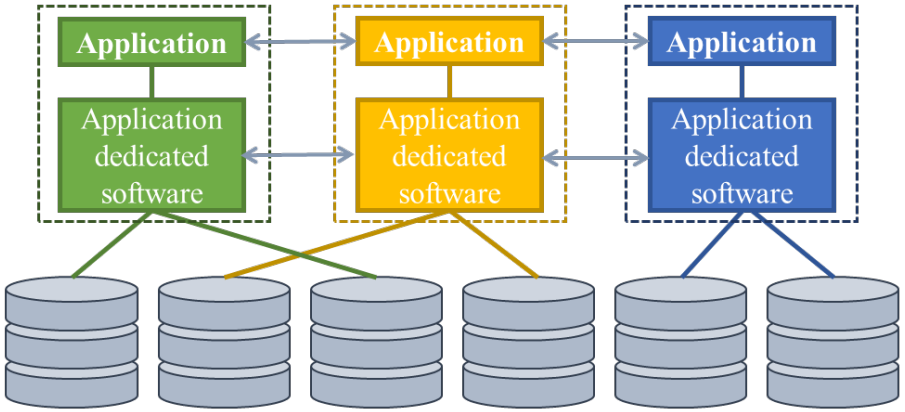
Overview of data format, data model and procedural metadata to support IoT

Nakyoung Kim

KAIST Korea Advanced Institute of Science and Technology
School of Electrical Engineering

Scattered IoT ecosystem

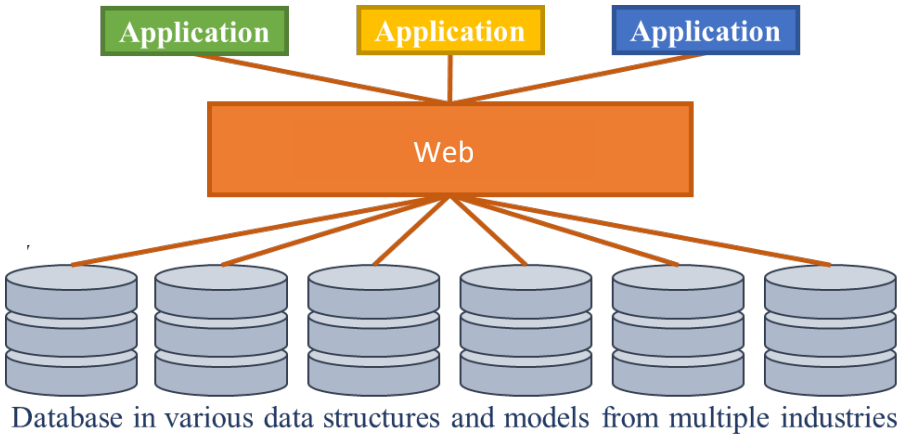
- **Service and application–dedicated data formats & models**
  - Data formats and models have been developed to suit the **specific requirements of each industry, service, and application**.
  - The current scattered-data ecosystem has been established, and it requires **high costs to process and manage data for service and application convergence**.
- **Needs of data integration and aggregation**
  - The boundaries between industries are gradually crumbling down due to domain convergence.
  - For the future convergence markets of IoT and smart city, **data integration and aggregation are important**.



Application-dedicated data management

Interoperability through the Web

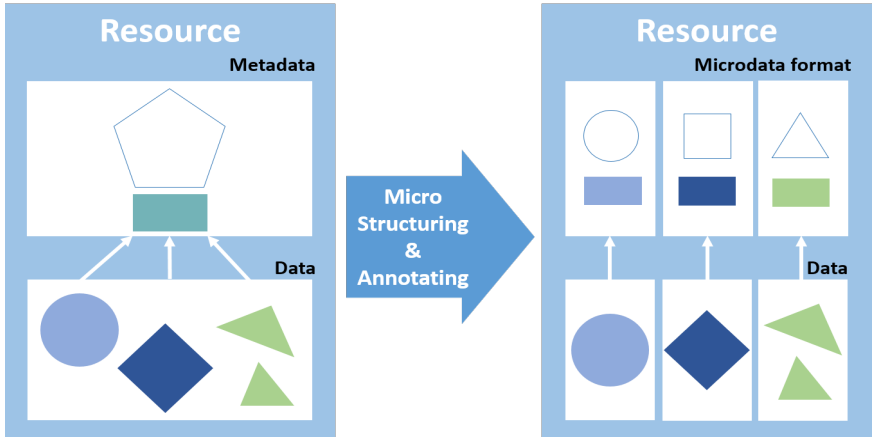
- **Bridging the data**
  - Data formats and models have been settled in the current forms over a period to resolve issues at each moment and fit case-by-case requirements.
  - Therefore, it is **hardly possible to reformulate or replace the existing data formats and models**.
  - **Bridging the data formats and models is feasible**.
- **Connecting via Web**
  - A semantic information space for heterogeneous data, platforms, and application domains, Web can provide technologies that support the interoperability of IoT.
  - **Enhanced interoperability in IoT ecosystems can be fulfilled with the concepts of Web of Things**.



Data-driven data management

Annotation and Microdata format

- **Structured data for annotation**
  - To exploit what the Web offers, **IoT data first needs to be structured and annotated** just like how the other data is handled on the Web.
- **Microdata format**
  - Microdata formats refer structured data markups **describing and embedding the meanings of resources on the Web along with their properties and relationships**.
  - They **utilize tags to convey additional metadata and other attributes** in web pages so that data items can be managed in small data units.



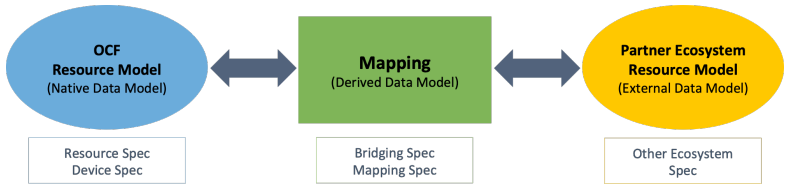
Structuring and micro-annotating a resource with microdata format

Microdata Format	Descriptions
Microdata	Microdata is a nested structured data within HTML content. It uses HTML tag attributes to name the properties of the structured data.
RDFa	RDFa is an HTML5 extension that uses HTML tag attributes that correspond to the user-visible contents to support linked data.
JSON-LD	JSON-LD is a JavaScript notation separate from the HTML body. The markup can be detached from the user-visible text and dynamically injected into the contents.

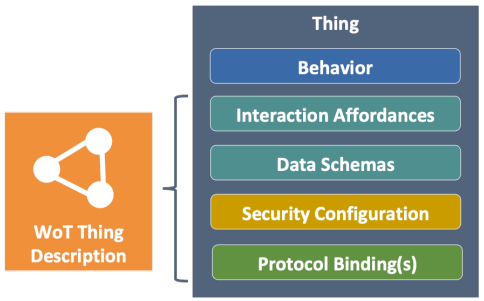
Leading Microdata formats & descriptions

Interoperability in IoT data and device

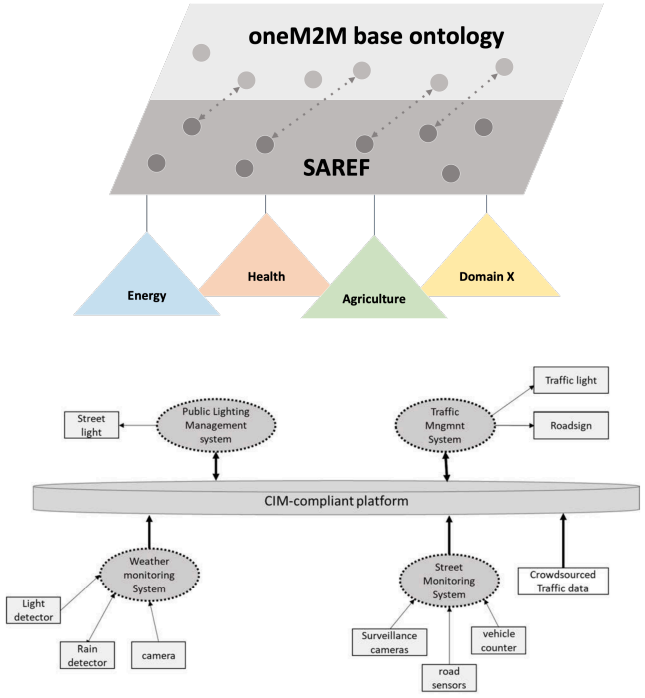
- **Approaches for data and device interoperability**
  - For semantic annotation and linkage of data and devices, some Web-based approaches have targeted **providing common descriptions of data, devices, and their working processes**.



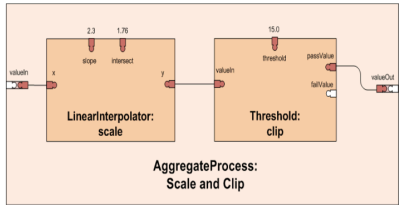
OCF – oneIoT & IoTivity



W3C – Thing Description Repository



ETSI – SAREF & CIM w/ NGSI-LD




```
<?xml version="1.0" encoding="UTF-8"?>
<sml:AggregateProcess gml:id="scaleAndClip06"
    xmlns:sml="http://www.opengis.net/sensorml/2.0"
    xmlns:swe="http://www.opengis.net/swe/2.0"
    xmlns:gml="http://www.opengis.net/gml/3.2"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns:xlink="http://www.w3.org/1999/xlink"
    xsi:schemaLocation="http://www.opengis.net/sensorml/2.0 http://schemas.opengis.net/sensorml/2.0/sensorML.xsd"
    definition="http://sensors.ws/process/linearInterpolator">
    <!-- ============================== -->
    <!--        Descriptions            -->
    <!-- ============================== -->
    <gml:description>
    A simple aggregate process that scales according to linear equation y = 2.3x + 1.76 and then clips if below 15.0
    In this example all processes are defined externally and configured.
    </gml:description>
    <gml:identifier codeSpace="uniqueID">urn:myCompany:swe:process:scaleAndClip02</gml:identifier>
    <gml:name>Scale and Clip Aggregate Process 06</gml:name>
    <!-- ============================== -->
    <!--          Aggregate Process Inputs          -->
    <!-- ============================== -->
    <sml:inputs>
        <sml:InputList>
            <sml:input name="valueIn">
                <swe:Quantity definition="http://sensorml.com/ont/swe/property/SimpleDN">
                    <swe:uom code="any"/>
                </swe:Quantity>
            </sml:input>
        </sml:InputList>
    </sml:inputs>
    <!-- ============================== -->
    <!--          Aggregate Process Outputs          -->
    <!-- ============================== -->
    <sml:outputs>
        <sml:OutputList>
```

OGC – sensorML & SensorThings API

Interoperability in IoT intelligence

- **Sharing common descriptions on complex intelligence**
  - IoT systems **interact with their environments** by actuation based on the processed results
  - It requires methods, such as **machine-readable pseudocodes or workflows**, to support interoperability of procedures **to make decisions and perform tasks** in connected heterogeneous devices and systems.

Technique	Example

Selection/Conditional Statement	IF *condition* THEN
	true alternative
	ELSE
	false alternative
	ENDIF

Iteration/Loops	FOR *i to/in ...*
	statements to carry out
	END FOR
	
	WHILE *condition* THEN
	statements to carry out
	END WHILE
	
	REPEAT *statement* UNTILL ...

Operators	> Greater Than < Less Than >= Greater than or equal to <= Less than or equal to == Equal To = Equals != not equal to
Input and Output	INPUT.... OUTPUT...
Function/Procedure	PROCEDURE *name* END PROCEDURE FUNTION *name* END FUNTION

Pseudocode Examples

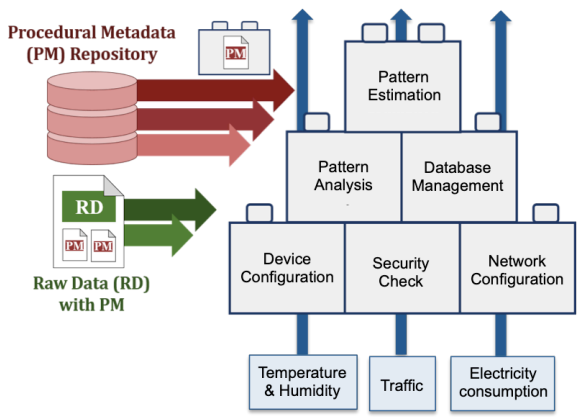
If {resource A}'s {value a} is in {condition 1} or {condition 2}
then {resource B} do {action α}
until {resource A}'s {value a} becomes in {condition 3}

Procedure to make a decision & perform a task

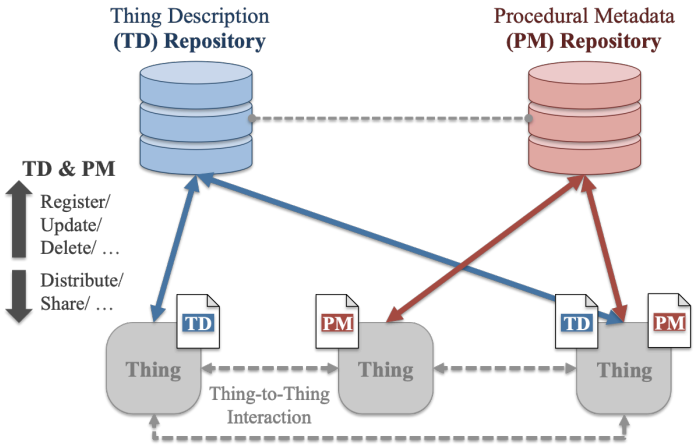
Procedural metadata: Subcategory of metadata

Metadata		Description	Primary Uses	Example Properties
Descriptive metadata		It includes basic information for finding or understanding a resource, such as title, author, subjects, keywords, publishers.	- Discovery - Display - Interoperability	- Subject - Genre - Publication data
Administrative metadata	**Technical**	It includes information to help manage the data resource. It provides information about not only how the data can be opened, read, used, etc. but also how it should be managed for future use and rights of the data.	- Interoperability - Digital object management - Preservation	- File type, size - Creation date/time - Compression scheme
	Preservation		- Interoperability - Digital object management - Preservation	- Checksum - Preservation event
	Rights		- Interoperability - Digital object management	- Copyright status - License terms - Rights holder
Structural metadata		It includes information on how the components of an object are organized or structured and describes the relationships of parts of resources to one another.	- Navigation	- Sequence - Place in hierarchy
Markup languages		It provides structural or semantic features within the data content by integrating metadata and flags (tags/indexes).	- Navigation - Interoperability	- Paragraph - Heading - List
Procedural metadata		It includes information for common descriptions on procedures and/or processes by utilizing metadata and flags (tags/indexes).	- Navigation - Interoperability - Procedure automation	- Function - Algorithm - Work instruction - Work flow

- **Procedural metadata**
  - Its goal is to achieve **interoperability in processing data**, with common descriptions on procedures.
  - It describes composable logic, functions, and workflows among data, devices, and systems, so that they can **interoperably and automatically engage together to make decisions and perform tasks**



Example of procedural metadata composing



Applying procedural metadata in W3C's TD concept

Thank You