https://doi.org/10.2352/issn.2168-3204.2017.1.0.34

Content-based Interoperability: Beyond Technical Specifications of Interfaces

Tobias Schweizer, Lukas Rosenthaler, Peter Fornaro; Digital Humanities Lab, University of Basel; Basel, Switzerland

Abstract

On the technical level, the International Image Interoperability Framework (IIIF) offers a common standard for referring to images or parts of images using a well-defined URL-based syntax. This allows transformations like scaling, rotation, mirroring, and format conversions to be applied. However, for the full integration of digital objects in other environments, interoperability has to go beyond mere technical specifications for access. Non-technical metadata such as image descriptions, annotations, and references to other objects (e.g., texts, transcriptions, and other visual representations) must be available as well. Using the example of digital facsimiles of the scientific notebook of Jacob I Bernoulli and its transcriptions, we show how XML together with the IIIF Image API allows for a seamless integration and merging of image and text sources in the same environment.

Introduction

Accessibility and interoperability of digital content is becoming increasingly important. The Image API of the International Image Interoperability Framework (IIIF) is an innovative development in this context. It solves the problem of different non-standardized APIs for accessing images.

Interoperability is crucial for bringing together various digital sources from different repositories into an infrastructure that allows researchers to annotate digital objects and supports them in gaining new knowledge.

However, the definition of a standardized access format is only one important step on the way to achieving this aim. Besides the definition of technical specifications to access resources properly, the ability to merge or compare digital objects themselves is needed. We call this next step "content-based interoperability". We will demonstrate the importance of this approach with a concrete use case from the Bernoulli-Euler Online (BEOL) project.[1] BEOL is an interdisciplinary collaboration between the Bernoulli-Euler Centre and the Digital Humanities Lab (both University of Basel), focusing on the mathematics influenced by the Bernoulli dynasty and Leonhard Euler.

Combining a VRE and an Image Repository

The Digital Humanities Lab is developing a flexible Virtual Research Environment (VRE) called Knora/SALSAH[2]. Knora refers to the server software, SALSAH to the web-based graphical user interface. Knora is based on Semantic Web technologies (an RDF triplestore and OWL ontologies) and offers a RESTful API for performing all necessary operations.[3] The Knora base ontology provides generic, abstract definitions that can be used and further extended in project-specific ontologies. SALSAH runs in a web browser and communicates with Knora using the API. Knora/SALSAH is part of the strategy of the Swiss Data and Service Center for the Humanities for making research data available on the long term, using a 'keep-alive' archiving strategy [3, 8]. Resources and annotations are stored in the VRE's RDF triplestore or are loaded from an external repository. A resource that can be worked with in Knora/SALSAH can be represented by digital images, and moving images and marked up texts are also supported.

Image-based content, however, is not stored or maintained by the VRE's triplestore itself. Instead, it is stored in an image repository called Simple Image Processing Interface (Sipi)[4] that puts a strong emphasis on preserving the images' metadata as required in image archives [4]. To make the VRE and the image repository interoperable, we rely on the IIIF Image API. With IIIF, the image archiving function of SIPI can be combined with the direct usability of the images by the VRE.

Use case: The Meditationes

Knora/SALSAH is being used to create a digital edition of the *Meditationes*, the scientific notebook of Jacob I Bernoulli (1654–1705) which is part of the Bernoulli-Euler Online project.[5] The edition is based on the definition of segments on the manuscript's pages that are transcribed individually. Each segment is represented by a region of interest (ROI) that is stored in Knora/SALSAH.

Figure 1 shows page 184 of the *Meditationes* with ROIs defined for *Meditatio* 151. The different colors indicate different categories of image segments: title text 'TT' and text 'T' (green), mathematical formulae 'M' (red) and mathematical figures 'F' (blue). The transcriptions are encoded in an XML-based format and refer to a ROI by its IRI. Listing 2 shows the XML transcriptions for three regions shown in Figure 1 (the two green regions on the top and the red region on the upper left as well as the blue region for the mathematical figure). Per entry – a *Meditatio* – one XML file is created that contains the transcriptions for all its regions. In the XML, each ROI is represented by a region element indicating its IRI and category as well as its name. The XML markup represents layout information, authorial changes of the text, expansions of abbreviations, and editorial interventions, and LaTeX is also used to encode the mathematical notation in the manuscript. Entities are used to encode special characters. The

[1] `http://p3.snf.ch/project-166072`

[2] `https://github.com/dhlab-basel/Knora`

[3] See documentation on `http://www.knora.org`

[4] `https://github.com/dhlab-basel/Sipi`

[5] The original manuscript with the full title *Meditationes, Annotationes, Animadversiones Theologicae & Philosophicae* is kept at the Manuscript Department of the Basel University Library (shelfmark L Ia 3, 367 pages). It is accessible (without transcriptions or commentaries) at `http://dx.doi.org/10.7891/e-manuscripta-43865`.

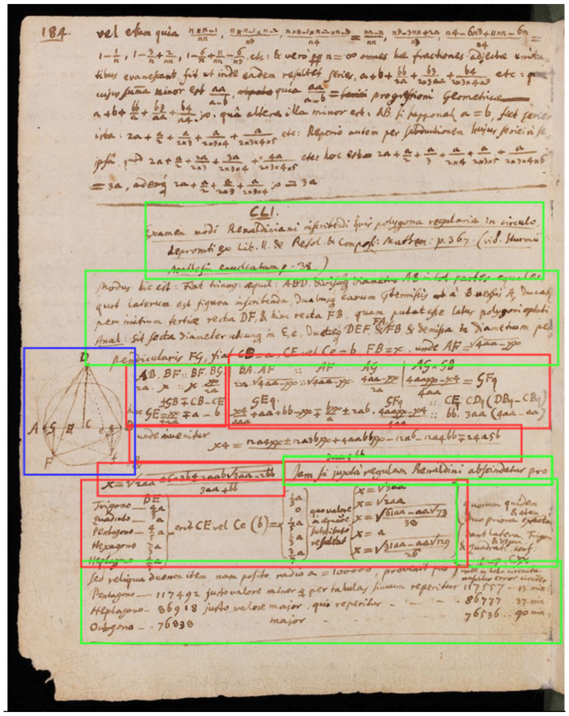

Figure 1. *Regions of Interest defined on page 184 of the Meditationes (Meditatio 151)*

XML does not conform to the TEI/XML guidelines, but TEI definitions are used whenever possible.[6] The region element is similar to the facsimile element defined by the TEI [9, section 11.1]. Since our approach is based on regions, the paragraph structure has to be encoded using milestone tags (‘<parab>’). Conventionally, paragraphs would constitute the primary hierarchy of a TEI/XML transcription.

With the XML transcriptions referring to Knora regions, it is now possible to merge text and image content automatically by converting Knora ROIs to IIIF URLs. Listing 1 depicts the title text region (green ROI at the very top: “M151-01-TT”) from Figure 1.

```
{
    "lineColor":"#33ff33",
    "lineWidth":2,
    "points":[
        {
            "x": 0.2478,
            "y": 0.2804
        },
        {
           "x": 0.9621,
           "y": 0.3888
        }
    ],
    "type":"rectangle"
}
```

Listing 1. *JSON representation of a Knora ROI*

Each Knora ROI is a rectangular shape defined by the positions of the top left and lower right corners as relative values ranging from 0 to 1. ROIs are rendered in the web-based SALSAH GUI as shown in Figure 1 by drawing on an HTML5 canvas element that is overlaid on the actual image (please note that the upper left corner of the canvas has the coordinates 0, 0). Relative values have the advantage that they are easy to apply to a scaled image.

When these segments have been defined and transcribed, we can associate the transcription of each segment not only with the corresponding manuscript page, but also with a particular section of the image. Given the rectangular shape of a ROI, a IIIF URL can easily be constructed. From the two positions defining the ROI depicted in Listing 1, the region parameter of a IIIF-URL can be created using percent values: `http://sipihost:port/prefix/Med151.jpx/pct:24.78,28.04,71.43,10.84/full/0/default.jpg`. Please note that the IIIF region parameter requires the region’s width and height [1, section 4.1], which can be determined by calculating the difference between the two given points.

The technical specification of a IIIF URL is thus combined with the domain knowledge we have from Knora, to present a region’s transcription aligned with its image section, as shown in Figure 2. In some cases, it may also be helpful to rotate an image section when displaying it to the user. Sometimes Jacob Bernoulli added text in the margin of the page, rotated 90°. When we identify these ROIs in Knora, the rotation parameter can be used in IIIF to make the image’s content easier to read. ROIs are categorized by their content (text, mathematical formulae, and figures). Knora can generate all IIIF URLs for ROIs of a certain category to facilitate image processing (e.g., automatic recognition of figures).

The IIIF Presentation API

A promising way to merge images with additional digital content is the IIIF Presentation API [2]. It makes use of the Shared Canvas Data Model [7], which allows multiple images of the same physical object to be mapped to a common virtual canvas. Thus digital facsimiles of the same physical objects can be aligned. A use case for this model is the medieval manuscript domain, where only fragments of the original physical page may exist. These can be combined to reconstruct the page as a whole. The authors of the Shared Canvas Data Model had also in mind the issue that the facsimiles may come from different repositories [6].

The IIIF Presentation API aims to make digital images available with additional digital content, e.g. the pages of a manuscript with their transcriptions. The specification thus standardizes the presentation of digital image collections with metadata. These could then be displayed by any viewer complying with that standard, e.g. the *Universal Viewer*.[7]

The use case introduced in the previous section could surely benefit from these standardization efforts. However, the IIIF Presentation API seems to focus on image-based resources and adds metadata to them. In the case of the digital edition of the *Meditationes*, both image segments and transcriptions are equally important. The transcriptions shown in Figure 2 are the result of applying XSL transformations to the XML source file shown in Listing 2. XML-based transcriptions are complex objects in themselves, and may need to be processed before the resulting texts can be

[6] We intend to offer an XSL transformation that converts our XML format to TEI/XML that may be used as an interchange format.

[7] `http://universalviewer.io`

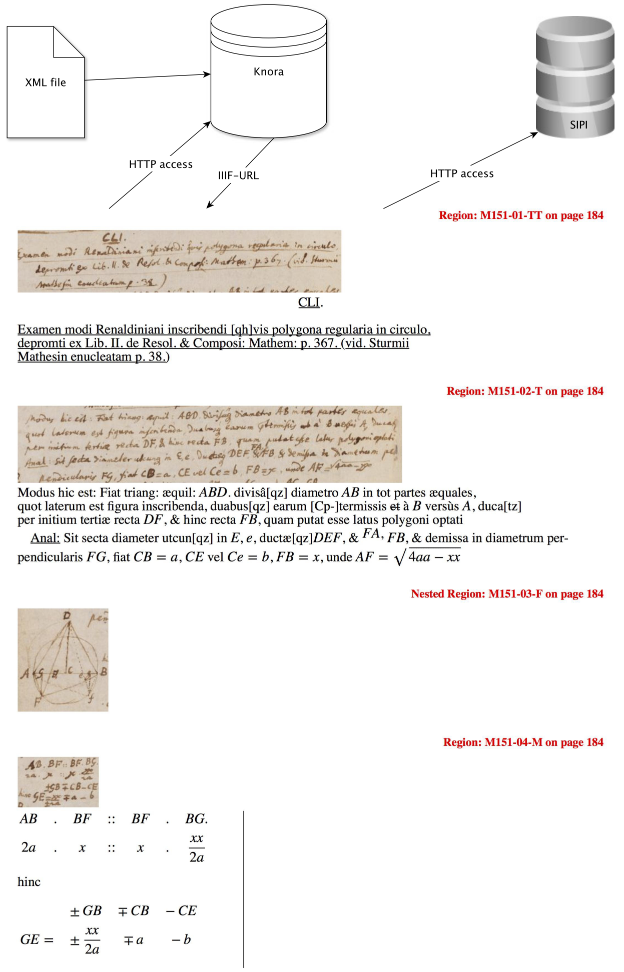

Figure 2. *Meditatio 151: Image Segments and their Transcriptions (partially)*

displayed. The IIIF Presentation API briefly discusses the inclusion of XPath expressions (XPointer) to extract a certain element from an XML source file [2, section 6.1].

For the use case of the *Meditationes*, not only XSL transformations haven to be applied to the XML source file before displaying it on the web as HTML, but also the mathematical formulae written in LaTeX have to be rendered. In order to do so, we rely on MathJax[8], which requires additional configuration options. The XSL transformations are applied using the Saxon XSLT Processor.[9]

Towards a generic API for querying textual Information

Content-based interoperability can be achieved by combining digital representations of objects, such as images and video, with text-based metadata and annotations. We have demonstrated the importance of this kind of functional interoperability using an example involving the IIIF Image API and XML in combination with the Knora/SALSAH VRE. While XML is a well-accepted technology for the transcription of manuscripts in the humanities, and this use has been standardized by the TEI community, the IIIF Image API offers a standardized way to access visual representations. They can be combined to integrate these two categories of sources, and this has already been standardized to some extent by the IIIF Presentation API. However, for textual information, standards for simplified access (without knowledge of the underlying technical implementation) are still lacking. We propose that for this kind of request, a standardized interface (API) should be established by analogy with the IIIF Image API. The diverse nature of such information from images makes the creation of such a standard more difficult.

An API for textual data could allow for the request of specified ranges of a text (indicated by character positions, by analogy to an image region) or elements identified by an ID. Such methods are also discussed in the context of the *Web Annotation Data Model* to refer to parts of external web resources [5, sections 3.2.3 and 4.2.5]. The user should also be able to specify the desired output format. However, the user should not need to know about the underlying storage format, since the proposed API should equally work for various text formats (e.g., XML, Markdown, and LaTeX). Therefore, an initial version of such an API, designed to support the lowest common denominator of various markup languages, could support only simple requests. If the underlying text format is known to be XML, queries involving XPath or even complex XSL transformations could be allowed. We propose that for any given text, a preliminary request could be made, to obtain information about the available request options as well as predefined IDs (e.g., elements with an ID in an XML or Markdown file, or elements with a label in LaTeX).

Conclusion

Our experience with the Knora/SALSAH framework shows that a generic interface standard like proposed above is feasible. We have already implemented a generic platform dealing with the creation, the querying and the manipulation of RDF resources and their properties, including versioning and permissions. Semantic Web technologies such as RDF, RDFS, OWL, and SPARQL allow for complex querying of data and their semantics. But mastering SPARQL in such an environment is difficult. In Knora/SALSAH, we therefore implemented a much simpler interface based on a RESTful API.

We believe that it would be valuable to add an API syntax for querying texts to Knora/SALSAH, complying with the proposed standard that has yet to be elaborated and defined. We have already added basic functionality for representing XML documents in an RDF triplestore, and we could add similar functionality for other markup languages as well. This would allow us to add a text-format-agnostic layer to Knora/SALSAH and still provide text-format-specific querying methods such as XPath. The main advantage of integrating textual information into Knora/SALSAH is the combination of complex marked up texts with other sorts of data, using the basic value types defined in the Knora base ontology. Knora/SALSAH offers ways to represent the semantic connections between resources of different types, and this makes it easier to integrate images and text-based resources. By using ontologies, these semantic connections can be adequately described and formalized.

In the case of the *Meditationes*, Knora/SALSAH could au-

[8]`https://www.mathjax.org`
[9]`http://saxon.sourceforge.net`

tomatically create the required information needed for a viewer complying with the IIIF Presentation API, by connecting images, ROIs, and parts of the transcription. This would involve the IIIF Image API as well as requests for textual information using the proposed standard API. Like the IIIF Image API, this API could also be used by itself, outside the context of the IIIF Presentation API.

References

[1] Michael Appleby, Tom Crane, Robert Sanderson, Jon Stroop, Simeon Warner, IIIF Image API 2.1, `http://iiif.io/api/image/2.1/#iiif-image-api-2-1`.

[2] Michael Appleby, Tom Crane, Robert Sanderson, Jon Stroop, Simeon Warner, IIIF Presentation API 2.1, `http://iiif.io/api/presentation/2.1/`.

[3] Lukas Rosenthaler et al., Final report for the pilot project "Data and Service Center for the Humanities", Tech. rep. Swiss Academy of Humanities and Social Sciences (2015), `http://www.sagw.ch/dms/sagw/laufende_projekte/DaSCH/FinalReport-DaSCH_print`.

[4] Lukas Rosenthaler, Peter Fornaro, The "International Image Interoperability Framework" and its Implication to Preservation, in Proceedings of IS&T Archiving 2016, pages 95–99 (2016).

[5] Robert Sanderson, Paolo Ciccarese, Benjamin Young, Web Annotation Data Model, `https://www.w3.org/TR/2017/REC-annotation-model-20170223/`.

[6] Robert Sanderson, Benjamin Albritton, Rafael Schwemmer, Herbert Van de Sompel, SharedCanvas: a collaborative model for medieval manuscript layout dissemination, in: Proceedings of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries, pages 175–184 (2011), `http://doi.acm.org/10.1145/1998076.1998111`.

[7] Robert Sanderson, Benjamin Albritton, Shared Canvas Data Model, `http://iiif.io/model/shared-canvas/1.0/`.

[8] Tobias Schweizer, Andreas Wassmer, Ivan Subotic, Lukas Rosenthaler, Long-term Access to Research Data as a Challenge to Migration, in: Proceedings of IS&T Archiving 2014, pages 28–33 (2014).

[9] TEI Consortium (eds.), Guidelines for Electronic Text Encoding and Interchange (2016-12-15), `http://www.tei-c.org/P5/`.

Author Biography

Tobias Schweizer is a research associate at the Digital Humanities Lab. Since July 2016, he works on the project Bernoulli-Euler Online that aims at integrating the works of the Bernoullis and Leonhard Euler in one platform. He also takes part in the development of Knora.

Lukas Rosenthaler is the head of the Digital Humanities Lab of the University of Basel. He also leads the Swiss national data center for providing long-term access to digital research data in the humanities, an initiative of the Swiss Academy of Humanities and Social Sciences. He is an expert for image analysis, image processing, data base systems, linked open data, virtual research environments and digital archiving and he is supporting open access initiatives.

Peter Fornaro is deputy head of the Digital Humanities Lab of the University of Basel. He is doing research in the field of digital archiving, imaging, cultural heritage preservation and computational photography. Fornaro is teaching at the University of Basel. Besides research and lecturing he is giving consulting to companies, archives and museums.

```
<?xml version="1.0" encoding="utf-8"?>
<!DOCTYPE meditatio
  SYSTEM "entities.dtd">

<meditatio xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" id="151"
     xsi:noNamespaceSchemaLocation="meditatio.xsd">

<region id="http://data.knora.org/beol/Med151-01-TT" name="M151-01-TT" regiontype="TT" pagenumber="184">
 <hi rend="center"><hi rend="underline">CLI</hi>.</hi><lb/>
 <hi rend="underline">Examen modi Renaldiniani inscribendi &quae;vis polygona regularia in circulo,<lb/>
    depromti ex <expan>Lib<am>.</am><ex>ro</ex></expan> II. de <expan>Resol<am>.</am><ex>utione</ex></expan>
         &amp; <expan>Composi<am>:</am><ex>tione</ex></expan> <expan>Mathem<am>:</am><ex>atica</ex></expan> p.
         367.</hi> (<hi rend="underline"><expan>vid<am>.</am><ex>e</ex></expan> Sturmii <lb/>
    Mathesin enucleatam p. 38.</hi>) <parab/>
</region>

<region id="http://data.knora.org/beol/Med151-02-T" name="M151-02-T" regiontype="T" pagenumber="184">
  Modus hic est: Fiat <expan>triang<am>:</am><ex>ulum</ex></expan>
       <expan>&ae;quil<am>:</am><ex>aterum</ex></expan> <formula notation="tex">ABD</formula>.
       divis&ahat;&que; diametro <formula notation="tex">AB</formula> in tot partes &ae;quales, <lb/>
  quot laterum est figura inscribenda, duabus&que; earum &prae;termissis <del>et</del> &agrav; <formula
       notation="tex">B</formula> vers&ugrav;s <formula notation="tex">A</formula>, duca&tur; <lb/>
  per initium terti&ae; recta <formula notation="tex">DF</formula>, &amp; hinc recta <formula
       notation="tex">FB</formula>, quam putat esse latus polygoni optati<reg>.</reg> <parab/>
  <region id="http://data.knora.org/Med151-03-F" name="M151-03-F" regiontype="F" pagenumber="184">
    <!-- Figure 151-03 as SVG here -->
  </region>
  <hi rend="underline"><expan>Anal<am>:</am><ex>ysis</ex></expan></hi><reg>:</reg> Sit secta diameter
       utcun&que; in <formula notation="tex">E</formula>, <formula notation="tex">e</formula>,
       duct&ae;&que;<formula notation="tex">DEF</formula>, &amp;
  <add><formula notation="tex">FA</formula>,</add> <formula notation="tex">FB</formula>, &amp; demissa in
       diametrum per&hyphen;pendicularis <formula notation="tex">FG</formula>, fiat <formula
       notation="tex">CB=a</formula>, <formula notation="tex">CE</formula> vel <formula
       notation="tex">Ce=b</formula>, <formula notation="tex">FB=x</formula>, unde <formula
       notation="tex">AF=\sqrt{\mathstrut}\,\overline{4aa-xx}</formula> <parab/>
</region>

<region id="http://data.knora.org/Med151-04-M" name="M151-04-M" regiontype="M" pagenumber="184"
     border="borderright">
  <formula notation="tex">
    \begin{array}{ccccccc}
      AB &amp; . &amp; BF &amp; :: &amp; BF &amp; . &amp; BG. \\
      2a &amp; . &amp; x &amp; :: &amp; x &amp; . &amp; \dfrac{\mathstrut xx}{2a}
    \end{array}
    \qquad
  </formula> <lb/>
  hinc
  <formula notation="tex">
    \begin{array}{lccc}
      &amp; \pm \,GB &amp; \mp \,CB &amp; -\,CE \\
      GE = &amp; \pm \,\dfrac{\mathstrut xx}{2a} &amp; \mp \,a &amp; -\,b
     \end{array}
 </formula> <parab/>
</region>

</meditatio>
```

Listing 2. *XML transcription of Meditatio 151 (partially)*