

IIIF Curation: the concept, its development, and its potential impact on search and discovery



Asanobu KITAMOTO

Director, Center for Open Data in the Humanities (CODH) and

National Institute of Informatics

<http://codh.rois.ac.jp/>

@rois_codh

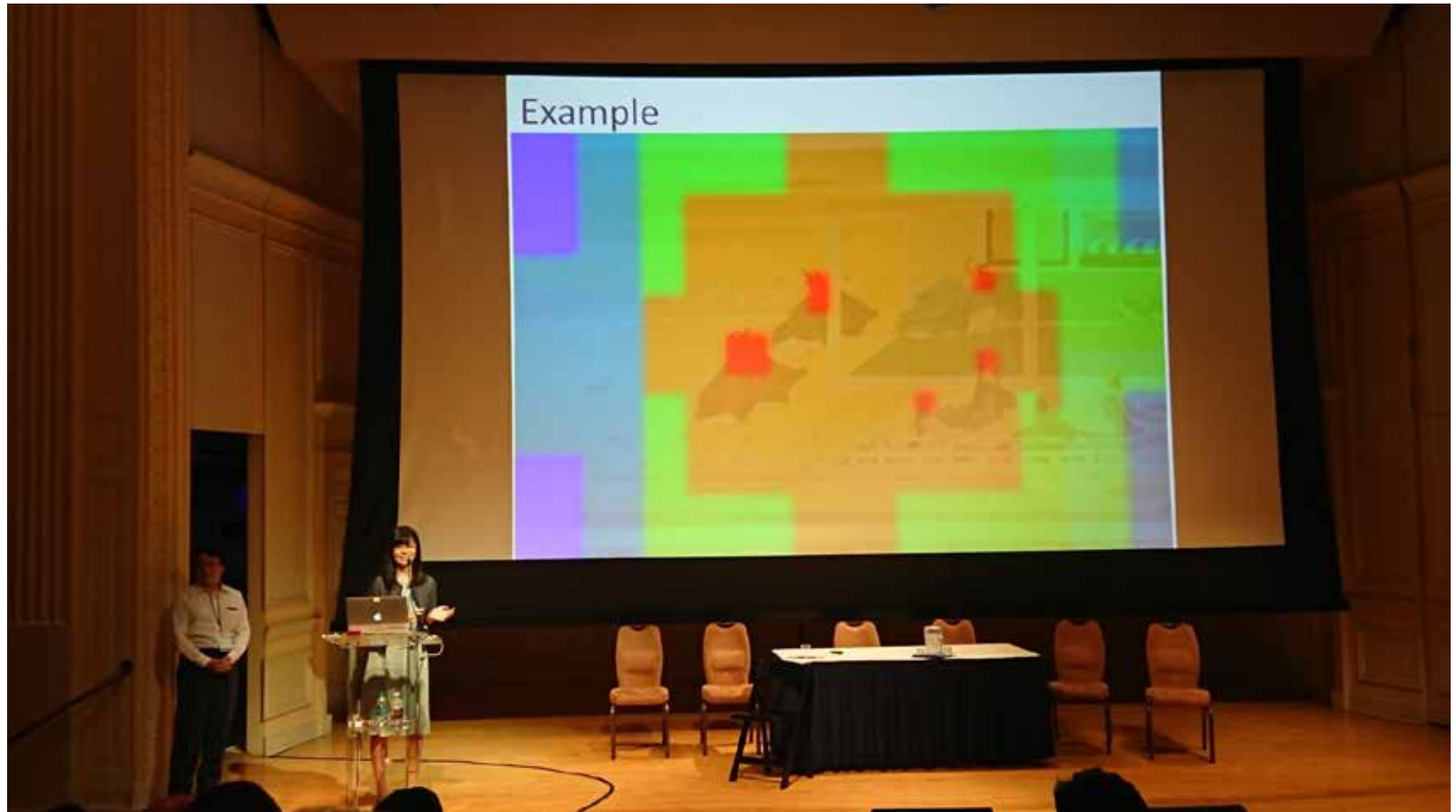
Co-authors:

Jun Homma (Flx Style)

Tarek Saier (NII Internship Student, Albert Ludwigs University of Freiburg)

Concept of IIF Curation

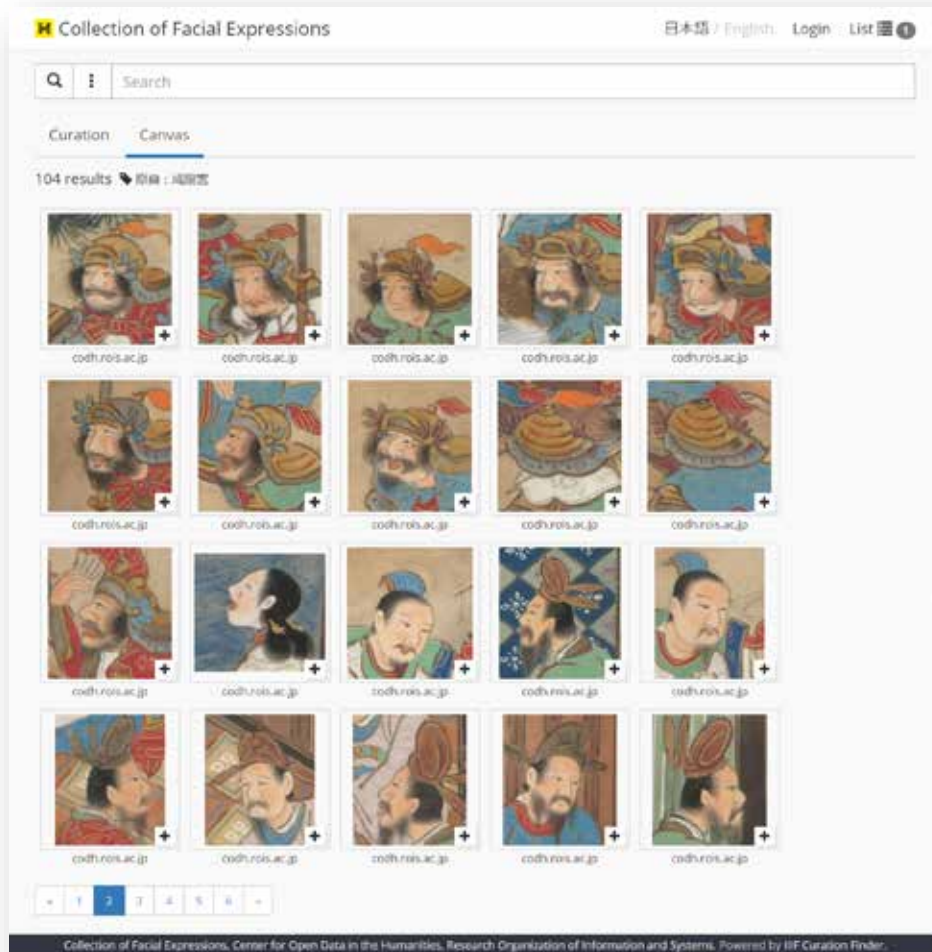
Unusual Access Patterns?



Chifumi Nishioka, Kiyonori Nagasaki, Visualizing which parts of IIIF images are looked by users, 2018 IIIF Conference

Artifacts Caused by Our Service?

<http://codh.rois.ac.jp/face/>



Released on
May 23, 2018

- Collection of Facial Expressions generates “fine-to-coarse” access patterns.

Units of Distribution


digirati

Looking up
and looking
down



Tom Crane, [Digirati](#), December 2017
This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#)
The [examples](#) used in the text can be found on [Github](#).

Looking up and looking down: IIIF Resources, Intellectual Objects, and Units of Distribution

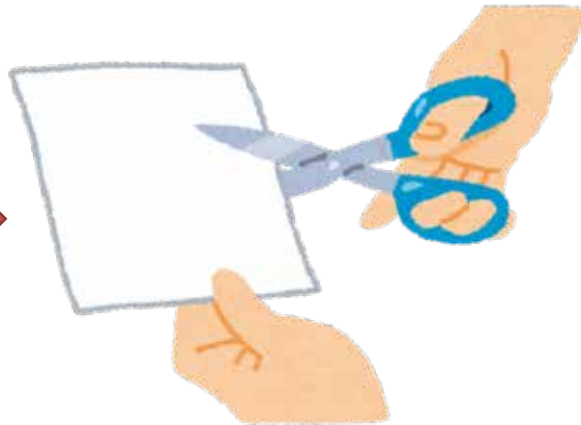
The IIIF Presentation API tells us that the Manifest is the unit of distribution of IIIF. If you want to share an object in IIIF space, describe it in a Manifest. Does that mean that a Manifest corresponds to an *Intellectual Object* - the thing we want to focus on right now? Well, sometimes, if the intellectual object we are thinking about corresponds to a convenient and human-friendly unit of distribution. It depends on the context, on our focus of attention. We can use the Presentation API to look up and look down from the level of the Manifest, to focus on whatever intellectual object is appropriate.

<https://resources.digirati.com/iiif/an-introduction-to-iiif/looking-up-and-down.html>

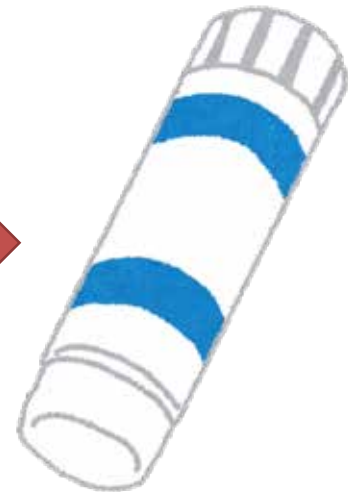
Good Old Analogue World



1 Scissors



2 Paste

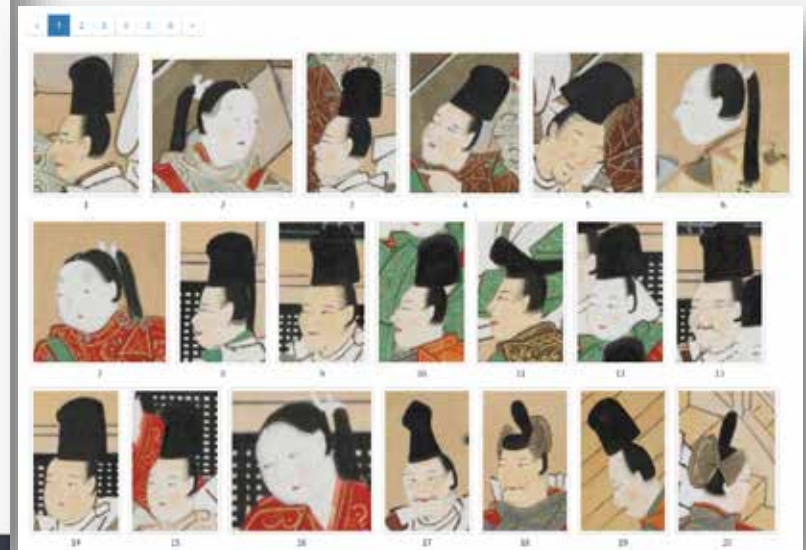


Source: いらすとや, <http://www.irasutoya.com/>

Frictionless Digital World



1. Draw a rectangle.
2. Add to favorites.



Collection of Facial Expressions

<http://codh.rois.ac.jp/face/>

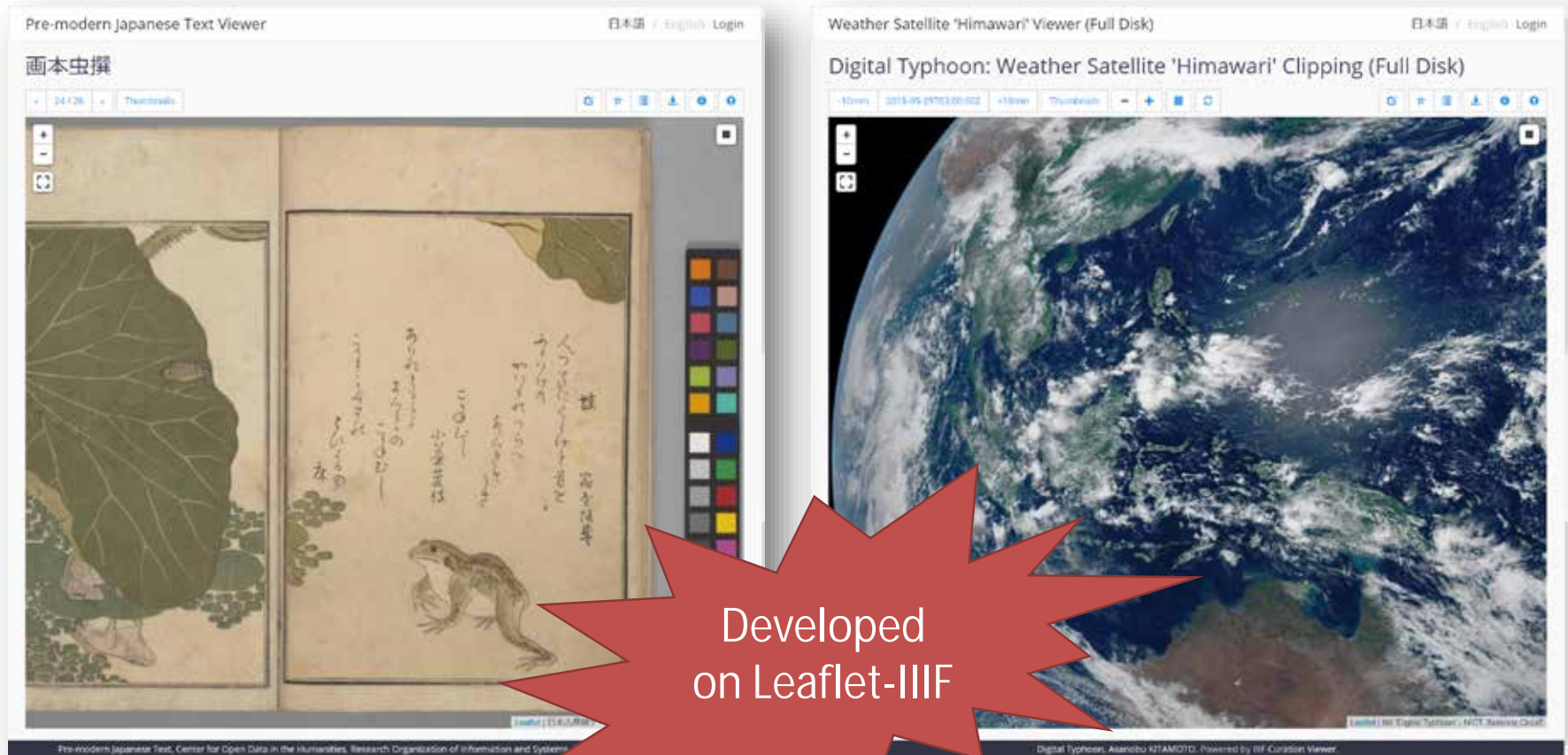
Tag: 「牛若丸」 (https://en.wikipedia.org/wiki/Minamoto_no_Yoshitsune)



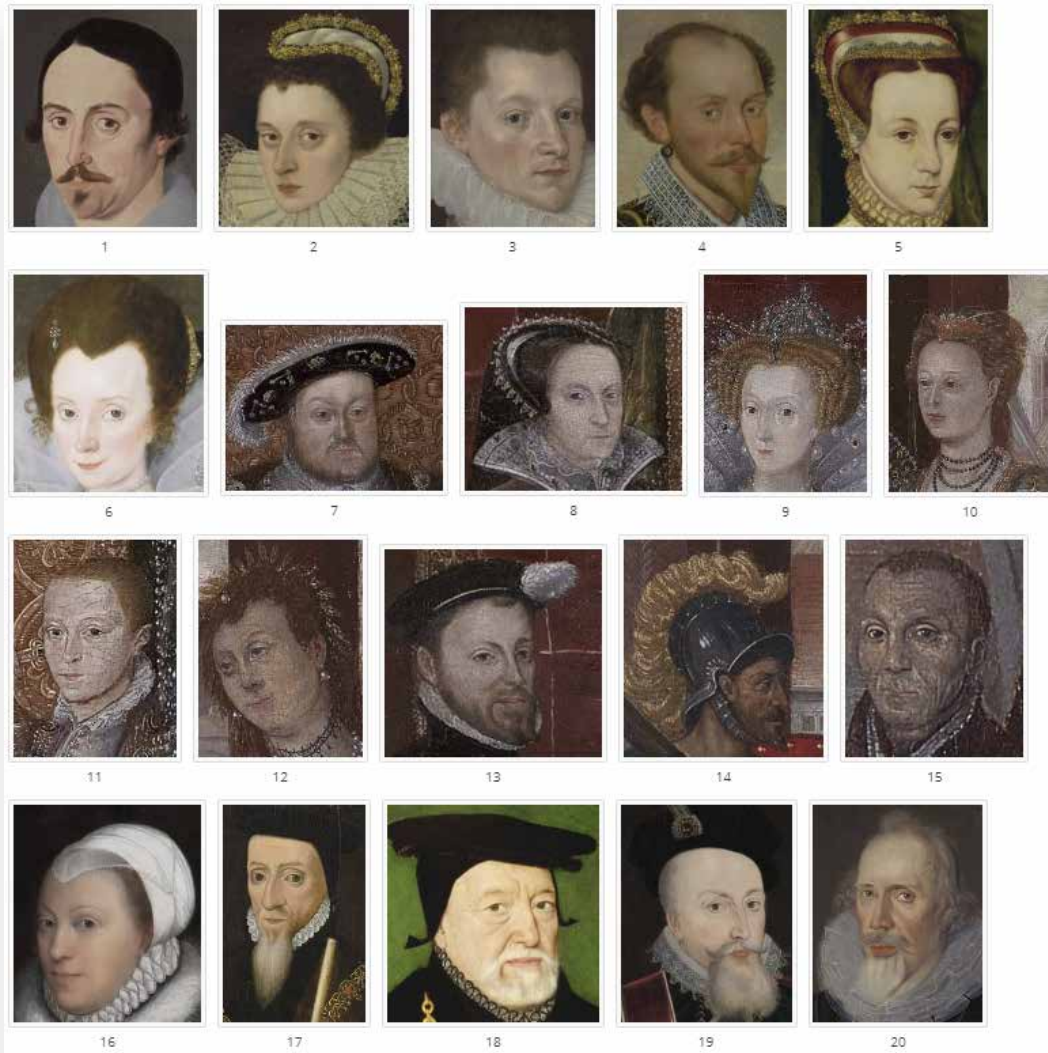
Curation of the same person appearing in different Japanese Emaki handscroll across IIF services.

IIIF Curation Viewer (for Timeline)

<http://codh.rois.ac.jp/software/iiif-curation-viewer/>



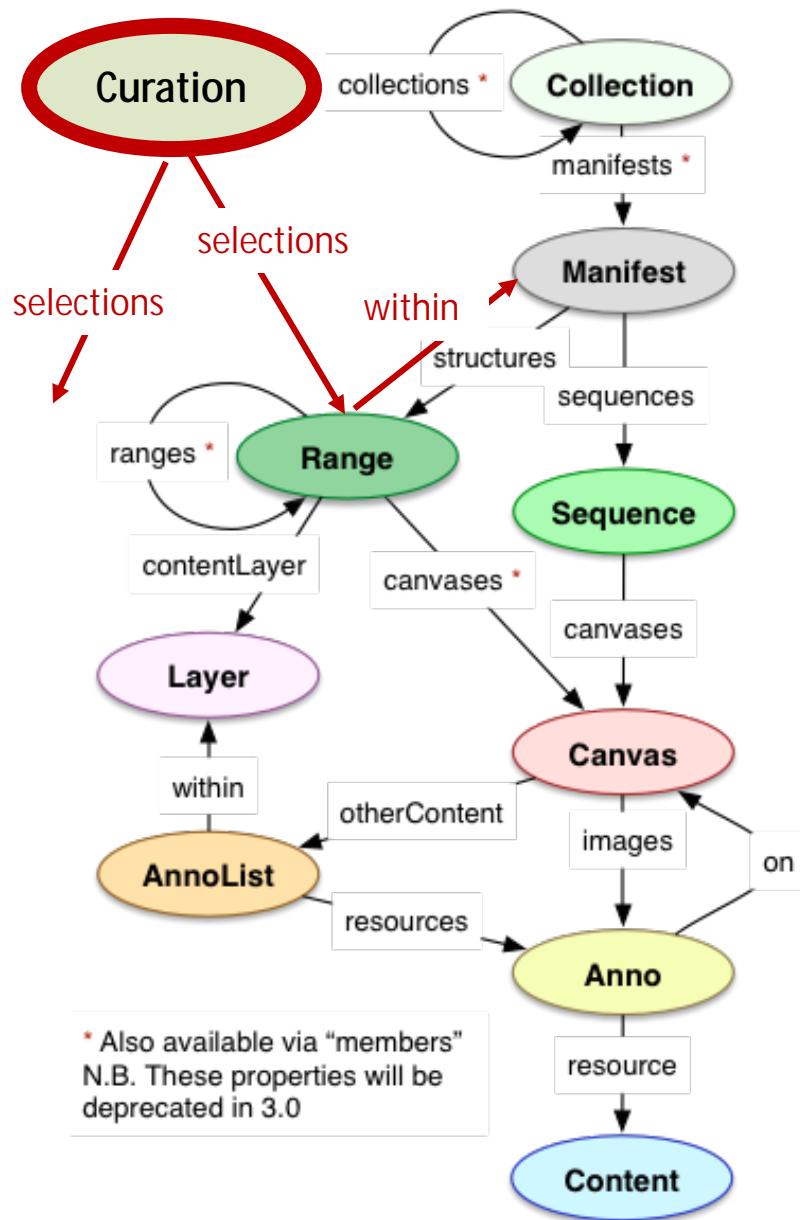
Sixteenth Century



Seventeenth Century



Facial expression data: Portrait <http://codh.rois.ac.jp/curation/exhibition/2/index.html.en>
Originally from Yale Center For British Art



Extension to Presentation API 2

- **Curation** is another upper element of Range.
- **Range** points to **Manifest** using "within" property.
- **Curation** is an array of ranges across multiple manifests.

Selection of (a part of) Canvases

<http://codh.rois.ac.jp/software/iiif-curation-viewer/>

```
"selections": [ {  
  "@id": "http://codh.rois.ac.jp/pmjt/book/200014778/range/r1",  
  "@type": "sc:Range",  
  "label": "Curated contents from 『画本虫撰』",  
  "canvases": [  
    "http://codh.rois.ac.jp/pmjt/iiif/200014778/canvas/00000",  
    "http://codh.rois.ac.jp/pmjt/iiif/200014778/canvas/00023",  
    "http://codh.rois.ac.jp/pmjt/iiif/200014778/canvas/00023#xywh=3435,2487,840,750" ],  
  "within": "http://codh.rois.ac.jp/pmjt/book/200014778/manifest.json" }, {  
  "@id": "http://codh.rois.ac.jp/pmjt/book/200003067/range/r1",  
  "@type": "sc:Range",  
  "label": "Curated contents from 『唐糸草紙』",  
  "members": [  
    { "@id": "http://codh.rois.ac.jp/pmjt/iiif/200003067/canvas/00000", "@type": "sc:Canvas", "label":  
      "p.1" },  
    { "@id": "http://codh.rois.ac.jp/pmjt/iiif/200003067/canvas/00010", "@type": "sc:Canvas", "label":  
      "p.11" },  
    { "@id": "http://codh.rois.ac.jp/pmjt/iiif/200003067/canvas/00010#xywh=2850,1000,730,680",  
      "@type": "sc:Canvas", "label": "p.11" } ],  
  "within": { "@id": "http://codh.rois.ac.jp/pmjt/book/200003067/manifest.json", "@type":  
    "sc:Manifest", "label": "唐糸草紙" } } ]
```

Manifest and Curation

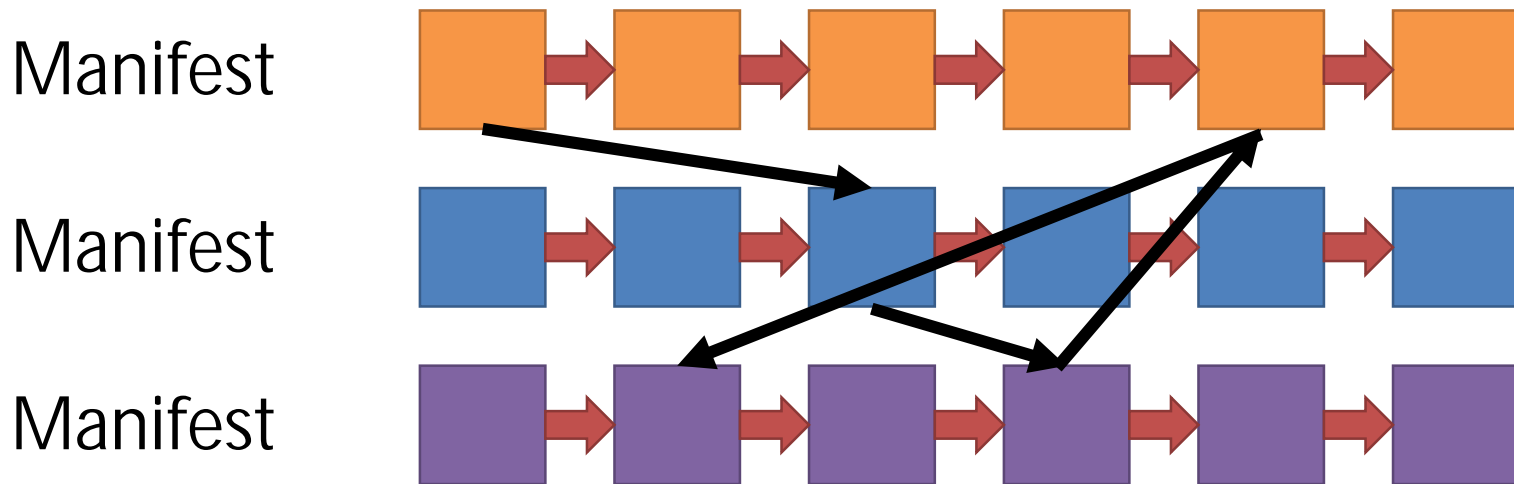
Manifest

- A digital representation of a **physical object** in the world.
- **Authoritative information** from content providers.
- Physical ordering.
- Coarse-to-fine viewing.

Curation

- A representation of a **semantic object** in the digital content.
- **Third-party information** from content consumers.
- Logical ordering.
- Fine-to-coarse viewing.

Fragmentation + Recombination



- A manifest is **fragmented** into canvases, and **recombined** by a curator.
- Physical ordering and logical ordering offer **two different contexts for interpretation.**

Development of IIF Curation Platform

IIIF Curation Platform

Software

1. November 2016: **IIIF Curation Viewer**
2. December 2017: **JSONkeeper**
3. **May 23, 2018**: **IIIF Curation Finder and Canvas Indexer**

Applications

1. **Collection of Facial Expressions**
2. **Bukan Complete Collection**
3. **Weather Satellite 'Himawari' Pictorial Collection**

IIIF Curation Platform



Image



Manifest

Developed by Mr. Homma
Developed by Mr. Saier

IIIF Curation Viewer

Curation JSON

JSONkeeper

Activity Stream

Curation JSON

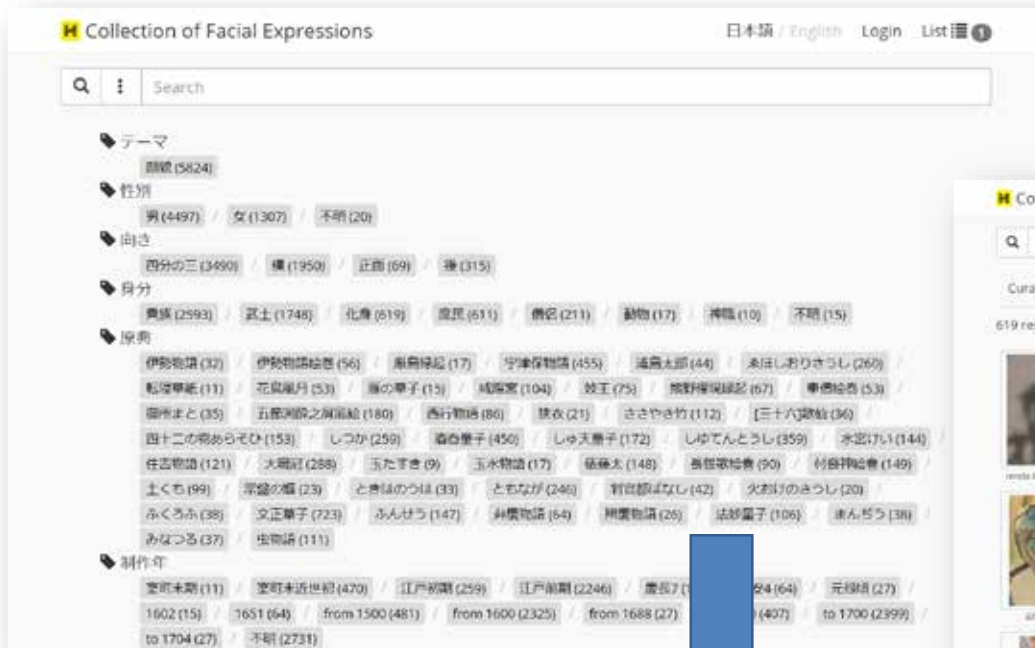
Canvas Indexer

IIIF Curation Finder

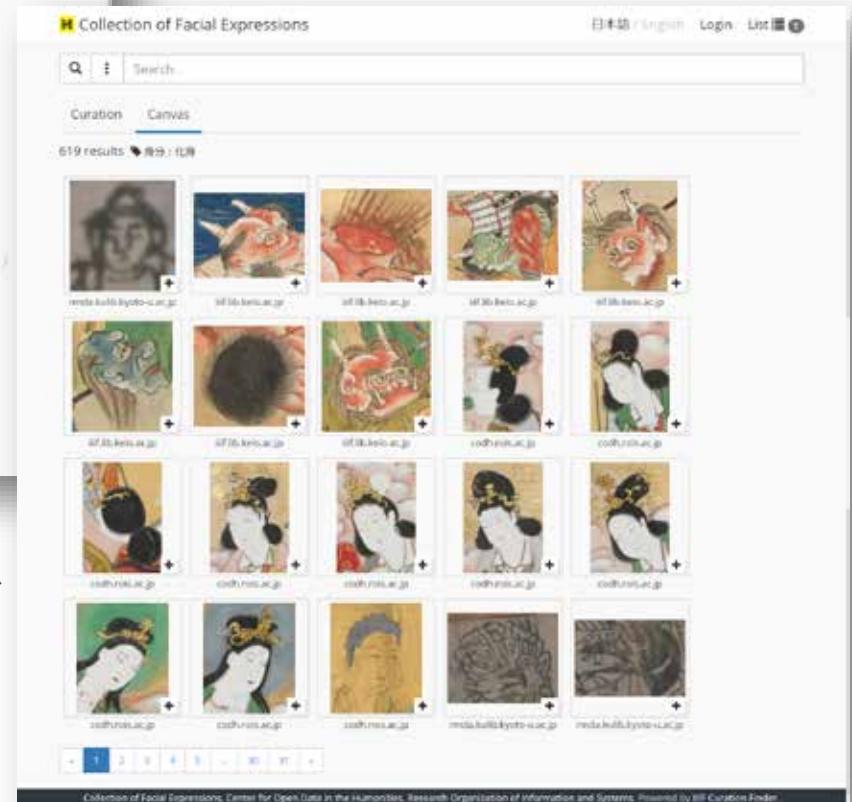
Image Tagging bot

IIIF Curation Finder

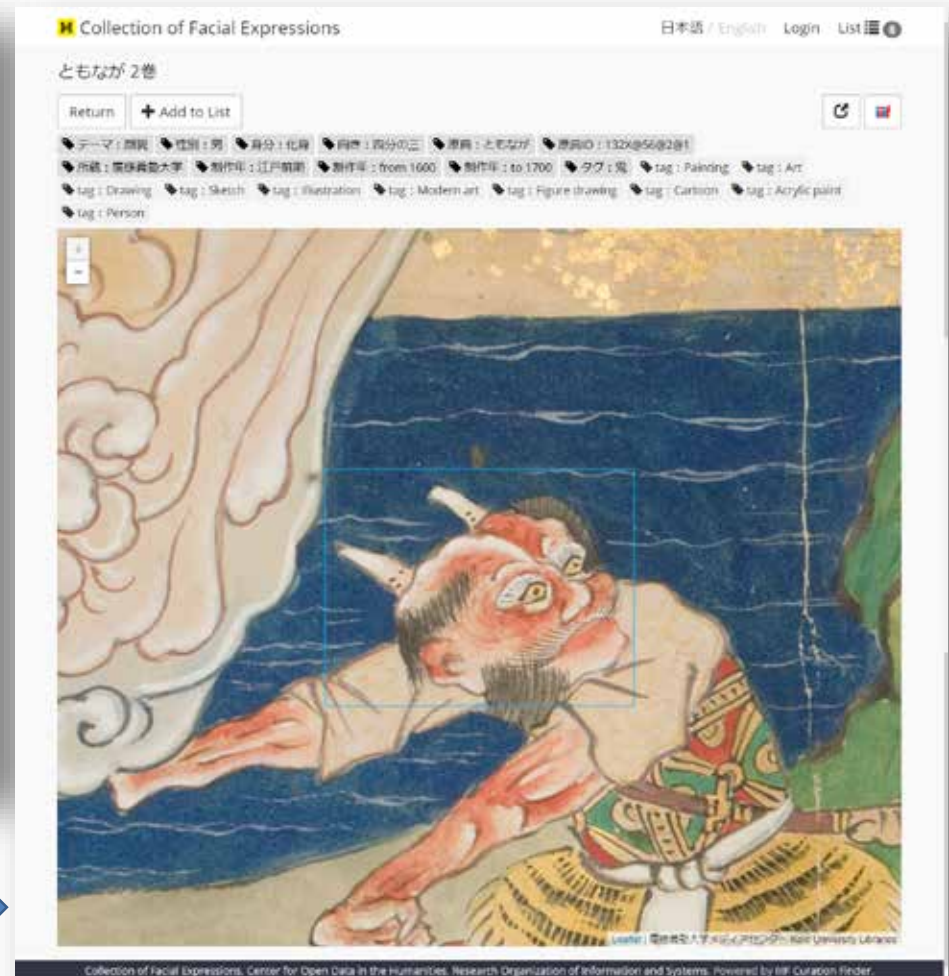
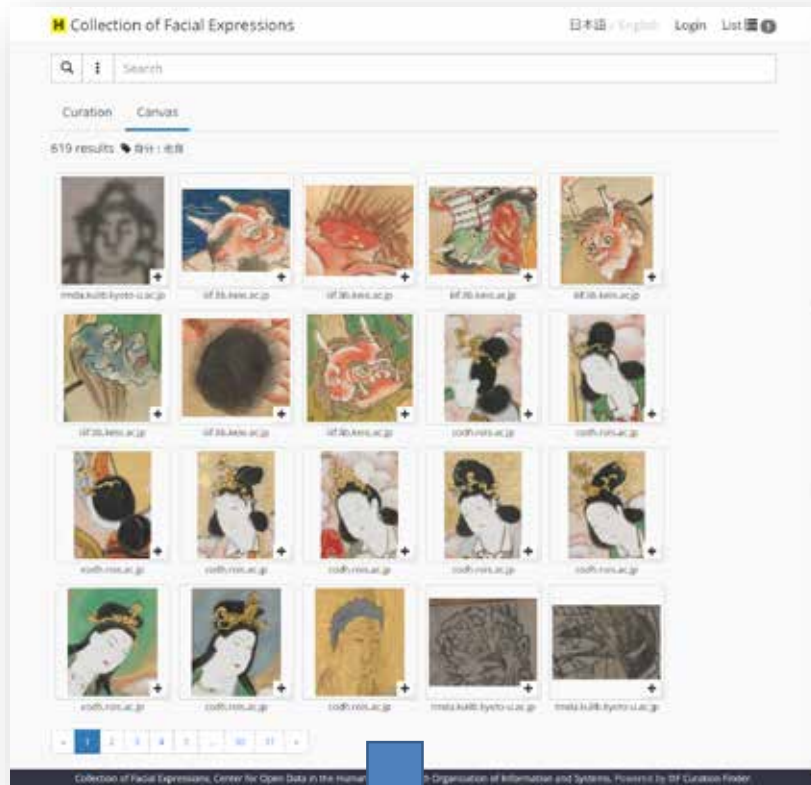
<http://codh.rois.ac.jp/software/iiif-curation-viewer/>



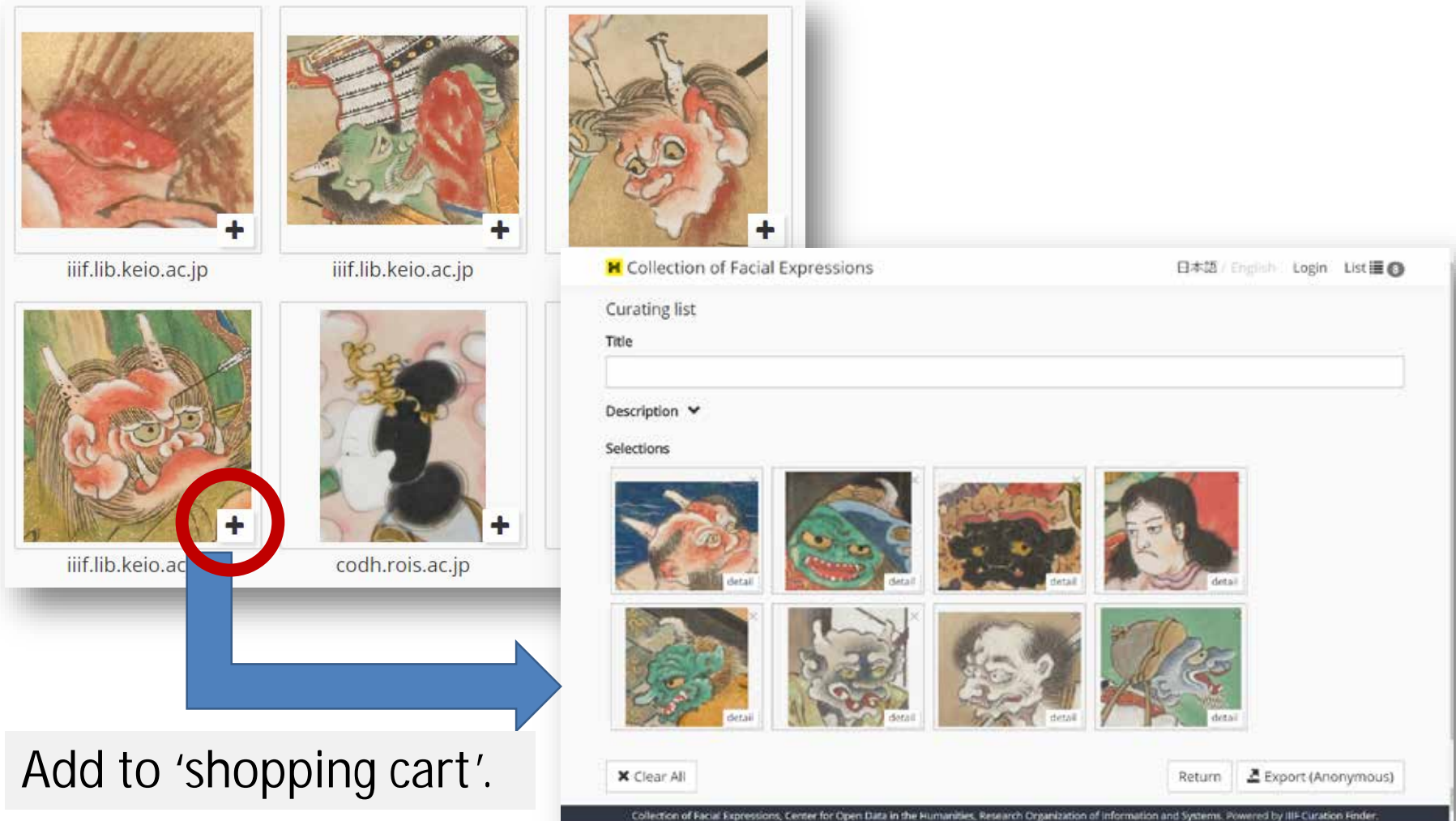
Click the tag, and get the list of a part of canvases annotated by the tag.



From Content to Context



Creating Derivative Curations



The image displays the IIIF Curation Finder interface. On the left, a grid of image thumbnails is shown, each with a '+' icon in the bottom right corner. A red circle highlights the '+' icon on the thumbnail from `iiif.lib.keio.ac.jp`. A large blue arrow points from this icon to the 'Add to 'shopping cart'.' text box below. The main panel on the right shows the 'Collection of Facial Expressions' curating list. It includes a 'Title' field, a 'Description' dropdown, and a 'Selections' grid of eight image thumbnails, each labeled 'detail'. At the bottom of the panel are buttons for 'Clear All', 'Return', and 'Export (Anonymous)'.

iiif.lib.keio.ac.jp

iiif.lib.keio.ac.jp

iiif.lib.keio.ac.jp

codh.rois.ac.jp

Collection of Facial Expressions

日本語 / English / Login List

Curating list

Title

Description ▾

Selections

detail detail detail detail

detail detail detail detail

✕ Clear All

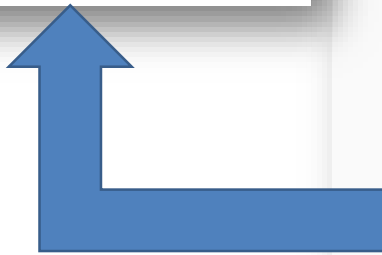
Return Export (Anonymous)

Collection of Facial Expressions, Center for Open Data in the Humanities, Research Organization of Information and Systems. Powered by IIIF Curation Finder.

Add to 'shopping cart'.

Human-Generated Metadata

```
"metadata": [  
  {  
    "label": "テーマ",  
    "value": "顔貌"  
  },  
  {  
    "label": "性別",  
    "value": "男"  
  },  
  {  
    "label": "身分",  
    "value": "武士"  
  },  
  {  
    "label": "向き",  
    "value": "四分の三"  
  },  
  {  
    "label": "原典",  
    "value": "しつか"  
  },  
]
```



IIF Curation Viewer


しつか (「しつか」『ふんせう』の類似顔貌)

資料 1 / 12 46 / 54

Curation Metadata (Canvas 1)

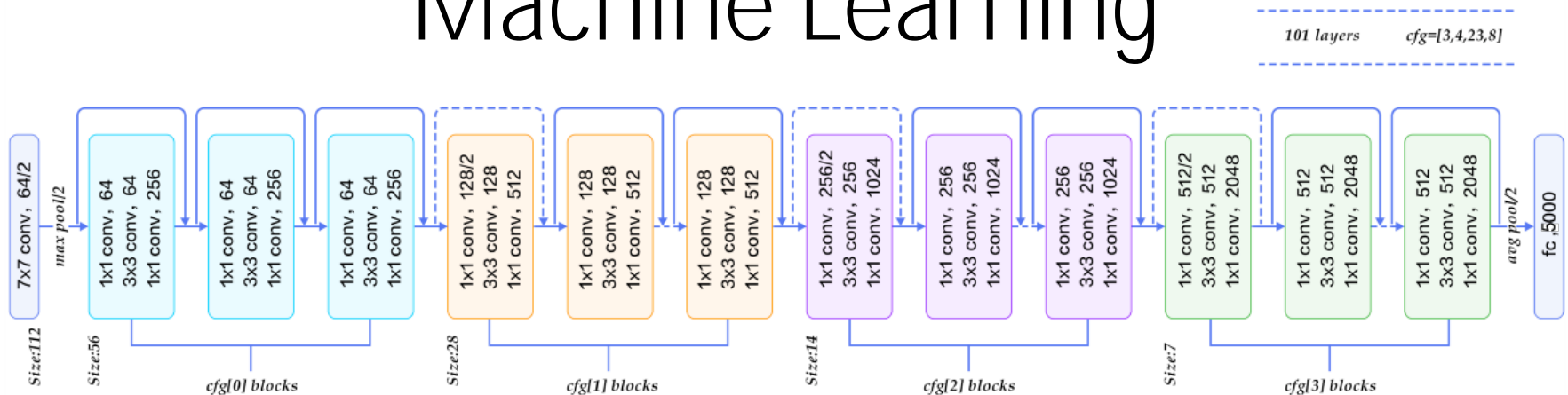
Metadata

label	value
テーマ	string * 顔貌
性別	string * 男
身分	string * 武士
向き	string * 四分の三
原典	string * しつか
原典ID	



Center for Open Data in the Humanities, Research Organization of Information and Systems, Powered by IIF Curation Viewer.

Automated Image Tagging by Machine Learning



https://medium.com/@siddharthdas_32104/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5

- ResNet 101 classifier learns 5000 tags from 9 million images (Open Images Dataset V2).
- We used the model already trained on general photographs, not on our dataset.

Machine-Generated Metadata

```
{
  "label": "tag",
  "value": "Sketch",
  "agent": "machine",
  "score": 0.4955078363418579
},
{
  "label": "tag",
  "value": "Drawing",
  "agent": "machine",
  "score": 0.480587363243103
},
{
  "label": "tag",
  "value": "Art",
  "agent": "machine",
  "score": 0.44167378544807434
},
{
  "label": "tag",
  "value": "Cartoon",
  "agent": "machine",
  "score": 0.3943924605846405
},
{
  "label": "tag",
  "value": "Illustration",
  "agent": "machine",
  "score": 0.3196071684360504
},
}
```

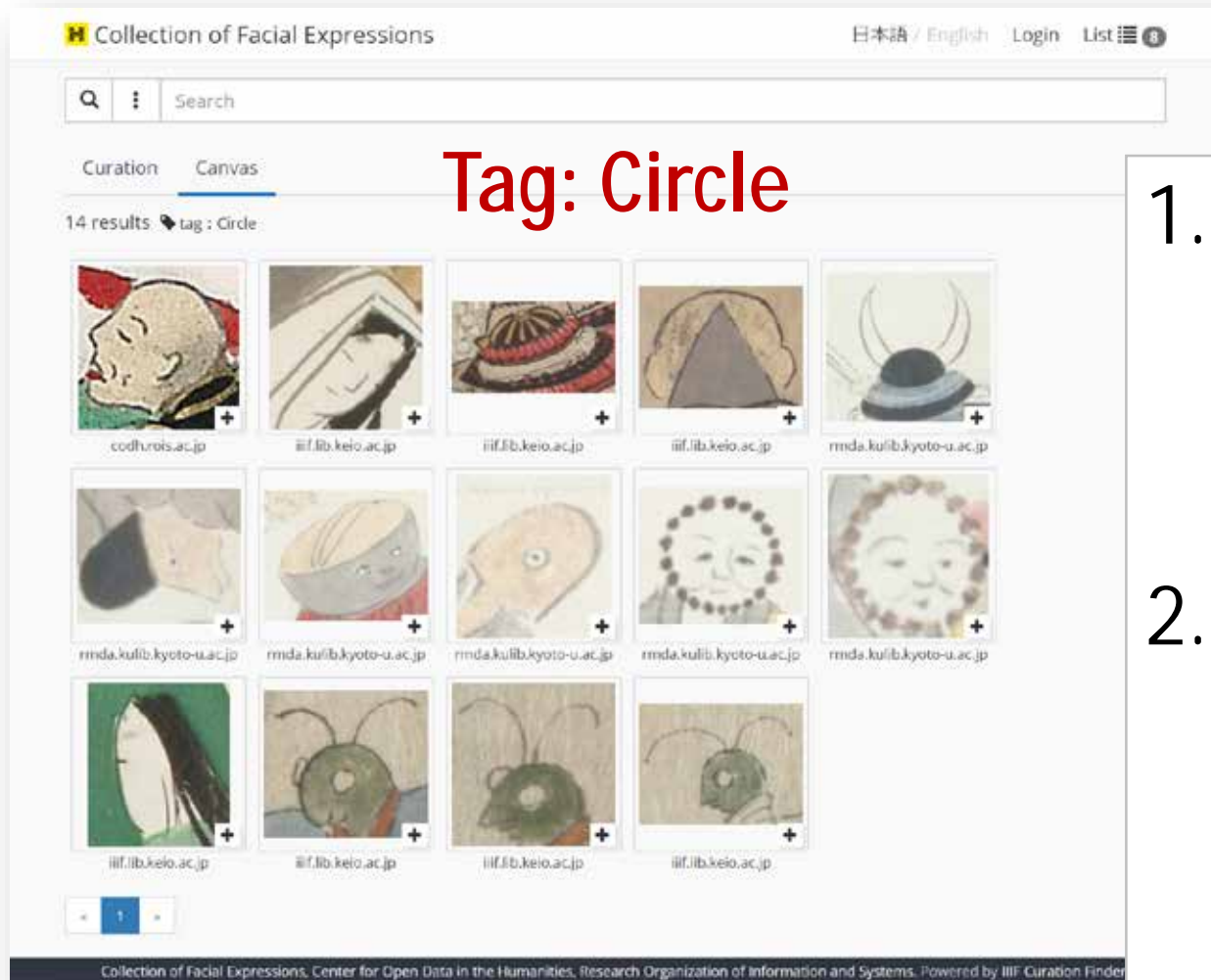
- “agent”: “machine” indicates machine-generated tags.
- “score” is the output of a machine learning system.
- Canvas Indexer accepts an **image tagging bot** for machine-generated metadata.

Ranking of Tags for Faces

tag

Art (5775) / Drawing (5503) / Sketch (5317) / Illustration (4893) /
Cartoon (4459) / Person (4280) / Painting (4071) / Poster (3127) / Face (2597) /
Clothing (2238) / Modern art (1777) / Comics (1081) / Mural (996) /
Figure drawing (850) / Head (804) / Animal (679) / Pattern (670) /
Carving (488) / Portrait (459) / Organ (442) / Ancient history (404) /
Sculpture (381) / Iron (362) / Comic book (306) / Graffiti (277) / Pet (257) /
Hand (236) / Relief (225) / Man (224) / Calligraphy (217) / Textile (209) /
Font (204) / Arm (195) / Album cover (189) / Bird (182) / Mammal (169) /
Food (162) / Text (155) / Wood (149) / Wing (127) / Vehicle (122) /
Green (114) / Human body (99) / Street art (98) / Invertebrate (92) /
Footwear (92) / Statue (89) / Money (89) / Currency (86) / Insect (84) /
Wall (83) / Toy (81) / Limb (73) / Stone carving (71) / Flower (69) /
Glass (68) / Hair (68) / Dog (67) / Number (59) / Fashion accessory (56) /
Nose (52) / Woman (51) / Design (49) / Leaf (48) / Sports (36) /
Hairstyle (35) / Monument (33) / Dessert (32) / Produce (32) / Cash (31) /

Serendipity by Machine Learning



1. Correct or not correct, that is not the question.
2. Serendipity: unexpected encounter with images.

Backend Software



JSONkeeper

A flask web application for storing JSON documents; with some special functions for JSON-LD.

Activity Streams
Change Discovery API 0.1
Conformance level 1

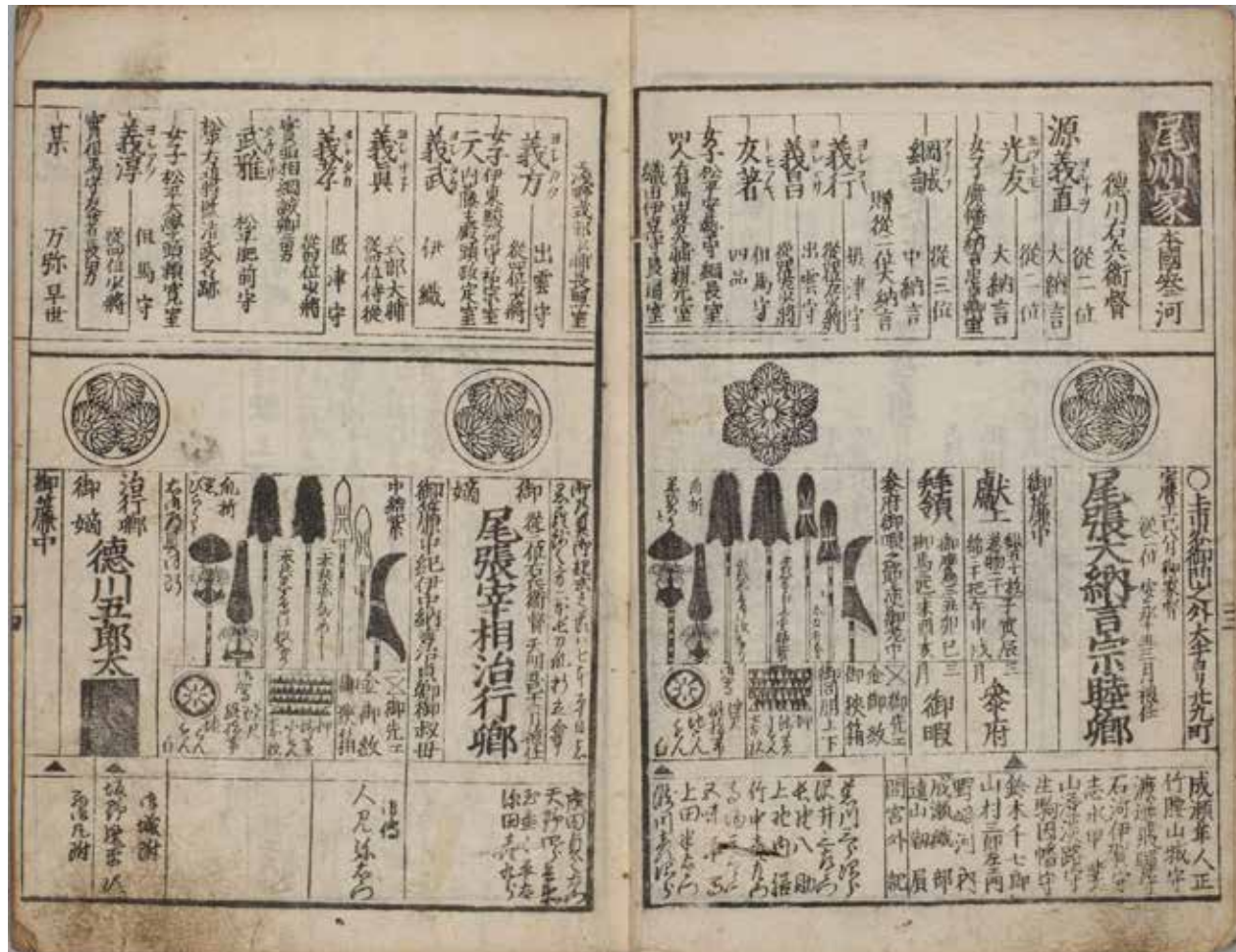


Canvas Indexer

A flask web application that crawls Activity Streams for IIF Canvases and offers a search API.

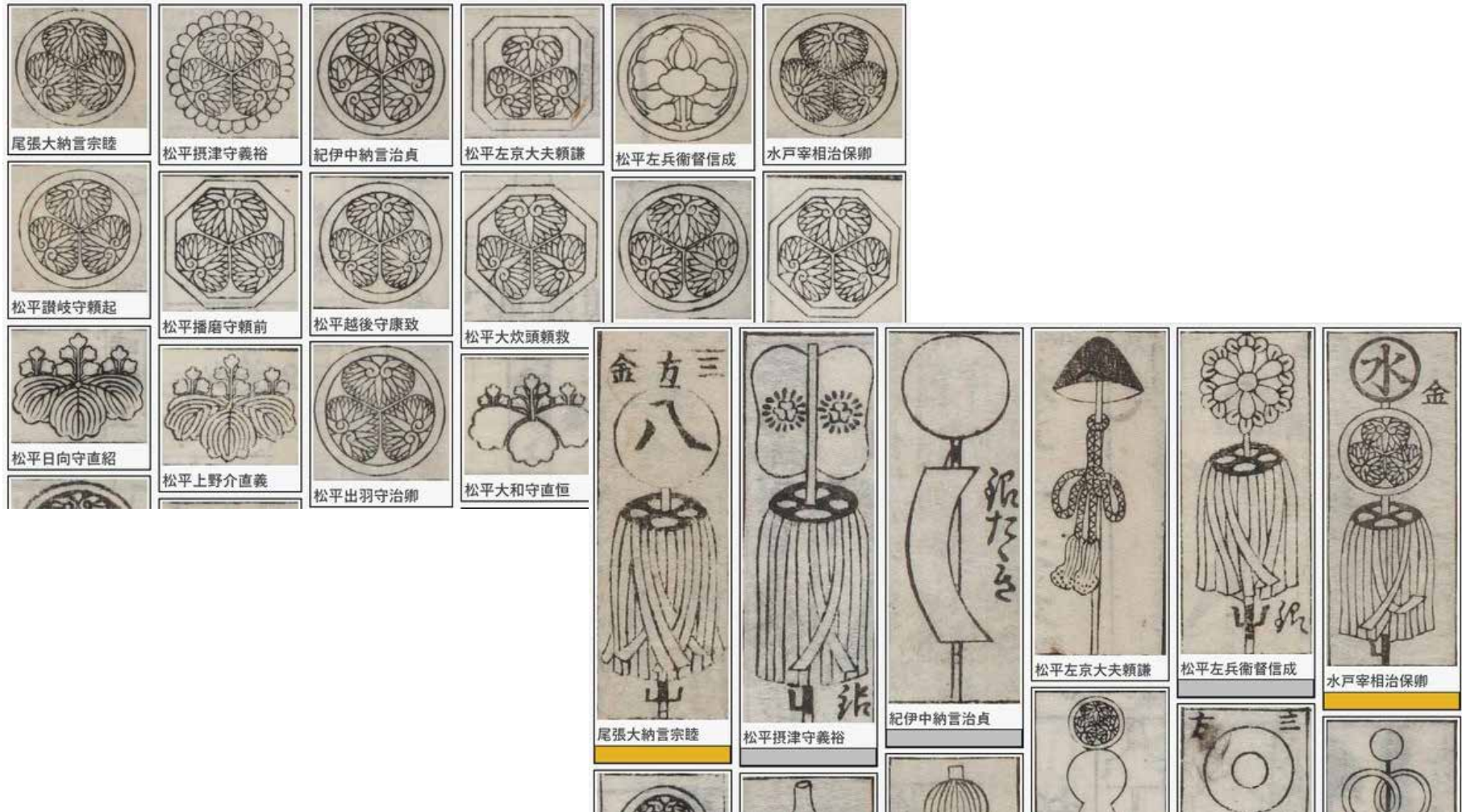
Bukan Complete Collections

<http://codh.rois.ac.jp/bukan/>



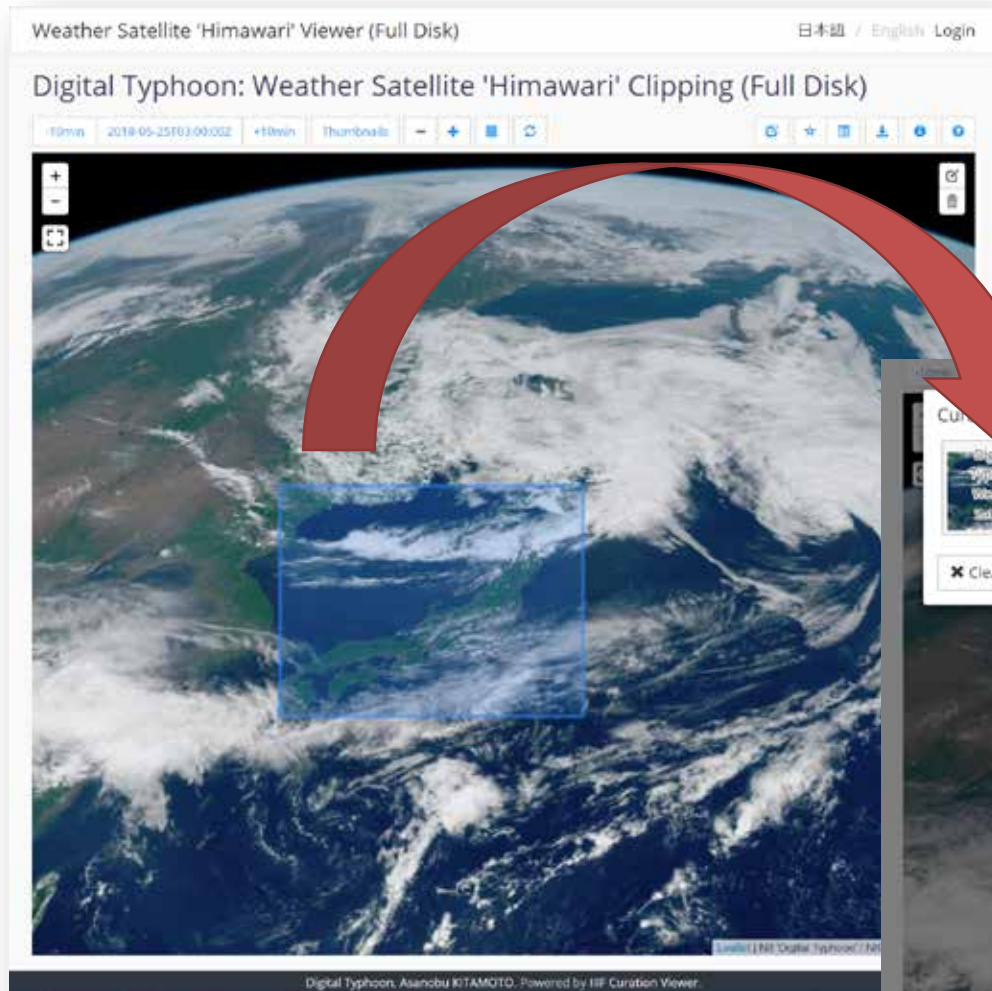
Bukan offers the directory of families of the state king (Daimyo) and bureaucrats of the central government (Bakufu) in the Edo period.
Source : 『寛政武鑑』 (1789), Dataset of Pre-modern Japanese Text (National Institute of Japanese Literature)

Curation of Emblems and Tools

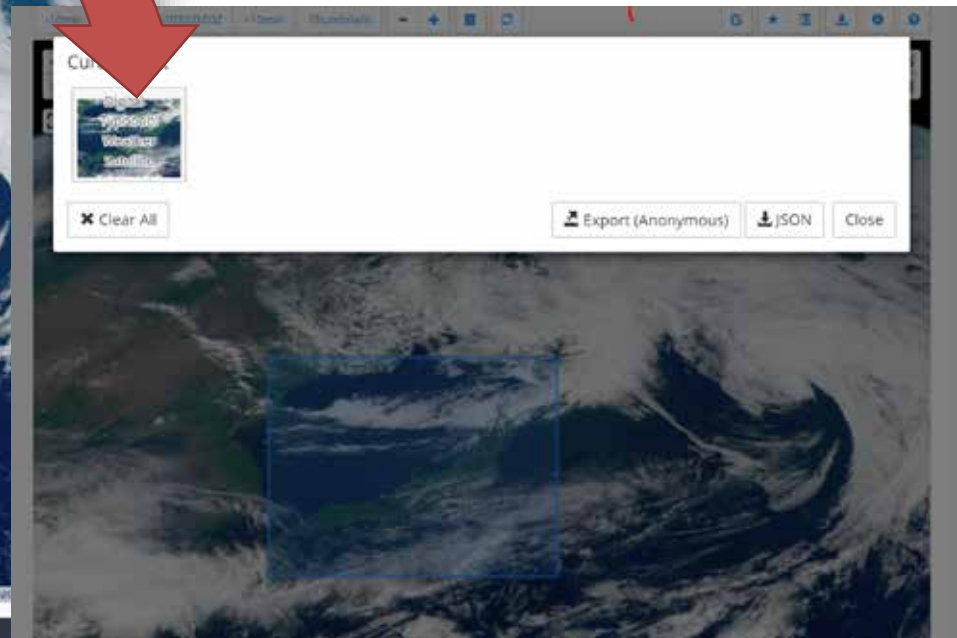


Scientific Applications

<http://agora.ex.nii.ac.jp/digital-typhoon/himawari-3g/iiif/>

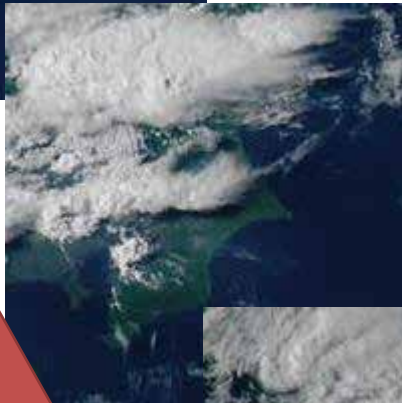
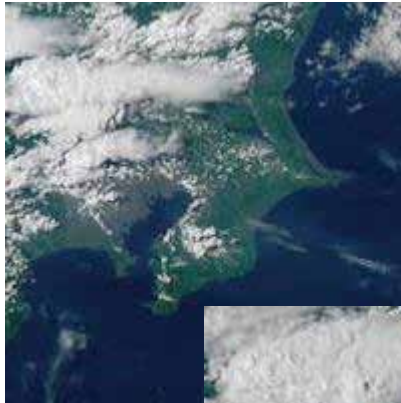


Clipping important events from an image.

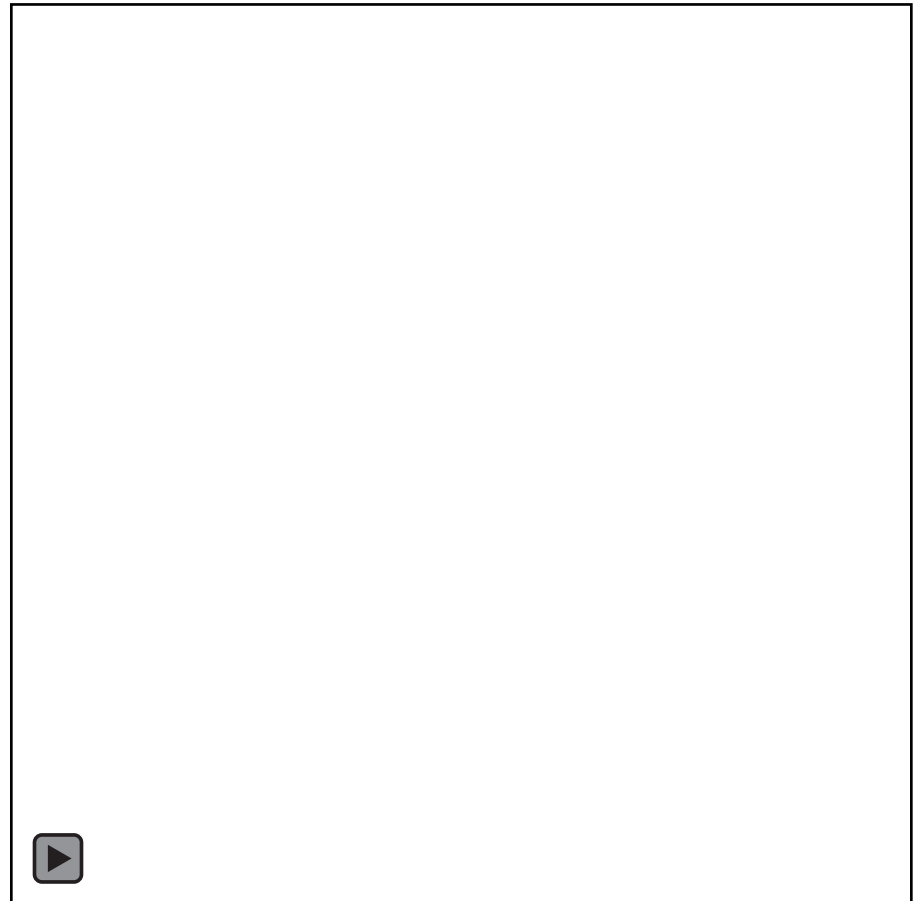


Static Images to Time Lapse

2016-07-14 Development of thunderstorms



Linear interpolation



Example of Time-Based Cursor

<http://codh.rois.ac.jp/iiif/timeline/>

```
"viewingHint": "time",
"cursors": [
{
"@id": "http://agora.ex.nii.ac.jp/digital-typhoon/service/iiif/Hsfd/cursor",
"service":
{
"@context": "http://codh.rois.ac.jp/iiif/cursor/1/context.json",
"@id": "http://agora.ex.nii.ac.jp/digital-typhoon/service/iiif/Hsfd/cursor",
"profile": "http://codh.rois.ac.jp/iiif/cursor/1/cursor"
},
"first": 1436234400,
"last": 1498867200,
"default": 1498867200,
"step": 600 }]
```

Request and Response

<http://codh.rois.ac.jp/iiif/cursor/>

Request for Cursor API

<http://agora.ex.nii.ac.jp/digital-typhoon/service/iiif/Hsfd/cursor?cursorIndex=1493596800>

Response from Cursor API

```
"within": "http://codh.rois.ac.jp/iiif/iiif-curation-viewer/sample/himawari/timeline.json",
"sequence": {
  "@id": "http://agora.ex.nii.ac.jp/digital-typhoon/service/iiif/Hsfd/cursor?cursorIndex=1493596800",
  "@type": "sc:Sequence",
  "label": "2017-05-01T00:00:00Z ~ 2017-05-01T23:50:00Z",
  "canvases": [ {
    "@id": "http://agora.ex.nii.ac.jp/digital-typhoon/iiif/2017/05/01/20170501000000",
    "@type": "codh:Canvas",
    "label": "2017-05-01T00:00:00Z",
    "height": 11000,
    "width": 11000,
    "cursorIndex": 1493596800,
    "images": [
```

Potential Impact on Search and Discovery

Curation for Search and Discovery

1. **Bottom-up approach for discovery**: (a part of) a canvas to a curation and a manifest.
2. **Third-party information**: user-generated curation can diversify interpretations.
3. **Metadata for the object level**, such as face, offers fine-grained search for IIF images.
4. **Humans and machines can help each other** to enhance metadata (tagging).

Remaining Challenges

1. **IIIF Curation Platform** has missing tools that need to be developed.
2. **Three extensions to IIIF specifications**: can we put them into the registry?
3. **Presentation API 3**: for a small group like us, adoption of the new spec is a huge burden.
4. **Digital humanities**: use this platform for collaborative research with art historians.

Acknowledgment

- Mr. Jack Reed for the development of Leaflet-IIIF.
- Mr. Chikahiko Suzuki (CODH) for creating curations of facial expressions.
- Mr. Hoàng Văn, Hà (NII internship student, Vietnam National University, HCMC) for automated image tagging.
- IIIF Curation Platform
 - <http://codh.rois.ac.jp/iiif-curation-platform/>