

IIIFによる デジタルアーカイブ の提供と利活用

令和6年度大学図書館職員長期研修

2024/7/2

永崎研宣

慶應義塾大学文学部図書館・情報学専攻 教授

一般財団法人人文情報学研究所 主席研究員

「デジタルアーカイブ」と私

- 学術的に利用しやすいデジタルアーカイブの構築
 - ⇒ 古典籍を対象としたデジタル研究環境の構築
 - 具体的には：
 - [SAT大蔵経テキストデータベース](#)他色々
 - サーバシステム（ハードウェア）の仕様書作成・発注
 - ネットワークシステムの仕様書作成・発注
 - 必要なソフトウェアの自力開発（1996年～今も現役）
 - 必要な国際標準規格の調査と適用・改訂
 - Unicode (ISO/IEC 10646), TEI Guidelines, IIIF, ISO 639, CIDOC-CRM等
 - 各地の貴重資料のデジタル化・デジタル化支援＋IIIF等に対応
 - 東京大学総合図書館[嘉興蔵](#)・[サンスクリット写本](#)
 - 増上寺[三大蔵](#)
 - 諸々の活動のためのファンドレイジング
 - こういった活動を学術的にオーソライズするための学会設立と運営
 - 同様の理由による日本学術会議での活動（デジタル時代の人文・社会科学に関する分科会委員長）
 - …

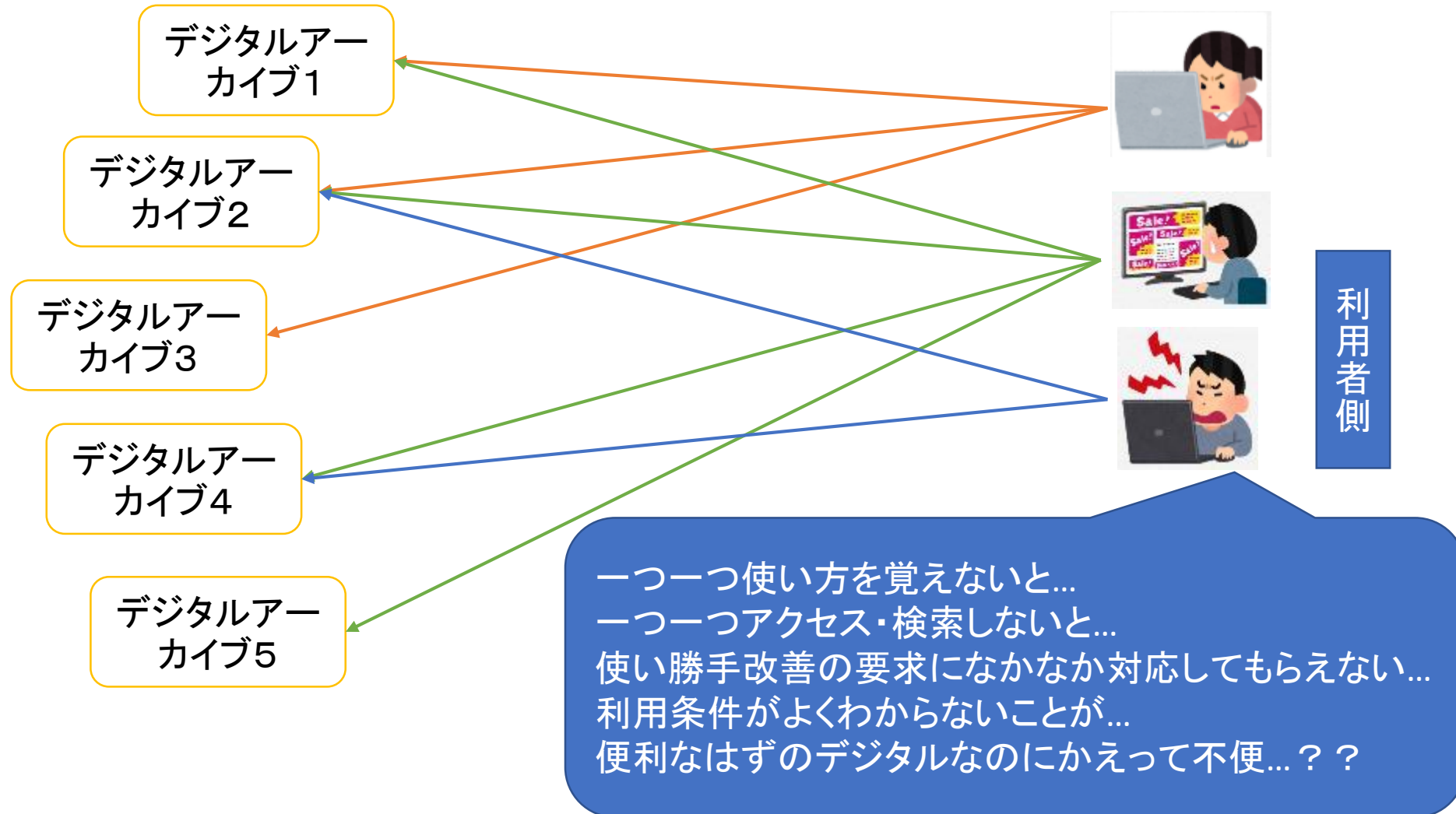
本件に関わる自己紹介

- 日本学術振興会 人文学・社会科学デジタルインフラストラクチャー構築推進センター 研究員 (2021-2022年度)
- 東京大学大学院人文社会系研究科次世代人文学開発センター人文情報学部門 客員研究員・非常勤講師 (2012年-)
 - その他、デジタル人文学の授業担当 (関西大・同志社大・立教大・筑波大・大阪大・広島大)
- 京都大学人文科学研究所共同研究班「人文学にとってのWebを再探する」班長(2013-15年)
- 国立国会図書館研究員 (委嘱) (2014年-)
- 国文学研究資料館客員教授 (2021-2023年度)
- 広島大学文学部客員教授 (2021-2023年度)
- 北米日本研究資料調整協議会 (NCC) CDDP Task Force (2020年-)
- 日本デジタル・ヒューマニティーズ学会議長 (2019年-)
- Alliance of Digital Humanities Organizations運営委員 (国際DH学会連合) (2019年より)
- 情報処理学会論文誌編集委員 (2017-19) / 人文科学とコンピュータ研究会運営委員(?より)
- Text Encoding Initiative Consortium 理事 (2017-18年)
- IIF Consortium TRCメンバー (2018-21年)
- 情報規格調査会SC2委員会委員 (2013年-)
- ISO/IEC JTC1/SC2 リエゾンメンバー (SATからの代表として) (2017年-)
- SAT大蔵経テキストデータベース研究会技術担当 (2005年-)
- 日本学術会議連携会員・デジタル時代における新しい人文・社会科学に関する分科会委員長 (第25・26期)

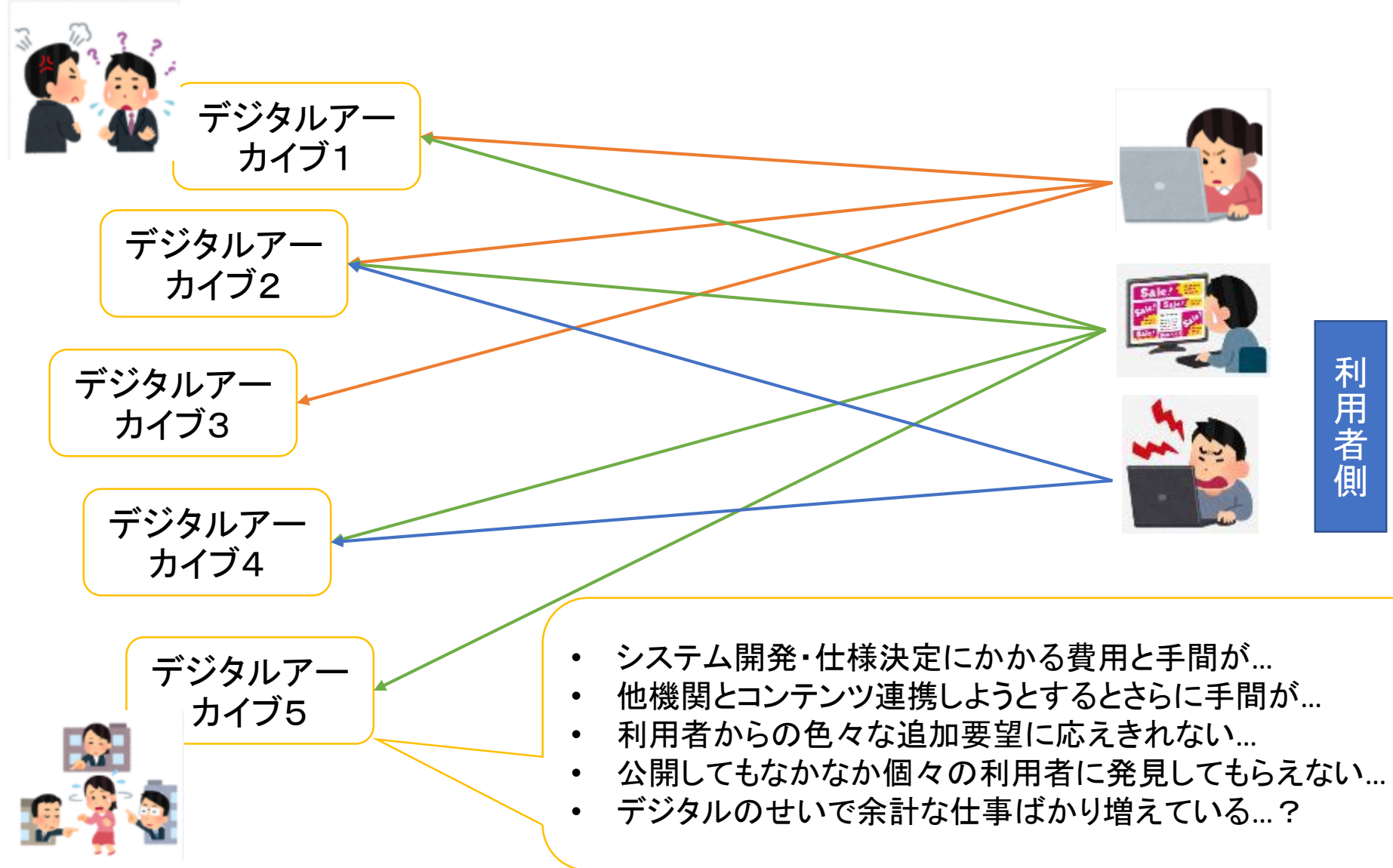
デジタル化した資料を活用したサービス

- 2018年頃までの主流
- 「デジタルコレクション」 「デジタルアーカイブ」
 - 貴重資料をWebサイトで見せる
- ⇒自分のサイトに「来て」「みて」もらうところまで
 - 利用者側：
 - それぞれに使い勝手の違う多くのサイトをめぐらねばならなくて大変
 - 再利用には申請が必要でなかなか大変
 - 基本的に、「見るだけ」
 - 公開者側：
 - 独自性を目指して頑張る or 既存のプラットフォームにとにかく載せる

これまでの「デジタルアーカイブ」



これまでの「デジタルアーカイブ」



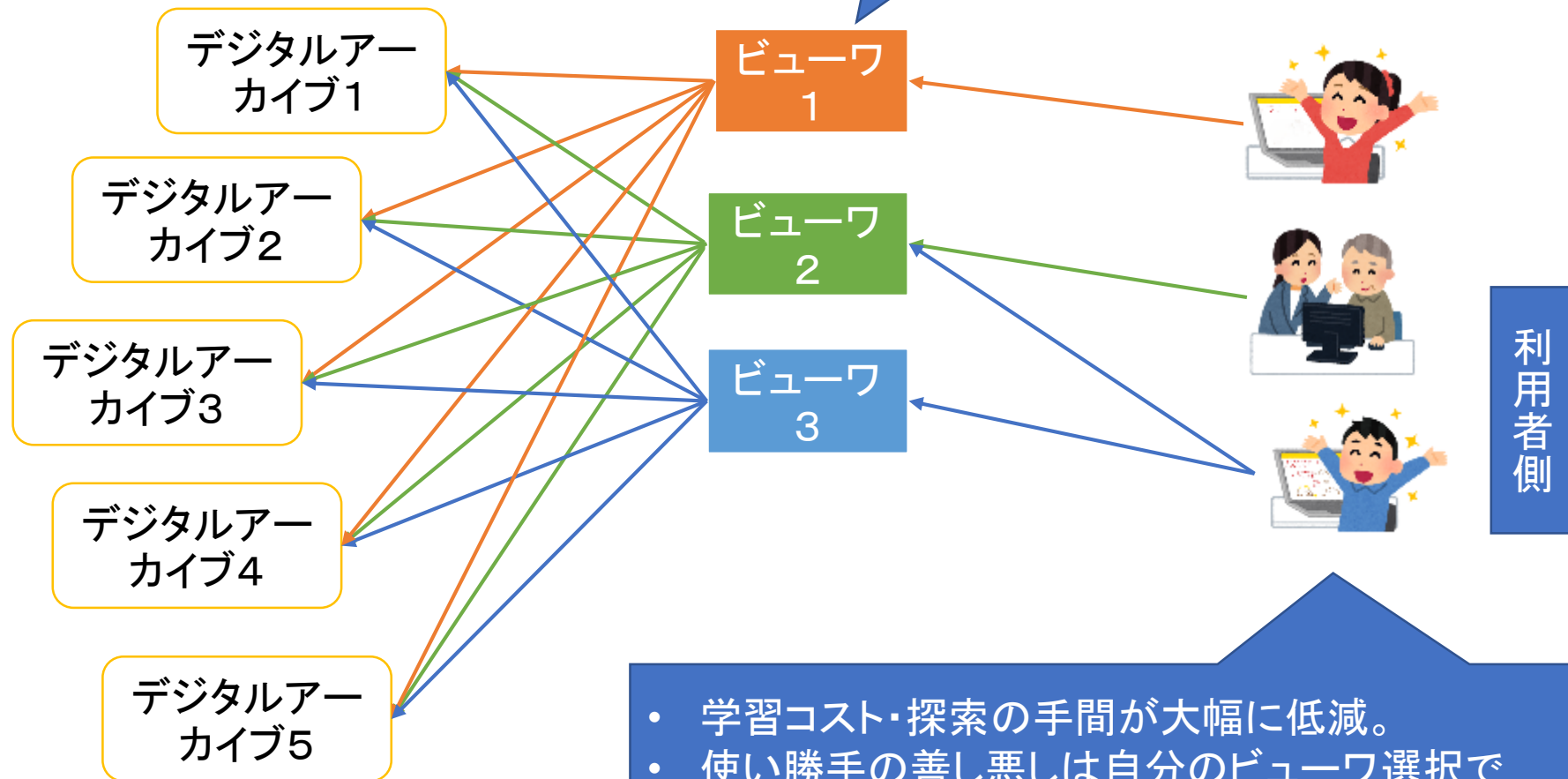
二つの転機： 利用条件のオープン化とIIIFの普及

- 2015年3月、東京大学総合図書館所蔵の万暦版大蔵経（嘉興蔵）のデジタル画像が**CC BY**で公開
 - ⇒ 自由な利用条件の明示により、自由な再利用が可能に
 - 先例：
 - 2014年3月の東寺百合文書WebのCC BY公開
 - 2014年5月1日、NDLデジコレは著作権保護期間満了の画像の転載申し込みを不要に
 - CC BYから「パブリックドメインだけど所蔵者表示希望」へ
- 2018年5月、NDLデジタルコレクションがIIIFに対応
 - 34万点のIIIF対応デジタルコンテンツの出現
 - ⇒ Webサイト間での画像の自由な横断的操作が可能に
 - ⇒ 対応機関 & コンテンツが増えれば増えるほどサービスの質が向上
 - 東大、京大、慶應大、国文学研究資料館等も相前後して対応
 - 現在は多くの大学図書館が対応

解決策

「画像のやりとりの仕組みを（APIで）共通化」

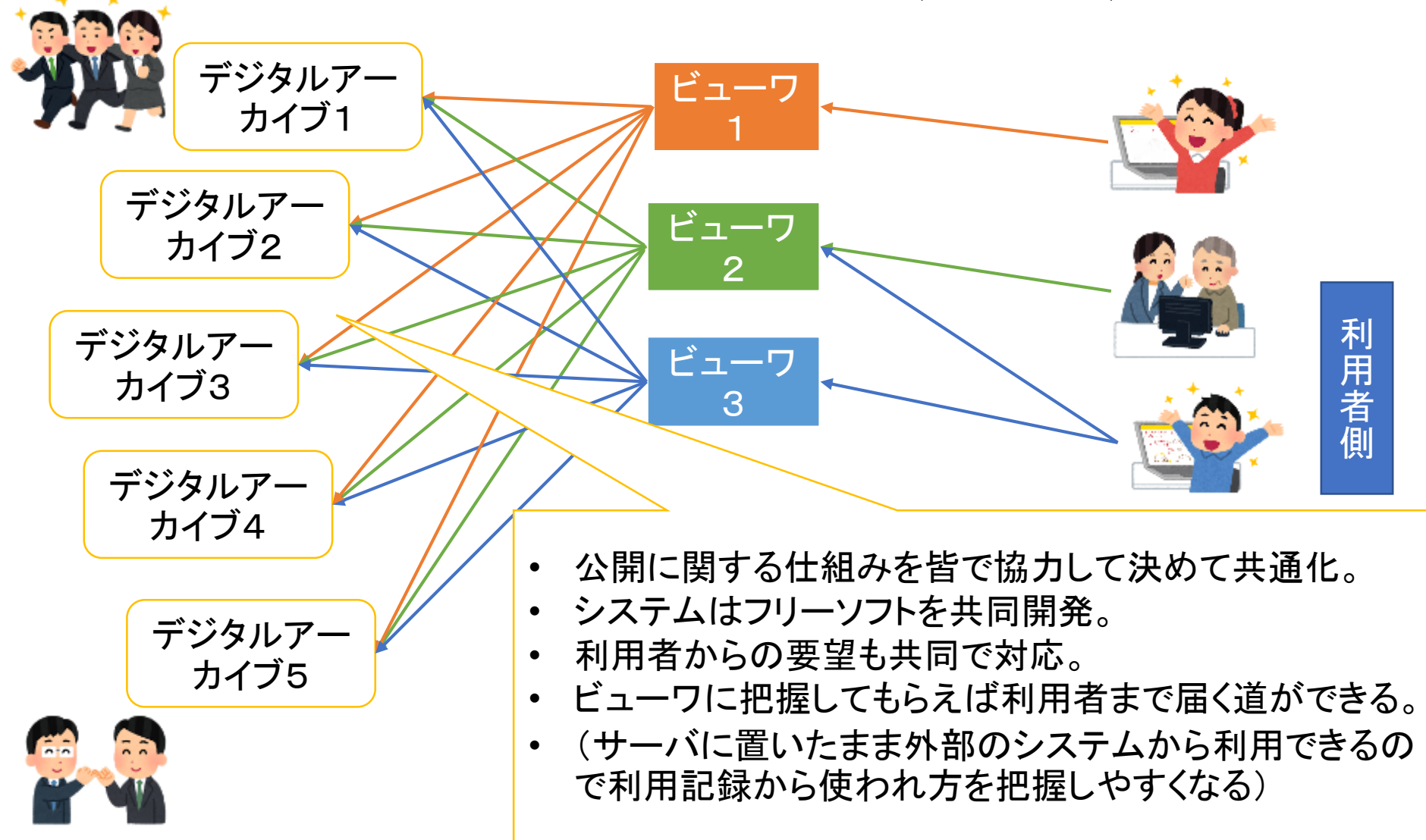
ビューワが裏側ですべてのデジタルアーカイブとやりとりしてくれる



- 学習コスト・探索の手間が大幅に低減。
- 使い勝手の善し悪しは自分のビューワ選択で
- 改善は自分が利用するビューワの開発者に要求
- ライセンス表示も共通なのですぐわかる

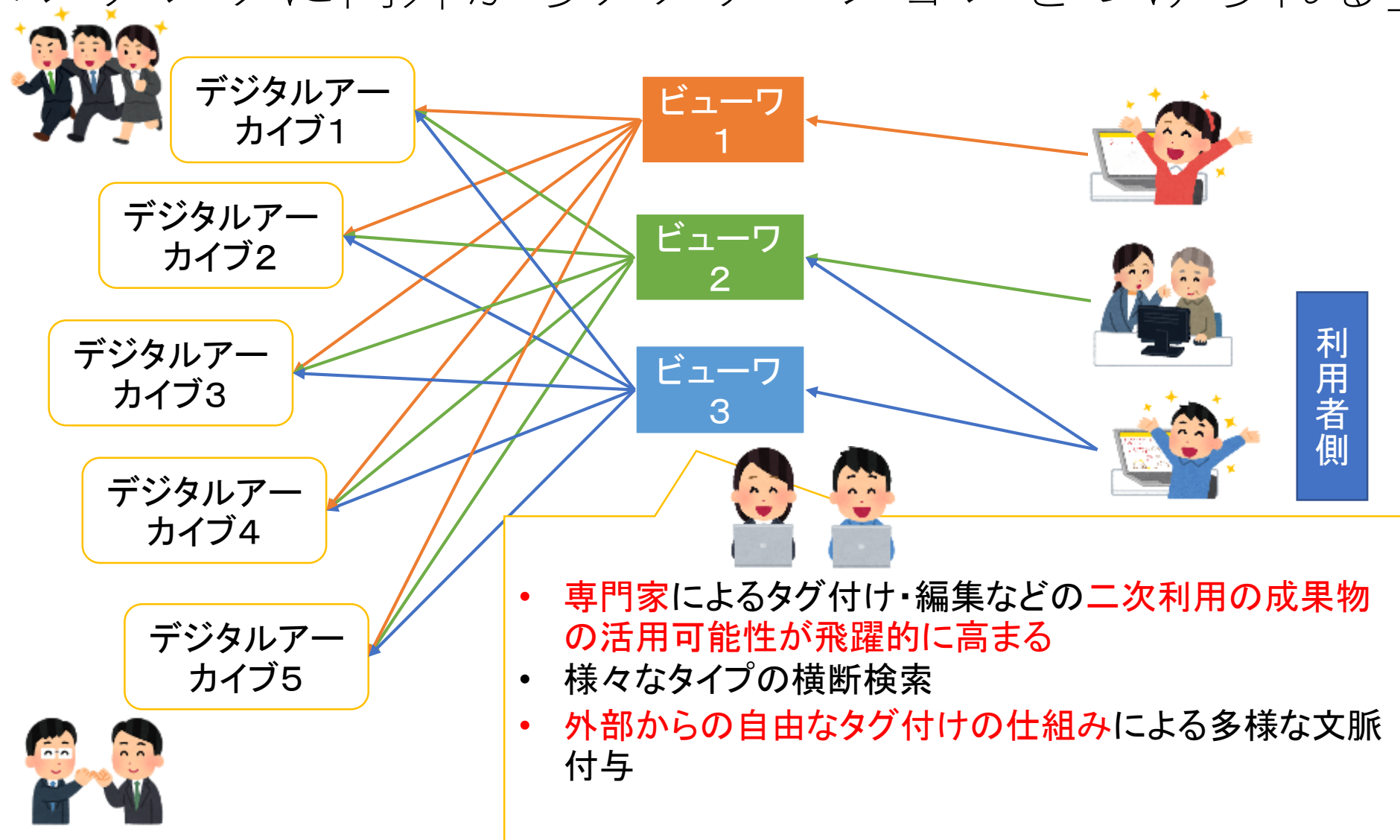
解決策

「画像のやりとりの仕組みを（APIで）共通化」

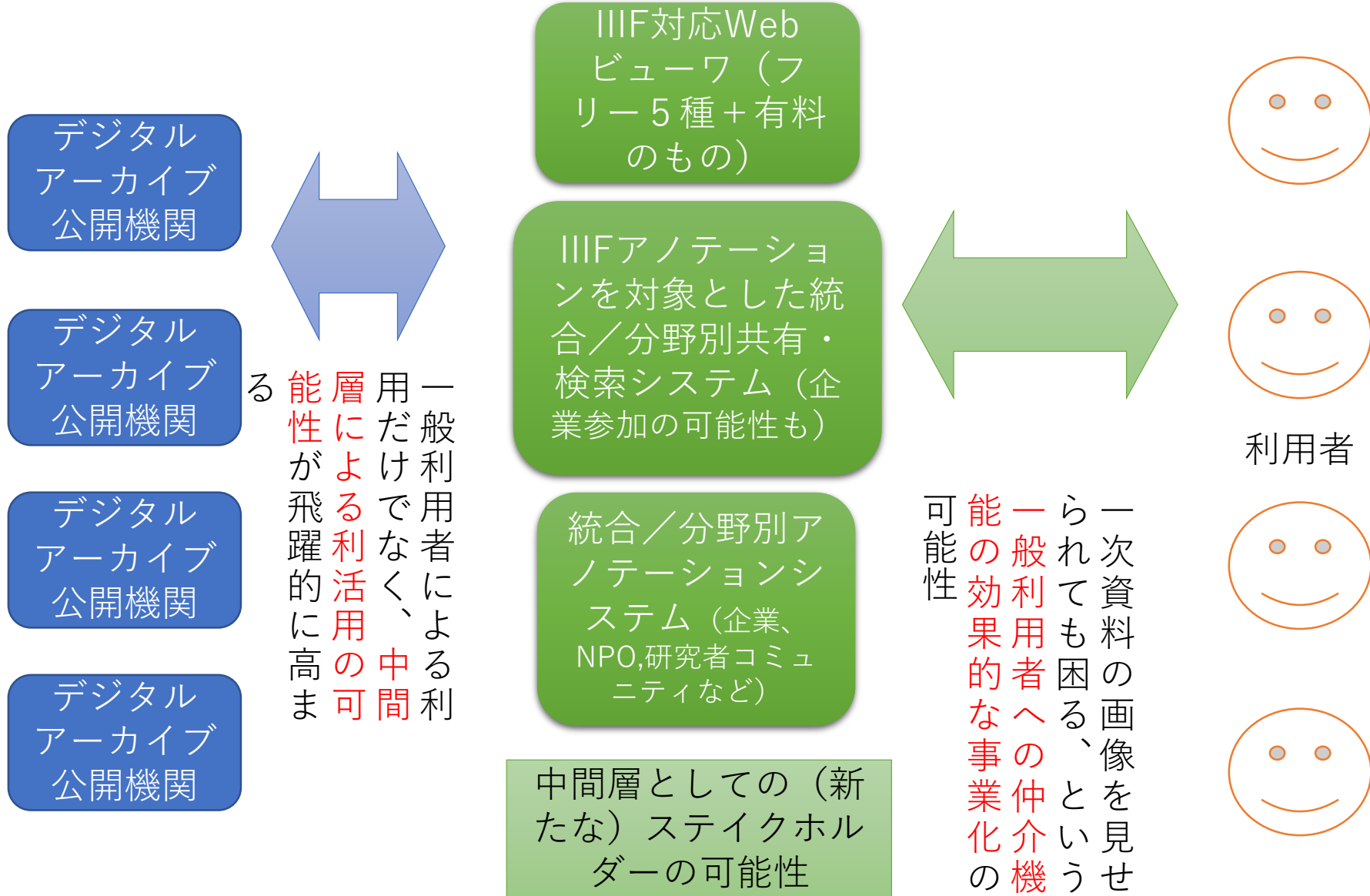


さらに：

「コンテンツに内外からアノテーションをつけられる」



新たなステイクホルダーの可能性



「公開したIIF対応画像の活用」による サービス展開

- 国内でも様々な展開がなされるように：
- 様々なアノテーションを付与&並置
- 各地の画像を持ってきて並置
- 各地の画像を別のサイトに表示させて新たなサービスを展開
- 各地の画像の任意の部分を切り出して並置したり分析したり

様々なアノテーションを付与

- <https://dzkings.l.u-tokyo.ac.jp/SATi/images.php>
 - SAT大正蔵図蔵DB
 - ポイント：
 - 国際標準規格IIIFに準拠した形式のテキストデータ（JSON-LD準拠のIIIF形式）を作ると、国際標準規格に準拠したビューワ（フリーソフト）でアノテーションとしてきれいに表示できる。
 - 複数の画像を並置して比較できる
 - 参照：<https://bungaku-report.com/sat.html>（OA本）
 - <http://id.nii.ac.jp/1001/00177361/>
- <https://demos.biblissima.fr/ovide-moralise/>
 - インキュナブラの挿絵比較 by Biblissima
 - オウィディウス『変身物語』の挿絵

各地の画像を持ってきて並置

- http://www.kulib.kyoto-u.ac.jp/rdl/digital_fujikawa/
 - 富士川文庫デジタル連携プロジェクト試行版
 - <http://id.nii.ac.jp/1001/00192389/>
- <https://genji.dl.itc.u-tokyo.ac.jp/>
 - デジタル源氏物語
 - <http://id.nii.ac.jp/1001/00206588/>

各地の画像を別のサイトに表示させて新たなサービスを展開

- <https://honkoku.org/>
 - みんなで翻刻
 - ⇒ <https://honkoku.dhii.jp/>
 - みんなで翻刻サーチ
 - <http://id.nii.ac.jp/1001/00208573/>
- <https://bauddha.dhii.jp/SAT/iiifmani/show.php>
 - IIF Manifests for Buddhist Studies
 - 京大図書館の画像に専門家がメタデータの一部を付与
 - Cf. <http://id.nii.ac.jp/1001/00192403/>
- <https://cultural.jp/>
 - Cultural Japan
- <https://jpsearch.go.jp/>
 - ジャパンサーチ

データ駆動型研究のプロセスがもたらす意義



データ駆動型研究の プロセスを通じた社会貢献

= データ作成期間中の
社会貢献

協働プラットフォーム



みんなて翻刻
<https://honkoku.org/>

少子高齢化社会に
おける共創の場

貢献

学び

人力+AI協働
の学びの場

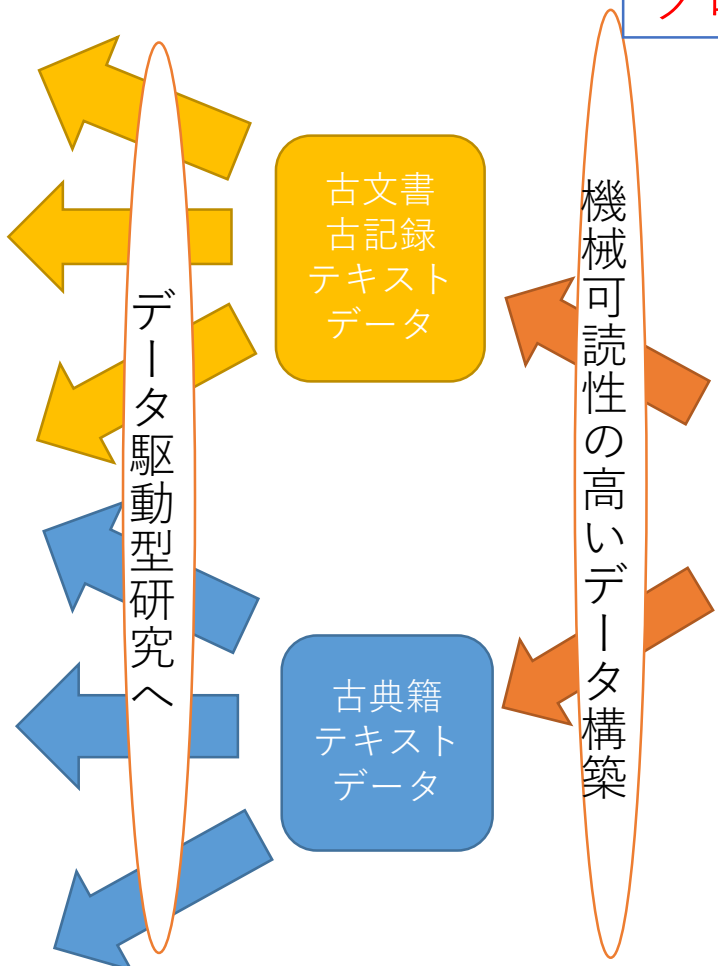
高齢者

若者

経験と学び

IT

日本文化



古気象研究
(地球温暖化)

古地震研究
(防災・減災)

地域史
(地方創生)

ジェンダー問題
(男女共同参画)

日本文化の
再発見

雉尾攷—日本書紀にみる赤気に関する一考察
<http://id.nii.ac.jp/1013/00005767/>

歴史のなかの地震・噴火: 過去がしめす未来
<https://ci.nii.ac.jp/ncid/BC05967143>

各地の画像の任意の部分を切り出して並置したり分析したり

- <http://codh.rois.ac.jp/face/>
 - 顔貌コレクション
 - <http://id.nii.ac.jp/1001/00208577/>
- https://base1.nijl.ac.jp/~collectors_seal/seal_script/?char=%E5%B1%B1
 - 篆字部首検索システム（テキスト検索版）
 - 切り出し君2017 IIIIF版によりデータ作成
 - <https://candra.dhii.jp/tesseract/iiifcrop/>

静止画像を超えて：

- テキストデータベースとの組み合わせ
 - <https://genji.dl.itc.u-tokyo.ac.jp/>（既出）
 - デジタル源氏物語 - 「似たページ」の自動検出
 - <https://21dzk.l.u-tokyo.ac.jp/SAT2018/master30.php>
 - 1億字のテキストデータベースと世界各地のIIIF画像の組み合わせ
 - <https://candra.dhii.jp/nagasaki/manyo/manyoviewer2021.html>
 - https://www.ku-orcas.kansai-u.ac.jp/news/20210129_758/
 - 廣瀬本萬葉集画像とTEI構造化詳細テキストデータの組み合わせ
 - <http://id.nii.ac.jp/1001/00209265/>
-
- 動画アノテーション
 - <https://dh.l.u-tokyo.ac.jp/activity/iiif/video-annotation>
 - 2017年制作の動画にアノテーションをつけることで最新に

最近の技術動向として

- <https://current.ndl.go.jp/ca1989>
 - **CA1989 - 動向レビュー：IIIFの概要と主要APIバージョン3.0の公開**

二つの方向性が可能に

- 自館のデジタル化資料をIIIF対応で公開することで利活用の環境を提供
 - IIIF対応システムの納入実績のある企業リスト
 - <https://digitalnagasaki.hatenablog.com/iiif>
 - フリーソフトの組み合わせだけでも対応可能。
 - ⇒腕に覚えのある司書／研究者が手伝ってくれば費用はあまりかからない
- 各地のIIIF対応のコンテンツを応用して新たなサービスを提供
 - 自館のコンテンツの強みをさらにアピールできることも
 - ⇒デジタル源氏物語、富士川文庫等、篆字部首検索システム／字体データセット、みんなで翻刻…
 - 自館のコンテンツは必ずしも必要ない
 - ⇒顔貌コレクションなど

参考情報

IIIFの導入

JSON-LDに変換

- メタデータ、タグ、アノテーション等
- 「1画像で1キャンバス」
- 「1オブジェクトに1マニフェスト」
- 「オブジェクトのまとまりで1コレクション」

提供者・
開発者向け

マニフェストファイルのURLさえあればどこのビューワでも表示可能（要合意形成）

- ### IIIF対応ビューワ
- OpenSeadragon
 - Mirador
 - Universal Viewer
 - Leaflet-IIIF
- And so on...

Presentation
API (JSON-LD)

既存のデジタル
アーカイブ

画像サーバ
上の画像URL
を記述（必須）

IIIF対応
デジタルアーカイブ
公開・活用システム

Image API
(URL)

要求URLに応じたサイズ・部分・
色数等で画像を返戻

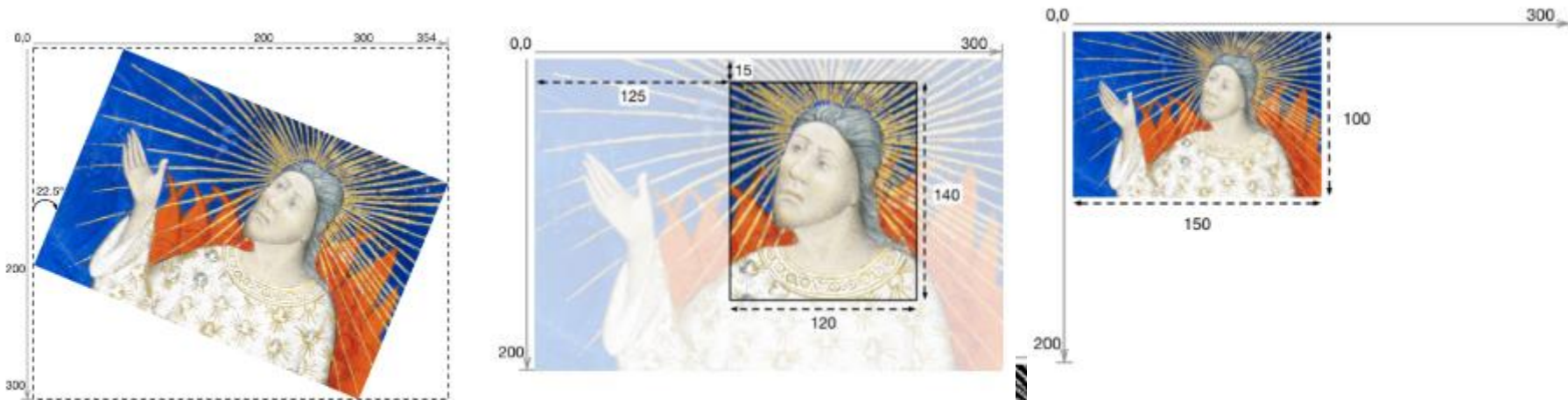
- ## 対応画像サーバ(フリーソフト有り)を用意
- Jpegをそのまま使えるLoris server, digilib等
 - Tiled Tiff, Jpeg2000対応のIIP Image server等

様々なシステム、
ビューワなど

IIIF Image API

<http://iiif.io/api/image/2.1/>

- URLで画像の形状・部分等を指定して取得するルール



4 rotation=22.5
.../full/full/22.5/default.png



3 quality=gray
.../full/full/0/gray.jpg

3 region=125,15,120,140
.../125,15,120,140/full/0/default.jpg



4 quality=bitonal
.../full/full/0/bitonal.jpg

2 size=150,
.../full/150,/0/default.jpg

IIIF Presentation API

- 一つの資料ごとに一つのManifest (file)
 - Manifestをまとめてコレクションを形成可能
- JSON-LDとWeb Annotationの規格を利用
 - メタデータ等もJSON-LDで格納
- Shared Canvasという概念を採用
 - Canvasの並び方とCanvasに載せるものを規定する
 - Canvasに載せるものはWeb Annotation (Open Annotation)、Media Fragment URI等のW3C標準のルールで記述

関連URL

- <http://iiif.io/> (IIIF公式サイト)
- <https://github.com/IIIF/awesome-iiif> (Awesome IIIF)
- <http://www.slideshare.net/azaroth42/introduction-to-iiif> (IIIF協会の中
心人物によるIIIF紹介スライド)
- <http://projectmirador.org/> (Mirador公式サイト)
- <https://github.com/UniversalViewer/UniversalViewer/wiki> (Universal
Viewerサイト)
- <http://codh.rois.ac.jp/icp/> (IIIF Curation Platform)
- <http://digitalnagasaki.hatenablog.com/iiif> (IIIFに関する日本語情
報の私的なまとめ)