12th World Telecommunication/ICT Indicators Symposium (WTIS-14)

Tbilisi, Georgia, 24-26 November 2014

Presentation

Document C/19-E
25 November 2014

English

SOURCE: Ministry of Internal Affairs and Communications, Japan

TITLE: Efforts for Big Data and Open Data in Japan

Ministry of Internal Affairs and Communications

Efforts for Big Data and Open Data in Japan

Director General for International Affairs,
Global ICT Strategy Bureau, MIC, Japan

MORI Kiyoshi

25 Nov. 2014

Arrival of data utilization society

Rapid evolution of ICT

Many kinds and large amounts of data

- is generated, distributed and accumulated
- can be easily handled by companies etc.

Utilizing of the data is important

1 Big Data

Many kinds and large amounts of digital data exist in society, market, etc.

Utilizing

- **creation of new value**
- **resolution of social issues**

2 Open Data

Open up public data owned by the nation and local public bodies to the private sector in formats suitable for secondary use.

Utilizing

- **Creation of new businesses and services**
- **Realization of public services by public-private collaboration**

″Data quality" in accordance with big data and open data

○ "Data quality" is a critical issue.

"Data quality" Ex.

Importance | Accuracy, reliability | Consistency, compatible potential

1 Big Data

There is a wide variety of data, so it is difficult to grasp the current situation.

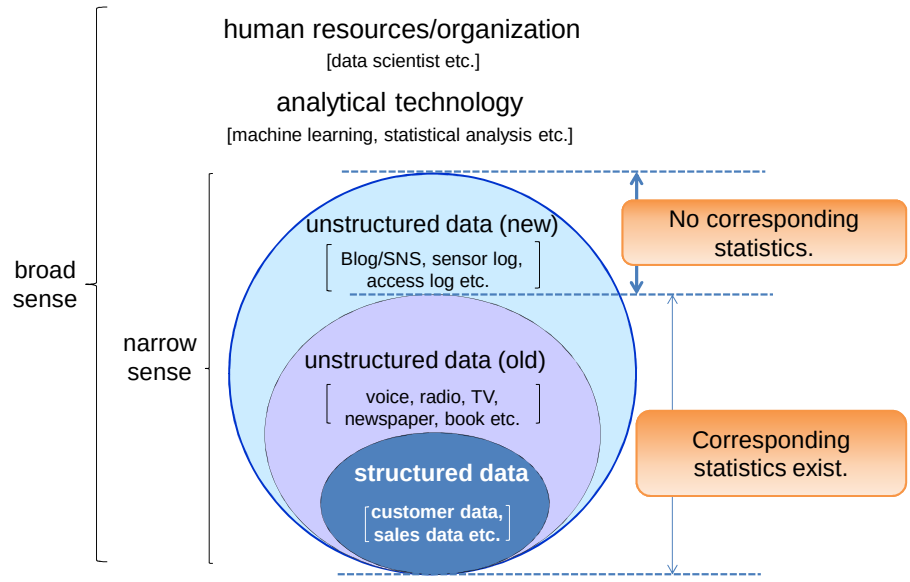
First, one of the issues is to measure the type and amount of data.

2 Open Data

Accuracy and reliability of the data itself is high in many cases. However, data of various fields exists in various data formats.

The important issues are standardization of data formats, rules for secondary use and API.

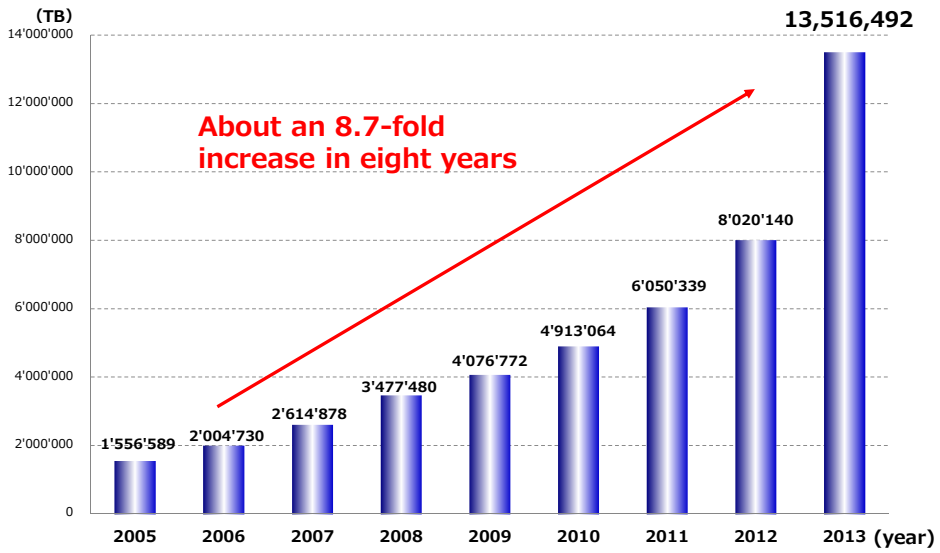
Measurement of distribution and accumulation amount of big data



(Ref.) The targeted media for estimation of Big Data distribution

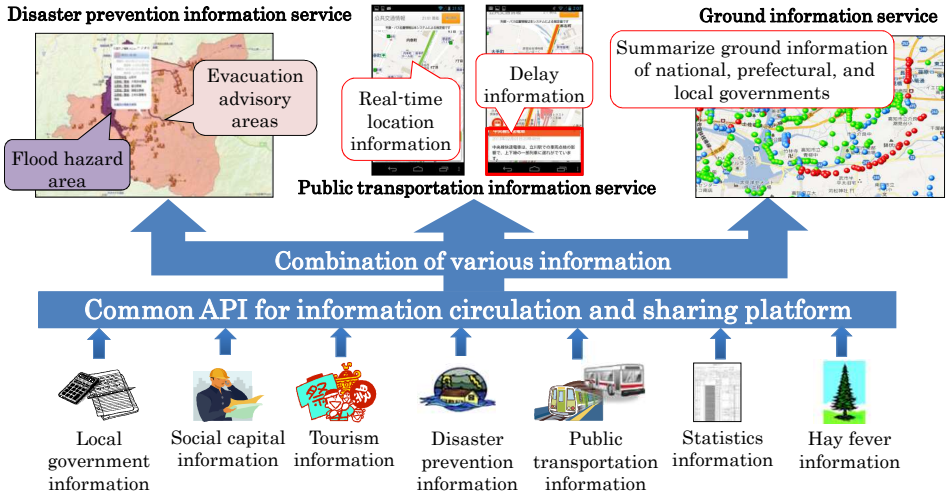
	Structured data		Unstructured data	
Business system	Customer DB	POS data (basket data at the point of sale)	Business diary (text)	【Medical】diagnostic imaging (image, still image, movie)
	【Medical】receipt data	Accounting data	【Medical】electronic health record (text)	CTI voice log data (voice)
Web service (EC, etc)	Sales log in e-commerce			
Sensor GPS M2M	GPS data	RFID data	Sensor log	Security and remote monitoring camera
	Weather data		Traffic-congestion information	
Media content			Video and image viewing log	
Personal media social media			Blog, SNS articles etc (text)	Mobile phone [including PHS,voice]
			Access log (various logs)	E-mail (text)
				Fixed IP phone[voice]

Estimation results: the volume of Big Data distribution in Japan

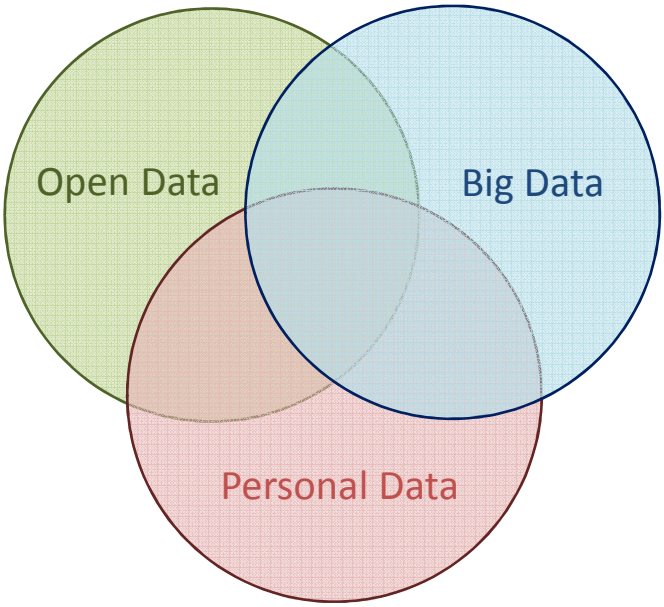


Promotion of Open Data: Demonstration Experiments

The MIC is conducting a demonstration experiment for environmental improvement related to open data. We aim to (1) establish the common API (Application Programming Interface), (2) develop rules for secondary use of data, (3) visualize the benefits of open data.



Three data areas



Challenge of data utilization

1. Utilization of personal data

(Enterprises)
High value of data related directly to users

(Consumers)
Privacy infringement, concerns of exploitation by a third party

A balance between protection and utilization of personal data is important.

2. Data scientist training

Securing human resources who derive useful knowledge from data and use it to solve problems is important. However, there are few such human resources ("Data Scientist").

Data scientist training is important.

Ministry of Internal Affairs and Communications

Thank you for your attention.