Open Government Data Projects in Japan

Shuichi Tashiro
Information-technology Promotion Agency, Japan
About IPA

Information-technology Promotion Agency （IPA）


• 100% funded by Japanese government.

• Working with METI, CAS and other government organizations to realize and ensure secure and reliable IT environment...
Dawn of Open Government Data

**Projects**

- **2010.9** databox (Statistics open data, API)
- **2010.10** ideabox (bulletin board for discussion)
- **2011.7** Energy conservation project (open data, API)
- **2011.10** Japanese character set (open data, API)
- **2012.1** Recovery and Reconstruction Support Program Database (open data, API)

**Policy Papers**

- **2008**
  - Public CIO Forum
  - Started discussion
- **2009.3**
  - Open government project was unveiled
- **2010.7**
  - Open government Lab. was opened
- **2010.5**
  - New IT Strategy
- **2011.3**
  - Principal for promoting eGovernment
- **2012.7**
  - Open government data strategy
  - Strategy for the Rebirth of Japan

---

**Great East Japan Earthquake**

- **2010.10** ideabox (bulletin board for discussion)
Before “Open Government Strategy” issued in 2012

• Many kind of government data had been opened.
The Government Portal Site “e-Gov” (since Apr.2001)

The e-Application System has been operated at the Government Portal Site “e-Gov” enabling to accept multiple administrative procedures such as applications and reports at the single gateway through Internet.

- The e-Application Systems in ministries have been integrated to “e-Gov” for one-stop services. Some systems which are not integrated yet are scheduled to be consolidated into “e-Gov” sequentially.
- The “e-Government Customer Support Centre” is opened to deal with various inquiries on how to use “e-Gov”.

Guide for how to use “e-Gov” in easy-to-understand
Support of utilization of e-Government
Provide cross-ministries administrative information in web sites
Web sites of government ministries
Web Search
Guide
Submit multiple applications at once
Kosumigosha WAN
Electronic application
How can I process an e-application?
I want to establish a new company. How can I do?
I want to collect information from various sources. It’s time consuming to visit and search many different Web sites.
Earthquake early-warning system

- Real-time data link
- All trains stopped safely before the Earthquake wave attack in 2011

(Map & Photo: JR EAST)

Discussion Board

Data box

Open Government Lab.

Idea box

Portal of official twitter

Text Analysis

3/14 11:30
Before “Open Government Strategy” issued in 2012

- Many kind of government data had been opened. But ..
  - Low coordination
    - Published by each ministry in each way
    - Developed individually by each ministry
  - Low interoperability
    - Data format
    - Semantics
    - Character code
  - Low reusability
    - Licensing
    - Technology
The Great Earthquake, March 11 2011

• 20,403 peoples died
• 112,612 houses were lost, 142,575 houses were seriously damaged
• Number of peoples in emergency evacuation centers
  – 386,739 (1 week after)
  – 147,536 (1 month after)
  – 88,361 (3 months after)
A lot of demand..

• Data
  – Who is where
  – Who needs what / who can provide what
  – Availability of road / energy / water / food …
  – Pollution level

…

• Collaboration
  – G/G G/B G/C B/B B/C C/C
Traffic record map to find “travelable road”

“ITS Japan”, NPO

Honda, Pioneer, Toyota, Nissan, Google

Courtesy of Hajime Amano
President of ITS JAPAN
航空機モニタリングの結果について（地表面から1mの高さの空間線量率（μSv/h））

2011.04.29※
2011.05.26
2011.07.02
2011.11.05
2012.06.28
2012.12.28

文部科学省 放射線量等分布マップ拡大サイトより（※2011.04.29 のみ凡例が異なります）
Public appeal for energy saving

- METI (IT Project Office) made a public appeal on its Twitter account for an application that transmit current energy usage data from TEPCO

- Many (more than 50) attractive applications were developed by volunteers
• This is a one-stop service site where users can search through a wide variety of support programs operated by the central and local governments.
Google Person Finder

Google Person Finder: 2011 Japan Earthquake

- An open source web application
- A registry and message board
- Search for information about each other’s status and whereabouts
- Google provide the service on the day.

Voluntary progression was made a week later.

Bulletin board at shelter
Someone take photos of survivors name list.
And upload it on the web.
Someone type it from hand writing to text.
Well coordinated spontaneously

Persons information
Search persons
どうらかを運送してください。
Residents record database vs. GIS
Character code problem
- ideographic variation -

Moji-Joho-Kiban Ideograph set

戸籍 (55,266)
Defined by MOJ

住民登録 (19,563)
Defined by MIC

JIS (10,000)

Daily use Ideographs (2,136)
Recommended by Ministry of Education
We recognized

- Importance of efficient data connection
- Importance of reusable open data

- Lack of interoperability
  - Data structure
  - Semantics
  - Character code
Policy paper of
Open Government Data Strategy

July 2012

IT Strategic Headquarters
Cabinet Secretariat, Government of Japan
Open Government Strategy

I. Objectives

1. Transparency and accountability
2. Participation of citizens
3. Innovation and economic growth
   - Private sectors -> Business opportunity
   - Public sectors -> cost down / efficiency
Open Government Strategy

II. Directions

1. Government shall actively release public data.
2. Public data shall be released in machine-readable formats.
3. The use of public information shall be encouraged whether for commercial or non-commercial purposes.
4. Specific measures shall be taken promptly.
Open Government Strategy

III. Specific Measures

1. Promoting the Use of Public Data
   i. Identification of Public Data Use Needs
   ii. Examination and Organization of Issues Relating to Current Data Provision Methods
   iii. Development of Private Sector Services

2. Development of Environments for the Use of Public Data
   i. Creation of Rules Necessary for the Use of Public Data
   ii. Development of Data Catalogs
   iii. Promotion of Standardization of Data Formats and Structures
   iv. Investigation of Support to Provider Agencies
Opening reusable data

- Datameti.go.jp
- Project started in November 2012
- Published 200 datasets in March 2013
- Using CC for licensing
- Raw data (but almost all are in xls / pdf)
<table>
<thead>
<tr>
<th>Code</th>
<th>2011FY Display Unit</th>
<th>石炭+</th>
<th>石炭製品+</th>
<th>原油</th>
<th>石油製品+</th>
<th>天然ガス+</th>
<th>都市ガス</th>
<th>再生可能・未事業用水</th>
<th>一次エネルギー</th>
<th>産業用</th>
<th>自家用</th>
<th>地域熱供</th>
<th>一般ガス製造</th>
<th>炭製品製造</th>
<th>他転換・品種振替</th>
<th>純転換部門計</th>
<th>自家消費・送配損失</th>
<th>他転換増</th>
<th>消費在庫</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>Prim</td>
<td>4665110</td>
<td>26181</td>
<td>7978559</td>
<td>2120563</td>
<td>4690377</td>
<td>0</td>
<td>844432</td>
<td>724337</td>
<td>885960</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1100</td>
<td>Int:国内出荷</td>
<td>0</td>
<td>0</td>
<td>29630</td>
<td>0</td>
<td>148937</td>
<td>0</td>
<td>844431</td>
<td>724337</td>
<td>885960</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1200</td>
<td>Int:輸入</td>
<td>4665110</td>
<td>26181</td>
<td>7978592</td>
<td>2120563</td>
<td>4541440</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1500</td>
<td>To総供給</td>
<td>-71</td>
<td>-29414</td>
<td>0</td>
<td>-1016476</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1700</td>
<td>St:供給在庫(+取崩/-積増)</td>
<td>0</td>
<td>-3689</td>
<td>-8753</td>
<td>15067</td>
<td>216041</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1900</td>
<td>DD:国内供給</td>
<td>4665039</td>
<td>-6922</td>
<td>7969806</td>
<td>1119154</td>
<td>4906419</td>
<td>0</td>
<td>844432</td>
<td>724337</td>
<td>885960</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>Ener:一次エネルギー</td>
<td>-4177276</td>
<td>1255624</td>
<td>-7933776</td>
<td>6164060</td>
<td>-4869221</td>
<td>1485642</td>
<td>-806461</td>
<td>-724337</td>
<td>-885960</td>
<td>3363100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2100</td>
<td>Pt:事業用発電</td>
<td>-1984388</td>
<td>-177602</td>
<td>-459125</td>
<td>-673828</td>
<td>-2971738</td>
<td>-63851</td>
<td>-132034</td>
<td>-573447</td>
<td>-885960</td>
<td>3233492</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2200</td>
<td>Au:家庭用発電</td>
<td>-231005</td>
<td>-16685</td>
<td>-59</td>
<td>-273292</td>
<td>-31247</td>
<td>-76503</td>
<td>-323266</td>
<td>-150890</td>
<td>0</td>
<td>497727</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2300</td>
<td>Im:産業用蒸気</td>
<td>-208858</td>
<td>-36321</td>
<td>-78</td>
<td>-253408</td>
<td>-22827</td>
<td>-68241</td>
<td>-345910</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2350</td>
<td>Di:地域熱供</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2400</td>
<td>Te:一般ガス製造</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2500</td>
<td>Cr:石炭製品製造</td>
<td>-1745056</td>
<td>1699224</td>
<td>0</td>
<td>-10970</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2600</td>
<td>Oi:石油製品製造</td>
<td>0</td>
<td>0</td>
<td>-7476538</td>
<td>7652921</td>
<td>4906419</td>
<td>0</td>
<td>844432</td>
<td>724337</td>
<td>885960</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2700</td>
<td>Ot:他転換・品種振替</td>
<td>24772</td>
<td>0</td>
<td>497727</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2800</td>
<td>Te:純転換部門計</td>
<td>-4144536</td>
<td>1368616</td>
<td>-7936201</td>
<td>6364428</td>
<td>-4740446</td>
<td>1534501</td>
<td>-807208</td>
<td>-724337</td>
<td>-885960</td>
<td>3727094</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2900</td>
<td>Ov:家庭消費・送配損失</td>
<td>-21642</td>
<td>-106453</td>
<td>-62</td>
<td>-270234</td>
<td>-117920</td>
<td>-48859</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-363994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3000</td>
<td>Ot:他転換増(+受入/-払出)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3500</td>
<td>Fs:消費在庫(+取崩/-積増)</td>
<td>-11099</td>
<td>-6539</td>
<td>2487</td>
<td>10276</td>
<td>-10855</td>
<td>0</td>
<td>747</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**IMI** (Infrastructure for Multilayer Interoperability)

Japanese word “IMI” means “semantics”

- **2012.11** Study project launched
- **2013.3 α** version of core vocabulary
- **2013.4** Started the project for actual system

---

**Business Architecture**
- Front office Information Share
  - Open Data (technology < Policy)
  - RDF, LOD
  - APIs

**Data Architecture**
- Back office Information Share
  - NIEM / JOINUP
  - Semantic interoperability

**Application Architecture**
- Character Set / Code

**Technology Architecture**

---

**NIEM (U.S.)**

- NIEM Core consists of data elements that are commonly understood across domains.
- Future Domains are added to NIEM as necessary based on an established need.
Define the meaning, usage, relationship, data type of basic words

<table>
<thead>
<tr>
<th>項目名（Type/Sub-properties）</th>
<th>英語名</th>
<th>データタイプ</th>
<th>項目説明</th>
<th>項目説明（英語）</th>
<th>キーワード</th>
<th>サンプル値</th>
<th>Usage Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>人</td>
<td>PersonType</td>
<td>PersonName</td>
<td>性别名</td>
<td>Name of a Person</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>性別</td>
<td>Gender</td>
<td>&lt;abstract element, no type&gt;</td>
<td>性別</td>
<td>Gender of a Person</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substitutable Elements:</td>
<td>GenderCode</td>
<td>CodeType</td>
<td>性別コード</td>
<td>Gender of a Person</td>
<td>1</td>
<td></td>
<td>APPLIC標準仕様V2.3データ基準</td>
</tr>
<tr>
<td>性別名</td>
<td>GenderText</td>
<td>TextType</td>
<td>性別名</td>
<td>Gender of a Person</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>現住所</td>
<td>PresentAddress</td>
<td>AddressType</td>
<td>現住所</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>本籍</td>
<td>AddressType</td>
<td>本籍</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Substitutable Elements:

| 項目名（Type/Sub-properties） | 英語名 | データタイプ | ...
|-------------------------------|--------|-------------|-----|
| 氏名                          | FullName | TextType | ...
| フリガナ                       | FamilyName | TextType | ...
| 姓                            | 文献名 | TextType | ...

...
Character code problem

- Ideograph set for “戸籍” (registry of citizens) (55,266 ideographs)
- Ideograph set for “住民登録” (registry of resident) (19,563 ideographs)

MOJ

MIC

JJIS X 0213:2003 (approx. 10,000)

PUA

Mandatory for public administration use

1,700 local governments operate their own coding systems
Character set infrastructure project

• 60,000 characters fonts in open license ... done (2011.5)
• Application to ISO standardization ... done (2012.11)
• Registration of ideographic variation selector codes to UTC ... on going
• RDF based character information database ... project started (2012.11)
Relationship of ideographic variation (Characters with similar shape)
Description of character structure
Description of Relationship (example of “regulation” and “commercial dictionary”)

Government regulation for usage in human name

Font provided by IPA

Commercial dictionary

http://mjdb.ipa.go.jp/GOJIZOKUISEIJI

http://mjdb.ipa.go.jp/DAIKANWA

MJDB : http://mjdb.ipa.go.jp/dcmjdb/terms/ (MJDBの語彙を表す仮URI)
Conclusion

- We are about to dawn of open government data
- We did many things but lack of coordination
- Semantics and character code interoperability are challenge in Japan
- We don’t want to re-invent existing technology
  → use open standards
  → do international collaboration