

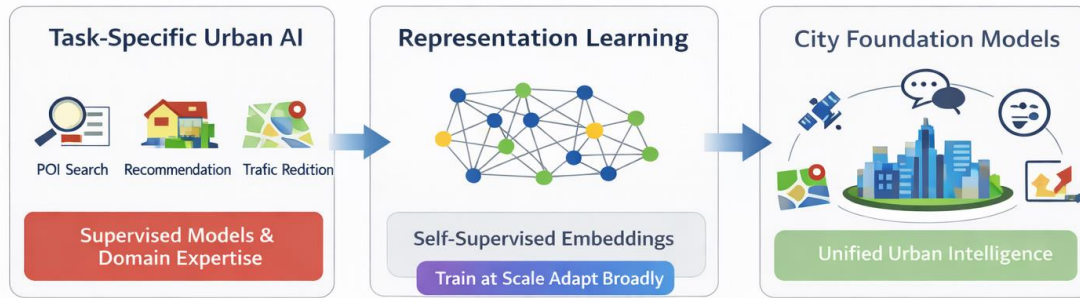
Toward City Foundation Models: From Task-Specific Spatial Approaches to Urban General Intelligence

Gao Cong

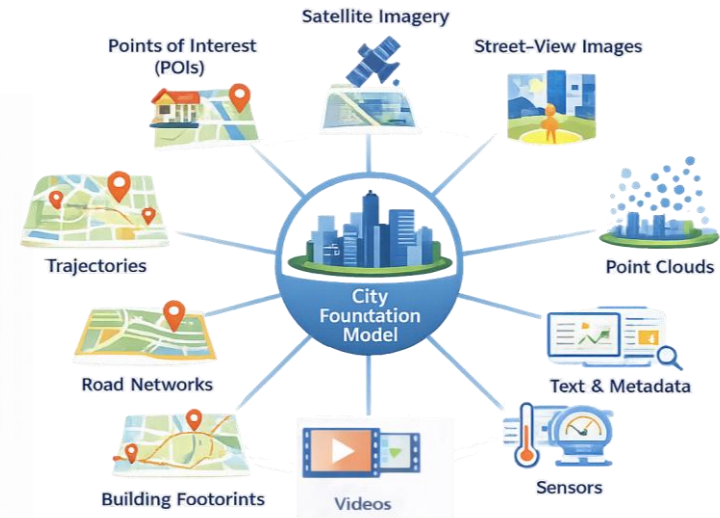
Frontiers of Computer Science, DOI: [10.1007/s11704-026-52054-3](https://doi.org/10.1007/s11704-026-52054-3)

City Foundation Model

The Perspective outlines a path toward **City Foundation Models**: general-purpose models that can learn from multimodal city data and support many urban tasks—such as retrieval, prediction, reasoning, and decision support—within a single, unified framework.



From Task-Specific AI to City Foundation Models



Geospatial models have largely been built for individual applications, producing embeddings that perform well for one task but transfer poorly to others. The field is now shifting toward self-supervised learning on large-scale unlabeled urban data to learn reusable representations from intrinsic city structure (e.g., road connectivity and trajectory patterns), enabling broader multi-task use and paving the way for City Foundation Models..

City Foundation Models (City FMs) as a new class of urban-focused foundation models designed to integrate spatial, temporal, and semantic signals across multiple modalities

Open Research Directions

The Perspective lays out a research agenda with six priority directions for making City Foundation Models reliable and useful in real urban systems.

Open Research Roadmap for City FMs



1. Defining capability scope and new applications
2. Multimodal urban data integration and scalable training
3. Spatially grounded reasoning with verifiable gains
4. Agentic capabilities and integration with existing spatial systems
5. Benchmark datasets and evaluation protocols
6. Ethical and societal considerations