1 Introduction

The Virtual Manuscript Room Collaborative Research Environment (VMR CRE)\(^1\) brings community and a toolbox of powerful research components to support all stages of research and production of a digital edition. Beginning with the popular open-source portal, Liferay, the VMR CRE integrates 30 DH components to naturally support: Cataloging witnesses; Managing and displaying images; producing well-formed TEI transcriptions using a web-based WYSIWYM editor and storing those transcriptions to a versioned transcription repository; community volunteer task assignment management; automatic realtime collation of witnesses; regularization and apparatus editing; online publishing of the final results-- as a traditional apparatus, or with interactive tools which let users choose different ways to visualize the data produced in the edition. Originally developed to facilitate the globally disparate teams editing the approximately 5700 Greek New Testament manuscript witnesses for the Editio Critica Maior\(^2\), the VMR CRE has been adopted by a variety of other projects, including The Digital Edition and Translation of the Coptic-Sahidic Old Testament\(^3\) The Canons of Apa Johannes the Archimandrite\(^4\), and The Avestan Digital Archive\(^5\). Each of these adaptations has presented unique challenges, with the result that the software, which was originally highly focused on New Testament manuscript research, has evolved into a useful tool for general digital edition research and electronic publishing.

The VMR CRE is an ecosystem of components, enabling projects to choose only the tools pertinent to their research. Many teams have contributed to the ecosystem, including the European Union’s COST initiative, Interedition\(^6\), providing the CollateX collation engine\(^7\) and initial work on variant graph display; University of Trier, providing a web-based WYSIWYM TEI editor for manuscript transcription\(^8\); and the Institute for Textual Scholarship and Electronic Editing (ITSEE) at the University of Birmingham, contributing the Apparatus Editor\(^9\).

---

1. http://vmrcre.org
2. http://ntvmr.uni-muenster.de
4. coptot.manuscriptroom.com/web/apa-johannes
A walk through the workflow at the Institut für neutestamentliche Textforschung (INTF), in their efforts to edit the Editio Critica Maior (ECM\textsuperscript{10}), provides opportunity to touch on many components available in the VMR CRE. Work can be dividing in 9 discrete stages, progressively: 1. witness cataloging; 2. witness selection; 3. image management; 4. indexing of folio content; 5. transcribing; 6. regularizing; 7. collating; 8. editing an apparatus; 9. genealogical analysis of the witness corpus.

2 Metadata and Feature Tagging

The VMR CRE stores with each manuscript a very limited set of descriptive data and reserves the primary metadata capture for a dynamic tagging facility called Feature Tagging. A Feature is any defined metadata information which might be captured for a manuscript or manuscript page. For example, an alternative catalog identifier, an external image repository, the canvas material, the ink type, the script type; these are all Features which might be captured for a manuscript; For individual pages, an illumination, a canon table, or even individual sample script characters might be tagged at Features. These Features must first be defined in the system, and the VMR CRE comes by default with a predefined set of Feature Definitions used at the INTF. A Feature Definition can specify that zero or more values should be captured with the Feature tag and what those values and value types should be. Once a Feature is defined, it can be used to tag manuscripts or manuscript pages, capturing individual Feature values for each tag, if necessary.

Every Feature Definition adds to the number of facets available in the catalog search facility. For example, one might search for all manuscript folio sides from Egypt which include Illuminations and any part of the Gospel of John. A Feature Tag can also include a region box, marking the area on a folio image where the Feature is present. If a region box is captured, a search query can specify to show the region box clips in the result. For example, one might choose to capture a set of representative letters for each manuscript and then perform a search for all double column manuscripts with a height of at least 20cm between the II and V centuries, and to ask the query results to show the representative α (alpha) clips.

3 Transcription and Reconciliation

Transcription work in the VMR CRE is done using a What You See Is What You Mean (WYSIWYM) web-based editor originally developed by the University of Trier in collaboration with the INTF and ITSEE in Birmingham. This transcription editor has been developed as a plugin for the popular TinyMCE\textsuperscript{11} HTML editor component. The editor includes menus and dialogs to assist the researcher with composing a transcription, without asking the transcriber to learn special markup codes. The content may then be obtained as EpiDoc\textsuperscript{12} influenced TEI.

The VMR CRE saves content in a versioned transcription repository backed by Git\textsuperscript{13}.

\textsuperscript{10}http://egora.uni-muenster.de/intf/projekte/ecm_en.shtml
\textsuperscript{11}https://www.tinymce.com
\textsuperscript{12}http://www.stoa.org/epidoc/gl/latest
\textsuperscript{13}https://git-scm.com
A user may have access to create and edit their own personal transcription, a project-wide transcription, or a site-wide (= published) transcription—each having version history.

The VMR CRE also includes a palaeography tool to assist a transcriber when encountering rare symbols, abbreviations, or ligatures. If a portion of the unknown text can be identified, the researcher can enter one or more letters and will be presented with images of text instances elsewhere which including these letters, offering possibilities. As more and more rare text items are tagged, the system grows more helpful.

Quality assurance for the ECM requires that a transcription for a manuscript be produced independently by two transcribers. The products are then compared to each other and differences are reviewed by a manager and reconciled to produce a final transcription. The VMR CRE provides tools to facilitate this reconciliation work.

4 Collation and Visualization

Collation is a key component to find differences in text witnesses when producing a critical edition. Collation facilities in the VMR CRE are performed by CollateX. Collation and regularization of uninteresting differences is an iterative cycle in digital editing and the VMR CRE ties these two actions together with an intuitive visual interface. Visualization of a collation, either during the editing process or for the reader, can be rendered as a variant graph, an alignment table, or as a traditional negative apparatus.

5 Web Services, Open Programmatic Access

The VMR Web Services API layer is primarily useful for exposing the functionality of the VMR to other research projects wishing to access the functionality or contribute to the dataset through their own systems and tools. The VMR CRE Web Services API generally uses noun/verb nomenclature organized by category. This means that the last 2 segments of an API URL will consist first of the type of object the call will affect, and second of the action to be performed on the object. Any path before the final two segments are merely for organizational purposes. This is different from a strict REST convention which confines the action to one of 6 HTTP verbs. The VMR CRE places no semantic meaning on the HTTP verb. Both GET and POST HTTP verbs are accepted as identical, relegating the verb, or semantic action to the final segment of the URL. This allows easy testing and examples for every action directly within a web browser. Parameters to a service request are passed as standard HTTP FORM POST parameters or as query string parameters. As an example, an API request to obtain a transcription of manuscript 30093 of the Apocalypse by Darius Mueller in TEI form with no stylesheet applied would be:

http://ntvmr.uni-muenster.de/community/vmr/api/transcript/get/?docID=30093&indexContent=Rev&fullPage=true&userName=darius.mueller&format=teiraw

Notice the final noun/verb segments of the URL: transcript/get
References

