



**HAL**  
open science

# Connecting online early music libraries and musicological resources: Experiments in ergonomics in the **Biblissima+** framework

David Fiala, Kévin Roger

## ► To cite this version:

David Fiala, Kévin Roger. Connecting online early music libraries and musicological resources: Experiments in ergonomics in the Biblissima+ framework. DLFM 2023: 10th International Conference on Digital Libraries for Musicology, Nov 2023, Milan (Italie), France. pp.128-131, 10.1145/3625135.3625140 . hal-04500077

**HAL Id: hal-04500077**

**<https://hal.science/hal-04500077v1>**

Submitted on 15 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Connecting online early music libraries and musicological resources\*

Experiments in ergonomics in the Biblissima+ framework

David Fiala<sup>†</sup>  
CESR/Biblissima+  
University of Tours  
Tours, France  
david.fiala@univ-tours.fr

Kévin Roger<sup>†</sup>  
CESR/Biblissima+  
University of Tours  
Tours, France  
kevin.roger@univ-tours.fr

## ABSTRACT

The French research infrastructure for written heritage Biblissima started in 2012. It now includes in its second phase, from 2021 onwards, a subgroup focused on musical written heritage, which aims at enhancing the use and accessibility of music data in Biblissima's ecosystem and, more specifically, in its updated central web portal. This paper presents first results of this research group. It offers an overview of available online complementary resources pertaining to pre-1600 music and music books, an assessment of their technological and scientific possible interactions, and first experiments on ways of merging them into interconnected user interfaces based on IIIF. An alignment algorithm has been developed for combining data on musical works and manuscripts (available in DIAMM) and online digitized manuscripts (of the BnF). Methods for collecting more detailed metadata on music manuscripts' contents through the use of an AI classifier are then discussed.

## CCS CONCEPTS

• Applied computing → Arts and humanities → Sound and music computing • Computing methodologies → Machine learning → Learning settings • Computing methodologies → Computer graphics → Graphics systems and interfaces • Human-centered computing → Human computer interaction (HCI) → HCI design and evaluation methods •

## KEYWORDS

early music, IIIF, datasets, metadata alignment, augmented visualization, digital annotation, manuscript, printed document

\* Connecting online early music libraries and musicological resources.

<sup>†</sup>Both authors contributed equally to this research.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2023 Copyright held by the owner/author(s).

## ACM Reference format:

David Fiala, Kévin Roger. 2023. Connecting online early music libraries and musicological resources: Experiments in ergonomics in the Biblissima+ framework. In *10<sup>th</sup> International Conference of Digital Libraries for Musicology, Milan, Italy*, <https://doi.org/10.1145/3625135.3625140>

## 1 INTRODUCTION

Biblissima+<sup>1</sup> is a multi-site digital infrastructure for research and service dedicated to the history of the transmission of ancient texts, from Antiquity to the Early modern period, in the West and in the East. Funded by the French Ministry of Research and Higher education (2021-2029), it aims to bring together digital resources and tools produced by 17 French academic research units, and to make them available through a central web portal, according to a common DMP[1].

Biblissima+ is organized in 7 areas of expertise or “clusters”. These expert groups bring together researchers, curators and engineers around specific objects or technologies. Its cluster 6<sup>2</sup>, coordinated by David Fiala (RicercaLab / CESR), is devoted to the digital scientific mediation of music books or scores, focusing in a first stage on pre-1600 polyphonic music [2]. It welcomes musicologists and information specialists working at enriching the use and accessibility of early music data. Its objective is to have music metadata and data properly included in Biblissima's ecosystem. More specifically, it enhances the quality and richness of musical metadata and provides multimedia solutions for the accessibility to early music digital data, both for eyes and ears.

## 2 SPECIFICITIES OF EARLY MUSIC BOOKS

The medieval and early modern periods are the historical core of the whole Biblissima program. The specificities of music heritage bring other explanations for Cluster 6's focus on this period: the musical language of this long period, modal counterpoint, forms a historically coherent continuity from the 10<sup>th</sup> to the 16<sup>th</sup> c. This way of creating and performing music resulted in a broadly coherent corpus of music books, namely in terms of page layout. The main

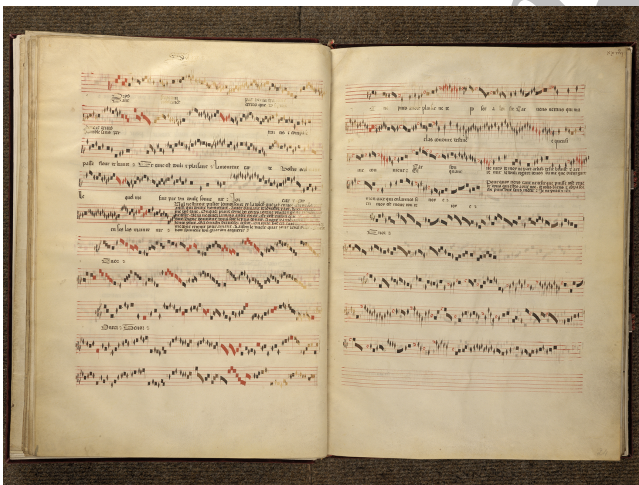
<sup>1</sup> <https://projet.biblissima.fr/en>

<sup>2</sup> <https://projet.biblissima.fr/fr/projet/clusters-biblissima/cluster-6>

characteristics of this musical language and these music books form a sharp contrast to what came next, which can be summarized in the invention of the musical score format, in which all lines of the polyphonic texture are graphically precisely aligned together one above the other, according to a common rhythmic measure (resulting, for example, in the apparition of vertical barlines).

The main challenges facing any digital treatment of pre-1600 music books arise from two main characteristics of their layout:

1. The page layout of pre-1600 music books presents the lines of the polyphonic texture one after the other. Though different types of books (choirbooks, partbooks, songbooks or chansonniers) corresponded to varied performing situations, all of them contain a succession of notated melodies, the appropriate grouping of which produced correct polyphonic performances. The way these notated melodies had to be grouped together in performance to produce the expected polyphonic combination was made more or less clear by various graphic clues, obvious to experimented performers but much less so to average readers, not to mention computers.
2. The medieval and early modern music corpus consists of short pieces (between 1 and 10 minutes in performance), transmitted in numbers (from 10 to hundreds) in physical anthologies.



**Figure 1: An opening (double-page of the open book) of a late 14<sup>th</sup> c. music manuscript containing two musical works. Chantilly, bibliothèque du musée Condé, ms. 564, f. 23v-24**

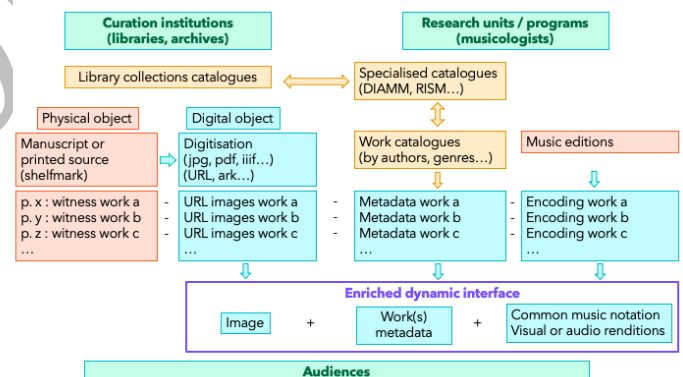
Music books produced before 1600 are collections of independent works, themselves constituted of the combination of several lines or voices (from 2 to 5 in most cases but up to 40 in exceptional ones), copied one after another. Fig. 1 shows two works for 3 voices copied each on one page. On the left page of this example, lines 1-

5 contain the upper voice, with text to be sung under the music notation, lines 6-7 a lower voice labelled *Tenor*, and lines 8-10 another lower voice labelled *Contratenor*.

In other words, early music books present a continuous succession of melodies which, once correctly connected together, produce a continuous collection of polyphonic works. Music librarians are experts in dealing with music books. Musicologists' main expertise focuses on the works transmitted by these books. A main challenge of early music study is to make sensible and significant this connection between books and works, which tend to live in separate spheres.

### 3 THE DIGITAL EARLY MUSIC LANDSCAPE

The cluster 6 general goal is to promote this repertoire through digital and augmented visualizations of the sources via the IIIF protocol, central to the Biblissima+ ecosystem, enriched by musicological resources<sup>3</sup>. It especially focuses on improving online visualization services as well as on creating new tools to fill certain gaps in musicological data. Indeed, in many cases, the contents of music manuscripts and prints have already been fully inventoried, analyzed and sometimes made available in various types of modern editions. This metadata, obtained after considerable human effort, is, however, most often stored in databases completely independent of the visualization services provided by digital libraries.



**Figure 2: Interoperability-based connection project between musicological and codicological resources**

In addition to connecting musicological resources, the cluster aims to improve the inventory and analysis of early musical sources by developing an artificial intelligence model trained to recognize significant features of mensural notations from the 13<sup>th</sup> to the 16<sup>th</sup> c. The aim is to generate the metadata of book contents to the level of works and even voices when they are not available, or, when available, to refine and align them with the online images of music sources. These metadata can then automatically be fed into IIIF annotation files, providing the musicological community with a

<sup>3</sup> International Image Interoperability Framework™: <https://iiif.io/>. Biblissima contributes actively since its beginnings to the development of this set of standards for the interoperability of digital images: <https://iiif.biblissima.fr/>.

dense, high-quality augmented visualization of early musical manuscripts and prints. In the end, this connection can provide the link between the image of the historical source and any modern musical edition or encoding, or indeed any corresponding audio file.

Together, these projects aim to build as many interoperable gateways as possible in the landscape of online musicological, codicological and historical resources. Fig. 2 proposes a general overview of the landscape of early music objects, both physical (in pink, namely books) and digital (in blue). Bibliographic and scientific catalogues (in yellow) provide the necessary links between library collections and musicological resources<sup>4</sup>. Biblissima's cluster 6 aims at designing technological means to merge these data and metadata into a user-friendly interface.

## 4 CONNECTING MUSIC RESOURCES

For many years now, Oxford University's Digital Image Archive of Medieval Music (DIAMM)<sup>5</sup> has hosted the metadata and some digitizations of numerous early music sources, apart from chant and prints. Little by little, this database has aggregated a colossal quantity of data, to the point of counting over 4,000 inventoried sources covering almost all of the medieval music heritage. However, for reasons of rights or even limitations imposed by partnerships, it is currently not possible to consult all sources' images and their metadata at the same time, in the same interface.

Cluster 6 proposes an innovative solution to this problem by establishing a connection between DIAMM and online music scores. To achieve this, the project relies on the power of the IIIF (International Image Interoperability Framework) protocol, which is highly valued by Biblissima+ for its technological capabilities [3]. The main aim of this comprehensive set of standards is to improve the interoperability of image distribution on the World Wide Web. The IIIF protocol is primarily based on an image API, which involves publishing images on dedicated IIIF servers. This approach enables image manipulation and retrieval simply by using the corresponding URL. In a further step, these images are associated with an IIIF manifest, which can be described as a JSON (JavaScript Object Notation) file. A wide range of metadata relating to a single image, or a sequence of images can be seamlessly integrated in it.

To achieve this, DIAMM provides an API for querying data in JSON. This approach produces a structured file containing the inventory of a music source. However, given the complexity and unique structure of DIAMM data, it needs to be handled with care. A Python program (DIAMMtoIIIF)<sup>6</sup> had to be developed to automate this extraction [4]. To develop DIAMMtoIIIF, Cluster 6 focused on the resources available at both DIAMM and the Bibliothèque Nationale de France (BnF). Using BnF ARK identifiers, the program regenerates IIIF URIs and integrates them into a IIIF manifest created using a Python library (IIIF-Prezi). In

practice, the most difficult aspect is not the retrieval of the DIAMM information, but the fine alignment of the correct folio or pagination with the IIIF-URI. Indeed, the discrete unit of a IIIF manifest is the digital image identifier, whereas music inventories refer to folio numbers. We designed an algorithm that aligns images and work metadata according to the one or more folio numbers corresponding to each work.

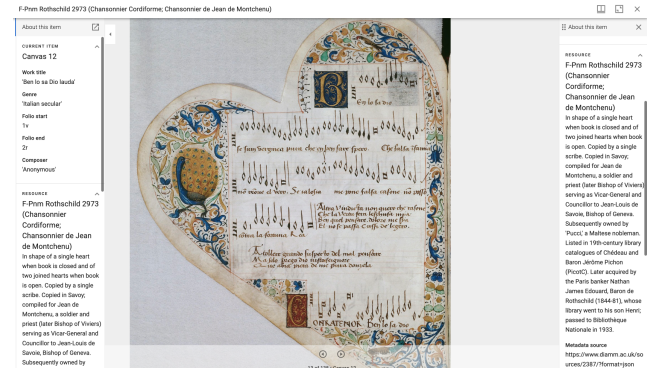


Figure 3: IIIF manifest of the *Chansonier Cordiforme* (BnF, Rothschild 2973) generated with DIAMMtoIIIF – Mirador viewer

The project's perspectives can be extended to other musicological resources, opening new possibilities in the field. One potential application is the integration of the project's methodology with other renowned music libraries and music datasets.

Yet, this experiment also revealed some problems that may in the end prove insurmountable. The quality of the page metadata is of course completely dependent both on the scientific and technological state of the database from which it is extracted. Even in the case of DIAMM, it appeared that some fields provided by the API could not be exploited with sufficient stability to hope for a systematic integration in the IIIF manifest without creating severe disturbance. Such a situation was due either to inconsistent data management in the database itself or to impossible alignment of some information. The use of an API directly extracting data from an external database thus needs to be limited to fields providing perfectly consistent information.

Another limitation of the DIAMMtoIIIF model concerns its potential application to other musicological resources (mostly those mentioned as work catalogues in fig. 2) from which could be extracted information unavailable through DIAMM. Not only very few of these resources provide APIs but some of them do not even keep the necessary data to build a connection to the sources' images. This is for example the case of the Catalogue de la Chanson française maintained by the RicercarLab<sup>7</sup> which includes disambiguated information on almost 10,000 French Chansons known in the 16<sup>th</sup> c. but no systematic record of all sources of each

<sup>4</sup> In this Fig. 2, on DIAMM, see below. RISM is an even older and much wider endeavor of the musicological community. Its digital transformation does not yet allow such experiments as the one described here with DIAMM. <https://rism.info/>

<sup>5</sup> <https://www.diamm.ac.uk/>

<sup>6</sup> <https://github.com/Biblissimacluster6/DIAMMtoIIIF/tree/main>

<sup>7</sup> <http://ricercar-old.cesr.univ-tours.fr/3-programmes/basechanson/index.htm>

chanson, not mentioning the shelfmarks or page numbers of their book copies. Collecting and recording such information would be unreasonably time-consuming in regard of its only real interest: aligning images of digital libraries with the database records. Such information is scientifically uninteresting but necessary for of alignments allowing the kind of enriched interfaces the Cluster 6 is aiming at. Other methods for automatically collecting these data should then be considered.

## 5 AI MODEL APPLIED TO EARLY MUSIC SOURCES

As the interoperability of pre-existing digital resources is in fact difficult to generalize and develop, AI seems to open up more possibilities and flexibility for collecting metadata. This is why cluster 6 has decided to develop an artificial intelligence model for the recognition of early musical sources (manuscripts and printed documents). The aim is, firstly, to automatically collect a large number of metadata that is difficult to recover from conventional inventories (for example, the nomenclature of voices and musical parts, or the style of notation used). Secondly, the data collected would be automatically integrated into IIF manifests to enrich the metadata, annotation files and layers used.

This approach differs from that of OMR systems dedicated to early music, such as Aruspix [5, 6] or the tools developed as part of Cantus Ultimus (SIMSSA) [7], for two main reasons, both technical and conceptual. First, the level of granularity required is different. It is not a matter of collecting information at the level of a musical character or other discrete values and converting it into a manipulable format (MEI, MusicXML). On the contrary, we capture the musical document at the natural, continuous level of page layout, voice distribution, and notation linearity. The primary goal remains the collection of metadata to improve the exploration and inventory of early sources. Second, the use of AI models for prediction and recognition has been little used in the field of early music. While some recent OMRs are based on AI (i.e. Poliphonia) [8, 9, 10], the possibilities offered by our project would make it possible to generate semantic, and less symbolic, data concerning the production of documents: type of notation, handwriting or layout similarities between corpora, chronological grouping, etc.

In collaboration with the BnF Data Lab<sup>8</sup>, Cluster 6 intends to assemble a substantial dataset integrating musical sources from the 13<sup>th</sup> to the 16<sup>th</sup> c. Some deep learning strategies have already been defined and executed in experimental versions of the model. The first datasets have been manually tagged using the image annotation tool LabelIMG. The first tasks trained were, among others, those corresponding to the tagging of musical parts and the distribution of staves (Fig. 4). Based on these data, a first statistical analysis is performed to evaluate the potential nature of the voices – *cantus*, *tenor*, *contratenor*, etc. (Fig. 5). The model learning was developed in Python, using the ImageAI library<sup>9</sup> and the YOLO recognition algorithm.

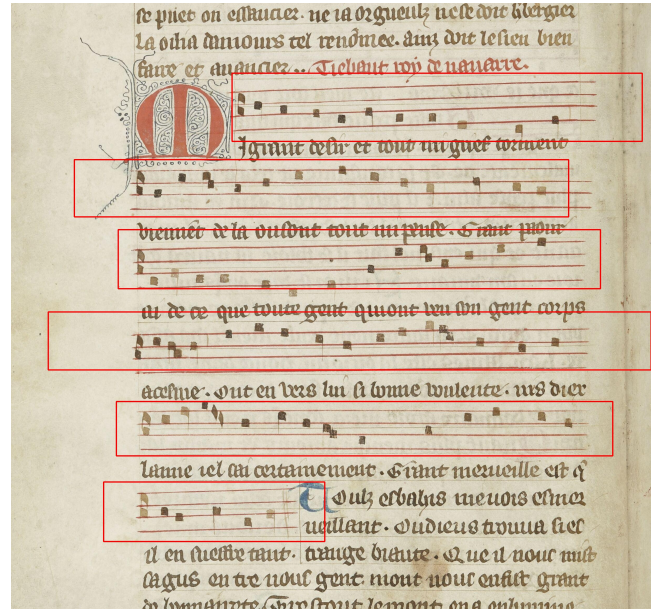


Figure 4: Automated recognition of staves (without subsequent treatment) in a 14<sup>th</sup> c. manuscript (BnF, fr. 1591)

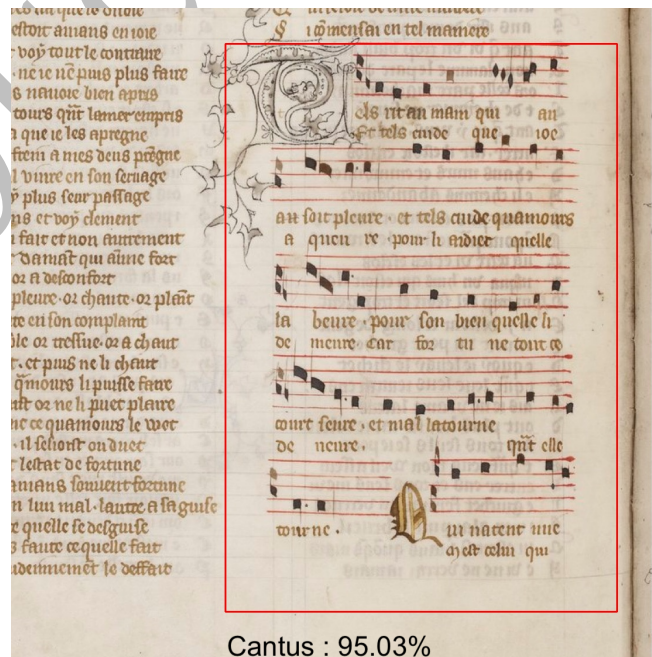


Figure 5: Automated recognition of voice nature (*cantus*) in a 14<sup>th</sup> c. manuscript (BnF, fr. 1584)

YOLO is an AI object detection model that efficiently detects and localizes features in images. It divides the image into a grid and makes predictions for each grid cell, including bounding box coordinates, object class probabilities, and confidence scores. Once the predictions are made, YOLO performs a post-processing to

<sup>8</sup> <https://www.bnf.fr/fr/bnf-datalab>

<sup>9</sup> <https://github.com/OlafenwaMoses/ImageAI>

filter out low-reliability detections and combine overlapping bounding boxes. At the moment, the mAP (mean Average Precision) terminal of our different models varies between 55% and 60%, for limited datasets (of the order of a few hundred images), which is quite encouraging.

However, initial testing has allowed us to re-evaluate our deep learning methodology. While our models are gradually assimilating the tagging of music notation and staves, voice recognition requires more detailed and massive learning of different layouts. In addition to the purely statistical approach to the different musical configurations, it is possible to improve the success rate by using other data. The cluster is currently evaluating the success of a character recognition model, such as Kraken<sup>10</sup>, on the corpus for extracting voice names when they are recorded in the sources.

In a second phase, we hope to train more advanced AI models for the recognition of different ancient musical notations and chronological grouping. Such models could lead to a set of tools surrounding a visual interface designed for musicologists, archivists and librarians, similar to what offers *e-scriptorium* for text annotation<sup>11</sup>.

## 6 CONCLUSION

Digital music libraries provide direct and immediate access to fundamental documents for the history of music. But they generally only accompany them with basic and general bibliographical information, limited to the physical object of the book. On the other hand, musicological resources collected huge amounts of detailed information on the works contained in these books. This paper proposed two generic approaches for filling the gap which separates these two digital worlds. It is probably by using and adapting a combination of these two approaches that the gap might finally be filled.

## ACKNOWLEDGMENTS

This paper was written with the support of the Biblissima+ research infrastructure. It benefitted of the many exchanges and discussions which occurred during the first meeting of Cluster 6 in Tours in 2022 [2]. All participants are warmly acknowledged. Biblissima+ benefits from State funding managed by the ANR within the "Programme d'investissements d'avenir" integrated in France 2030, reference ANR-21-ESRE-0005.

## REFERENCES

- [1] Emmanuelle Morlock, Régis Robineau, Eduard Frunzeanu, & Kevin Bois. (2022). *Plan de gestion des données de l'observatoire des cultures écrites Biblissima+ (1.08)*. <https://doi.org/10.5281/zenodo.7081674>
- [2] Kévin Roger, & David Fiala. (2023, March 13). *Compte rendu des journées inaugurales du cluster 6 Biblissima+*. <https://doi.org/10.5281/zenodo.7727961>
- [3] Sajdak, Cécile, & Daugas, Mathilde. (2023, June 30). *IIIF pour la mise en ligne des archives numérisées - Formation*. <https://doi.org/10.5281/zenodo.8099431>
- [4] Kévin Roger. (2023, June 19). *Computer vision applied to musical documents - The Biblissima+ cluster 6: "The Challenges of Musical Heritage"*. <https://doi.org/10.5281/zenodo.8055626>

- [5] Laurent Pugin and Tim Crawford, 'Evaluating OMR on the Early Music Online Collection'. In *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR 2013)*: 439–44.
- [6] Laurent Pugin. 'Editing Renaissance Music: The Aruspix Project'. In *Digitale Edition zwischen Experiment und Standardisierung: Musik – Text – Codierung*, ed. Peter Stadler and Joachim Veit, Berlin, New York: De Gruyter, 2009, pp. 147-156. <https://doi.org/10.1515/9783110231144.147>
- [7] See Kate Helsen, Jennifer Bain, Ichiro Fujinaga, Andrew Hankinson, and Debra Lacoste. 'Optical Music Recognition and Manuscript Chant Sources'. In *Early Music* 42 (2014): 555–58.
- [8] Jorge Calvo-Zaragoza, Alejandro H Toselli and Enrique Vidal. 'Handwritten music recognition for mensural notation with convolutional recurrent neural networks'. In *Patterns Recognition Letters* 128, 115–21.
- [9] Arnau Baró, Pau Riba and Alicia Fornés. (2022). Musigraph: Optical Music Recognition Through Object Detection and Graph Neural Network. In *Frontiers in Handwriting Recognition*, 171–184.
- [10] Christoph Wick. 'Optical Medieval Music Recognition'. (2020) PhD thesis, University of Würzburg.

<sup>10</sup> <https://kraken.re/main/index.html>

<sup>11</sup> <https://escriptorium.paris.inria.fr/>