportations, pedestrian traffic, safety, emergency management
- Advice and help to citizens, For instance, seeking out properties featuring solar panels on their roofs, or evaluating a home's energy efficiency based on comprehensive insights gleaned from aerial imagery, including the potential suitability of roofs for solar panel installation, among other factors.
- Many other applications

# Transportation Foundation Models

Variety of types of models and modeling tasks

10 years ago only big tech companies could model traffic, now this task is available to smaller companies, startups and municipal organization

TransWorldNG by CAST Lab, S. A. C. Labs, 2023

TrafficGPT, (traffic foundation models + chatGPT), Transport policy May 2024

TengYun, CASIA etc, 2022

Beyond traffic modeling, vision models to serve intelligent transportation,(see Open-TransMind, Baidu 2023)

# Transportation Foundation Models

Traffic monitoring and prediction, optimization of traffic

Smart parking management

Deployments on the edge, driver assistance

Transportation planning (including public transportation)

Disaster management, Emergency response

Text analysis for deriving insights from social media on urban traffic

Smart fleet management (grocery and food delivery, public transport, taxis )

# Environmental, Weather and Climate Foundation Models

- Extreme weather forecasting by Google, GraphCast system
- Prithvi, 100M NaATA/IBM model, various weather tasks on geosat images
- FengWu, atmospheric dynamics and predict the future land and atmosphere states at 37 vertical levels on a  $0.25^\circ$  latitude-longitude resolution
- ClimaX, UCLA, MS, 2023
- FourCastNet (2022), short to medium-range global predictions at  $0.25^\circ$  resolution, extremely fast inference
- W-MAE, 2023, UESTC, etc
- PanguWeather, 2023, Huawei Cloud
- CLIMATEBERT <- text processing for weather

# Environmental, Weather and Climate Foundation Models

Nowcasting (Precipitation)

Forecasting

Data Assimilation

Climate text analysis, Weather pattern understanding

Extreme weather prediction

# NLP and ASR Models

Multilingual speech recognition has become markedly simpler compared to just five years ago.

Whisper (OpenAI) model is very popular for this task. Besides open source, there are many cloud offerings (Google Cloud: Chirp, Amazon, Azure, etc)

Speech generation: SpeechT5, Bark, MMS, other

NLP and Conversational AI: a wide abundance of models , (LLama 2, Mistral, many open source models, many commercial models, (Google, Amazon, OpenAI, MS Azure, Cohere, Anthropic ) etc etc)



# NLP and ASR Models

Support of Emergency Response services

Accessibility features

Smart in-vehicles conversational and visual interfaces

Interactive guidance and navigation, informational assistance, safety and emergency information, civil engagement and community information, personalized interactions

Call Intelligence solutions for all city call centers and conversation-based services, coupled with privacy-preserving smart conversational interfaces designed to enhance citizen engagement within urban environments

Plethora of new opportunities within the context of smart cities (NLP/ASR opportunities)

# Other foundation models for Smart Cities

Foundation Models that will be useful in Smart Cities

Time series, strong progress in forecasting for time series

Document Understanding, other forms of Document Intelligence

Wearable Sensors data

IoT Sensor data

Text to Image

Reinforcement learning

Robotics

# Conclusion

There is a strong need in massive collections of publicly available data to develop more foundation models: images, video, sensor data, audio, transportation data, to enable wide scale development of Generative AI applications for Smart Cities by private businesses and government organizations

Having foundation models will enable many new AI based use cases and products for Smart Cities

# Conclusion

We provided a brief overview of Generative AI, Foundation Models and their potential application for Smart Cities

The overview is not comprehensive, there are many more applications are possible