

Creative Scissors and Paste: IIF Curation Viewer and Inquiry in Digital Humanities



Asanobu KITAMOTO

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

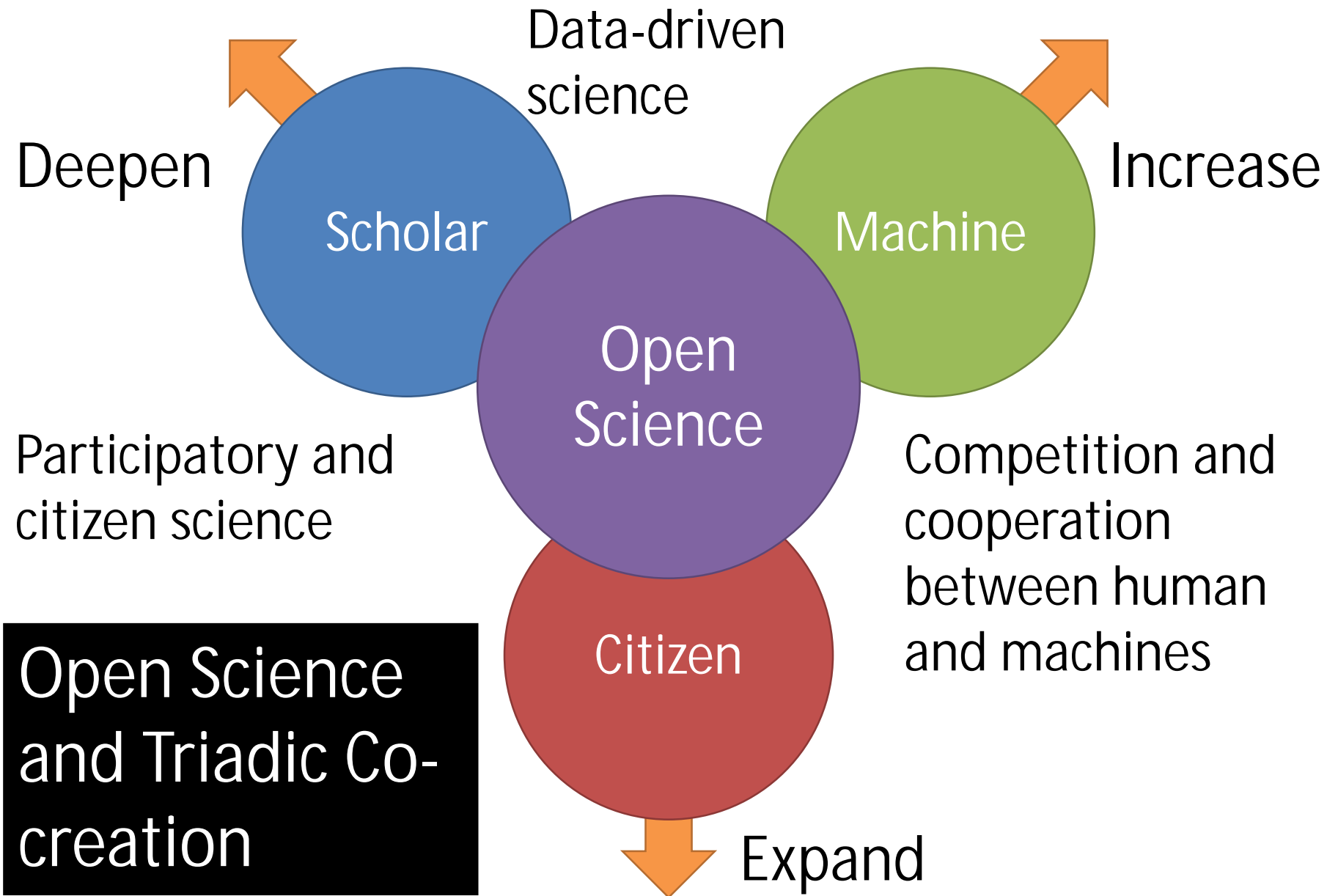
National Institute of Informatics

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

@rois_codh

What is CODH?

- April 1, 2017: Officially launched.
- One of the centers in Joint Support-Center for Data Science Research.
 1. Humanities research using technology and methodology of informatics and statistics.
 2. Informatics and statistics research using humanities data.
- **NIJL-NW Project is one of our collaborators.**



Open Data for Scholars

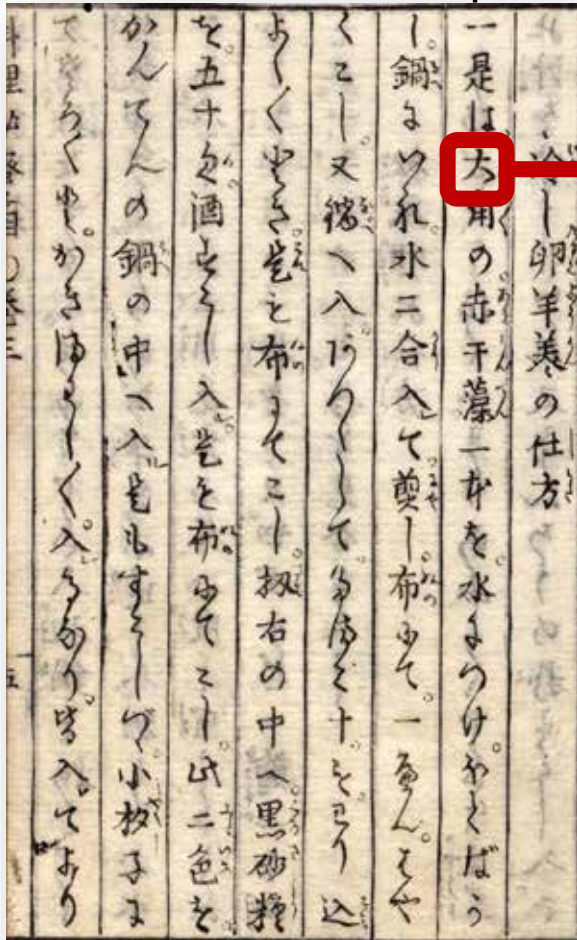
<http://codh.rois.ac.jp/pmjt/>



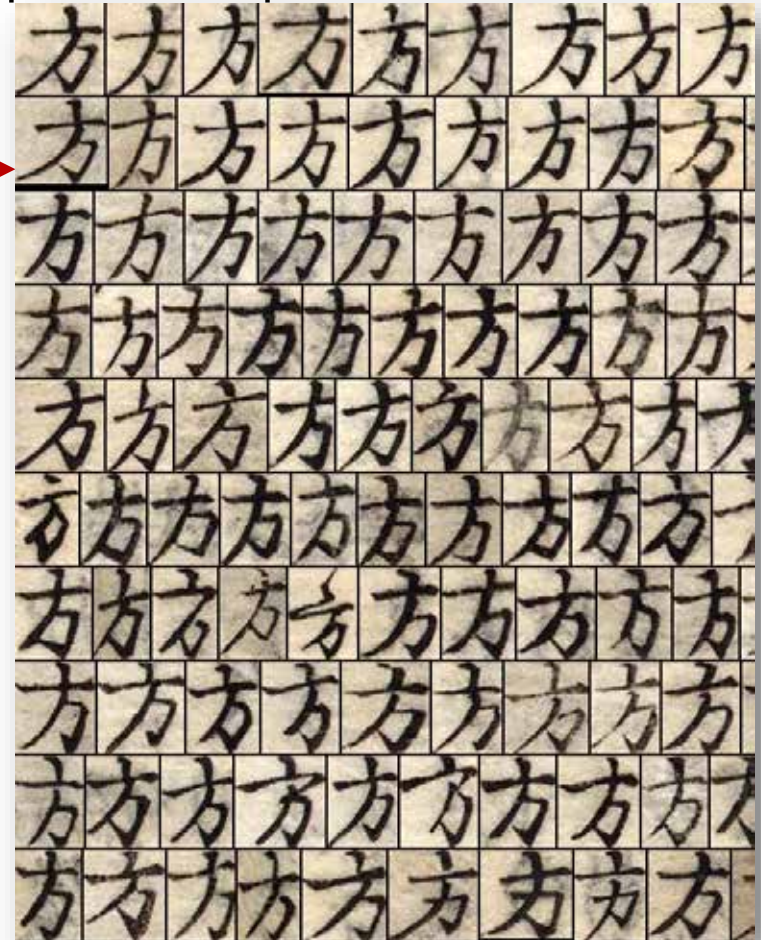
Pre-Modern Japanese Text Dataset (from NIJL)

Open Data for Machines

<http://codh.rois.ac.jp/char-shape/>



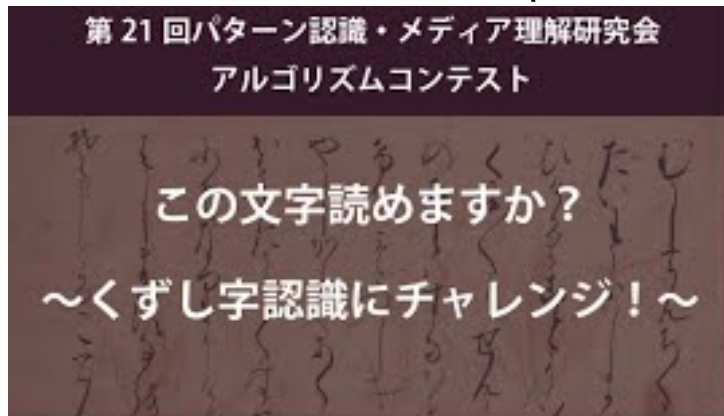
PMJT Dataset (from NIJL)



PMJT Character Shape Dataset
(from NIJL and processed by CODH)

Kuzushiji Challenge!

<http://codh.rois.ac.jp/char-shape/>



古典籍画像の指定領域に含まれるくずし字を認識して、各字の Unicode を出力する課題です。

使用言語：C++ / Python
開発・実行ができる
仮想環境を配布します



詳細はアルコン HP にて

サンプル公開：2017年4月末
応募開始：2017年5月31日
応募締切：2017年8月31日



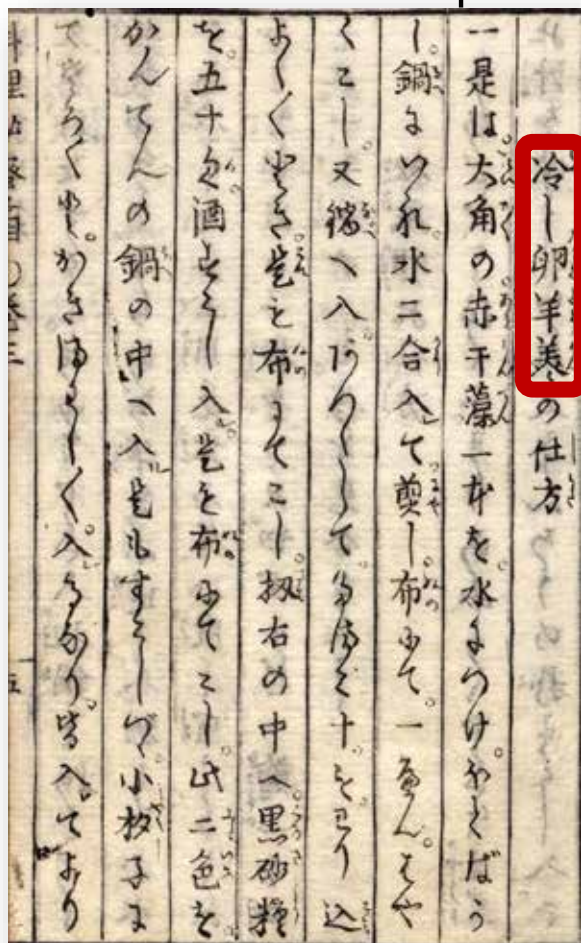
@alcon2017prmu

主催：電子情報通信学会 / パターン認識・メディア理解研究会
後援：国文学研究資料館
情報・システム研究機構 / データサイエンス共同利用基盤施設 / 人文学オープンデータ共同利用センター
情報処理学会 / 人文科学とコンピュータ研究会
データ提供：国文学研究資料館 / 人文学オープンデータ共同利用センター

- Optical Character Recognition (OCR) does not work.
- Can AI (artificial intelligence) read old characters?
- Hold a competition to invite computer scientists to participate.

Open Data for Citizens

<http://codh.rois.ac.jp/edo-cooking/>



PMJT Dataset
(from NIJL)

冷し卵羊羹



Edo Cooking Recipe Dataset
(Created by CODH)
Adapted Material on NIJL Dataset
(from NIJL)

Edo Cooking Recipes at Cookpad

Deposit recipe data to “Cookpad Edo Dishes,” run by Cookpad and The Japan Society of Home Economics.



Recipe with cooking steps, photographs, original old text, and a link to the CODH website.

クックパッド サービス一覧 お弁当におすすめのどんぶり料理 プレミアム

毎日の料理を楽しみに 253 レシピ

検索 料理名・食材名 × 目的・用途 レン

ほうれん草 干き椎茸 親子丼 ダイコンの葉 ちつぽ

クックパッド江戸ご新のレシピ

江戸時代のスイーツ 甘さスツキリ冷卵羊羹

江戸の料理本から見つけた和風スイーツです。プリンのごですが、牛乳不使用でさらっとした菓糖の甘さがやみつきになります。

クックパッド江戸ご新

材料 (2~4名分)

卵	5個
寒天(赤)	1本(4g)
菓砂糖	100g
水	180cc
片栗粉	適量
酒	適量

カロリー:276kcal/人 塩分:0.3g/人

実器編 自動中 設定

1. 寒天を水に付けて、ふやかします。
2. 卵をよく溶きまぜます。
3. 溶いた卵液を布でこします。
4. ③に入れ、再びこします。
5. ④を③に入れ、再びこします。
6. ⑤を布などでこし、再び④に入れて熱めます。
7. ⑦の熱した寒天の中に、⑤の卵液を少しずつ入れます。

<http://cookpad.com/recipe/4153357>

Big Impact from the Release

人文学オープンデータ共同利用センターさんがリツイート

うずら @caille2006 · 11月26日
このプロジェクトがすごいのは、古文書の情報をさらに現代の生きた情報にするために、クックパッドにアカウントを開設してレシピを公開し「つくれぽ」も受け付けていること。江戸ご飯とつくれぽというこの未来感バネい。 cookpad.com/kitchen/146046...



クックパッド江戸ご飯 のキッチン

プロフィール

トップ	レシピ 32	つくれぽ 0	献立 0
-----	-----------	-----------	---------

レシピを検索

7317 retweets

<https://twitter.com/caille2006/status/802575840819089409>

2017/8/30

人文学オープンデータ共同利用センターさんがリツイート

NII 国立情報学研究所(NII) @jouhouken · 11月24日
[プレスリリース]
江戸の文化を現代に取り込む「江戸料理レシピデータセット」を整備～江戸時代の料理本を「レシピ化」し、クックパッドでも公開～
nii.ac.jp/news/2016/1124



1 1,074 971

1052 retweets

<https://twitter.com/jouhouken/status/801693251052781568>

2017 EAJS Pre-Event

Photographs by Cooking Experts

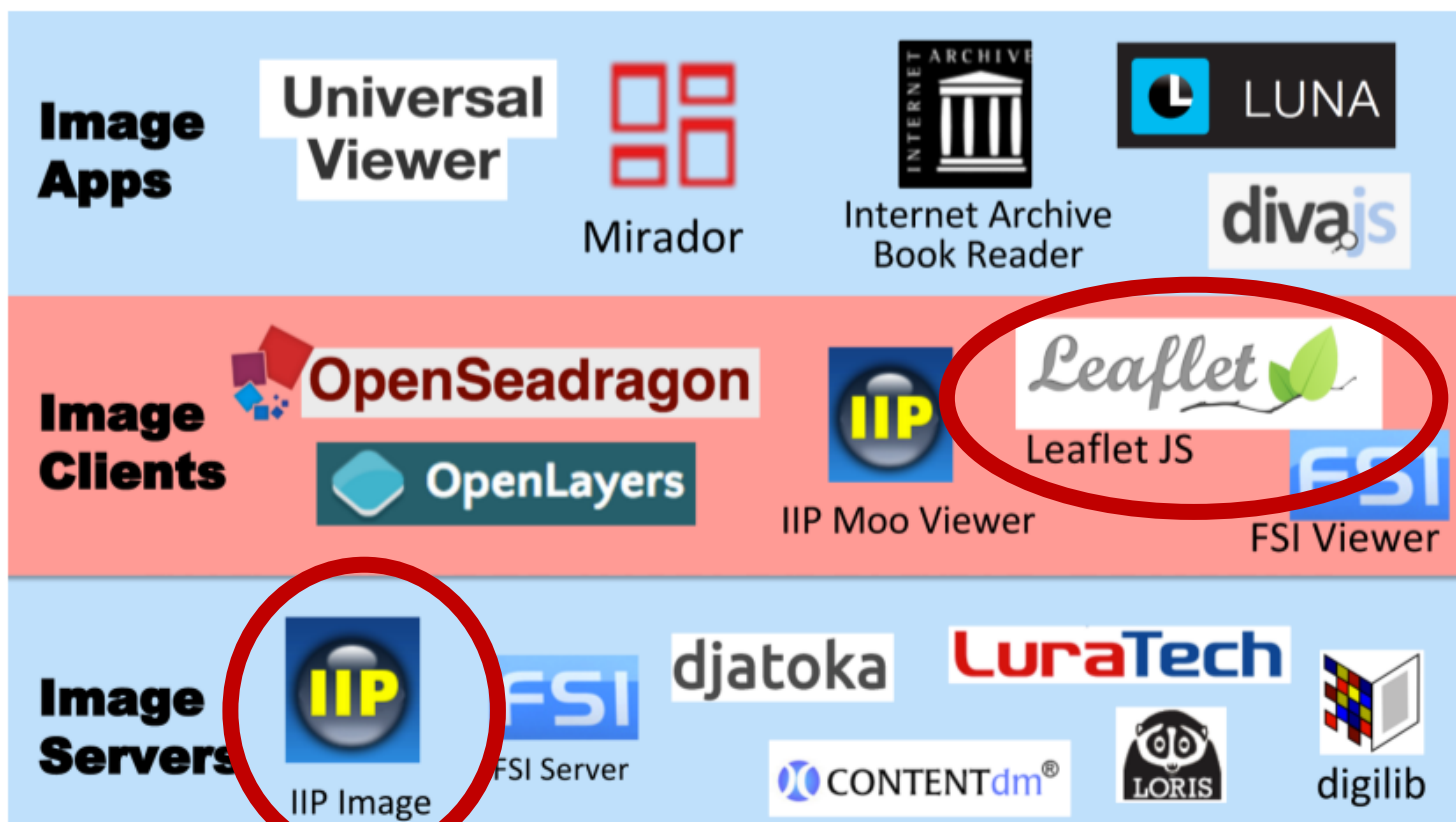




IIF-based Image Delivery

- IIF (International Image Interoperability Framework) is now widely used in humanities-related communities.
 1. **Image API**: Delivery of single images.
 2. **Presentation API**: Delivery of a set of images (e.g. books) with metadata
- Interoperable APIs allow people to develop and use **digital tools that fit all**.

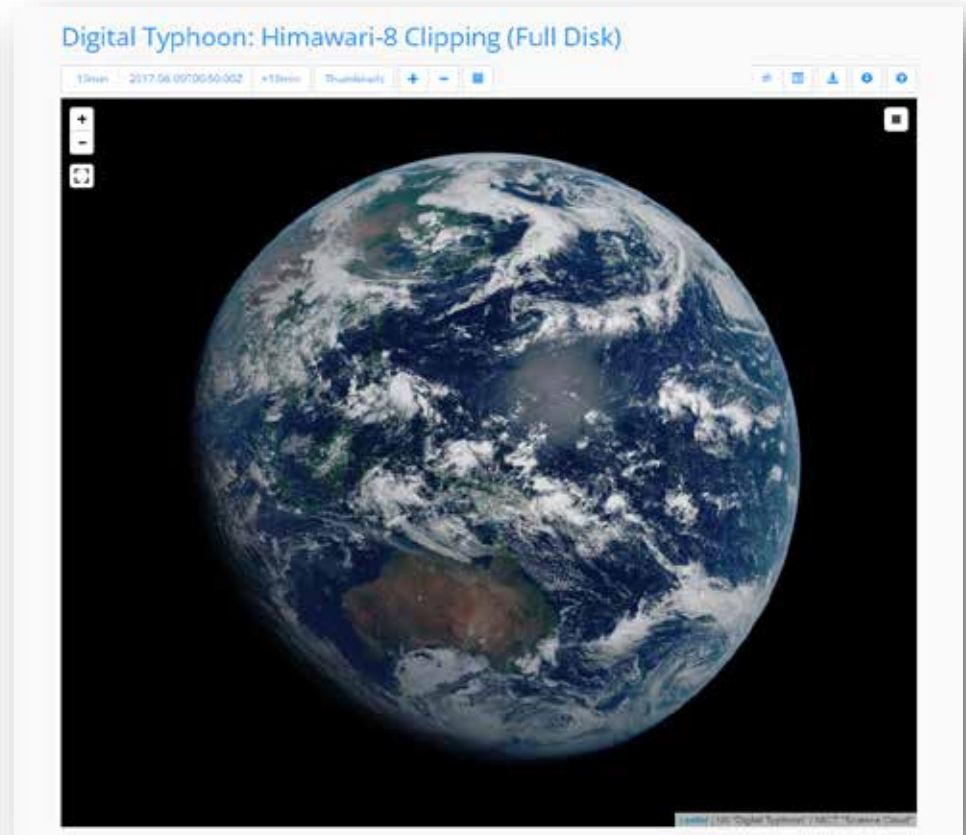
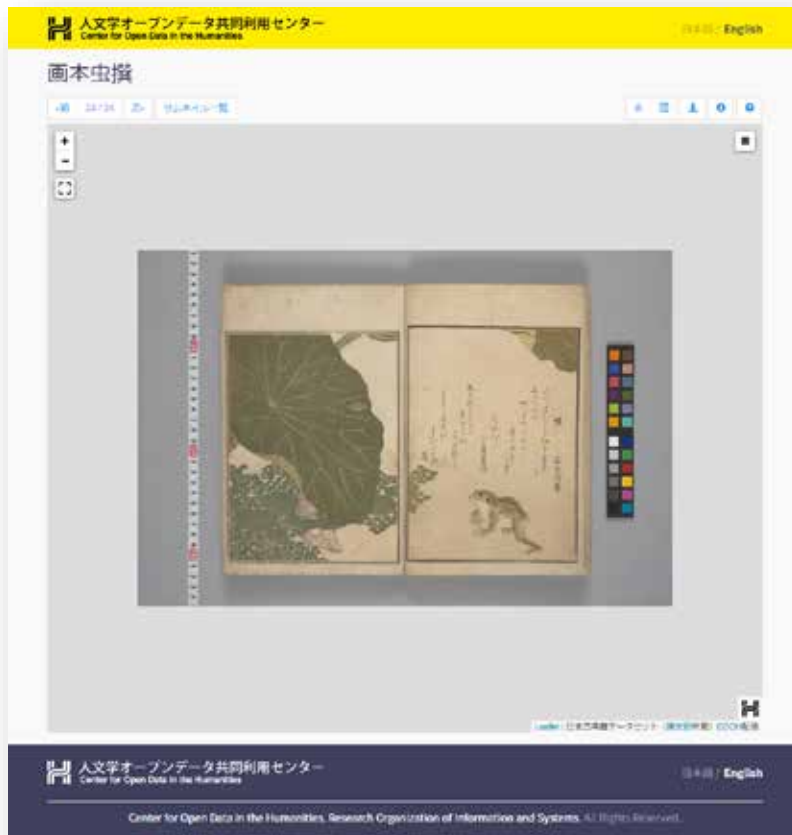
Commercial & Open Source Communities



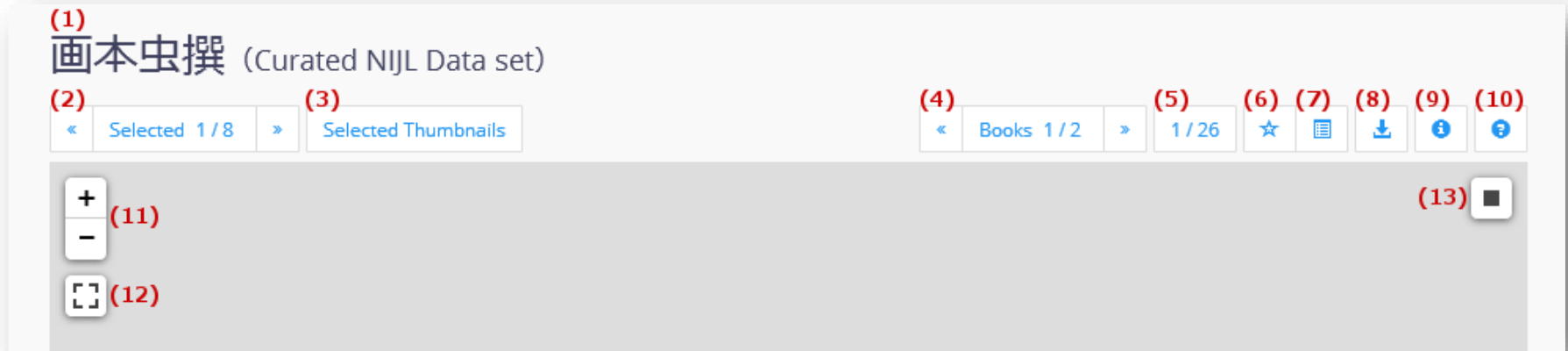
Sheila Rabun, IIF Community Groups & Engagement, IIF Conference 2017.

IIIF Curation Viewer (for Timeline)

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



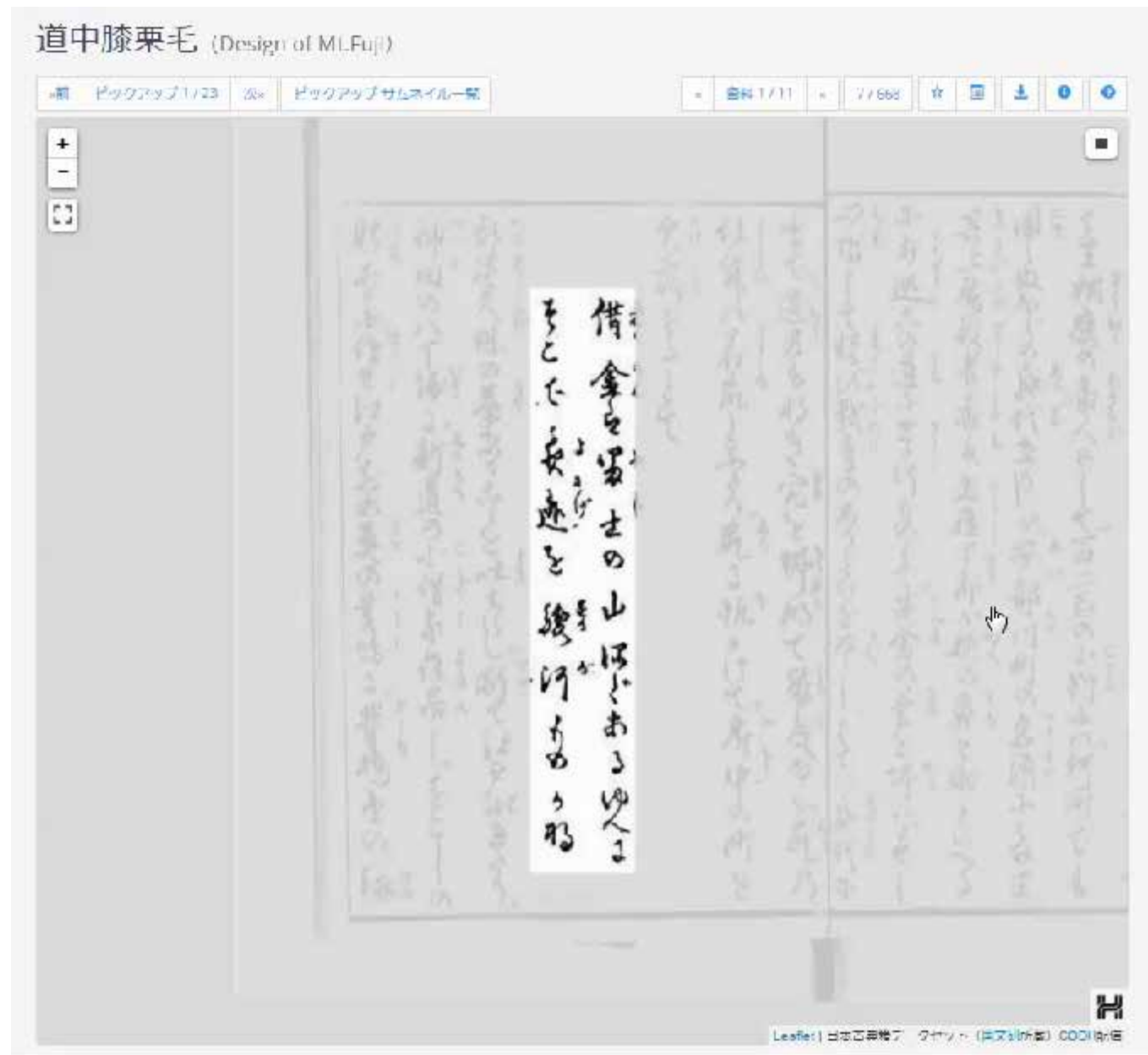
Curation on the Viewer



- We define curation as selection and ordering of interesting objects from the collection.
- ' ' (13) is a tool to draw a rectangle on a canvas to select the region of interest.
- ' ' (6) is a “favorite” button to keep interesting objects (a whole image or a region)

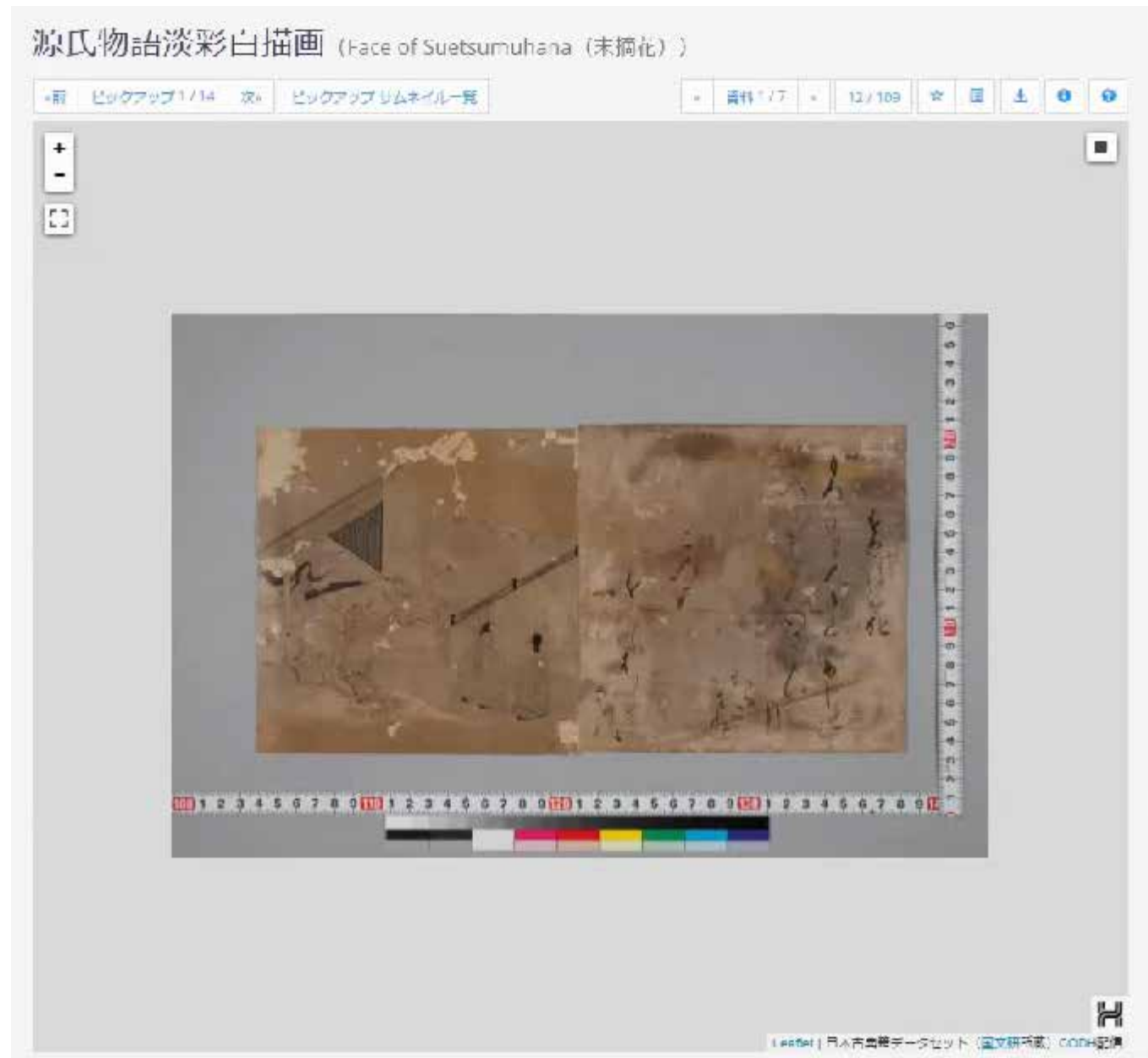
Design of Mt. Fuji

<http://codh.rois.ac.jp/iiif/iiif-curation-viewer/index.html?curation=/software/iiif-curation-viewer/demo/fuji.json>



Face of Suetsumuhana

<http://codh.rois.ac.jp/iiif/iiif-curation-viewer/index.html?curation=/software/iiif-curation-viewer/demo/suetsumu.json>





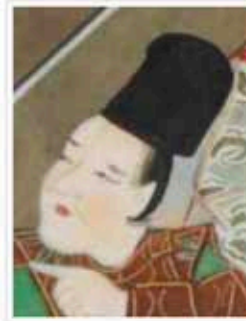
1



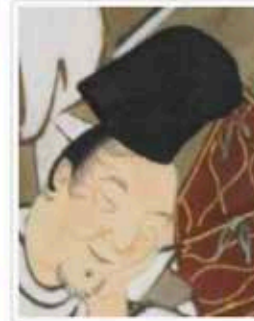
2



3



4



5



6



7



8



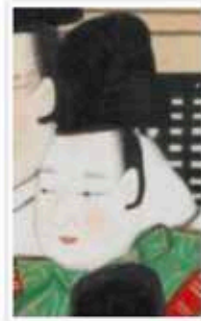
9



10



11



12



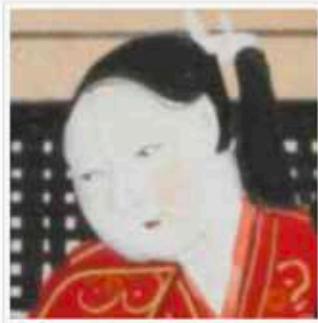
13



14



15



16



17



18



19



20

『宇津保物語』日本古典籍データセット（国文研所蔵）CODH配信



61



62



63



64



65



66



67



68

咸陽宮/30



69



70



71



72



73



74



75



76



77



78



79



80

『咸陽宮』日本古典籍データセット（国文研所蔵）CODH配信

Impact of Digital Humanities

1. It is not only about data, using digital sources instead of analogue sources.
2. **It is also about tools, answering new questions by digital tools on digital data.**
3. IIF Curation Viewer is a **tool to reduce the “friction”** of the analogue world.
4. By reducing pain, the tool allows scholars to **focus on the creative part of inquiry.**

粟田口隆光作品	『石山寺縁起絵巻』	粟田口隆光作品	『石山寺縁起絵巻』
			
豊田宗廟 上巻第一段	第二段	豊田宗廟 上巻第二段	第一段
			
豊田宗廟 上巻第二段	第一段	織通念仏 上巻第十段	第三段
			
豊田宗廟 上巻第一段	第三段	織通念仏 上巻第十段	第三段
			
豊田宗廟 上巻第四段	第一段	豊田宗廟 上巻第二段	第一段
			
豊田宗廟 上巻第二段	第一段	豊田宗廟 上巻第一段	第一段



神図5 『織通念仏縁起絵巻』 清涼寺



神図6 『豊田宗廟縁起絵巻』 豊田八幡宮



『石山寺縁起絵巻』



『織通念仏縁起』

神図7 相沢正彦氏の複製した『織通念仏縁起』と『石山寺縁起絵巻』

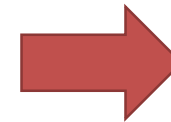
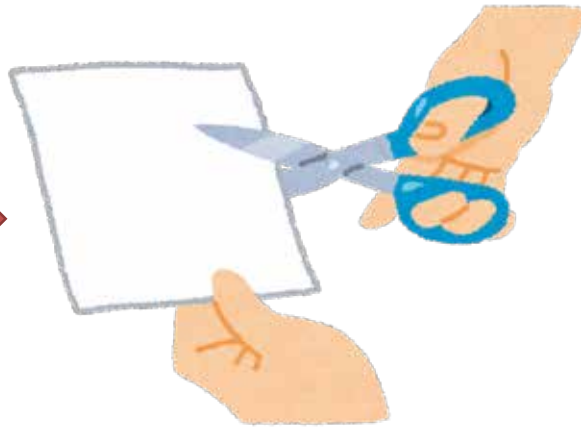
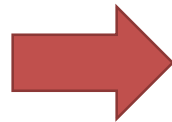
相沢正彦『石山寺縁起絵巻集成 論考・資料編』中央公論美術出版（2016年）P.20

Good Old Analogue World



Scissors

Paste

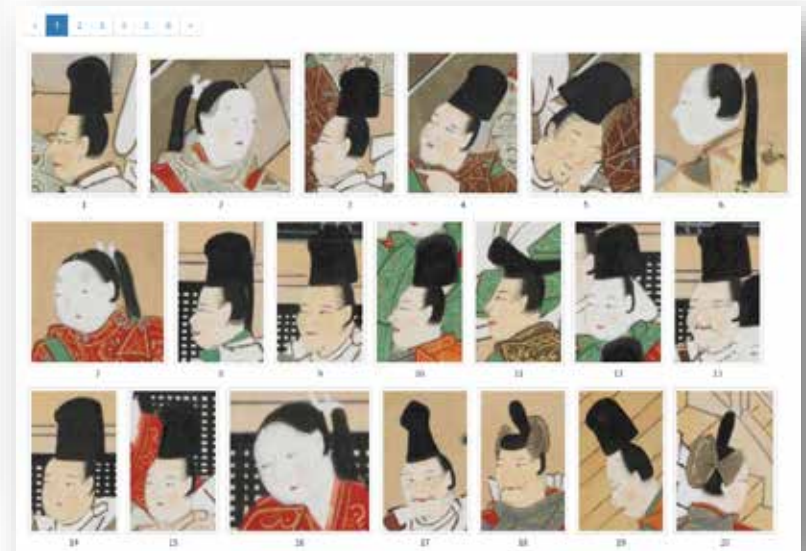


Source: いらすとや, <http://www.irasutoya.com/>

Frictionless Digital World



Simply add to Favorites





21



22



23



24



25



26



27



28



29



30



31



32



33



34



35



36



37



38



39



40

『宇津保物語』日本古典籍データセット（国文研所蔵）CODH配信

Imagine World-Wide Curation

Gallica

World Digital Library

Harvard art museum

Biblissima

Stanford University Library

Welcome Library

Bodleian Library

Qatar Digital Library

University of Illinois

Heidelberg University

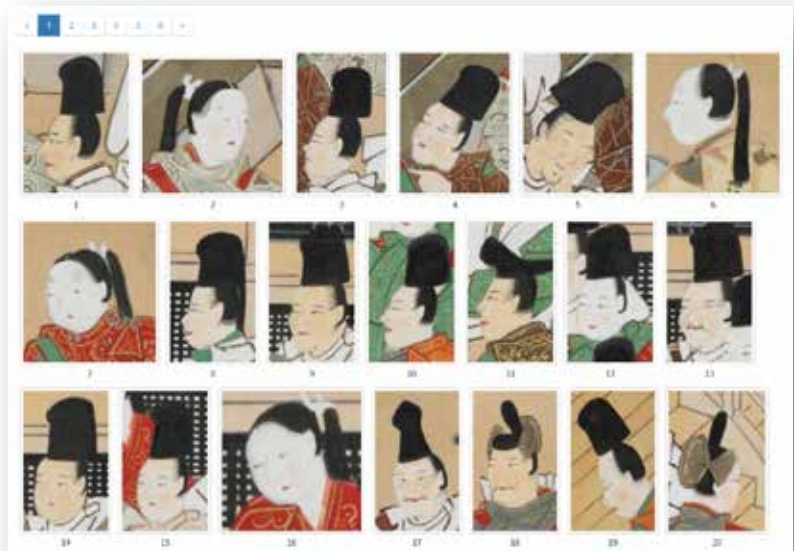
Yale University

Getty Museum

Vatican Digital Library

e-codices

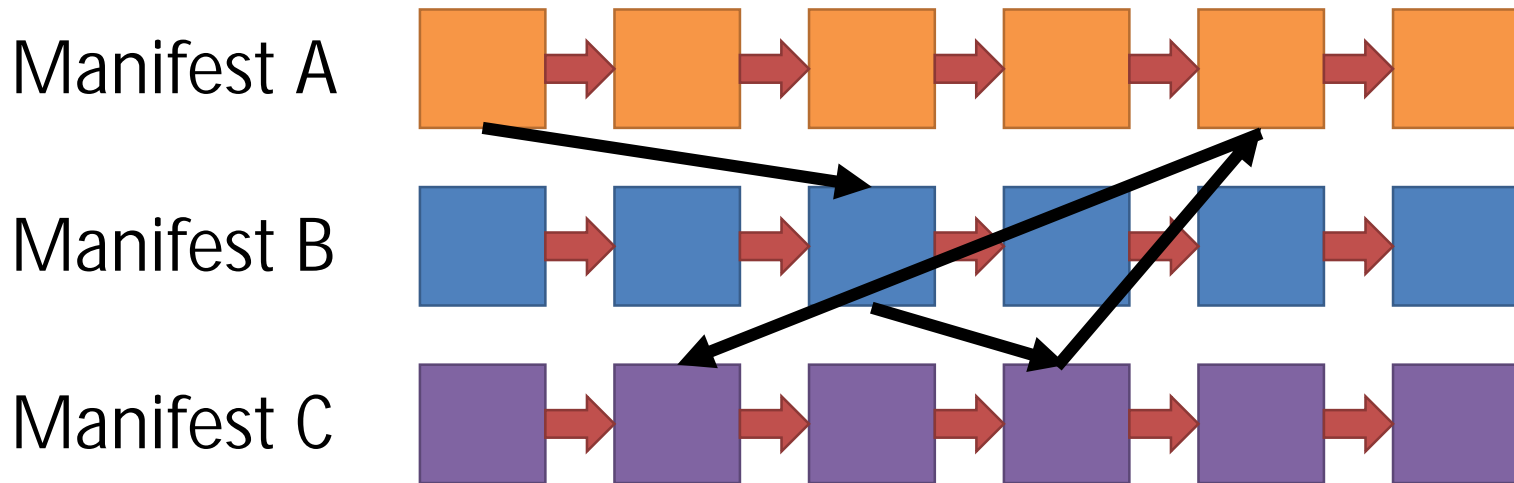
Simply add to
Favorites



Inquiry in Art History

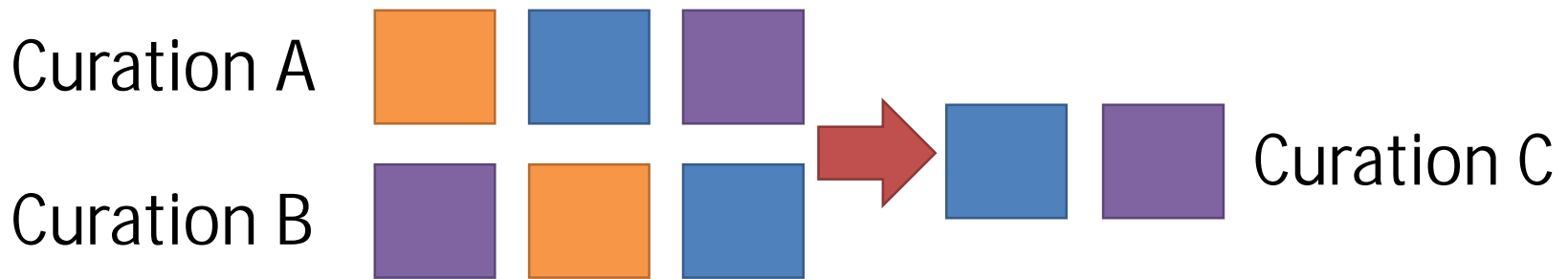
- Our great predecessors used analogue “scissors and paste” to study the style of art.
- Recent trend moves away from this tradition, probably because it requires “labor.”
- Digital version of “scissors and paste” allows scholars to focus on their creativity.
- Digital scissors work virtually and do not break the link to the original sources.

Fragmentation + Recombination



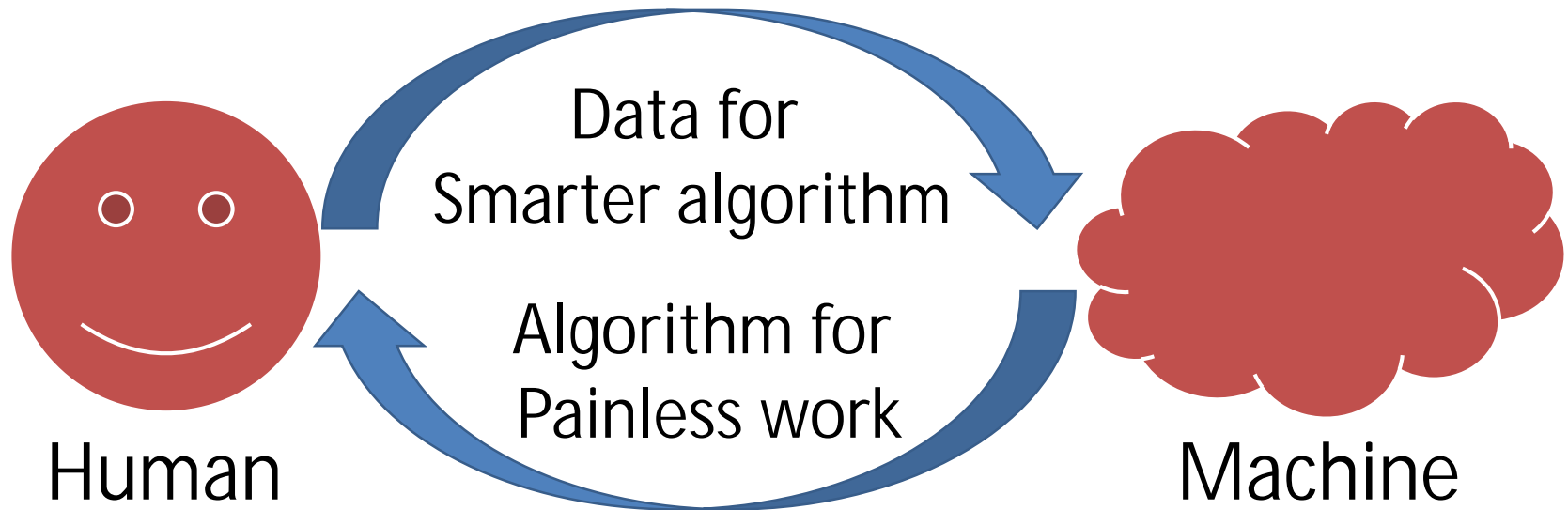
- A manifest is **fragmented** into canvases, and then **recombined** by a curator.
- **From content to context**, in which a viewpoint of a curator creates new meaning.

Evolution of Curation



1. **Chain of curation** = network of interesting objects in digital archives.
2. **Registry of curation** = alternative entry point to the world of IIF.
3. Curation with description will be **a new form of citable scholarly publication.**

Human-Machine Co-Evolution



- Curation = **annotation about interesting regions with simple metadata (tagging)**.
- Curation = **training data for machine learning (e.g. face recognition)**.

Digital Silk Road: Digital Archive of Toyo Bunko Rare Books

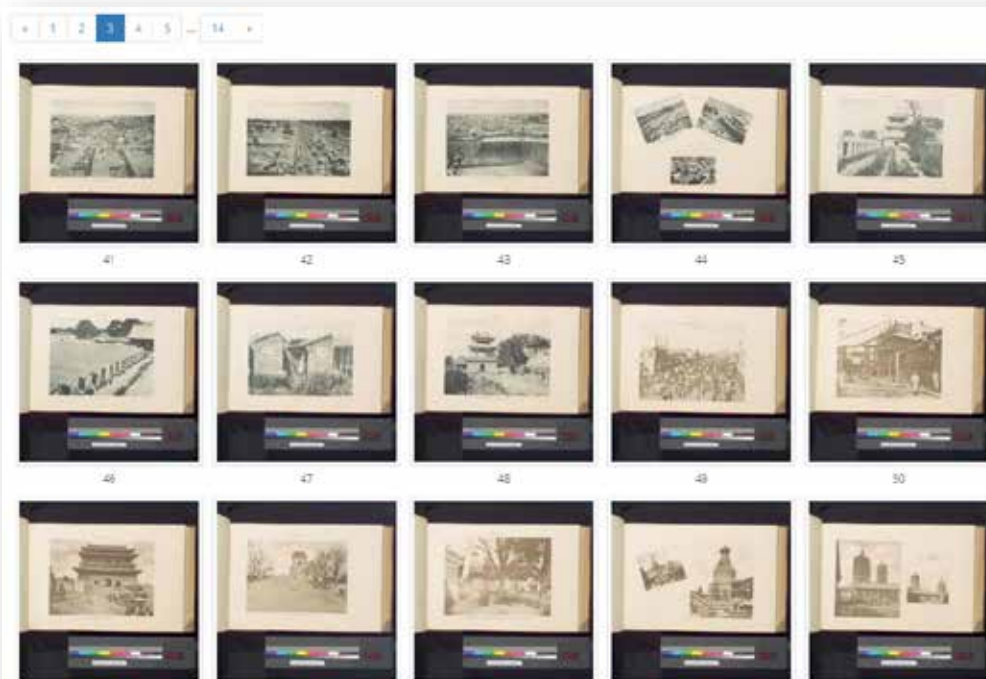
April 2004

Color images were restricted to 900 pix.



July 2017

Color images were upgraded to a few thousand pixels.



<http://dsr.nii.ac.jp/toyobunko/viewer/index.html?pages=La-161/V-1>

IIF-based for image delivery and **DOI-based permanent access** = a standard practice of the future.

Summary

- **IIF-based image delivery** is released for NIJL-PMJT and Toyo Bunko Rare Books.
- **IIF Curation Viewer** has “add to favorites” button based on our proposed Curation API.
- **Curation** allows selection of interesting pages and regions across books (and archives).
- **Registry of curations** may become a hub of research and educational materials.

Acknowledgment

- Mr. Jun Homma (@2SC1815J) as a core contributor for IIF Curation Viewer.
- Dr. Chikahiko Suzuki (CODH) as a creator of PMJT curations.
- IIF Curation Viewer is copyright of Center for Open Data in the Humanities (CODH), and MIT licensed.
- This work is partially supported by NIJL-NW project.

Please visit our Website
<http://codh.rois.ac.jp/>



Asanobu KITAMOTO

Director, Center for Open Data in the
Humanities (CODH) and
National Institute of Informatics

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

@rois_codh