Archeologia e Calcolatori
35.2, 2024, 167-176
doi 10.19282/ac.35.2.2024.18

EXTENDED MATRIX MANAGER: AN OPEN TOOL FOR EM BASED KNOWLEDGE GRAPHS MANAGEMENT

1. Introduction

The following contribution aims at presenting a new free and open-source software, EM Manager, that has been created to assist archaeologists at approaching virtual reconstruction projects with the Extended Matrix method. EM Manager represents the most recent update of the Extended Matrix Framework and, due to its features, it will expand the user base. The Extended Matrix (EM) was conceptualized in 2015 (Demetrescu 2015, 2018; Demetrescu, Ferdani 2021; https://www.extendedmatrix.org) as a method with the main purpose of mapping metadata and paradata connected with a virtual reconstruction. EM is a node-based formal language, with rules and specific nodes, developed by the Digital Heritage Innovation Lab (DHILab) of Rome CNR. The method allows to graphically map all the sources used within a virtual reconstruction process, using an oriented graph (the so-called Extended Matrix) that 'extend' a real Harris matrix, and it visually represents the reliability of a reconstruction, using a color scale connected to the nodes of the method. EM is based on standards such as stratigraphy, computer graphics, knowledge networks, and CIDOC-CRM (https://cidoc-crm.org) to register the documentation process of a context or the documentation process of a reconstruction. In one sentence, since we map data provenance of the reconstructive information, the EM is about scientific-driven content creation (Fig. 1a).

A different meaning and a different scope are fulfilled by the Extended Matrix Framework (EMF). This framework represents a large open source ecosystem where tools are shared with other communities in order to support the application of the Extended Matrix method (Fig. 1b). As example, EM tools and 3D Survey Collection (3DSC; https://github.com/zalmoxes-laran/3D-survey-collection), two addons of Blender (https://www.blender.org), has been created and developed with the intent to manage EM data within Blender (EM tools; Demetrescu *et al.* 2023; https://github.com/zalmoxes-laran/EM-blender-tools) and visualize-represent-edit-export topographical and photogrammetric data in Blender (3DSC). For dissemination, presentation and sharing purposes ATON and its web-app EMviq allow user to visit online a 3D scene, with the possibility of multiple interactions, multiple navigation mode, portability on different devices (all these features are natively available on ATON; Fanini *et al.* 2021), and to explore the sematic layer provided by the EM (feature available through EMviq; Demetrescu *et al.* 2023).

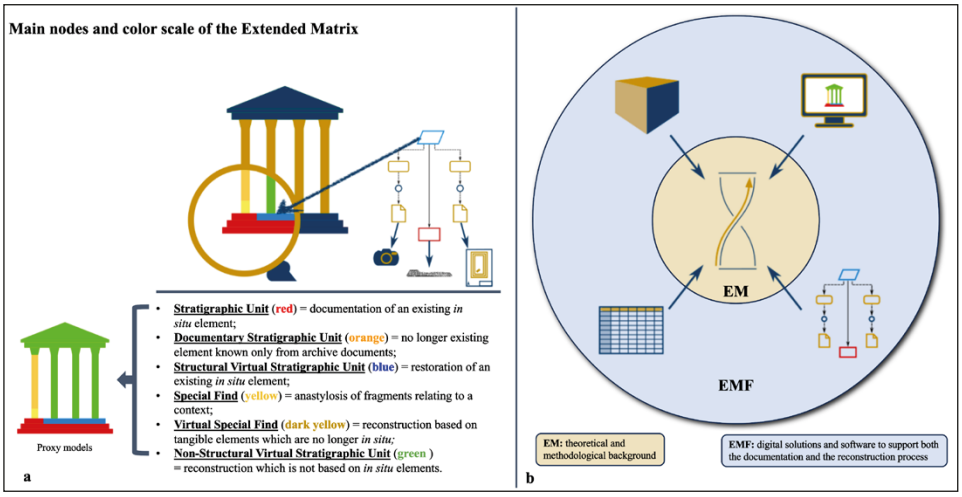

Fig. 1 – a) Schematic representation of the Extended Matrix with its main nodes and the color scale explained; b) schematic representation of the relation between the Extended Matrix and the Extended Matrix Framework.

While the EM represents the theoretical and methodological back ground, which can be deployed even outside a digital environment (for example, an EM can be sketched out on a piece of paper with a pencil, as well as with a software tool). On the contrary, the EMF is a framework that includes digital solutions and software to support the documentation process behind a virtual reconstruction. If EM is about scientific-driven content creation, as already mentioned, the EMF, instead, is about technological-driven solutions for assisting EM. Within the EMF user can find tools useful to create an EM and edit, manage, update, link, share reconstructive data.

S.B., E.D.

2. State of the art

In 2021, within the EMF ecosystem another step towards the EM was made with pyArchInit (Cocca 2016; Montagnetti, Mandolesi 2019). This plugin of QGIS (https://qgis.org/), realized in 2005, allows to store within an organized and up to date database the archaeological documentation of an excavation (excavation sheets, topographical data, images, etc.) and automatically export all the required information with default (or customized) layout. One of the key features of pyArchInit consists in the possibility of automatically creating a Matrix, starting from the information stored within its database. Considering this advanced feature of the

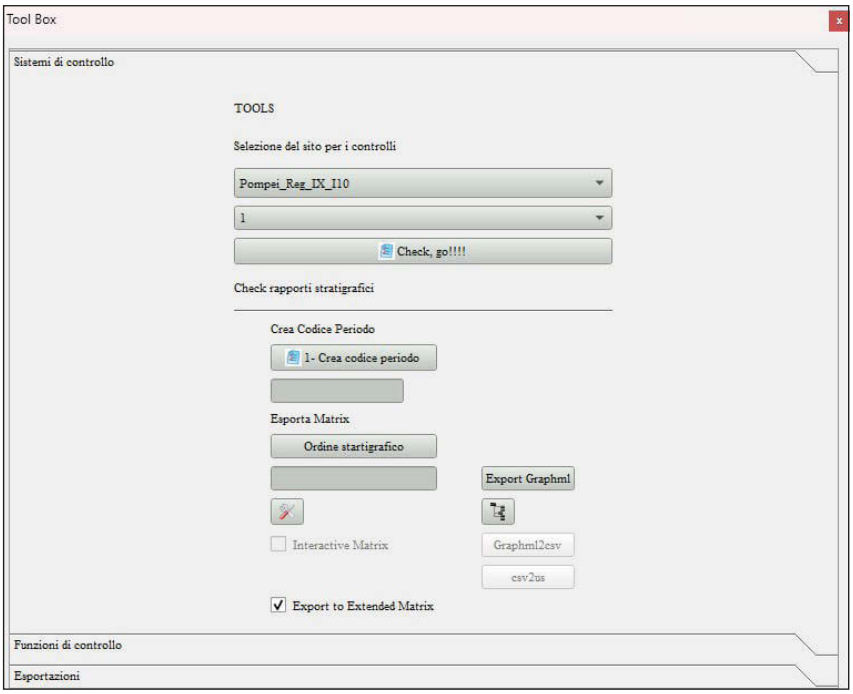

Fig. 2 – The 'Export GraphML' button on the 'Control systems' panel of pyArchInit, within QGIS. This function allows to automatically generate an Extended Matrix using the documentation stored in pyArchInit.

plugin, in 2021 an updated version of pyArchInit was released to allow the automatic generation of an EM, with all the necessary nodes defined by the 'Extended Matrix palette' of the method, using all the information stored in the database (Fig. 2). This fundamental step, occurred within the EMF, established a real connection between the environment of QGIS, where pyArchInit operates, and the 3D environment of Blender, where all the reconstructive information of an EM can be easily visualize in 3D, using proxy models (that is, simplified 3D geometries that allow to represent in a three-dimensional shape, not automatically generated, data which are stored within an EM). This step of the 'development phase' creates the opportune conditions for formulating the idea of EM Manager. As already mentioned, EM Manager was created to assist the creation of an Extended Matrix using an automatic process, in other words EM Manager automatically generate an EM graph from a table.

In general, in order to create a complete EM-project we need at least '7 key roles' within a perfect 'EM-Team'; in this case, the EMF provide a

dedicated software, or a dedicated workflow, for each of these 'roles'. The first role is the 'Archaeologist', who is fundamental for documenting stratigraphic data during the field work. The second is the 'Bookworm' (the so called 'Source-Hunter'), whose role is valuable when the reconstruction proposal requires comparisons due to the lack of other data. The third is the 'DB manager', whose role is important for storing and managing data (such as: an expert in database, an expert in GIS, etc.) that could be useful for assisting the virtual reconstruction process. The fourth role is the 'Expert in survey techniques', who is the one responsible of collecting information (on the field or in the laboratory) and managing geometrical data regarding the archaeological context or the object that needs to be documented and reconstructed. The fifth is the 'EM drawer', who has the role of collecting and organizing all the data functional for the reconstruction, using the formalism of the Extended Matrix. The sixth role is the 'Proxies Modeler', this member of the team does not need specific 3D skills for drawing proxy models, since these geometries are simples-low poly 3D models named as the Stratigraphic nodes of the EM. Proxies are the bridge between the EM graph and Blender. The seventh role, the last one, is the 'Expert in 3D modelling', who is the 3D modeler that will reconstruct the context or the object, with detailed 3D models, using photorealistic materials, appropriate light conditions, etc.

Once defined these '7 key roles', we focused our attention on the 'EM drawer'. This fourth role requires some specific skills for the purpose of drawing a complete EM graph, such as: familiarity with the formalism of the EM (knowing both the meaning of each node, their color scale, and how to use them, etc.); capability of drawing an EM using the software yED (at the moment EM needs yED's 'swimlane' option for the creation of its final layout that maintain chronological division in horizontal sections); knowing how to collect data useful for both the documentation and the reconstruction phases. In most of the cases, when we are dealing with the reconstruction of an archaeological context, 'archaeological data' are frequently stored in databases, and this means tables with data acquired on the field, or elaborated in a second moment (such as: numbers of stratigraphic units; their description; their images; their stratigraphic connection; references to planimetries; references for comparisons; etc.). During the creation of an EM, it is also important to 'optimize time' with the intention of providing a balance between the time spent searching data, for documenting and reconstruing a case study, and the effective time for drawing and filling an EM (maybe at first as a handmade sketch and then with yED).

The final requirement that completes the process of creation of an EM is the 'error checking' step, when the 'EM drawer' controls all the nodes of

the graph to avoid the presence of potential errors. Since most of the time the 'EM-Team' lacks the presence of one or more of the 7 roles, previously described, the idea of creating EM Manager as an assistant for the creation of an EM graph grown up with more motivation. The software would have overcome the potential absence of the 'EM drawer', or simply it would have assisted the 'EM drawer' during the creation process of an EM.

S.B., E.D.

3. EM Manager

EM Manager is a standalone software, free and open source, coded in Python (https://www.python.org), already available for download on GitHub (https://github.com/enzococca/EM-Manager; Fig. 3a). The fact that EM Manager has been coded using Python language ensures the possibility of future implementation with other Python based platforms (such as, QGIS and Blender). At the moment, the software allows to:

1) read archaeological data stored in a 'specific' table (Fig. 3b, c). This table follows: the rules of stratigraphy; the rules of a Stratigraphic Unit sheet; the rules of pyArchInit, since EM Manager maintains a native connection with pyArchInit, even if it is a standalone software;

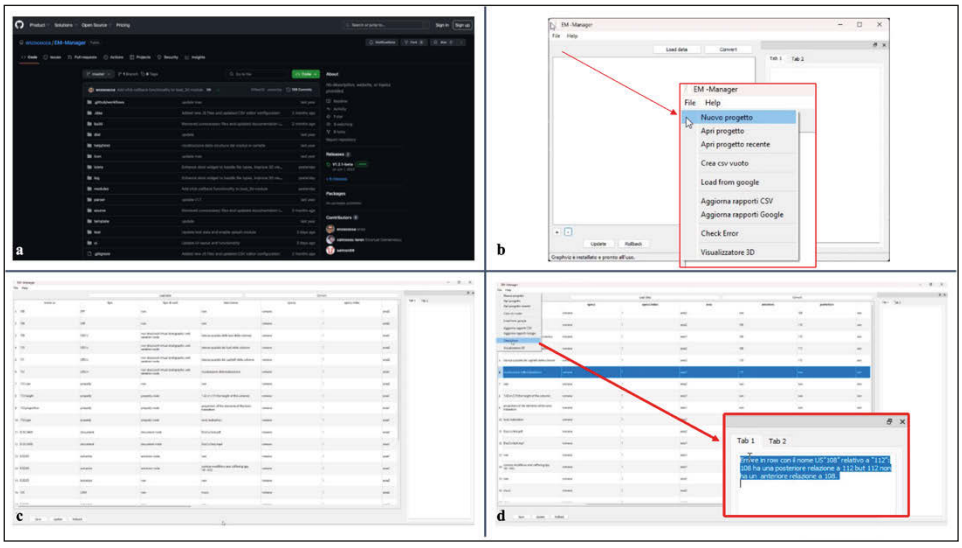

Fig. 3 – a) View of the GitHub online repository where all the versions of EM Manager are stored; b) the 'File' menu of EM Manager where all the functions of the software are displayed; c) panoramic view of how EM Manager displays imported data stored in a table; d) example of how the 'Check Error' function of EM Manager reports errors to user.

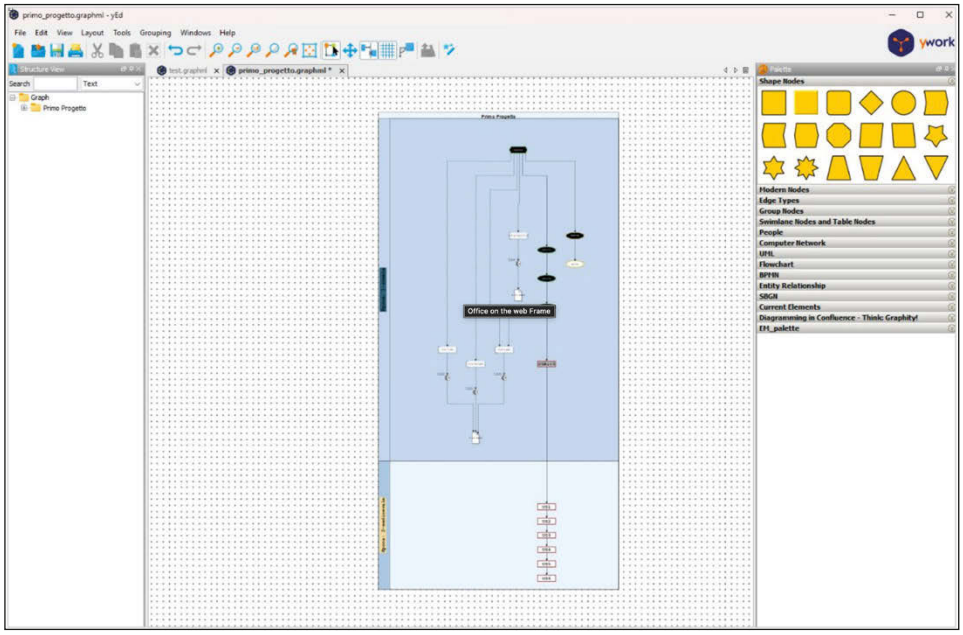

Fig. 4 – View of the Extended Matrix graph in yED after the 'Swimlane' option has been applied. In the example nodes are displayed with respect to the formalism of the Extended Matrix.

2) check errors with respect to the rules of stratigraphy and the rules required for storing data within the table (Fig. 3d). This means that after importing a table it can automatically indicate where the table presents some errors, if any;
3) automatically convert a table into a complete EM (with all the necessary nodes and vectors required by the formalism). In a second step, within yED, the application of the 'Swimlane' option will complete the final layout of the graph (with all the nodes spatially organized in chronological periods; Fig. 4);
4) represent the EM and its data in a 3D environment embedded in the software (Fig. 5).

This feature gives the possibility to visualize a 3D representation of the EM with also all the proxy models. This specific feature, the one that allows visualizing the 3D models in EM Manager, requires the presence of proxy models within a particular folder. This precise condition describes an advanced step of a reconstructive process, the one where EM and proxy models are already made.

Delving into its core, EM Manager is a sophisticated software tool that leverages a variety of Python libraries to handle complex data management

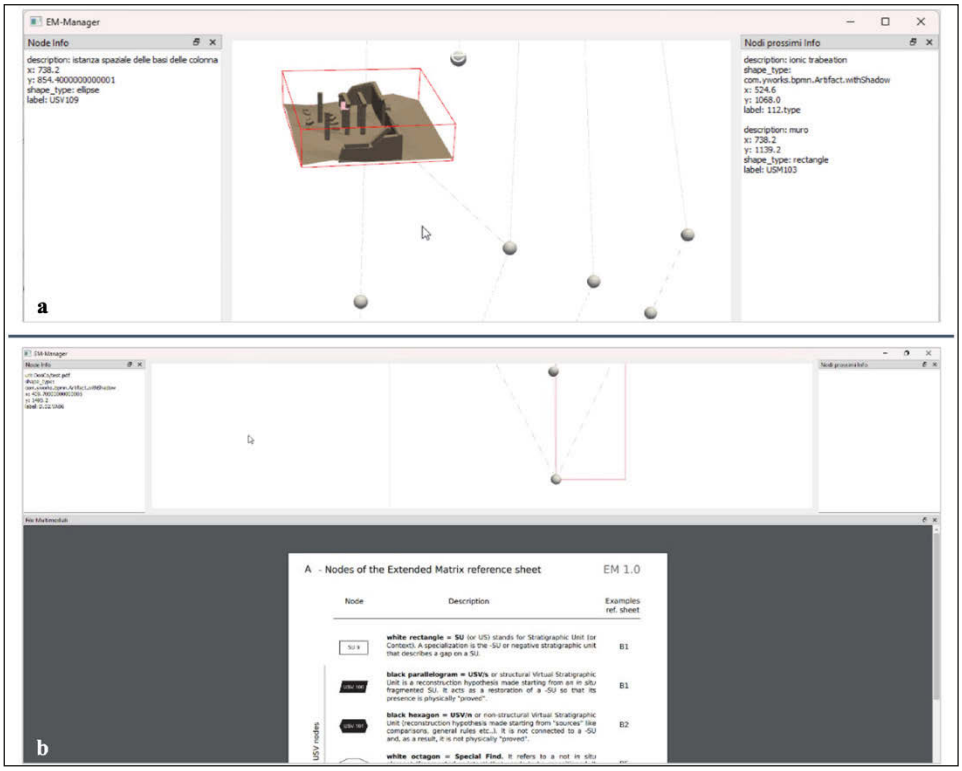

Fig. 5 – a) 3D viewer of EM Manager with the three-dimensional representation of both the Extended Matrix and Proxy models; b) example of data visualization in EM Manager when files are connected to the nodes of the EM.

and visualization tasks in a user-friendly manner. By integrating powerful libraries such as PyQt5, pyVista, pandas, and OpenCV, EM Manager provides an extensive suite of functionalities that cater to professionals who need robust and efficient data processing capabilities. Starting with PyQt5 (https://pypi.org/project/PyQt5/), EM Manager utilizes this library to craft graphical user interfaces (GUIs). PyQt5 is renowned for its ability to create professional and aesthetically pleasing interfaces, making it an ideal choice for applications that require a high level of interaction with users. This allows users to navigate through complex data sets with ease, enhancing the overall usability of the software.

For 3D data management, EM Manager incorporates pyVista (https://docs.pyvista.org/version/stable/), a library known for its advanced 3D visualization and mesh analysis tools. This integration enables users to

interact with three-dimensional datasets, perform geometric transformations, and visualize data in ways that are not possible with two-dimensional representations. The ability to manipulate and analyze 3D data is crucial for users who deal with spatial data, providing insights that are critical for informed decision-making. When it comes to managing tabular data, EM Manager utilizes pandas (https://pandas.pydata.org), a library that excels in data manipulation and analysis. With pandas, users can efficiently organize, sort, and filter large datasets, perform statistical analysis, and export data to various formats. The library's powerful data handling capabilities make it an indispensable tool for managing complex datasets with multiple attributes.

For media manipulation, EM Manager employs OpenCV (https://opencv.org), a library designed for real-time computer vision applications. This integration allows users to process images and videos within the software, perform image enhancements, and extract meaningful information from visual data. OpenCV's functionality extends the capabilities of EM Manager into the realm of media analysis, making it a versatile tool for handling a wide range of data types. Beyond these Python libraries, EM Manager also utilizes Graphviz (https://graphviz.org) for specific graph-related tasks. Graphviz is an external software that excels in the creation of visual representations of networks and graphs. EM Manager uses Graphviz to generate Harris matrix in the DOT file format, which is then transformed into GraphML using a custom-written parser. This feature is particularly useful for users who need to visualize and analyze relational data, providing a clear and structured representation of complex relationships.

Overall, EM Manager represents a comprehensive solution that combines multiple Python libraries and external software to address a broad spectrum of data management needs. By integrating these tools, EM Manager not only enhances its functionality but also provides a seamless and efficient workflow that helps users manage and analyze data effectively.

E.C., E.D., S.B.

4. Conclusions

The software described in this contribution represents an important step in the development of the EMF since it offers a software solution, free and open source, that can assist EM users to approach the EM even if they do not cover all the 7 key roles of the 'EM-Team'. At the moment, as already mentioned, the software is available on GitHub, but its development is still in progress. Future development of the software will include: a stable release for Windows OS; an installer for Mac OS; a general update of the online

documentation, following the one realized for 3D Survey Collection (3DSC) and EM tools; an official 3D layout of EM nodes to be visualized in EM Manager; an improved 3D viewer for the visualization of proxy models; an automatic converter from EM to table.

S.B., E.C., E.D.

Simone Berto, Emanuel Demetrescu
Istituto di Scienze del Patrimonio Culturale - CNR
simoneberto@cnr.it, emanuel.demetrescu@cnr.it

Enzo Cocca
AdArte s.r.l
enzo@adarteinfo.it

REFERENCES

Cocca E., Mandolesi L. 2016, *Analisi statistiche e geostatistiche con PyArchInit: prima sperimentazione*, in A. Caravale, P. Basso, P. Grossi (eds.), *ARCHEOFOSS. Free, Libre and Open Source Software e Open Format nei processi di ricerca archeologica. Atti del IX Workshop (Verona 2014)*, «Archeologia e Calcolatori», Suppl. 8, 132-140 (https://www.archcalc.cnr.it/journal/id.php?id=876&sup=true).

Demetrescu E. 2015, *Archaeological stratigraphy as a formal language for virtual reconstruction. Theory and practice*, «Journal of Archaeological Science», 57, 42-55 (https://doi.org/10.1016/j.jas.2015.02.004).

Demetrescu E. 2018, *Virtual reconstruction as a scientific tool: The Extended Matrix and source-based modelling approach*, in S. Münster, K. Friedrichs, F. Niebling, A. Seidel-Grzesińska (eds.), *Digital Research and Education in Architectural Heritage*, *5th Conference, DECH 2017, and First Workshop, UHDL 2017 (Dresden 2017)*, Communications in Computer and Information Science, Cham, Springer International Publishing, 102-116 (https://doi.org/10.1007/978-3-319-76992-9_7).

Demetrescu E., Fanini B., Cocca E. 2023, *An online dissemination workflow for the scientific process in CH through semantic 3D: EMtools and EMviq open source tools*, «Heritage», 6, 1264-1276 (https://doi.org/10.3390/heritage6020069).

Demetrescu E., Ferdani 2021, *From field archaeology to virtual reconstruction: A five steps method using the Extended Matrix*, «Applied Sciences», 11, 11, 5206 (https://doi.org/10.3390/app11115206).

Fanini B., Ferdani D., Demetrescu E., Berto S., d'Annibale E. 2021, *ATON: An open-source framework for creating immersive, collaborative and liquid web-apps for cultural heritage*, «Applied Sciences», 11, 22, 11062 (https://doi.org/10.3390/app112211062).

Montagnetti R., Mandolesi L. 2019, *QGIS, pyArchInit and Blender: Surveying and management of archaeological data with open source solutions*, «Archeomatica», 10, 4, 128-138 (https://doi.org/10.48258/arc.v10i4.1706).

ABSTRACT

The following contribution aims at presenting a new free and open source software, EM Manager, created to assist archaeologists at approaching virtual reconstruction projects with the Extended Matrix (method developed at the Digital Heritage Innovation Lab - DHILab - of the CNR-ISPC of Rome). EM Manager is a free and open source standalone software, already

available on GitHub, that allows to convert a table into an Extended Matrix. The software represents a remarkable update of the Extended Matrix Framework and, due to its features, it will expand the user base, since it helps EM users to cover one, or more, of the '7 key roles' considered by the Extended Matrix method. In addition, the fact that EM Manager is Python based ensures the possibility for future implementation with other platforms (such as QGIS and Blender) based on the same language.