Urban Intelligence: IoT data integration and movement analysis

Name: Kyoung-Sook Kim

Title: Team Leader

Organisation: Artificial Intelligence Research Center

National Institute of Advanced Industrial Science and Technology

Email: ks.kim@aist.go.jp

Self-Introduction

- Education: B.S., M.S., and Ph.D.
  - Pusan National University, South Korea
- Work Experiences
  - 2007.11 ~ 2014.03: Researcher, NICT, Japan
  - 2014.04 ~ 2017.08: Researcher, AIST, Japan
  - 2017.09 ~ present : Team Leader, AIST, Japan
- Research Projects
  - Spatial and Spatiotemporal Data Management
  - Big Data Analysis and Visualization
  - IoT-based Smart City Applications
  - etc.
- International Standardization
  - A co-chair of Moving Features SWG of the Open Geospatial Consortium(OGC)

AIRC@AIST

AIST
NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY (AIST)

Energy and Environment | Life Science and Biotechnology | Information Technology and Human Factors | Materials and Chemistry | Electronics and Manufacturing | Geological Survey of Japan | National Metrology Institute of Japan

HOME > About AIST

AIST:AI

The National In
organizations ir
industry and so

For this, AIST is
comprehensive

AIST, as a core
research and d
formulated with

AIST is also act
comprehensive

Initiativ

Research Unit

- Information Technology Research Institute
- Human Informatics Research Institute
- Intelligent Systems Research Institute
- Automotive Human Factors Research Center
- Robot Innovation Research Center
- Artificial Intelligence Research Center

President
Budget
ender Equality
s
location count.

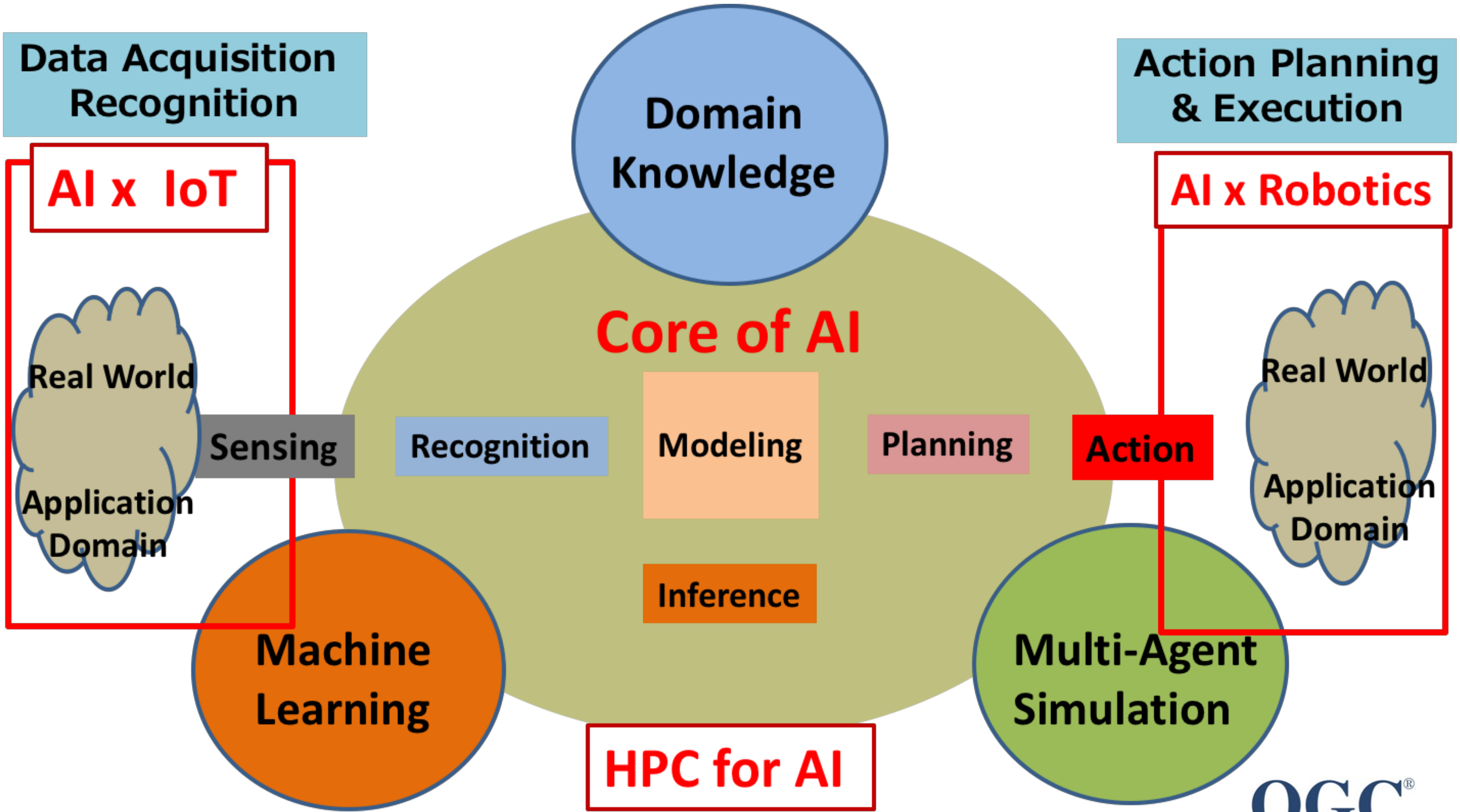
Artificial Intelligence Research Center (AIRC)

AI Embedded in the Real World
- from the Internet to the Real World -

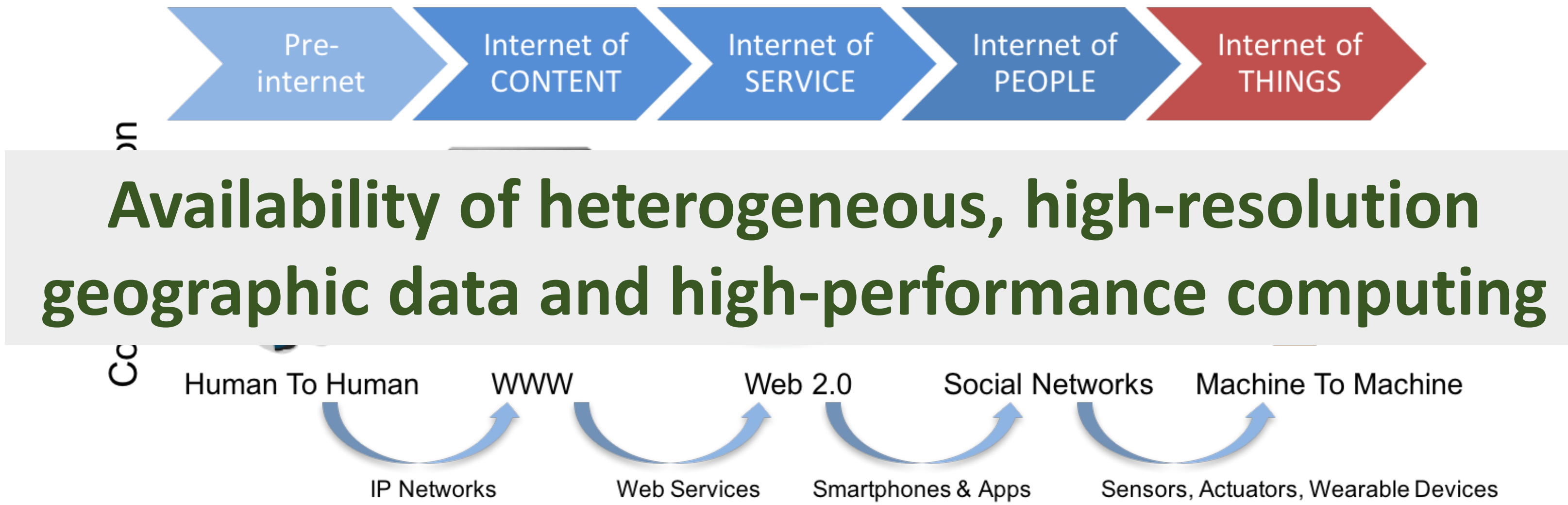
AI which cooperates with Human

Cooperative Autonomy, Explainable AI

AI Embedded in the Real World



IoT + AI + Geo



Availability of heterogeneous, high-resolution geographic data and high-performance computing

- An infrastructure of interconnected objects, people, systems and information resources together with intelligent services to allow them to <u>process information of the physical and virtual worlds and react</u>. (ISO/IEC JTC 1/SWG 5 AHG1)
- The fusion of the physical (reality) and virtual (perception) worlds accelerates **geospatial capture, coordination, and intelligence** in unprecedented ways.

HPC for IoT + AI + Geo

as the world's first large-scale OPEN AI Infrastructure

ABCI AI Bridging Cloud Infrastructure

- Top-Level SC compute & data capability, 0.55 EFLOPS (HP), 37 PFLOPS (DP)
- Open, Public & Dedicated infrastructure for AI & Big Data algorithms, software, and applications
- Platform to accelerate joint academic-industry R&D for AI in Japan

- 0.55 EFLOPS (HP), 37 PFLOPS (DP)
- 2.3 MW
- < 1.1 Avg. PUE

Operational 2018 August

ABCI

AI Infrastructure for everyone

Ultra-dense IDC design from ground-up

- Custom inexpensive lightweight “warehouse” w/ substantial earthquake tolerance
- X20 thermal density of standard IDC

Extremely green

- Ambient warm liquid cooling, and high-efficiency power supplies
- Commoditizing supercomputer cooling technologies to Cloud IDC (70kW/rack)

De facto HW&SW architecture

- State-of-the-art AI and HPC software properties built for de facto commodity arch.
- Rapid technology transfers to commercial clouds and companies

Leveraging software ecosystem

- Container & repository support, to ease development & deployment, as well as to share/reuse codes among community

Data happy but secure

- Multi-PB scale storage & data platform for sharing/publishing your own data
- Both comm. channel and data can encrypted by de facto encryption methods

Capable to serve for >100 research projects, >1000 researchers & engineers

ABCI HW

High-Performance Computing System
0.55 AI-EFLOPS, 37.2 PFLOPS
476 TiB Mem, 1.74 PB NVMe SSD

Computing Nodes (w/ GPU) x1088

GPU	NVIDIA Tesla V100 SXM2 x4
CPU	Intel Xeon Gold 6148 x2
Memory	384GiB
Local Storage	1.6TB NVMe SSD
Interconnect	InfiniBand EDR x2

Multi-platform Nodes (w/o GPU) x10
- Intel Xeon Gold6132 (2.6GHz/14cores) x2
- 768GiB Memory, 3.8TB NVMe SSD

Interactive Nodes x4

Management and Gateway Nodes x15

Large-scale Storage System
22 PB GPFS

DDN SFA14K (w/ SS8462 Enclosure x 10) x3
- 12TB 7.2Krpm NL-SAS HDD x2400
- 3.84TB SAS SSD x216
- NSD Servers x12

Object Storage
- Now in development
- Available from 4Q 2018

Interconnect (Infiniband EDR)
- Mellanox CS7500 x2
- Mellanox SB7890 x229

Gateway and Firewall
- Nexsus 3232C x2
- FortiGate 1500D x2
- FortiAnalyzer 400E x1

100Gbs

SINET5

Service Network (10GbE)

ABCI SW Stack

Software

Operating System	CentOS, RHEL
Job Scheduler	Univa Grid Engine
Container Engine	Docker, Singularity
MPI	OpenMPI, MVAPICH
Development tools	Intel Parallel Studio XE Cluster Edition, PGI Professional Edition, Python, Ruby, R, Java, Scala, Perl
GPU SDKs & Libraries	CUDA compiler, IDE, Debugger, Profilers, cublas, cufft, nvgraph, cudnn, …
Deep Learning	Caffe, Caffe2, TensorFlow, Theano, Torch, PyTorch, CNTK, MXnet, Chainer, Keras, etc.

Container support

- Containers enable users to instantly try the state-of-the-art software developed in AI community
- ABCI supports two container technologies
  - Docker, having a large user community
  - Singularity, recently accepted HPC community
- ABCI provides various single-node/distributed deep learning framework container images optimized to achieve high performance on ABCI

ABCI Grand Challenge Program

FREE

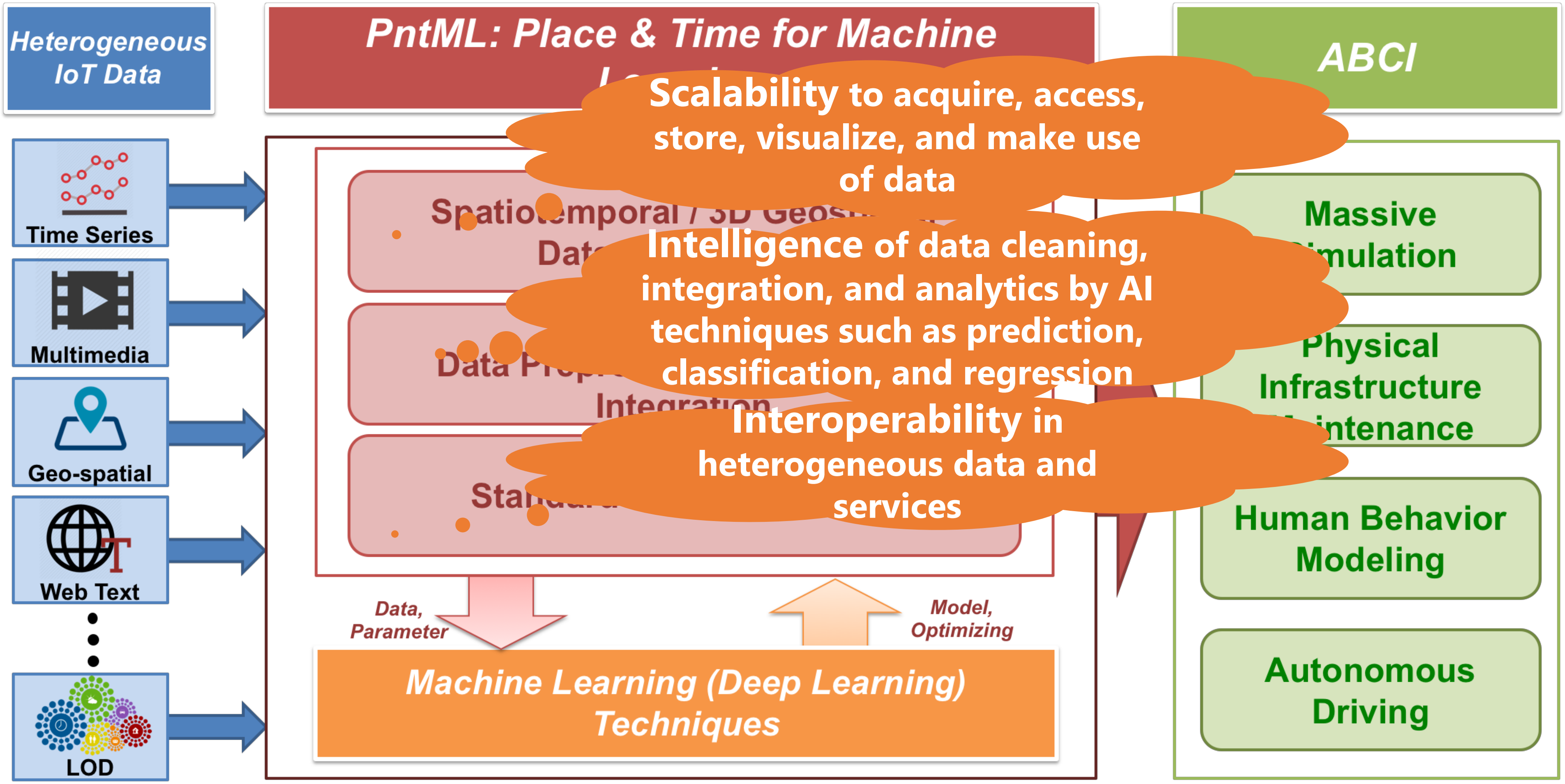
http://abci.ai

Resource: 1,088 nodes（4,352GPU） 24 hours

	Application Period	Notification of Selection	Challenge Execution Period
No.1 (Finished)	2018.4.1-30	2018.5.30	2018.7.23-27
No.2 (Application Closed)	2018.8.1-31	2018.9.30	2018.10.23-26
No.3	2018.11.1-30	2018.12.21	End of 2019.1. (Tentative Date)

- Themes are open. An individual or a group may propose a theme.
- About 2 themes will be selected for each ABCI Grand Challenge program period.
- Joint research themes with AIST may be accepted if both parties are in agreement.
- Testing opportunities will be available with small-scale trials before the challenge.
- ABCI is free of charge including during testing.
- Challengers will be asked to follow the ABCI Agreement.

Data Platform for IoT + AI + Geo



Seamless 3D Geospatial Management

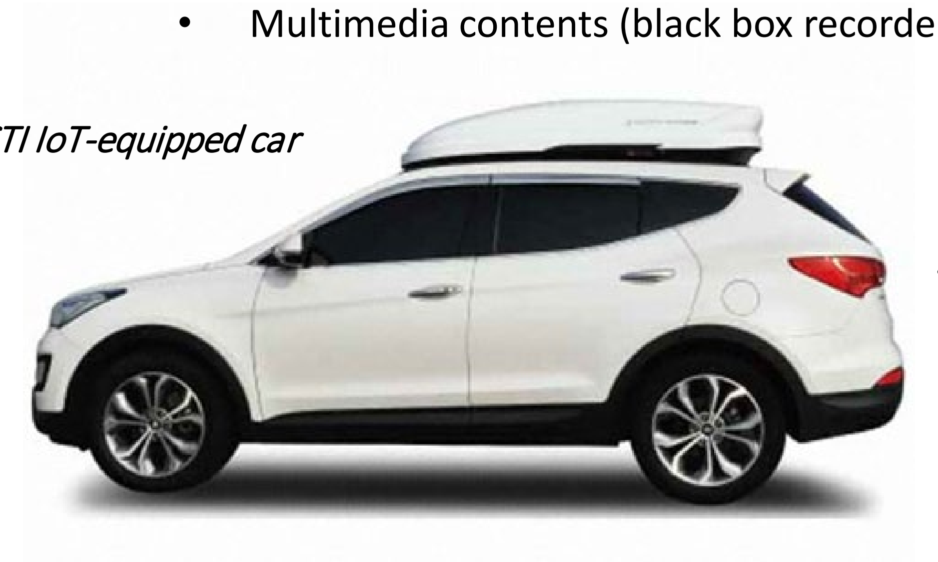
Use Case: KISTI Urban Sensing Data by Taxis

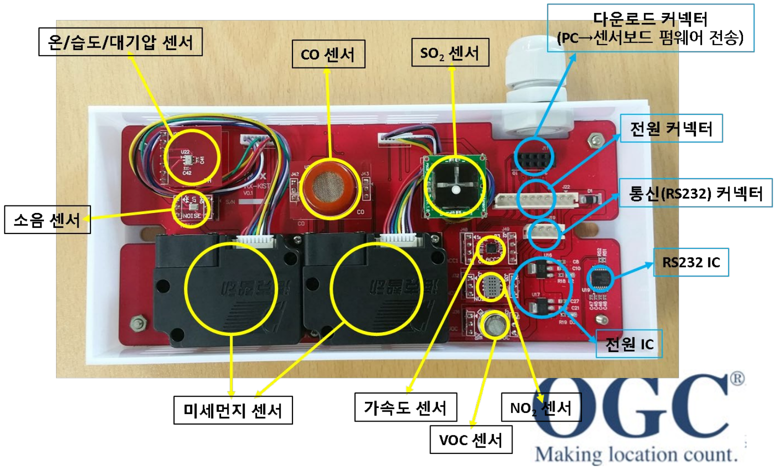
- Mobile urban sensing dataset
  - Fifteen-type raw data
    - Moving object properties
      - time, geolocation, and vibration
    - Environment information
      - temp, humid, particle matters(PM2.5, 10) NO_2, SO_2, CO, VOC, noise
    - Multimedia contents
      - Video image from Black box recorder, webcam
  - Utilizing parts of KISTI data sets
    - Moving object properties (time, geolocation)
    - Multimedia contents (black box recorder)

Korea Institute of Science and Technology Information
www.kisti.re.kr

Collaboration with Dr. Ryong Lee

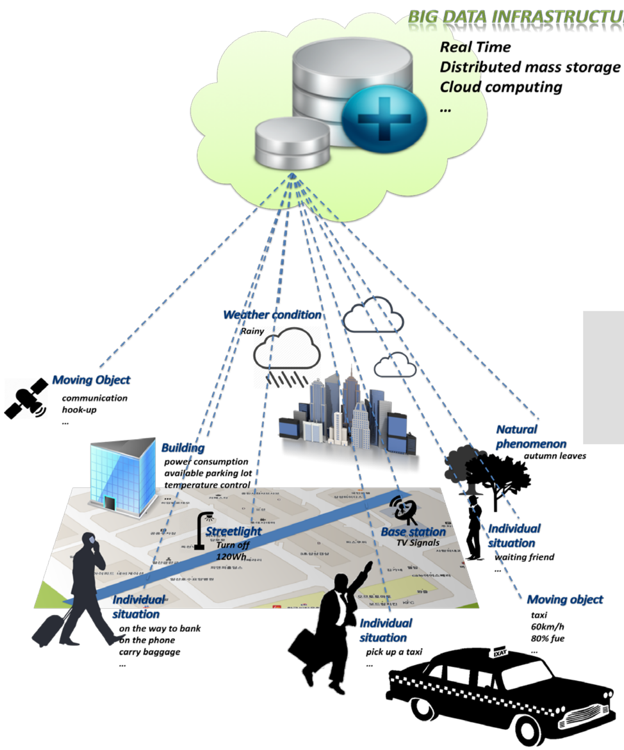
KISTI IoT-equipped car





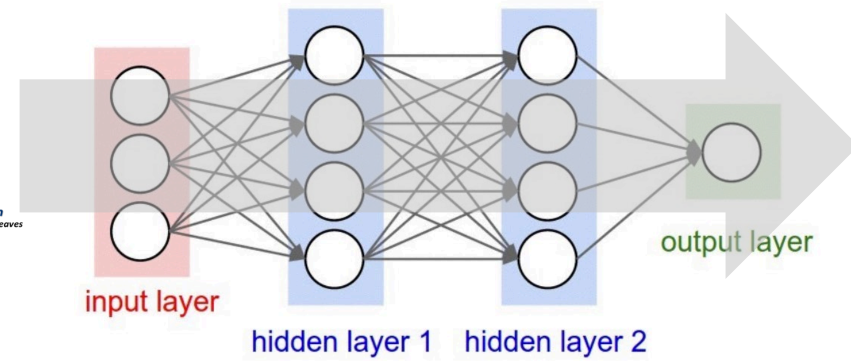
OGC®
Making location count.

Benefits of OGC Moving Features



Different spatial and temporal granularity

Machine Learning Network



How to align data at an arbitrary time and location?

OGC Moving Features

Time instance t in T= [a, b], geometry g in G

$mf(t) : T \rightarrow G, mf(g) : G \rightarrow T$

Making location count.

OGC Moving Features Modules

Service Interface Specifi-cations

16-140r1 JSON (RESTful API)
(for handling moving feature data over HTTP)

16-120r3 Moving Features Access
(guideline for implementing interfaces to support moving feature data)

Encoding Specifi-cations

14-084r2 Simple CSV (compact encoding for massive moving points)

16-114r3 netCDF
(compact binary encoding)

16-140r1 JSON
(for encoding trajectories, linestring, polygon with dynamic non-spatial attributes)

14-083r2 XML Core
(for encoding trajectories)

Data Model

Moving Features 0D
(points)

Moving Features 1D/2D
(lines, curves, polygons, etc.)

Moving Features 3D
(cubes, spheres, 3D model, etc.)

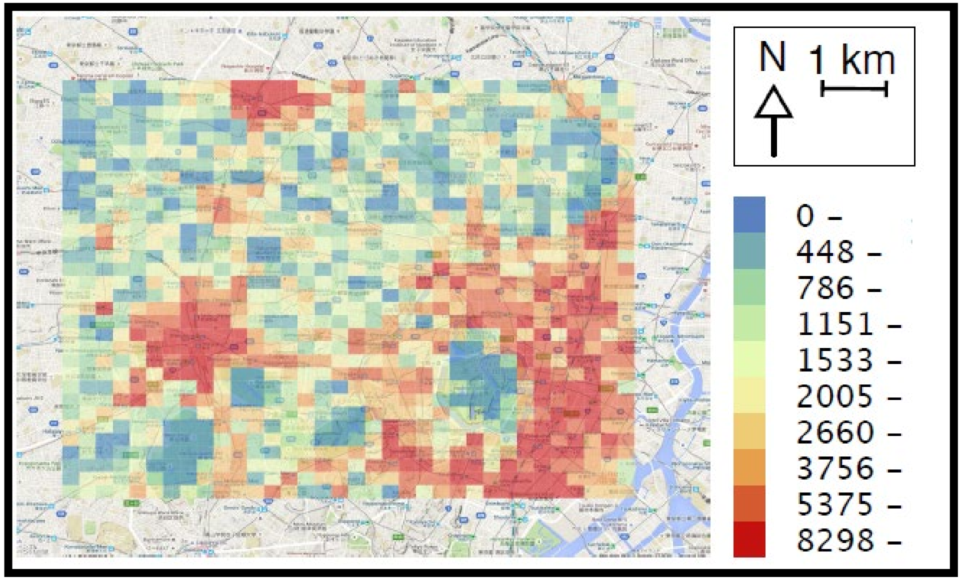
OGC®
Making location count.

Data Integration

Spatio-temporal Data Analysis

Movement Prediction

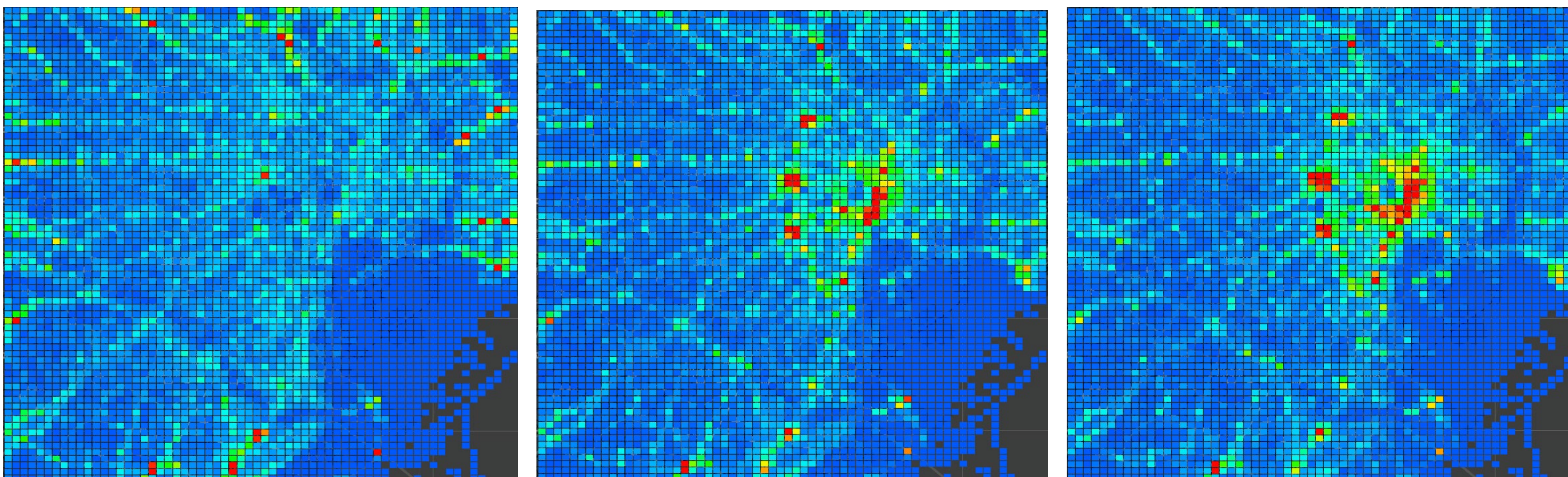
- Mesh-based crowd density presentation → Density Map
  - ✓Some web services like “its-mo navi” can provide crowd density presentation of the current or past.



Prediction function for the future (e.g. 2 hours later) is still challenging and not included in such kind web service.

Urban-scale Mobility

- Density map can be obtained/observed at fixed timestamps (e.g. every 10 minutes, 08:10 -> 08:20 -> 08:30 -> …).
- Given current observed multiple steps of density map, predict or simulate next multiple steps.



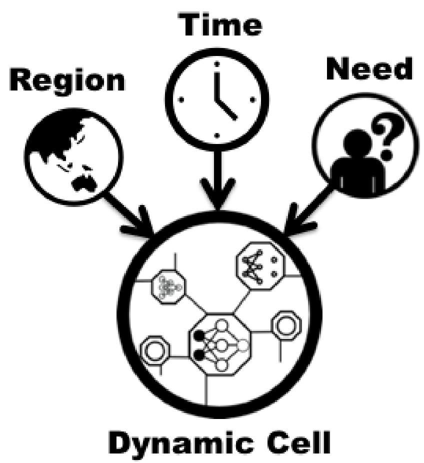
08:10 density map → 08:20 density map → 08:30 density map

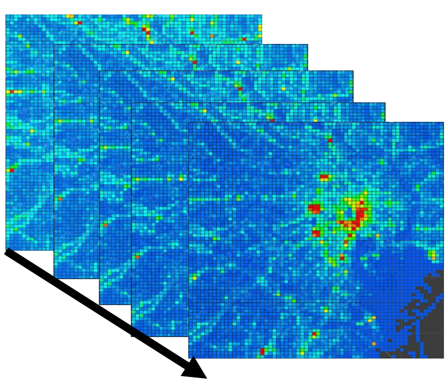
→ How the density map will look like at 08:40, 08:50 and 09:00 ?

OGC® Making location count.

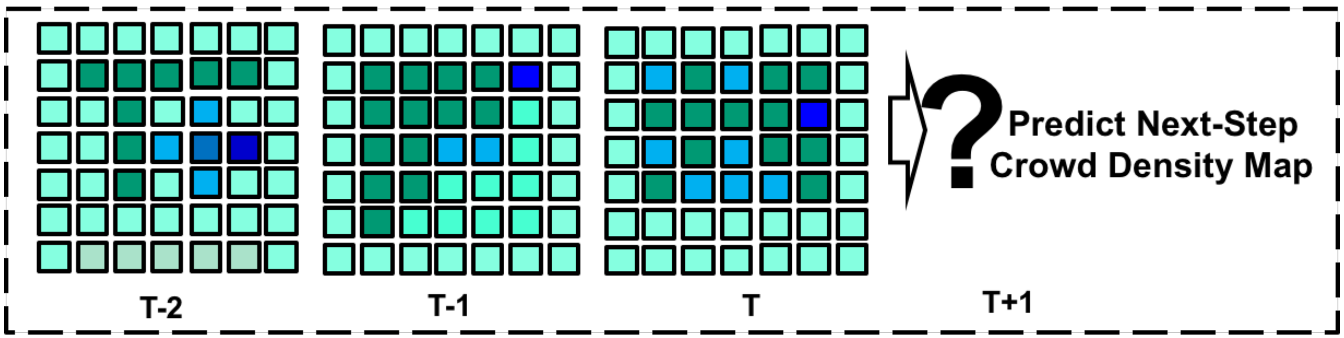
Dynamic Cell Modeling of Urban-scale Mobility

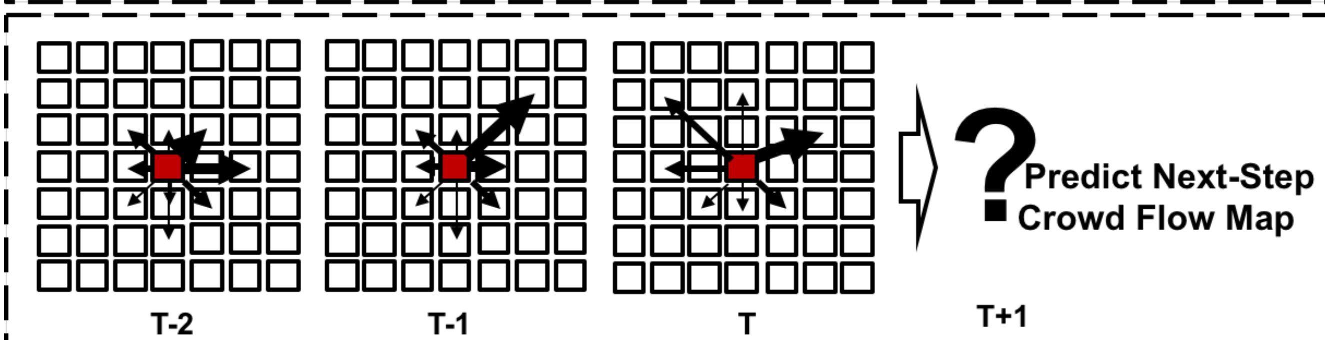
Dynamic cells: a new model and framework that can simultaneously predict movement density and flow to seamlessly connect between 3D indoor and outdoor.





Crowd density/flow prediction is a challenging but important topic in the field of urban computing and simulation.

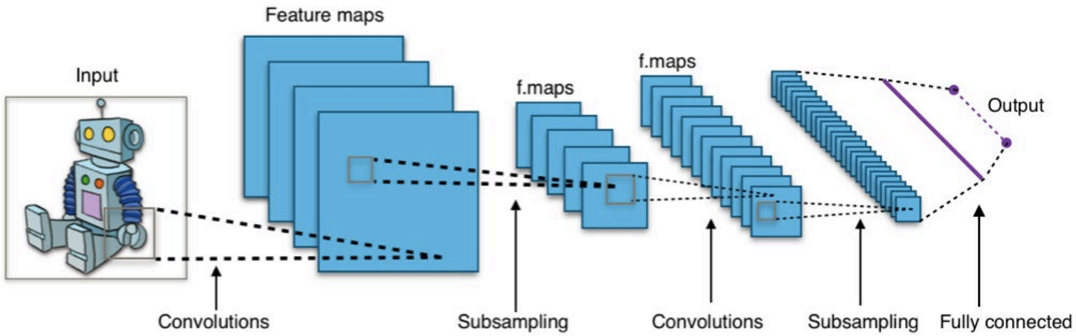




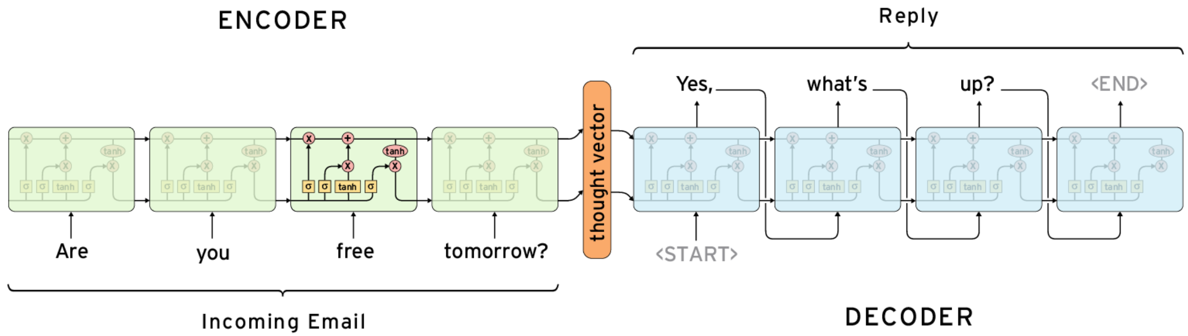
Making location count.

Movement Prediction based on Deep Learning

- Deep learning technologies have demonstrated superior performances on various datasets.
  - Image recognition: Convolutional Neural Network (CNN).



  - Sequential data (e.g. speech, text): Recurrent Neural Network (RNN).



→ **deep-learning technologies to predict crowd density**

Preliminary Results

- Two deep video models (1) CNN and (2) Convolution LSTM.

✓CNN Result

```
Model MSE for multi step, 0, 7.669393
Model MSE for multi step, 1, 11.859134
Model MSE for multi step, 2, 16.573191
Model MSE for multi step, 3, 21.872316
Model MSE for multi step, 4, 27.496741
Model MSE for multi step, 5, 33.319080
```

Average Model MSE : 19.798309

✓Convolutional LSTM Result

```
Model MSE for multi step, 0, 7.014705
Model MSE for multi step, 1, 10.258397
Model MSE for multi step, 2, 13.578582
Model MSE for multi step, 3, 17.053038
Model MSE for multi step, 4, 20.590691
Model MSE for multi step, 5, 24.188083
```

Average Model MSE : 15.447249 ← **Better than CNN**

➔MSE is really good at the citywide level for this short-video model.

➔(RMSE < 4.0 for each 500-meter mesh-grid.)

Questions for the Audience

- What is the biggest barrier for sharing mobility data of crowd?
- How do we share the mobility data beyond the concern of privacy?
- The mobility data is closely related to the physical infrastructures. How do we predict the urban mobility when a disaster happens and it destroys the physical infrastructures?