Applying Mobile Big Data for Transport Planning

Relationship between Big Data use and Personal Information Protection

Japan Transport and Tourism Research Institute

Research Fellow
Toshiaki MUROI

Top Questions on Big Data Utilization

① What is the Act on the Protection of Personal Information in Japan?
（What is the definition of “Personal Information"?）

➡ Introduce Act and definitions of personal information and “guideline”

② How has Japan’s carrier (such as NTT Docomo) cleared the challenge of protecting personal information when providing their data to third parties?

③ Have you ever had a problem handling the personal information in Japan?

➡ Analyzes success and failure cases and their factors in Japan, and introduces the key points of success in Japan

Act on the Protection of Personal Information in Japan

How can you comply with Act on the Protection of Personal Information ?

Personal information, personal data, retained personal data
(Article 2 paragraphs 1, 4 and 5)

The term “personal information”... shall mean information about a living individual which can identify the specific individual by name, date of birth or other description contained in such information.
(including such information as will allow easy reference to other information and will thereby enable the identification of the specific individual.)

Details about this personal information are described in the guidelines set by the Ministry of Internal Affairs and Communications (MIC).

MIC Guidelines on the Protection of Personal Information in the Telecommunications Business

Guideline issued by the Ministry of Internal Affairs and Communications in Japan

電気通信事業における個人情報保護に関するガイドライン
（平成１６年８月３１日総務省告示第６９５号）

最終改正　平成１７年１０月１７日総務省告示第１１７６号

第１章　総則

（目的）

第１条　このガイドラインは、電気通信事業の公共性及び高度情報通信社会の進展に伴い個人情報の利用が著しく拡大していることにかんがみ、通信の秘密に属する事項その他の個人情報の適正な取扱いに関し、電気通信事業者の遵守すべき基本的事項を定めることにより、電気通信サービスの利便性の向上を図るとともに、利用者の権利利益を保護することを目的とする。

（定義）

第２条　このガイドラインにおいて、次の各号に掲げる用語の意義は、当該各号に定めるところによる。

一　電気通信事業者　電気通信事業（電気通信事業法（昭和５９年法律第８６号）第２条第４号に定める電気通信事業をいう。）を行う者をいう

Act on the Protection of Personal Information was revised in 2017.

Specifying basic matters that telecommunications carriers should observe for proper handling to protect personal information.

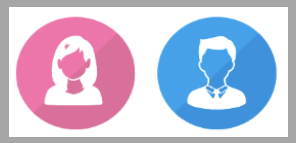
Guideline revision history

- Enacted in 1998
- Revised in 2005
- Revised in 2009
- Revised in 2010
- Revised in 2011
- Revised in 2013
- Revised in 2015
- Revised in 2017

Personal Information and Anonymously Processed Information

Original Data (including Personal Information)


ID Number


Gender


Date of Birth


Name


Phone number


Detailed location

Deletion and encryption of some data


ID Number


Encrypted


Gender


D.O.B


Name

Deleted


Detailed location


Phone number

Deleted

Personal Information

Deletion, encryption, aggregation, and concealment of some data


ID Number

Deleted


Gender


D.O.B

Tabulated & Concealed


Name

Deleted


Phone number

Deleted

Anonymously Processed Information

Relationship between Personal Information and Anonymously Processed Information

① Operational data

● Data with which people can be identified
➢ Mr. Tato Yamada

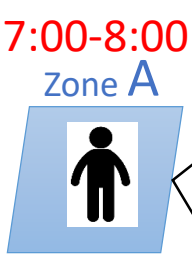
- ID
- Name
- Phone Number
- Gender
- D.O.B.
- Detailed location/time
- ⋮

Personal information

Non-identification process

② Data without personal ID

● Somebody's data
➢ Mr. X located in Zone A



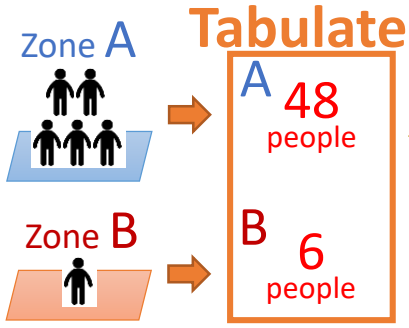
- Encrypted ID
- Name
- Phone Number
- Gender
- Birth year
- Zone
- Around time

Anonymously Processed Information

Tabulation process

③ Estimated population

● Some group's data
➢ 40's, 30's, … in Zone A etc.

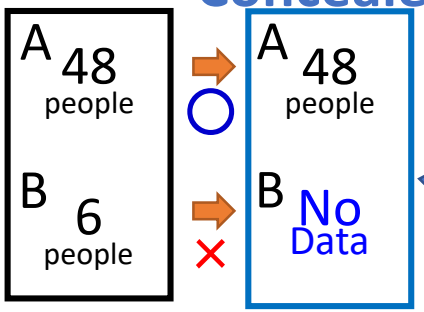


Group information

Concealing process

④ Mobile spatial statistics (NTT Docomo)

● Data of a group including a specific number of people

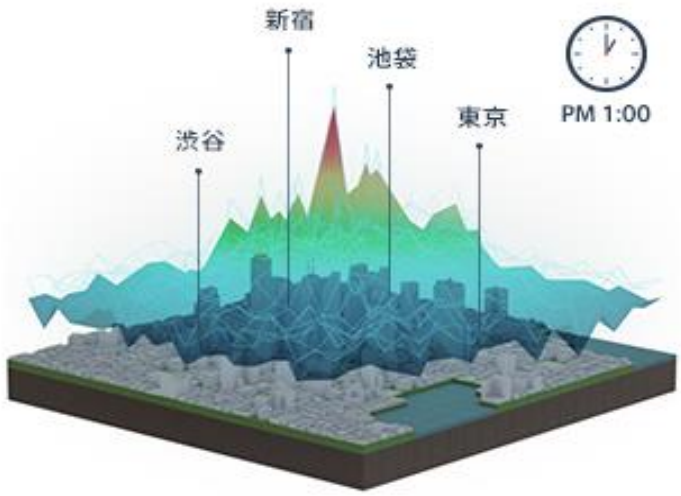


Statistics (Non-personal information)

Two examples of using big data in Japan

NTT docomo

CASE 1



Mobile Spatial Statistics
Cell phone location information
Number of costumers 79,203,000(2018)
market share 44.5% (2018)
Research started in 2008
Starting in 2013, this service has been implemented

JR東日本

CASE 2

SUICA（Super Urban Intelligent CArd）
Information when passing the gate
SUICA release quantity 69,420,000(2018)
Num. of Passengers 17,000,000/day(2018)
Research started in 2009
Announcing commercialization in 2013, Canceled due to public opinion

Study group on examining issues around ICT services from the user perspective by NTT Docomo

Base station location info.

"Personal information" means information that can identify the specific individual.

De-identification
Aggregation
Concealment

The act of causing a loss of identification with anonymization is not the use of personal information.

Mobile big data

Determined that the Act on the Protection of Personal Information will not apply because of loss of identification

Precautionary principle

Prompt businesses to pay social consideration, not just comply with the Act.

Study Group met four times
(from Sep. 24 to Dec. 10, 2009)

Law scholar (Chair)	Masao Horibe Prof. Emeritus, Hitotsubashi Univ.
Information economy expert	Akihiko Shinozaki Prof., Kyushu Univ. Graduate School
Statistics expert	Takeshi Hiromatsu Prof., Institute of Information Security
Consumer perspective advocate	Sawako Nohara President & CEO, IPSe Marketing, Inc.
Lawyer	Tsunemichi Yokoyama Mori Hamada & Matsumoto

Legal compliance + Precautionary principle + Third-party committee

Compliance with precautionary principle (example)

(1) Promote public relations	⇒	Issue press releases and disclose information on the website to eliminate users' anxiety and discomfort.
(2) Secure transparency	⇒	Draw up guidelines on big data creation which clarify the purpose, basic principle, creation procedures, three-step processing, etc., and make them available on the website.
(3) Secure opportunities for user involvement	⇒	Ensure that providers can accept, at anytime, users' requests that base station location information should not be provided to third parties.

Independent Committee report (excerpt)

Organized by docomo

- Automated de-identification, aggregation and concealment may make personal identification impossible.

(1) Does not violate the Act on the Protection of Personal Information, (2) Comply with the precautionary principle, and (3) conclusion of the Independent Committee

⇒ (Telecommunications carriers will) provide third parties with data when approved by Ministry of Internal Affairs and Communications.

Diversity of mobile big data utilization fields

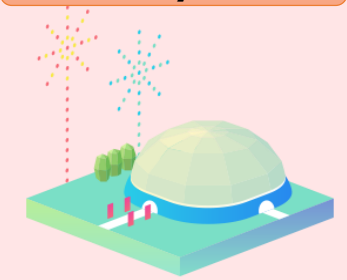
Disaster prevention

Public transportation

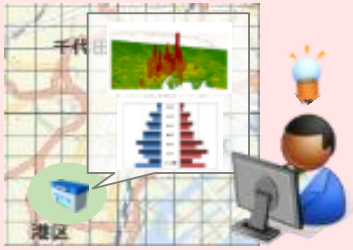
Urban planning

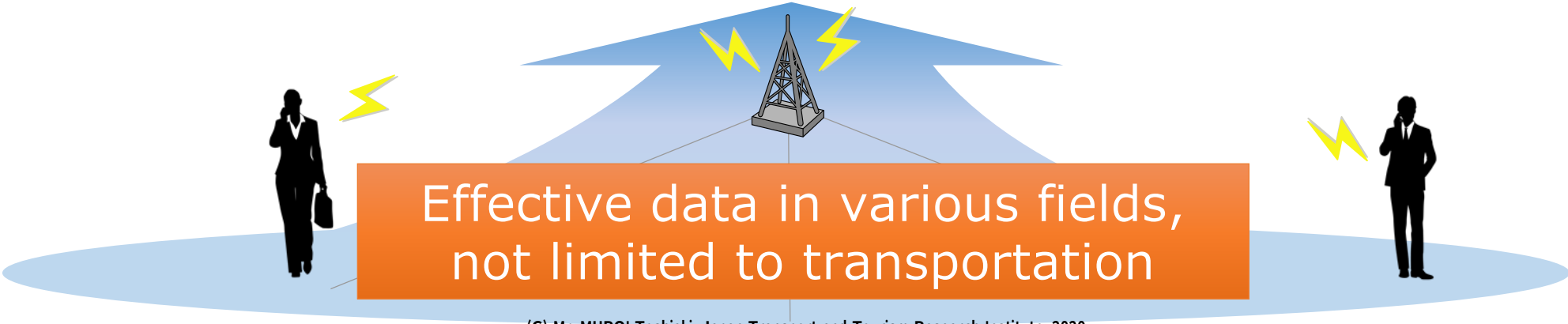
Tourism encouragement

Event analyses

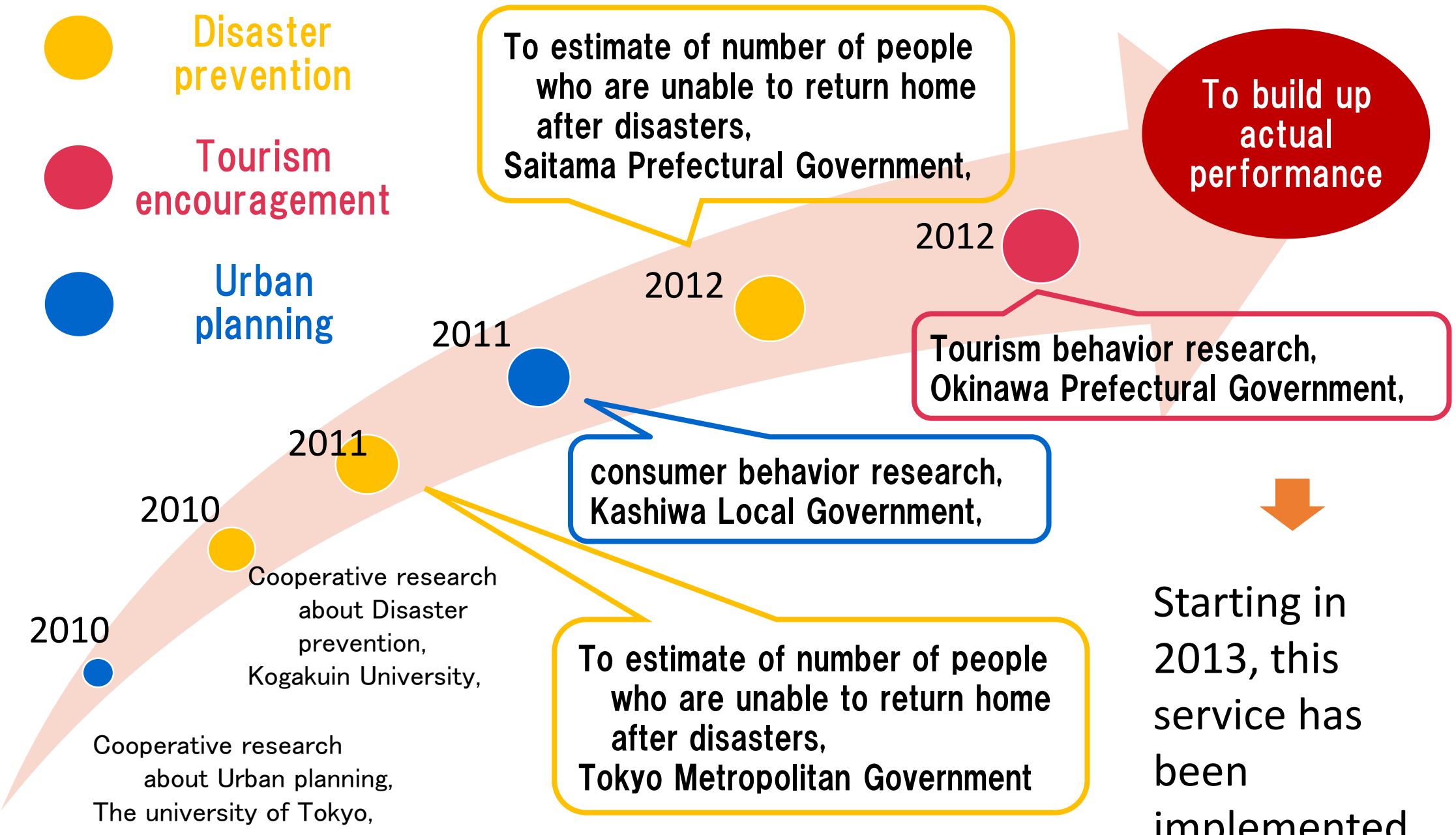


Customer information





cooperative research conducted by NTT DoCoMo

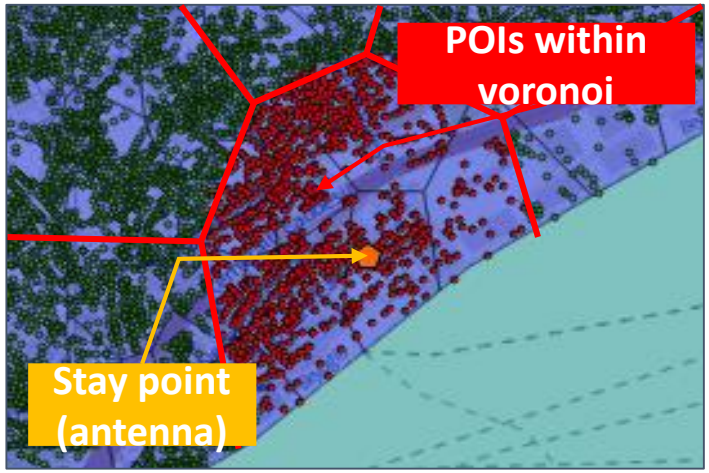


Differences in handling data between the two companies

	NTT docomo		JR東日本	
(1) Compliance with laws and guidelines	○	Both complied	○	Both complied
(2) Prior explanation and information disclosure to users	○	Establishment of a Independent Committee and Dedicated website opened	×	Hardly explained
(3) Dissemination of opt-out procedure	○	Instructions on website	×	(at first) there was no announcement
(4) Data accuracy is appropriately “rough”	○	Only hourly & 500m mesh count results are provided	×	Provide data without aggregation
(5) Start use from public purpose	○	Start use for public purposes by academics and governments	×	Start with private business use
(6) Benefits returned directly to users	○	Providing wide benefits to society	×	Unable to show benefits to general users

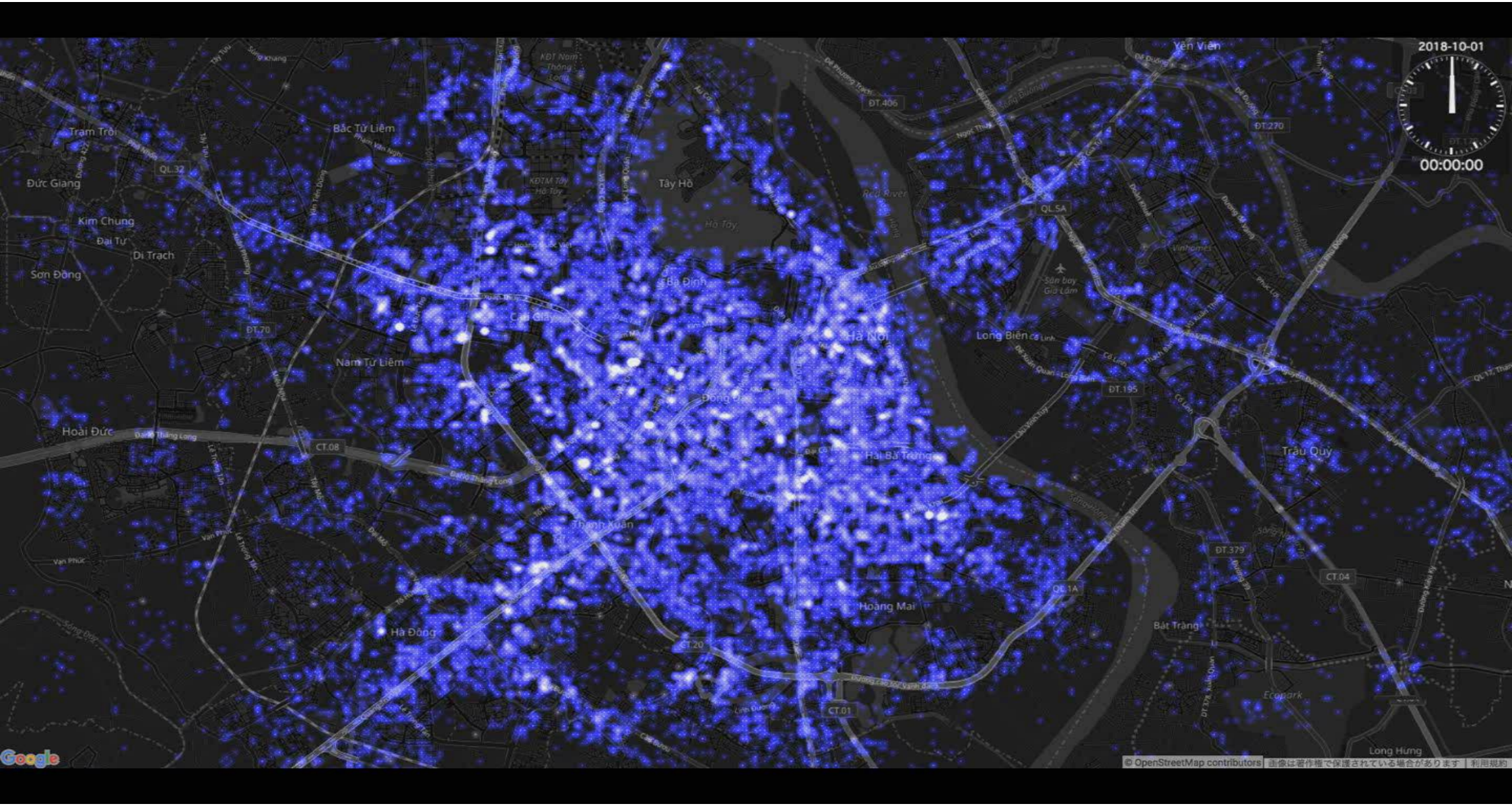
Market Failure Possibilities and Risks

There is a high demand for private operators to want to use mobile big data.
However, data that the private sector wants to use mainly for marketing is generally a unit for each building or facility, such as each red dot shown in the figure.
This is too detailed in terms of space and time, giving the impression that an individual has been identified and may be at high risk of social rebound.



Matter	Private Use	Public Plan
Main Application	Marketing	Transportation Plan City Master Plan
Space and Time Resolution	Store, Building Narrow（50m²） Short（5min int.）	Zone Wide（500～1,000m²） Long（1hr, 1day）
The risk that an individual is identified	Large	Small
Social resilience risk	Large	Relatively restrained
Publicity	Low	High
Data Publish	Hard (Due to contract between private sector)	Easy (Due to data acquire public groups such as governments)

Estimated People Movement based on VinaPhone CDRs

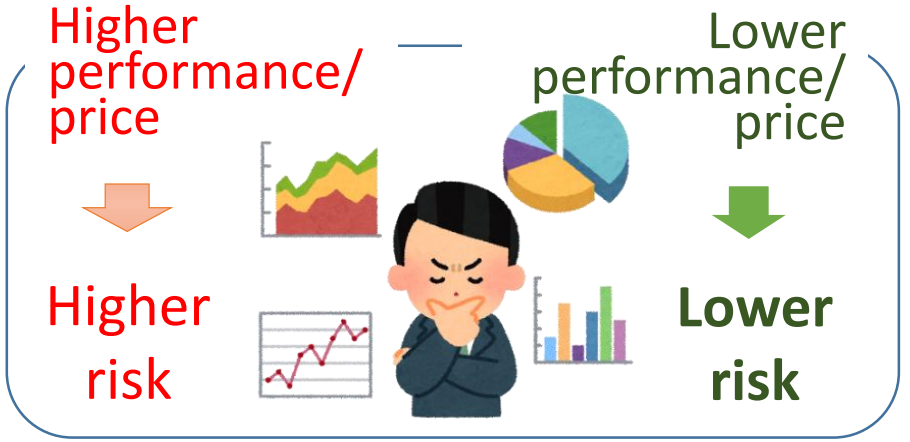


SHIBASAKI & SEKIMOTO LAB. CSIS/IIS/EDITORIA. UTOKYO

Future MBD use in transportation plans

- Growing expectations for PPP/PFI
  - ASEAN governments are having financial difficulties.
- Current demand forecast has issues with accuracy.

- Risk aversion when demand risk is high
  - Investors call for unfailing recovery of their investment
  - Decision makers consider the balance between improvement and investment



- Too much focus on demand risk (price) will result in

Highly accurate demand forecasts and statistics development contribute to truly addressing traffic problems

Thank you for your attention

This survey was conducted with support from the Nippon Foundation.

Supported by THE NIPPON FOUNDATION