

repoVizz: a Framework for Remote Storage, Browsing, Annotation, and Exchange of Multi-modal Data

Oscar Mayor, Quim Llimona, Marco Marchini, Panos Papiotis, Esteban Maestre
Music Technology Group, Universitat Pompeu Fabra
Barcelona, Spain
{oscar.mayor, marco.marchini, panos.papiotis, esteban.maestre}@upf.edu
quim.llimona01@estudiant.upf.edu

ABSTRACT

In this technical demo we present repoVizz¹, an integrated online system capable of structural formatting and remote storage, browsing, exchange, annotation, and visualization of synchronous multi-modal, time-aligned data. Motivated by a growing need for data-driven collaborative research, repoVizz aims to resolve commonly encountered difficulties in sharing or browsing large collections of multi-modal data. At its current state, repoVizz is designed to hold time-aligned streams of heterogeneous data: audio, video, motion capture, physiological signals, extracted descriptors, annotations, et cetera. Most popular formats for audio and video are supported, while Broadcast WAVE or CSV formats are adopted for streams other than audio or video (e.g., motion capture or physiological signals). The data itself is structured via customized XML files, allowing the user to (re-)organize multi-modal data in any hierarchical manner, as the XML structure only holds metadata and pointers to data files. Datasets are stored in an online database, allowing the user to interact with the data remotely through a powerful HTML5 visual interface accessible from any standard web browser; this feature can be considered a key aspect of repoVizz since data can be explored, annotated, or visualized from any location or device. Data exchange and upload/download is made easy and secure via a number of data conversion tools and a user/permission management system.

Categories and Subject Descriptors

E.5 [Files]: Organization - structure; E.2 [Data Storage Representations]

Keywords

repository, exchange, remote, HTML5, multimodal

¹<http://repovizz.upf.edu/>

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). Copyright is held by the author/owner(s).

MM'13, October 21–25, 2013, Barcelona, Spain.

ACM 978-1-4503-2404-5/13/10.

<http://dx.doi.org/10.1145/2502081.2502247>

1. INTRODUCTION

Multi-modal data stream acquisition and analysis constitute a major pursuit in much of today's data-driven scientific research. The lack of common formats and structured storage tools for heterogeneous data forces researchers to use ad-hoc formatting schemes and software tools developed *in house*. Despite the obvious potential offered by collaborative, exchange-driven research, difficulties often found in sharing or browsing data sets are impeding research collaboration and cross-fertilization. As a response to the needs imposed by the challenging nature of international research collaborations, including the EU FET project SIEMPRE grounded on multi-modal data capture, analysis and exchange, we aimed at (i) defining the structure of data coming from multi-modal recordings, (ii) creating a repository to store and share multi-modal datasets collected in research experiments dealing with music performance, and (iii) implementing a tool to visualize acquired data and make annotations in a user-friendly and flexible manner. The goal was to open research results and share data with the research community. This drove us to develop a flexible framework for data formatting, storage, and browsing which allowed the integration of any kind of data or annotations to come, and facilitate access through a dedicated portal.

2. OVERVIEW OF REPOVIZZ

The repoVizz framework[2][1], which is conceptualized over a classical client-server configuration, comprises of (i) a formatting scheme that enables flexible multi-modal data structuring and aggregation, (ii) a number of tools that allow the user to convert data to/from formats in agreement with the repoVizz convention, (iii) a persistence server used as a (remote-access) repository to host and exchange data, and (iv) a HTML5-based (client) interface through which it is possible to browse, visualize, annotate, download, etc. (multi-modal) data from any location by using a web browser.

Although not mutually exclusive, two main types of users can be differentiated. A first type of user, which could be called *Producer*, gathers multi-modal data and pushes it into repoVizz. A second type, the *Consumer*, accesses repoVizz in order to visually explore, retrieve (download), annotate, and/or process existing data. The second type of user would also include those who, after downloading a dataset (or a portion/segment of a dataset), apply some (user) analysis/processing and upload their results as additional streams or annotations (thus editing the dataset). The data lifecycle in repoVizz, which mixes its use from the two aforementioned user types, is illustrated in Figure 1.

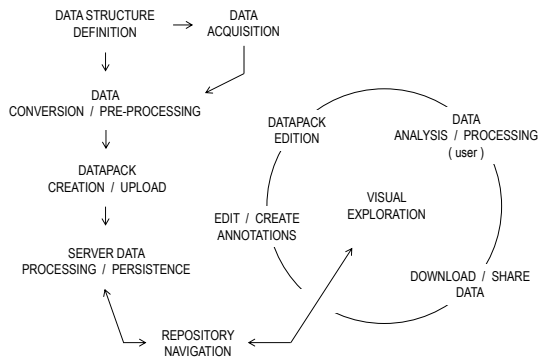


Figure 1: Overview of the data lifecycle in repoVizz.

2.1 Data formatting and structuring

The first and most important step for pushing data into repoVizz comes when defining the structure of the multimodal dataset. In repoVizz, data structuring is carried out via a customizable XML file, called *repoVizz Struct*, that serves as a flexible, tree-like skeleton for organizing the contents of the dataset. The nodes of a *repoVizz Struct* enable hierarchical organization of data streams and annotations, and are used to hold text descriptions and pointers to data files. repoVizz provides, via its portal, the *repoVizz Designer*, a java application that allows to create and edit *repoVizz Structs* in a visual manner, so the user won't need to manually edit an XML file, but rather enter the appropriate contents into fields defining category tags, data types, file names, etc. Altogether, a *Struct* and all pointed data files constitute a *repoVizz Datapack*, which is uploaded by the user as a .zip file (automatically generated by the *repoVizz Designer*) via the *repoVizz* portal.

2.2 The repoVizz Server

Once a *repoVizz Datapack* has been uploaded, some pre-processing and/or feature extraction is carried out in the server, before files are stored and indexed for subsequent access. The *Producer* can grant access to other (or all) users to access/visualize, annotate, or even edit and download the data. From that moment, *Consumers* can access the portal and navigate through the existing *Datapacks*, visualize signals, provide annotations, add signals, etc. by means of a powerful, flexible HTML5 web interface that allows interacting with the data without the need to download.

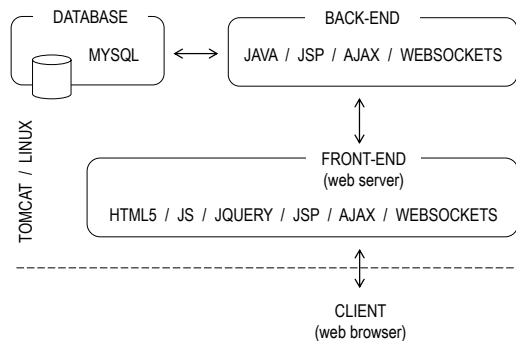


Figure 2: Schematic illustration of the *repoVizz* server structure.

2.3 The repoVizz Visualizer

The web interface has been built following a multi-track configuration where signals of different natures and sample rates can be synchronously visualized, played back, re-arranged, etc. Moreover, users may select a data segment directly from the multi-track interface and download a *Datapack* containing all synchronized data streams and annotations within the chosen time range. The portal also provides tools to convert downloaded data to popular formats such as Matlab structs.

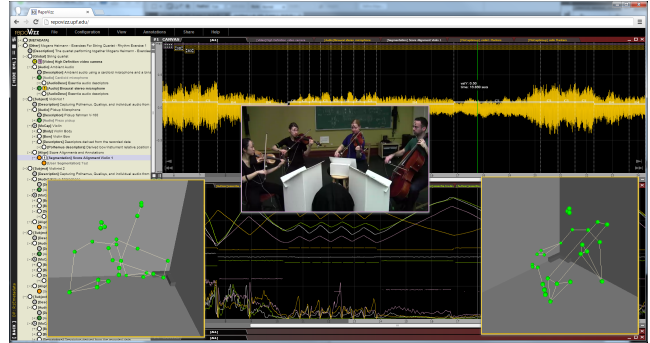


Figure 3: Screenshot of the *repoVizz* Visualizer, featuring some signals from computed descriptors, an audio file, annotations, video and motion capture.

In Figure 3 the *repoVizz* visualizer shows an example of how a multimodal recording, in this case of a string quartet, can be visualized within a web browser, audio and video streams can be played in sync, motion capture data can be visualized using the 3D visualization widget and features extracted from audio and instrumental gestures can be visualized simultaneously. More details about the visualizer can be found in the demo video[1].

3. ACKNOWLEDGMENTS

This work was partially supported by the EU FET-Open Project SIEMPRE (FP7-ICT-2009-C-250026) and the Catalan Government through an AGAUR Fellowship and TECNIO grant (TECCIT12-1-0003) from ACCIÓ. We would like to express our gratitude towards a number of people who have directly or indirectly contributed to the development or refinement of ideas behind *repoVizz*: Javier Jaimovich, Nick Gillian, Miguel Ortiz, Paolo Coletta, Donald Glowinski, and Alfonso Pérez. We also thank Marcelo Wanderley for his offering to collaborate on data acquisition for the string quartet recordings. Finally, special thanks go to Alexander Jensenius for inspiring discussions early in 2007 and 2008, and to Jordi Llop for his initial involvement in early developments during 2010.

4. REFERENCES

- [1] *repoVizz* video. http://repovizz.upf.edu/misc/repovizz_ACMM2013.mp4.
- [2] O. Mayor, J. Llop, and E. Maestre. *Repovizz: a multimodal on-line database and browsing tool for music performane research*. In *Proceedings of Int. Symposium for Music Information Retrieval*, 2011.