GA – PROJECT NUMBER:	**101158046**
PROJECT ACRONYM:	AUTOMATA
PROJECT TITLE:	AUTOmated enriched digitisation of Archaeological liThics and cerAmics
CALL/TOPIC:	HORIZON-CL2-2023-HERITAGE-ECCCH-01-02
TYPE OF ACTION	HORIZON RIA
PRINCIPAL INVESTIGATOR	Prof Gabriele Gattiglia, UNIPI
TEL:	+39 050 2215228
E-MAIL:	gabriele.gattiglia@unipi.it

This project has received funding from the European Union's HORIZON RIA research and innovation programme under grant agreement N. 101158046

D 2.1 Methodologies, scenario, and user requirements

Version: 1.0

Revision: First release

Work Package: 2 - Needs analysis, methodologies, specification, design
Lead Author (Org): Martina Naso (UNIPI), Claudia Sciuto (UNIPI)
Contributing Author(s) (Org): Rémy Chapoulie (UBM), Nevio Dubbini (MIN), Bruno Dutailly (UBM), Gabriele Gattiglia (UNIPI), Xavier Granier (UBM), Amala Marx (INRAP), Riccardo Persichini (QB), Heeli Schechter (HUJ), Holly Wright (UoY)
Due Date: M4
Date: 31/12/2024

Dissemination Level		
P	**Public**	**X**
C	**Confidential, only for members of the consortium and the Commission Services**	

Revision History

Revision	Date	Author	Description
0.1	07.11.2024	Claudia Sciuto	Structure content
0.2	11.11.2024	Claudia Sciuto	Draft content added and structure modified
0.3	12.11.2024	Martina Naso	Draft content added

0.4	13.11.2024	Martina Naso	Draft content added and structure modified
0.5	14.11.2024	Claudia Sciuto	Draft content added and structure modified
0.6	18.11.2024	Heeli Schechter	Content added in sections 1.2.1, 1.2.2, 1.3.1
0.7	20.11.2024	Riccardo Persichini	Revision text section 3.1
0.8	21.11.2024	Xavier Granier	Content added in sections 1.3.1 and 4
0.9	21.11.2024	Martina Naso and Claudia Sciuto	Draft content added and structure modified
0.10	22.11.2024	Gabriele Gattiglia	Revision of section 1.1 and 1.2, text added
0.11	24.11.2024	Claudia Sciuto	Draft content added and structure modified
0.12	25.11.2024	Martina Naso	Draft content added and structure modified
0.13	04.12.2024	Holly Wright	Draft content added in section 1.5
0.14	04.12.2024	Rémy Chapoulie	Draft content added in section 4.3
0.15	05.12.2024	Martina Naso	Draft content added and structure modified
0.16	06.12.2024	Xavier Granier	Draft content added in section 1.5
0.17	09.12.2024	Heeli Schechter	Draft content added in sections 1.2.1.1, 1.3, 1.3.1, 2.2, 4.3.1
0.18	10.12.2024	Amala Marx, Heeli Schechter, Martina Naso	Draft content added in sections 1.3, Draft content added and structured modified in 4.1, Draft added in section1.4
0.19	13.12.2024	Bruno Dutailly	Draft added in sections 1.3.1, 1.6
0.20	23.12.2024	Gabriele Gattiglia, Martina Naso, Claudia Sciuto	Content added, structure modifies and general revision

Disclaimer

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Contents

Abbreviations

WP	Work Package
M	Month
UNIPI	Università di Pisa
UBM	Université Bordeaux Montaigne
UB	Universitat de Barcelona
UoY	University of York
HUJ	The Hebrew University of Jerusalem
INRAP	Institut national de recherches archéologiques préventives
KCL	King's College London
MIN	Miningful srls
IIT	Fondazione Istituto Italiano di Tecnologia
QB	QBrobotics Srl
AMZ	Arheoloski Muzej u Zagrebu
CL	Culture Lab

Executive summary

Deliverable D2.1, *Methodologies, Scenario, and User Requirements*, provides a comprehensive overview of the AUTOMATA system, focusing on user requirements and the operational context in which the system will function. Specific details about the system's development will be elaborated in subsequent deliverables: D2.2 *State of the Art on Enhanced Digitisation*, D2.3 *System Specification*, and D2.4 *Ethical Guidelines for Trustworthy AI*.

This document begins with an archaeological needs analysis, outlining the scenarios in which the AUTOMATA system must operate (Sections 1.1 and 1.2). It then explores the processes involved in handling ceramic and lithic artefacts, from retrieval and processing to digitisation (Section 1.3), archaeometric analyses (Section 1.4), and metadata management (Section 1.5). The project's rationale is to ensure that the AUTOMATA system effectively addresses real-world challenges, including accessibility and legal, political, and ethical considerations (Section 2.1).

The system aims to integrate various sensors for 3D modelling and archaeometric analysis into a single portable and scalable robotic framework (Section 2.2). These tools, currently used independently by different operators across varying contexts and timelines, will be unified to enable streamlined operation. Section 2.3 delves into how the collected data will be integrated into enhanced 3D digital models, ensuring data quality and adherence to curation standards while maintaining alignment with European Collaborative Cloud for Cultural Heritage (ECCCH) objectives.

The culmination of these efforts is a structured and efficient workflow (Section 3) that facilitates the enhanced digitisation of archaeological artefacts, specifically ceramics and lithics. This workflow will allow for semi-automated data collection, requiring minimal input from a single operator. By consolidating data from various sensors into a cohesive process, the AUTOMATA system will enable the collection of Big Archaeological Data, transforming the current fragmented approach into a unified, efficient methodology. A key consideration in designing this workflow has been the establishment of a target average digitisation time capable to represents a well-calibrated compromise between operational performance and archaeological requirements, ensuring that the system is efficient on meeting the rigorous demands of research and analysis. By balancing these factors, the AUTOMATA project ensures a sustainable and practical approach to large-scale digitisation.

1 Working scenarios in archaeology

The archaeological practice involves both academic research institutions (universities, museums, etc.) and professional archaeology companies, which differ mainly in their objectives, timelines, and resources. Academic archaeology typically operates at a slower pace, with excavation seasons spanning a few months and more time allocated for post-excavation documentation and analysis. These projects often involve multidisciplinary teams that include specialists and students, and they benefit from dedicated workspaces for washing, drying, cataloguing, and analysing artefacts. The centralisation of artefact processing in university or museum departments also facilitates in-depth study and cross-site comparisons. Time constraints are tighter in professional archaeology, which is linked primarily to rescue or development-led projects. Due to connections with the construction industry, teams are composed mainly of trained archaeologists, with a limited focus on extended post-excavation analysis. Artefact processing typically prioritises retrieval and preliminary spot dating during excavation. Advanced archaeometric analyses are typically conducted by research institutions rather than professional archaeologists. These analyses, including techniques for material characterisation and provenance studies, require specialised expertise and equipment often unavailable to professional teams. This division highlights the importance of bridging workflows between academic and professional contexts to ensure comprehensive documentation and analysis of archaeological artefacts. This section examines these scenarios, providing the context for understanding how the AUTOMATA system can enhance current workflows by integrating digitisation and archaeometric tools within diverse archaeological environments.

1.1 The archaeology of ceramics and lithics, between established protocols and new research questions

The study of ceramics and lithic artefacts is at the core of archaeological research. These materials are frequently retrieved in large quantities, providing information about chronology and revealing essential insights into ancient lifestyles, trade networks, and technological practices. Considering their abundance and importance for understanding the archaeological record, protocols have been refined to optimise the documentation and study of pottery sherds and lithics.

Shared protocols (with minor variations) for treating ceramic and lithic artefacts aim to gather information to answer general research questions: What is the chronology of the context? What is the cultural horizon or trade network in which the context is embedded? How were the objects produced, and what materials were used?

After the field excavation protocol, cleaning, numbering, classifying, reassembly, and drawing of the artefacts are some of the essential steps usually undertaken when dealing with these materials. Most of the steps in these pipelines involve sorting and preparing the objects. On the other hand, the documentation of the objects often takes place analogically through observation and hand drawing. Object classification is usually done by experts who know how to associate specific chronological or contextual observations with the pieces.

Different scenarios have emerged in recent years linked to advances in the digitisation of documenting and classifying artefacts. The dematerialisation of the documentation process can take place mainly through two steps: the representation of artefacts in two or three dimensions and the characterisation of the physical characteristics of the pieces through non-destructive geochemical and geophysical analyses. The representation in two dimensions consists of reproducing the 2D profile of artefacts and can be achieved either through manual drawing, which can subsequently be digitised, or directly through digital methods, such as laser-aided systems (fig. 4). Three-dimensional models, instead, reflect geometry, morphology and appearance of artefacts. The application of digital techniques for the acquisition of three-dimensional data containing information on the geometry of the appearance of individual fragments has made it possible to refine certain specific observations, and the adoption of methods for processing 3D models makes it possible to explore even aspects invisible to the human eye.

Conversely, the miniaturisation of sensors for elemental and molecular analysis has made it possible to collect data on the composition of materials without altering the artefacts.

3D models and archaeometric analyses constitute two important aspects of artefact digitisation, the dissemination of which has transformed the study of archaeological materials, with specific regard to ceramics and lithics. The spread of these techniques has also led to the formulation of new research questions, which sometimes lead archaeologists to revisit legacy data. By creating detailed 3D models, databases associated with objects and digital archives, researchers are enabled to analyse and share artefacts and information more widely. Yet, according to the European Commission, only a fraction of Europe's cultural collections have been digitised, and even fewer items are captured in 3D, a limitation that hinders the potential for detailed study and comparison.
If combined with physico-chemical analyses, digitisation becomes even more valuable, as it allows researchers to link detailed visual and spatial data with precise information on material composition, provenance, and production techniques. Digital models of ceramic or lithic artefacts can be paired with chemical data on mineral content, trace elements, or residual compounds, allowing scientists to investigate production sources, technological processes, and even past usage. This integration supports archaeometric analysis on a larger scale, as researchers can remotely access both digital and chemical data, comparing artefacts across different collections without physical handling, reducing risks of damage or degradation. Through this convergence of digitisation and material analysis, scientists can reconstruct the lifecycle of artefacts — from raw material selection to production and use — ultimately enriching our understanding of historical societies and facilitating the preservation and sharing of cultural heritage across diverse audiences.

Despite these conspicuous advantages, digitisation of individual ceramic and lithic fragments occurs in very few cases and very rarely includes both 3D modelling and archaeometric characterisation using more than one technique. The vast majority of artefacts found in excavations are still processed manually. This project will give us the possibility to reflect on how to digitise and document large quantities of artefacts.

1.2 Ceramics and lithics: from retrieval to documentation

An archaeological assemblage is defined by its connection to a specific context, such as a stratigraphic unit (SU). Essential aspects of assemblages include composition, fragmentation, and integrity. Composition refers to the categorisation of items, such as by ceramic type or style, with each type representing a different class within the assemblage. Fragmentation and integrity reflect the current state of artefacts, which are typically incomplete due to the formation processes of the archaeological record. Measuring these properties can reveal insights into how artefacts were used and discarded.

1.2.1 Retrieval, cleaning and sorting

In archaeological excavations, ceramic and lithic fragments are typically retrieved by hand, with the goal of collecting all material fragments across various periods and categories without discarding any items. The quantity and quality of the materials gathered depend primarily on two factors: the excavation strategy and the characteristics of the deposit. Excavation strategies vary significantly, especially in rescue or preventive archaeology, where mechanical excavators are sometimes used. This method can lead to larger, fewer fragments but also risks damaging artefacts. In contrast, trowel excavation is more delicate, allowing for the recovery of smaller pieces with minimal impact on artefact preservation. The deposit's composition — its colour, texture, and compaction — also affects retrieval; for instance, clay-rich soils may form clumps that conceal fragments, whereas dark, muddy soils make it hard to spot smaller pieces of pottery. Sieving or sub-sampling can help reduce recovery biases caused by varying excavation methods. Archaeologists may conduct a preliminary review of pottery during excavation to estimate the site's chronology and avoid potential errors in the stratigraphic interpretation, as correct dating relies on careful excavation methods. When surveying, factors such as sediment type, clump size in ploughed fields, and sunlight conditions can affect the likelihood of finding fragments. Once gathered, materials are labelled with their context numbers and stored in transparent bags or trays, marked directly if necessary.

Regarding ceramic materials, after collection, fragments are cleaned with scrubbing brushes and left to dry, preventing the formation of mould. Once dry, each potsherd is marked with ink using a fine pen[1]. The fragments are then bagged or stored in trays by context and placed on shelving units. A preliminary inventory is created to log the finds, noting the quantity and location of bags and trays for each context. This list includes details like the number of plastic bags for each type of material, the tray number containing the finds, and the tray's location. These steps may differ slightly depending on working practice and context. For example, some bags may contain materials from mixed classes but from the same SU. Alternatively, materials could be stored in a tray representing an SU but divided within the bags according to the ceramic classes. In summary, the practice is not the same for all contexts; rather what remains the same is that the finds are bagged and stored in boxes, while the criteria for the subdivision and labelling of the containers may vary.

The procedure is similar for lithic finds. Lithics retrieved by hand during excavation or after sediment sieving are washed and brushed in water, and left to dry. Initial sorting in the field is usually confined to the widest techno-typological categories (Debris, Debitage, Tools, etc.), while finer typological identification (tool type, debitage industry, etc.) and separation are only occasionally achieved. Most often, this advanced step is only performed much later in the analysis process, disconnected from the fieldwork. Sorted lithic artefacts are stored in contextually marked bags, whether paper or plastic and deposited in boxes of various sizes and materials. This procedure may vary in detail depending on the excavation director, excavation type, context, etc. For example, individual lithic artefacts may occasionally be marked with ink pens, though usually not. The organisation of stored boxes is also variable and may be based on excavation season, field area, SU and so on, usually in a hierarchical manner (i.e. all boxes from a specific SU placed together, sorted by excavation season).

1.2.1.1 Encrustations and patinas

The surface of archaeological artefacts is often partially or entirely covered with encrustations and patinas, which provide tangible evidence of the environment in which they were buried, from the moment of burial until the moment they were unearthed.

These deposits often embed mineral grains of different types, plant fragments, organic substances, and tiny pieces of miscellaneous materials (Ogburn D. *et al.* 2013; Frahm E. *et al.* 2020). Other surface alterations affecting the buried fragments may be due to post-depositional processes triggered in archaeological deposits. Soil washing, for example, can lead to the displacement of certain elements, which then tend to reaccumulate within the soil pores or on the surface of the objects buried within it, forming patinas.
Suppose an artefact has been submerged in groundwater in direct contact with the water table for centuries. Salts dissolved in the water gradually deposit on its surface, forming encrustations that vary in thickness and colour. Depending on the concentration of mineral salts in the water and the duration or intensity of the process, these encrustations or patinas can be more or less thick. Thickness can also vary across the surface of a single artefact, influenced by its burial position and the degree of water circulation around it.

Climate plays a role as well: in arid regions or areas prone to extended dry periods, capillary action can bring groundwater to the artefact's surface repeatedly, alternating wet and dry periods that cause layered depositions and sometimes colour changes (Maritan L. 2020).
For example, in Mediterranean contexts, encrustations are often composed of calcium carbonate or iron oxide due to the prevalence of calcareous water and iron-rich soils. In areas with igneous or metamorphic rocks, like Lazio or Campania (Italy), silica-rich percolating waters may form microcrystalline quartz or amorphous chalcedony, resulting in siliceous encrustations with smooth, glassy appearances. By contrast, calcareous

[1] This last step was implemented systematically until a few years ago but is now less used. It is not certain that the pieces stored in the warehouses are actually all marked.

encrustations are rough, opaque, and non-homogeneous, with a chaotic structure of pits and growths. For example, artefact coatings titled Desert Varnish appear in desert areas of the Southern Levant and other regions. The varnish is a brown/black coating that evolves on rock surfaces in hot and slightly moist conditions. When moisture interacts with aeolian particles on the surface of a rock, a chemical reaction occurs. As the moisture evaporates, it leaves behind a residue, typically rich in iron or manganese. This residue forms a very thin coating on the rock, visible as a change in colour (Dorn R.I. 2024; Reneau *et al.* 1992).
Beyond these regions, the diversity of environmental conditions creates distinct encrustations and patinas. In Northern Europe's temperate climates, moist, acidic soils produce iron-rich encrustations on artefacts, particularly in peat bogs where anaerobic conditions facilitate sulphide reactions. Arctic regions, with their freeze-thaw cycles, deposit fragile cryogenic salts like gypsum, while arid Sub-Saharan climates foster thick salt deposits such as halite or gypsum, often layered due to alternating wet and dry conditions. In South Asia's floodplains, artefacts are often coated with dense calcium carbonate crusts from seasonal flooding, while tropical coastal zones in Australasia produce encrustations enriched with marine minerals and organic debris.

The presence of these surface alterations effectively interferes with the work of recognising and characterising the artefacts. On one hand, they obscure certain surface features of the pieces, such as marks that could help understand the techniques used in shaping them; on the other hand, they may influence the results of archaeometric analyses if conducted on areas affected by these alterations (see paragraph 1.4.1.2.1).

1.2.2 Artefacts' recording: classification, dating and documentation

To establish a rough date for the pottery and lithics findings as early as possible, archaeologists perform an initial assessment (spot dating) during or shortly after the first steps.

Specialists review key ceramic traits, such as surface treatment, decoration, and fabric type, to make a preliminary identification. Identifying the form is the next step, aided by comparison to reference catalogues. The pottery expert analyses the profile of each fragment, comparing it with known vessels, and noting any unusual pieces. In case a selection of artefacts happens, no potsherd is discarded until fully examined by the specialist responsible.

Lithics are similarly characterised and identified by complementing sets of traits. These classification avenues are usually internally hierarchical (i.e. category > sub-category > type) and may follow raw material, technological, typological, functional, etc. criteria. Traits that may be used for rough dating (i.e., a typical raw material, fossil directeur tool type, unique technological industry) are noted on site and used to assess SU chrono-cultural assignments. Basic classification on all levels may be performed on site and is usually further implemented by specialists in more advanced stages of analysis. No lithic find is discarded during the process. It is important to note that classification definitions vary and may not fully correlate between researchers, schools or contexts.

Quantification, or the systematic counting of items, is a foundational step in archaeological studies. Various methods are used for ceramics, including sherd counts, weight measures, and estimated vessel equivalents. These techniques, widely discussed in archaeological literature, remain essential tools for assemblage analysis. Traditionally, the calculation for ceramics was based on weight. However, with today's technologies, it would be more relevant to calculate different quantitative parameters, such as the surface area of the sherds, the volume of flint fragments, or the angles of the flaking. Fragments can be recorded individually or grouped by shared characteristics, such as ceramic type, function (e.g., cooking, serving, storage), or even by vessel parts like rims, handles, bodies and bases. Following Orton C. and Hughes M. (2013)'s guidelines, pottery can be classified into seven categories, from individual sherds to groups defined by shared fabric type, class, or form. By organising items from broad categories down to specific details, archaeologists can add depth to their classification, allowing each group to reflect unique properties like function or origin. Lithic finds are counted individually and often weighed and measured. The smallest subset (i.e. ‘chips’) may be weighed as unclassified groups.

In traditional documentation, visual recording generally involves a combination of drawing and photography. Archaeologists follow standardised techniques to create drawings that meet precise graphic standards composed of a 2D profile and a frontal view of the main features of the appearance, using specific tools such as pencils, callipers, profile gauges, and millimetre paper. Once drawn, images are scanned and digitally inked (or vectorised) for publication. The time needed for each drawing depends on the fragment’s size, condition, formal complexity and the draftsperson’s skill level. Although the method is straightforward, producing accurate illustrations requires practice, with experienced draftspersons generally working faster and more precisely. The act of drawing is a powerful tool for analysis and provides an understanding of the object; the draftsperson translates this understanding using graphic standards and renderings. Digitisation usually occurs post-excavation, where drawings are converted into vector drawings suitable for publication. Selected fragments are photographed, and these images, sometimes restricted in colour for print publications, are cropped, renamed, and archived digitally.

Drawing is still the preferred method for documenting pottery and lithic artefacts, as it captures technical details that aid specialists. Recently, laser-aided systems have been developed to streamline the 2D representation of artefacts. However, some features — such as colour, surface texture, and precise 3D shapes — can be challenging to render in drawings. Photographs help fill this gap by showing certain aspects of the artefact’s appearance, but they are often limited in capturing its complete 3D form. The addition of 3D acquisition and modelling has helped overcome these limitations, offering a comprehensive digital record of colour, texture, and shape. This integration of 3D models allows researchers to view artefacts from multiple angles and in varying detail, complementing traditional methods and providing richer visual documentation. By attaching the documentation on the 3D models, the interaction with documentation can be as intuitive as in front of real objects.

1.3 Digitisation of pottery sherds and lithics

Digital visualisation technology has rapidly developed into a vital sub-discipline within archaeology. Depending on the archaeological research questions or goals posed, and the size and morphological intricacy of the artefact, the model resolution or quality, and thus the 3D data acquisition technology, needs to be determined and adjusted. It includes advanced 3D techniques that enable archaeologists to represent archaeological remains dynamically and interactively, facilitating a deeper understanding of past societies. With modern 3D scanning technology, which relies on optical methods like laser or structured light triangulation and time-of-flight techniques, cultural objects and archaeological findings can be recorded with extraordinary precision. Advances in 3D graphics technology enable the creation of accurate digital representations, preserving the unique morphology and texture of artefacts with micrometre-level precision. This non-invasive and remote approach supports both scientific study and public engagement, enhancing institutions’ abilities to interpret and share cultural heritage with high fidelity.

Indeed, digital visualisation as documentation has become integral to archaeological and cultural heritage research, making a return to traditional methods highly unlikely. Beyond this, there are analytical methods, where the development of computer algorithms, driven by research questions, are based on the acquired 3D digital models. 3D modelling produces numerical data (i.e., the location of each point in the point-cloud that makes up the 3D model is represented by three numbers - coordinates - expressing its spatial location on a three-axis grid) that may be mathematically analysed. 3D modelling may thus be used for calculating various geometric-morphological traits of archaeological artefacts and assemblages, producing accurate, objective and reproducible results. This type of digital-data-based computational approach has been utilised to address a wide range of innovative as well as traditional archaeological questions encompassing technology, society and cognition (i.e., Dubinsky L., Grosman L. 2024; Harush O., Grosman L. 2021; Muller A. *at al.* 2023; Valletta F. *et al.* 2021; Yashuv T., Grosman L. 2024).

Thus, a critical approach is necessary when using 3D data capture technologies to ensure that they serve specific research goals and purposes effectively. To optimise the usefulness of visual representations and the digital data they are based on, it is essential to select 3D data collection and imaging techniques and scales that align with the specific information required, including scale and resolution, as well as visual appearance. Scientific imaging enables objects to be visualised at multiple levels, from microscopic details to macroscopic overviews. Different techniques provide distinct types of information, much like maps at various scales offer perspectives suited to particular needs — whether identifying a local feature or understanding global geography. Similarly, cultural heritage imaging ranges from capturing an artefact's overall form and context to documenting fine details of its surface and condition. For example, the computation of the 3D digital data may require high-resolution data for lithic scar tracing, while low-resolution data is sufficient for the centre of mass calculation (Grosman L. *et al.* 2022). Aligning the method and scale to research objectives ensures that 3D digitisation and visualisations effectively support the understanding, preservation, and communication of archaeological and cultural heritage.

The following paragraphs are organised according to the different scales of analysis that can be applied to the study of archaeological ceramics and lithics. At the macroscopic level, imaging focuses on the entire object, providing a clear view of its overall shape, condition, and state of preservation. Moving to finer scales, the analysis shifts to the microscopic and atomic levels, where analysis reveals increasingly detailed portions of the object. These finer-scale results are essential for scientific studies, offering specific information that aids in the interpretation and deeper understanding of the artefacts.

1.3.1 Creation of digital 3D models

As mentioned above, 3D models are used for various purposes, ranging from visualisation to morphometric and spatial analysis, in both cultural heritage as well as academic research. Following is a short overview of some of the 3D data acquisition technologies that are commonly used to create 3D models of archaeological artefacts and may be considered for incorporation in the AUTOMATA project, underlining the strengths and weaknesses of each acquisition method (Table 1).
These technologies capture the 3D data in different ways, yet all ultimately produce a point-cloud, where each point is identified by a three-coordinate location in space. The points, or vertices, are connected by a maximum set of non-crossing lines to form polygonal surfaces, and the angles between these polygons define the volumetric shape and outer surface of the visual 3D model. The angles between these polygons, or between their normal vectors, form the base for geometric morphological calculations.

In **photogrammetry**, specialised software identifies overlapping parts of multiple images taken from different angles of an object and stitches the photographs together along these shared points to construct a 3D model. This method excels in capturing a wide range of colours and textures with realistic visual results, particularly when using high-resolution cameras that can reproduce subtle colour variations. Photogrammetry is also highly effective for creating detailed 3D models of small objects, often achieving high-resolution results suitable for intricate items. For instance, using macro lenses or extension tubes, resolutions can reach micrometric levels, with magnifications higher than 2x, making photogrammetry a viable alternative to contact measuring systems for small objects (Galantucci L.M. *et al*. 2018). However, photogrammetry relies heavily on controlled and uniform lighting conditions, as uneven lighting can introduce shadows and distortions in the captured colours. While it excels in colour and texture reproduction, the process requires significant computing power and time for image processing. Additional limitations include difficulties in segmentation (separation of an object from the background) and need for normalisation calculations and scaling procedures due to the pixel-based nature of the data (Harush O. *et al.* 2020).

Laser scanning uses laser beams to measure distances to a surface by calculating the time it takes for the laser to bounce back, and, combined with angular information, it produces the spatial location of each point in the point-cloud. This technology is very versatile in the size and geometric complexity of scanned artefacts, and 3D data capture time is quick, though model processing time may vary. The produced point-cloud density varies yet may be considered as the middle range between photogrammetry and structured light, and high-quality hardware may be very expensive.
Laser scanning prioritises precise geometric measurements over colour fidelity. While some laser scanners integrate RGB sensors or cameras to capture colour, the resulting data is generally less detailed compared to photogrammetry or structured light scanning. They are particularly reliable for capturing uniform surfaces, but the quality of colour capture is moderate at best. Reflective or transparent materials can also create significant challenges, leading to artefacts in the scanned model.

The **structured light** approach analyses deviations between projected and captured light patterns. A structured light scanner consists of a projector, which projects a pre-set light pattern onto a 3D object, and cameras that capture the light as it bounces back. By calculating the inevitable distortion or deviation of the light pattern, the system triangulates the spatial location of each point in the resulting point-cloud. This method excels in delivering high-resolution detail, making it particularly suitable for small objects. Additionally, the data collection process is comparatively fast. Structured light scanning offers a well-balanced solution, combining precise geometric capture with accurate colour representation. Integrated RGB cameras allow the acquisition of colour data alongside geometric details, enabling the creation of visually uniform and realistic models. When lighting conditions are properly controlled, structured light scanners can achieve exceptional texture uniformity and faithful colour reproduction. These attributes make the technology especially effective for small-to-medium-sized objects, as it captures intricate details in both geometry and texture. However, this method is not without challenges. It struggles with shiny, transparent, or very dark surfaces, which can disrupt accurate data capture, although pre-treatment of the object with dulling sprays or other materials can mitigate these issues. Additionally, structured light scanning hardware is generally more expensive than other 3D scanning technologies, which can limit its accessibility.

	Photogrammetry	**Laser**	**Structured light**
Artefact size scale	Wide range - on the larger scale	Wide range - on the larger scale	Smaller
Resolution / accuracy	Low	Low	High
Necessary lighting conditions	Bright even light	Dim to dark	Dim to dark
Artefact surface limitations	Reflective	Reflective, transparent	Reflective, transparent, very dark
Capture speed	Slow (if dependent on one camera)	Quick	Relatively quick
Software processing speed	Slow	Medium	Fast
Hardware cost	Range. Can be low	Range. May be very high	High
Requires calibration	No	Yes	Yes
Artefact color detail	Excellent	Moderate to Good	Good to Excellent
Artefact texture uniformity	Moderate (lighting sensitive)	Good	Excellent
Handling difficult surfaces	Poor for shiny/translucent materials	Low to Moderate	Low to Moderate

Table 1. A simplified comparison table of different 3D data capturing technologies.

It is important to remember that all the compared parameters in all the above methods have ranges; hardware prices, for example, vary tremendously, yet high-quality equipment, usually paralleled with high-density point-clouds, is always more expensive. The suitable artefact size varies both between and within each method. Even among structured light scanners, some are specifically designed for smaller items, while others are better suited to larger objects. However, accurately capturing the smallest items remains a challenge, requiring specialised equipment and techniques to ensure precise results.

For example, below is a comparison of 3D models of handaxes; a very high-resolution scan (structured light Polymetric PT-M4) on the left and a very low-resolution scan (iPhone 15-pro LiDAR system) on the right.

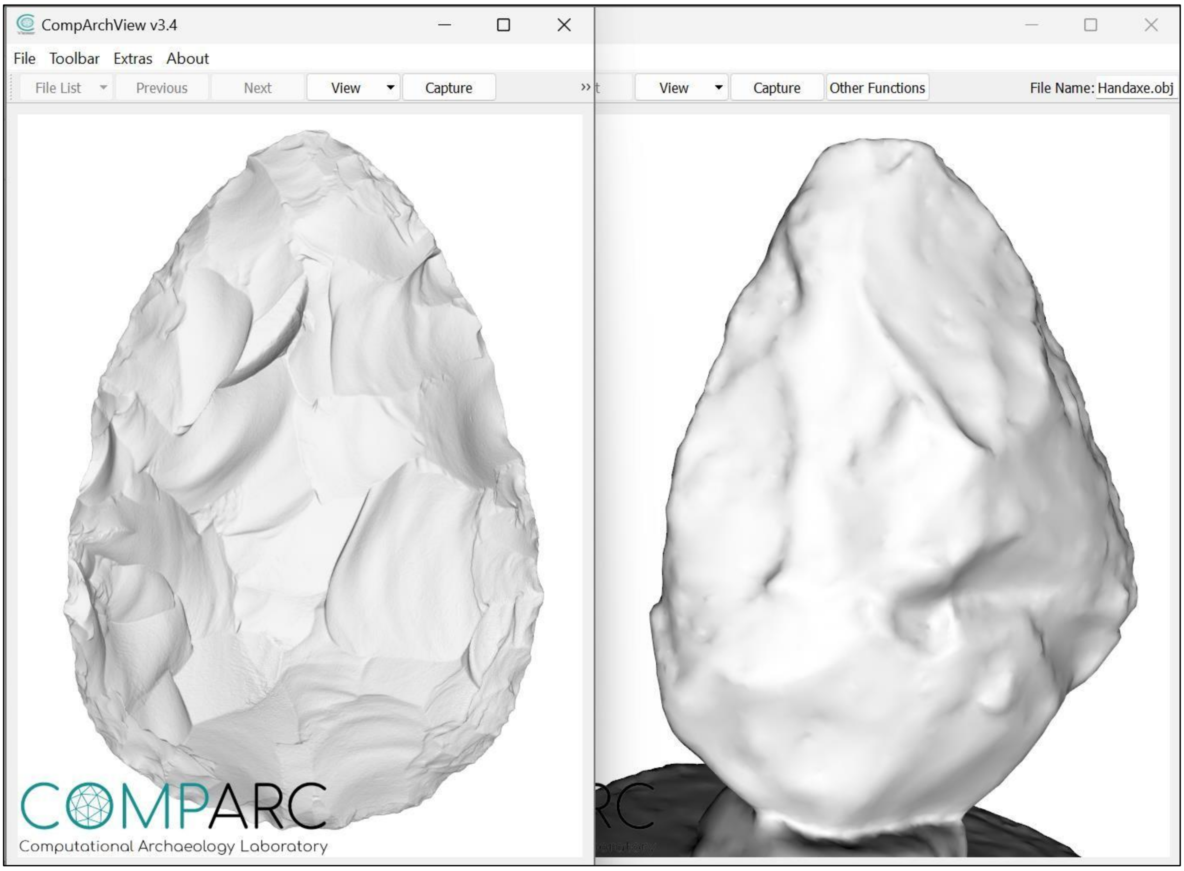

Fig. 1. Full surface.

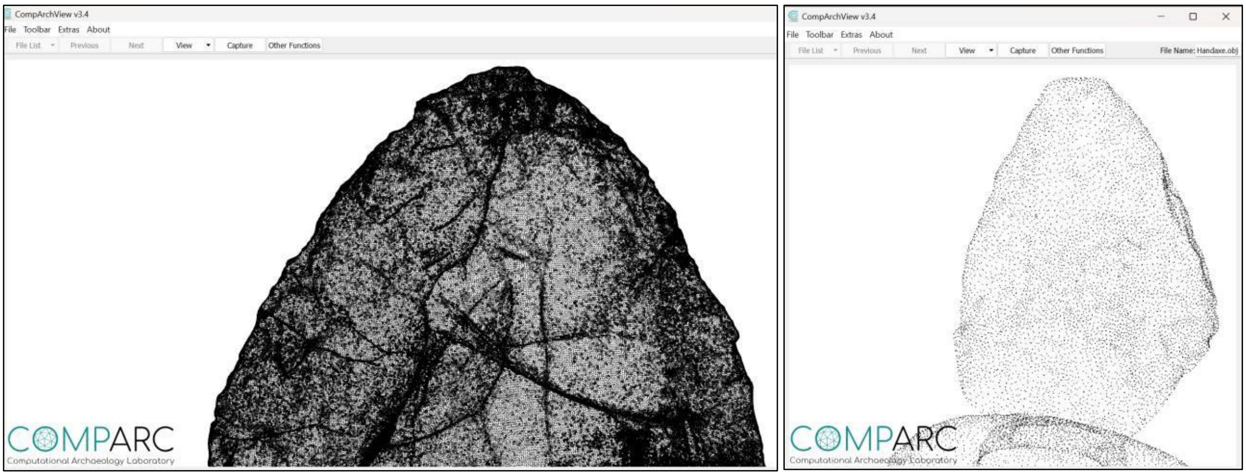

Fig. 2. Point cloud.

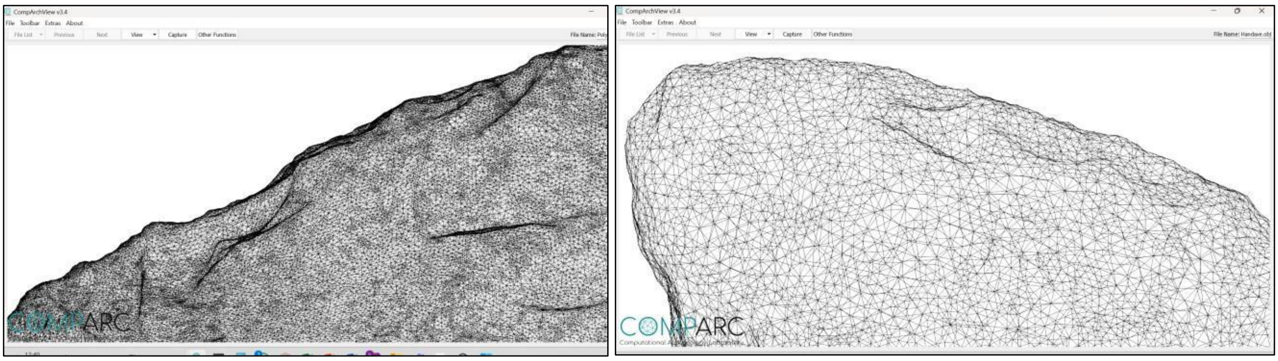

Fig. 3. Mesh.

An additional important factor is the processing software that converts the point-cloud into polygons by triangulating the 3D surface of the object. There are various computational methods available for this phase of 3D model creation, leading to differences in quality between software. For example, the distribution of surfaces can either be uniform in size or vary based on the density of points in the model, reflecting the complexity of the surface. These programs tend to be expensive, and their quality or suitability for archaeological research varies. To best fit the wide range of morphological intricacies of archaeological artefacts (from a practically flat platter or lithic flake to the delicate details of a floral relief or scaled retouch on a lithic tool) the software should have the ability to discriminate between different types of surfaces as based on their micro-topography, and the flexibility to manipulate point-cloud density accordingly.

Additionally, all methods require controlled lighting conditions that differ between methods. Artefact photogrammetry requires bright even lighting to eliminate shading and hue variation to produce visually uniform models. Laser and structured light scanning, on the other hand, require dim to dark lighting conditions so that the return laser/light is properly distinguished and accurately captured. Lastly, resource consumption is extensive for all 3D data capture technologies - 3D scanning takes time and the output is digitally 'heavy', necessitating strong computational power and large storage resources.

1.3.1.1 Publication of pottery and lithic assemblages using 3D models

3D models can be effectively used to publish archaeological assemblages, providing detailed and accurate representations that enhance the dissemination of research findings. The practices of the Computational Archaeology Laboratory (HCAL) at the Hebrew University of Jerusalem (HUJ) exemplify this approach, particularly in publishing ceramic and lithic assemblages, demonstrating how 3D models can support archaeological research and communication.

1.3.1.1.1 Ceramic

The publication of pottery assemblages is a common and routine motivation for 3D scanning in archaeology. The technical illustration of pottery for publication follows established conventions, including an outline drawing of the sherd, its profile, and scale. To enhance this traditional illustration, two views of the 3D model (side and top) are often included (fig. 4). For this purpose, a sample of pottery sherds is selected for scanning, usually focusing on those with an indicative part (rim, base, handle). Thus, only one sherd per vessel is selected and not all sherds of an assemblage are scanned. When working with fragmented pottery sherds, they are usually scanned in bulk (6-8 sherds hoisted on a frame contraption), using a Polymetric PT-M4 structured light scanner with a 16 mm lens, with a scanning run taking approximately five minutes. More complete sherds or entire vessels are scanned individually in at least two positions (two scanning runs) to achieve a comprehensive view of the artefact. The 3D models are constructed using the compatible software (QTSculpture, Polygon Technology by Polymetric GmbH) and analysed using software developed at HCAL - Pottery 3-D (Karasik A. 2010; Karasik A., Smilansky U. 2008). Both the model composition and the illustration processes are a combination of automated and manual steps, with each model requiring an additional processing time of approximately 10-30 minutes. On average, the completion of the entire process takes an hour for 4-5 artefacts.

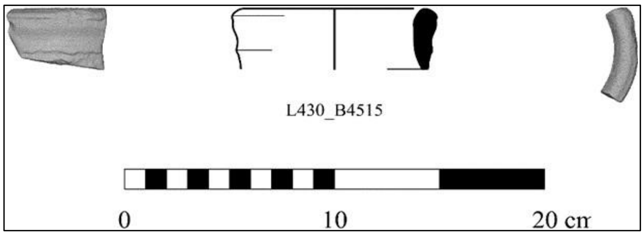

Fig. 4. An example of a technical illustration of a pottery sherd from Lachish in Pottery3-D.

Pottery 3-D calculates a set of metrics and mathematical measurements (i.e. radius, curvature), which may be exported as a spreadsheet, and offers a generalised typological classification of an assemblage employing discriminant clustering analysis (Karasik A., Smilansky U. 2011). This approach requires "medium" resolution and can also be achieved using photogrammetry when real appearance visualisation is requested or when other limitations, such as artefact/scanner accessibility, are imposed (Harush O. *et al.* 2020).

1.3.1.1.2 Lithic artefacts

To produce detailed illustrations, including cross-sections and standard metric measurements (fig. 5), lithic artefacts are typically scanned individually using a structured light scanner (e.g. a Polymetric PT-M4) with lens selection (55-75 mm), varying based on artefact size (smaller artefacts require larger lenses). Pre-treatment of artefacts, such as the application of dulling spray, powder or paint, is often necessary to alleviate scanning limitations caused by reflective, transparent or very dark surfaces. Artefacts are generally scanned in two to four positions to ensure full coverage and accurate capture of thin sharp ridges and edges. A complete scanning run can take approximately 10-20 minutes per artefact. The 3D models are individually constructed using the compatible software (QTSculpture) and analysed using dedicated software(Artifact3-D, Grosman *et al.* 2022, https://sourceforge.net/projects/artifact3-d/). This process combines automated and manual steps and takes an additional 5-10 minutes per artefact, allowing for the completion of 3-4 artefacts per hour in high-resolution mode.

Beyond illustration, the 3D modelling software has multiple analytical functions, such as calculating the centre of mass, volume, surface area, asymmetry, scar segmentation, ridge pattern tracing, and mean edge-angle calculation (fig. 6). This data can be exported as a spreadsheet and analysed to address archaeological questions such as manufacturing technology, engineering, cultural transmission, or cognition (e.g., Muller A. *et al.* 2022, 2023; Richardson E. *et al.* 2014; Valletta F. *et al.* 2021; Yashuv T., Grosman L. 2024).

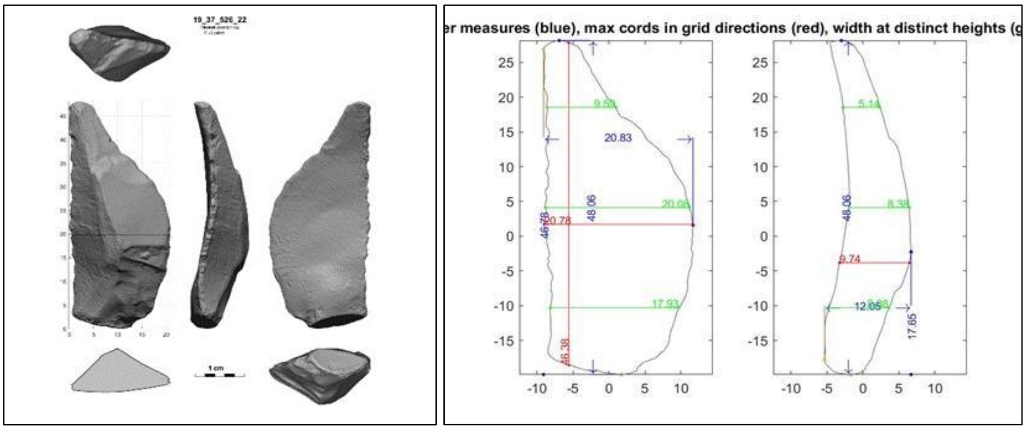

Fig. 5. Scaled illustration, section, and metric measurements of a flint artefact from Tel Beit Shemesh in Artifact3-D.

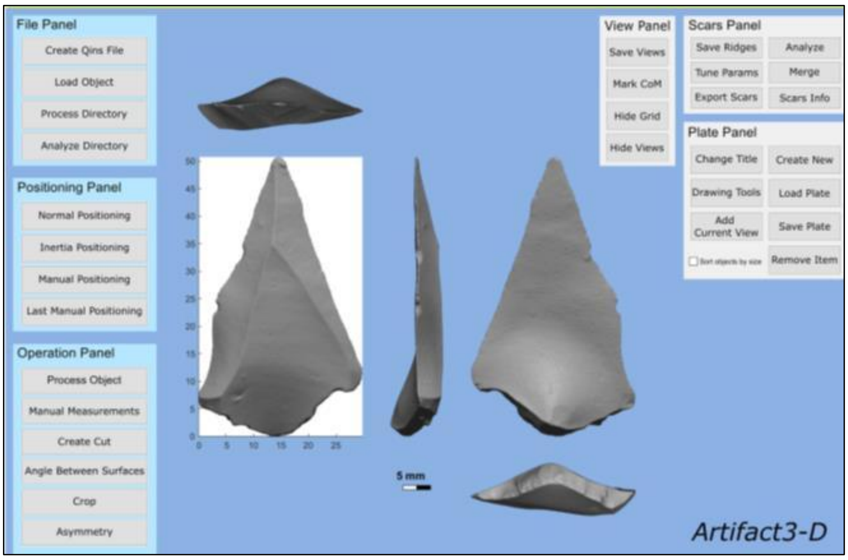

Fig. 6. Artifact3-D software interface showing multiple possible functions.

1.4 Use of portable, non-invasive analytical instruments

Archaeometric analyses of ceramics and lithics play a crucial role in archaeological research, offering insights beyond traditional observational methods. These analyses aim to characterise artefacts' physical, chemical, and mineralogical properties, providing detailed information on their composition, provenance, and technological processes. By examining these properties, archaeologists can address key research questions related to raw material sourcing, manufacturing techniques, trade networks, and the functional use of artefacts. For ceramics, archaeometric analyses can reveal details about the clay composition, firing conditions, and surface treatments, helping to reconstruct production technologies and cultural practices. Similarly, archaeometry of lithic artefacts helps understanding the selection of raw materials, (and therefore reduction techniques) as well as fire treatments or the use of substances for surface treatment or as adhesives.

To understand the importance of these techniques, it is helpful to distinguish between destructive versus non-destructive, and invasive versus non-invasive analyses. Destructive analysis involves diagnostic methods that can compromise the structural and functional integrity of the material, often requiring sampling from the object itself. In contrast, non-destructive analysis avoids material removal, preserving the object entirely intact. In practice, non-invasive analytical techniques frequently employ various forms of energy or particle beams to interact with an artefact, either assessing surface characteristics or, in some cases, penetrating deeper for bulk analysis. The type of energy or particles used and the specific properties measured determine the analysis depth. Techniques focusing on surface analysis, however, must account for potential contamination or historical surface alterations that could obscure an artefact’s proper composition (see paragraph 1.2.1.1). This risk requires careful application and interpretation, as misleading results may occur if surface modifications have affected the original materials. For example, analyses involving particle beams, such as X-rays or laser beams, come with additional considerations, including the depth of beam penetration, the escape depth of signals (which may be reabsorbed within the material before reaching the surface), and the geometric setup of the detection system.

The debate around the use of portable sensors revolves mainly around the question of their conscious use. Users must be informed about what information can be gained from the use of these instruments and what data must instead be discarded because they are unreliable or collected using fallacious procedures. Among the most frequently used portable techniques in conservation are X-ray fluorescence (XRF) and Raman spectroscopy. These methods are popular due to their mobility and ease of use. However, both XRF and Raman spectroscopy are sensitive to surface contamination and can be affected by geometrical constraints, which may lead to more qualitative rather than quantitative results. Careful application and interpretation are necessary to account for these limitations, as they can impact data quality. Despite these challenges, the rapid, accessible insights provided by portable XRF and Raman instruments have become invaluable, enabling researchers to conduct detailed, site-specific analyses while safeguarding cultural heritage artefacts for future generations.

1.4.1 Methods and techniques: principles and setting requirements

For the enhanced digitisation of artefacts within the AUTOMATA project, specific non-invasive techniques will be employed. The primary method is hyperspectral imaging (HSI), utilising either miniaturised sensors or commercial devices such as the Specim IQ camera (Sciuto C. *et al.* 2022). Data from the initial HSI screening will refine the analytical model and guide decisions on further analysis, considering factors such as the artefact's surface characteristics, the area targeted for examination, and the capabilities of the available sensors. If these conditions are met, additional analyses may be carried out. For example, portable X-Ray fluorescence (p-XRF) can provide insights into the composition of ceramic or lithic materials, while Raman spectroscopy can be used to identify surface pigments.

1.4.1.1 Hyperspectral Imaging (HSI)

A hyperspectral camera works by capturing the reflectance spectrum at each pixel of an image, revealing the wavelengths of visible and near-infrared light reflected from the object. High reflectance results in strong intensity peaks, while greater absorption shows lower intensities in the spectrum. Comparing these reflectance spectra to established databases often allows for the identification of specific material. HSI is a portable, non-invasive technique that preserves the integrity of samples, though it requires a controlled lighting environment for optimal imaging conditions. Its effectiveness is especially notable with flat surfaces, which aid in the accurate acquisition and processing of images.

HSI offers significant benefits in non-invasive analysis, making it particularly suited for cultural heritage applications. HSI in the spectral range spanning from the Near Ultraviolet (NUV) to the Near Infrared (NIR) is a widely recognized technique employed across various scientific and industrial domains, including pharmaceutics, agriculture, food production, and geology. The interaction of molecules with infrared radiation can be observed in distinct regions of the infrared spectrum: near-, mid-, and far-infrared. VIS-NIR spectroscopy (700–1700 nm) is particularly advantageous for bulk sample analysis, requiring minimal or no preparation. The VIS-NIR spectra feature broad absorption bands resulting from molecular overtones and combination vibrations. Organic molecules typically exhibit vibrations in their second and third overtones, making VIS-NIR spectroscopy especially effective for analysing organic materials across multiple applications. Additionally, accurate mineral classification is often achieved in the Short-Wave Infrared (SWIR) range, extending up to 2500 nm. The technology's speed and ease of use allow it to efficiently capture data from large surfaces and numerous objects. HSI is especially effective for studying decorated surfaces, as it facilitates mapping and distinguishing artistic materials and polychromatic patterns. Each pixel in an HSI image records reflectance spectra at different wavelengths, creating a detailed "data cube." Extracting meaningful insights from this data-intensive format requires multivariate analysis. However, recent advancements in data processing routines have enhanced both visualisation and classification, reducing computation time and complexity (Liu L. *et al.* 2023).

1.4.1.1.1 Setting: analytical and physical issues

The Specim IQ camera that will be used in the AUTOMATA project (see paragraph 2.2.3) is marketed as the first ultraportable compact hyperspectral camera. While it has primarily been used in agriculture and food analysis (Behmann J. *et al.* 2018), it has recently been adapted for applications in archaeology (Sciuto C. *et al.* 2022). The camera weighs 1.3 kg and has 207 x 91 x 74 mm dimensions. It features a CMOS sensor capable of capturing images in the 400–1000 nm wavelength range, with a spectral resolution of 7 nm and a spatial resolution of 512 x 512 pixels per image. The camera is versatile and suitable for both indoor and outdoor use, relying on controlled lighting such as halogen lamps or natural sunlight. Calibration is conducted using a

Spectralon tile placed near the target during image acquisition. The IQ camera's software enables recording a white reference in one image and applying it to subsequent acquisitions, which is especially beneficial in controlled lighting conditions. The artefacts can be positioned on a surface with two halogen lamps set at a 45-degree angle and the IQ camera in the middle. Images should be ideally acquired at a distance of about 50 cm from the target (but it could be less) and the spectralon tile should be always placed close to the objects (see fig. 10 and paragraph 2.2.3). Usually, one image is enough for each sample, and acquiring one image requires less than 30 seconds, considering the lighting conditions described above (see paragraph 3.2).

1.4.1.1.2 Data processing

After processing the monochromatic layers of the datacube — a multidimensional matrix in which each pixel corresponds to a continuous spectrum — the final hyperspectral image is reconstructed, with each pixel containing the reflectance spectrum of the corresponding area of the object.

Reflectance spectra acquired using an HSI camera can be processed to generate conventional images, such as RGB or False-Color Infrared, or subjected to advanced statistical treatments tailored to the target material and research objectives. Examples include algorithms for enhancing wall paintings (Legnaioli S. *et al.*, 2013; Grifoni E. 2019; Triolo P. *et al.*, 2020) or noise reduction techniques to address scattering effects (Sciuto C. *et al.* 2018). The images produced by ultraportable HSI cameras often result in large file sizes, requiring various approaches, software, and routines for effective processing and analysis. While Specim's IQ Studio software, available for free download, offers basic tools for exploring spectral information within the hyperspectral dataset, using third-party software or custom programming routines for practical applications is typically necessary. Processing HSI data involves several stages to optimise file size without compromising spectral accuracy. Multivariate Image Analysis (MIA) is a widely used method for analysing hyperspectral data, with Principal Component Analysis (PCA) being one of the most common models. PCA reduces the multidimensional hyperspectral data to a two-dimensional space, classifying individual pixels based on their spectral features and relationships. The resulting data is visualised as a colour composite, where each pixel is assigned a new value corresponding to the highlighted PCA components. Commercial software like Evince by Prediktera supports multivariate analysis of hyperspectral images, while programming workflows using libraries such as Spectral Python allow for importing and processing spectral images in ENVI format, offering flexibility for advanced analysis.

1.4.1.2 Portable X-ray fluorescence (p-XRF)

Portable X-ray fluorescence (p-XRF) represents a transformative tool in archaeological studies, particularly for ceramics and lithics analysis. Enabled by advancements in miniaturised computer and detector technology, p-XRF devices are compact, handheld, and, most importantly, non-invasive. Their portability allows them to be brought directly to objects in the field, providing elemental composition data without damaging the artefacts. By exposing the object's surface to an X-ray beam, p-XRF instruments produce a spectrum where the emitted X-ray energies identify specific elements present.
In ceramic studies, p-XRF can determine elemental concentrations in surface layers, such as glazes or paints (e.g. Belfiore C.M. *et al.* 2021). However, due to shallow X-ray penetration, it may not provide information on the underlying ceramic body. Interference is a concern, particularly with thin or unevenly applied surface layers, common in ancient ceramics, where X-rays might also excite atoms from underlying materials, potentially skewing results. The technique works best on homogeneous, fine-grained surfaces, though these conditions are rarely met in archaeological samples.

The field has seen tremendous improvement in p-XRF's capabilities over recent years, with enhanced component materials and detector technology expanding the range of elements measurable. Advances like the use of gold anodes and graphene windows allow higher voltages and detection of lighter elements, while modern detectors process X-rays far faster than older models. Such developments mean that contemporary p-XRF units can, in some cases, outperform older benchtop systems, making them invaluable in archaeology. However, debates persist on p-XRF's precision, accuracy, and comparability across devices. These concerns are not so much about the technology itself as about user expertise. Analysts require a solid understanding of analytical procedures, calibration standards, and elemental data interpretation. This has sparked discussions within the archaeological science community about research design, proper calibration, and the importance of disclosing all analytical protocols to ensure replicability (Frahm E. 2024).

Despite these challenges, p-XRF has democratised elemental analysis in archaeology, offering accessible, quick, and non-destructive insights into the composition of artefacts. Although some experts urge caution regarding overreliance on p-XRF in heterogeneous materials and encourage rigorous training, the technique is widely acknowledged as a valuable addition to archaeological science when applied thoughtfully.

1.4.1.2.1 Surface Issues

For XRF analysis to work effectively, X-rays must escape the specimen's surface without being reabsorbed. Only the X-rays within a certain depth from the surface can contribute to the measurement, making XRF a surface-sensitive technique. The depth of X-ray penetration depends on the material's density and the energy of the X-rays; for example, X-rays from heavier elements escape from deeper layers than those from lighter ones, but this penetration is generally shallower in denser materials like metals.

This characteristic of the analytical technique can lead to the production of hybrid spectra on coated ceramics. In the case of glazes, for example, the beam will likely penetrate beneath the surface layer, capturing information that also pertains to the ceramic body itself. The same applies to painted or slip-coated ceramics. A possible solution is to lower the energy used for the beam so that the laser's penetration into the material is also lowered.

Another issue related to the surface condition of the pieces is the formation of encrustations or patinas (see paragraph 1.2.1.1). These surface alterations persist even after washing the pieces and are particularly common on ceramic fragments due to their porosity. The presence of surface alterations can complicate or even render surface analysis impossible unless they can be removed mechanically (by scraping). The following images provide examples of two particularly significant pieces to illustrate the issues mentioned. In fig. 7, a glazed ceramic fragment with a surface patina on its exterior is shown. In this case, an analysis of the area affected by the patina (fig. 7, a) would yield geochemical data influenced by the patina itself and not reflective of the composition of the ceramic body. On the other side of the fragment (fig. 7, b), an analysis of the glaze would likely produce a hybrid spectrum, containing elements of both the glaze and the ceramic body.

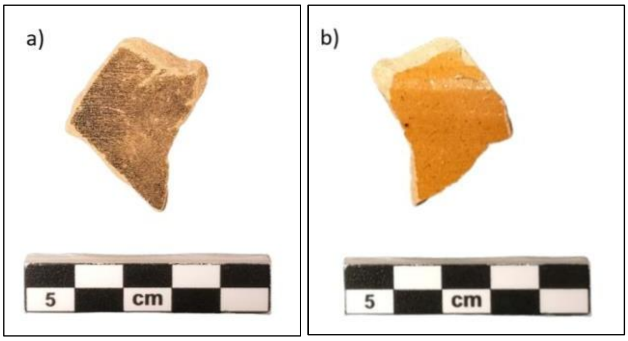

Fig. 7. Glazed ceramic fragment from S. Sisto excavation (Pisa, 2022-2023): a) exterior b) interior.

In fig. 8, another fragment is also of glazed ceramic. While the glaze-covered surface presents the same issue observed in the other fragment, the outer surface of this fragment shows a section covered by patina and a cleaner portion. In this second case, for XRF analysis to yield reliable results on the composition of the ceramic body, it must be conducted in an area unaffected by surface alterations.

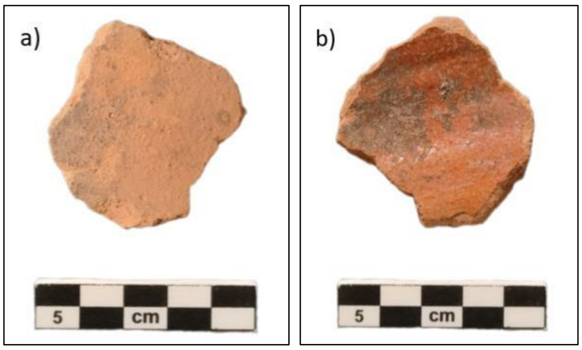

Fig. 8. Glazed ceramic fragment from S. Sisto excavation (Pisa, 2022-2023): a) exterior, b) interior.

1.4.1.2.2 Placement of Objects and Shape Issues

Optimal results require specimens to be placed carefully relative to the X-ray source and detector. Angling affects measurement, as the X-ray beam enters at a tilt, not perpendicularly, which alters the effective depth of analysis. Concave surfaces or uneven placement can disrupt accurate measurements, while convex or flat surfaces yield better results by maintaining a more consistent positioning relative to the detector. In general, it is essential to remember that the analyser window must be completely covered during analysis to prevent beam dispersion. This requirement has practical implications, limiting the size of objects that can be analysed. Specifically, objects smaller than approximately 2 x 1.5 cm — the average size of an XRF analyser window — are excluded. This applies, for instance, to many lithic artefacts that do not reach these dimensions or to small ceramic fragments. It also restricts the analysis of ceramic sections, as their thickness rarely reaches the necessary size to cover the analyser window fully. Collimators can narrow the beam for small areas, though they reduce X-ray intensity and require specific calibration adjustments.

Specimen shape can significantly impact XRF accuracy. Flat, smooth surfaces are ideal for minimising error. Tests have shown that concave or irregular surfaces can lead to decreased X-ray intensities, particularly for lighter elements, as surface morphology disrupts the ideal interaction with the X-ray beam. Surface corrections can partially address these issues, but analysing irregularly shaped artefacts remains challenging.

1.4.1.2.3 Number of Measurements

Accurate results may require multiple measurements to account for the material's heterogeneity. Using instruments with larger beam diameters can produce more representative readings, especially for coarse or heterogeneous samples. Corrective factors can help standardise results across different beam diameters and taking a higher number of measurements allows for an average composition that approximates bulk analysis. Generally, taking three measurements for each sample is preferable to obtain a reliable dataset and quickly identify any outliers.

1.4.1.3 Raman spectroscopy

Raman spectroscopy is a powerful technique for analysing molecular structures through vibrational movements within molecules. The Raman effect, involving the scattering of radiation and subsequent wavelength shifts, provides detailed insights into material composition. This non-invasive method has become essential in cultural heritage research, allowing the identification of materials directly on artefacts without invasive sampling (Rousaki A., Vandenabeele P. 2021; Vandenabeele P. 2013).

Portable and handheld Raman spectrometers have significantly enhanced the ability to perform on-site analysis of archaeological artefacts (Vandenabeele P. *et al.* 2004). These devices use laser illumination to generate Raman photons, revealing a material's molecular composition. They are commonly used for rapid verification by comparing the sample's spectrum with reference libraries. These instruments are widely used to study pottery and lithic materials providing information about pigments, glazes, and mineral compositions. For ceramics, Raman spectroscopy helps determine firing temperatures, characterise different clay sources, and distinguish production techniques. For instance, it has been successfully used to identify hematite, magnetite, and other iron oxides in ancient pottery decorations, providing insights into the firing conditions and pigment choices of past artisans (Jehlička J., Culka A. 2022; Colomban P. *et al.* 2006). For lithic studies, Raman spectroscopy helps identify the mineralogical composition of stone materials, as well as alteration caused by heat treatment or environmental exposure (Hernández V. *et al.* 2012; Jehlička J., Culka A. 2022).

1.4.1.3.1 Placement of Objects and Shape Issues

As with XRF, Raman spectrometer analyses also face challenges related to sample positioning to avoid beam dispersion. However, Raman analysis offers a wider variety of portable instruments and detector configurations. Nevertheless, for this technique as well, it is essential that the sample is in contact with the analyser window, fully covering it to ensure accurate readings. This requirement highlights the importance of correct sample alignment and complete coverage, as incomplete coverage can compromise data reliability by allowing beam loss, thus potentially skewing the results.

1.5 Metadata

The effective use of metadata in cultural heritage plays a fundamental role in promoting data sharing and reuse, which are essential for advancing research, preservation, and public engagement. In digital archaeology, metadata provide a structured way to document and interpret archaeological ceramics and lithics, facilitating their integration into broader analytical frameworks. These descriptive and procedural datasets enable researchers to preserve information about artefacts' contexts, digitisation processes, and analytical results, thereby ensuring the interoperability and longevity of digital assets. As metadata practices evolve, they increasingly support innovative approaches, such as digital twins, enhancing the accuracy and accessibility of archaeological knowledge while addressing challenges associated with data completeness and quality.

The term "digital twin" originates from industrial and CAD applications, where it describes a virtual prototype enabling full control over parameters. However, Heritage Digital Twins (HDT) differ significantly. An HDT involves creating a digital copy of a physical object, often without complete visibility or control over the captured information, raising challenges in ensuring the completeness and quality of the data. This requires careful decisions, particularly regarding resolution. A critical element distinguishing an HDT from a basic digital copy is its metadata scheme, which must be precisely defined and interoperable, as demonstrated by ARIADNE, CIDOC-CRM, and frameworks developed by the Consortium 3D for Digital Humanities and the National 3D Repository for Humanities (Quantin M. *et al.* 2023; Tournon S. *et al.* 2021). Metadata must include self-assessments covering the archaeological and historical context, the digitisation process with all associated parameters (paradata), and the choices made during the process, as referenced in the VIGIE Report. Metadata play a pivotal role in the FAIR principles (Wilkinson M.D. *et al.* 2016) and Heritage Digital Twin frameworks (Jouan P., Hallot P. 2020; Niccolucci F. *et al.* 2022, 2023), making them central to the ECCCH's ECHOES project. Specific analytical methods require tailored metadata, such as the CIDOC-CRM ontology proposed in ARIADNE for archaeological data (Fihn Marberg J. *et al.* 2022) or the Europeana Data Model. The consortium "3D for Digital Humanities" has introduced an extendable metadata scheme for the National 3D Repository for Humanities (Quantin M. *et al.* 2023), compatible with the Archaeology Data Service (ADS). These frameworks assist practitioners in planning, data collection, processing, and long-term curation, offering detailed guidelines for format selection, equipment settings, and documentation.

2 AUTOMATA requirements

The following paragraph reflects the outcome of discussions held with all consortium members, combining input from both technical and archaeological partners. These collaborative exchanges were essential for identifying user requirements and defining the needs related to the documentation and digitisation of ceramic and lithic artefacts within the project framework.

Given the working conditions and protocols generally used for documenting and studying ceramics and lithics, in this paragraph we concentrate on defining the project's specific requirements. These requirements are primarily driven by research questions: What is this object, and what material is it made of? For instance, the resolution or quality of the data, and consequently the technology used for data acquisition, will be determined and adjusted based on the archaeological research questions or goals, as well as the artefact's size and morphological intricacy.

A shared perspective and common specifications are essential, even if our research contexts differ. AUTOMATA aims to automate a digitisation pipeline that integrates various material treatments using different sensors into a single process. This process, managed by a robotic arm and supported by AI, seeks to streamline and enhance the digitisation workflow. Additionally, the user requirements must account for contextual or institutional constraints, such as those posed by laboratory or museum environments. These include technical, legal, and ethical constraints identified by project partners, contributing to a shared reflection on the system's design requirements. Such constraints are fundamental to ensuring the system's adaptability and effectiveness across diverse working contexts.
Artefact accessibility and legal constraints are key topics, as institutions face varied bureaucratic and legal restrictions on accessing and moving artefacts. The level of operational autonomy is therefore considered, assessing how much human oversight is necessary and what training staff may require to operate the system effectively. Concerns regarding artefact handling and movement address the risks of transporting artefacts during digitisation. This connects to portability and environmental requirements, exploring the design of a movable system. Technical aspects such as system setup — whether a conveyor belt or robotic arm is more suitable — and sensor specifications are also discussed, including the types and mobility of sensors needed to capture data effectively. Additionally, quality assurance criteria ensure the accuracy of digitised data, while data volume estimation considers storage needs for enriched 3D models.

It is important to highlight a distinction between the traditional users who typically conduct artefact analysis and digitisation and the broader group of stakeholders identified by the AUTOMATA project. Traditionally, these tasks are carried out by specialised professionals who invest many hours in the manual or semi-manual creation of 3D models and archaeometric analysis of objects. In contrast, the AUTOMATA project aims to automate parts of this process, thus expanding the system's potential user base beyond these specialists. By introducing automation, AUTOMATA broadens access to high-quality digitisation and analysis, allowing a more diverse range of stakeholders to engage with the system.

Key stakeholders for AUTOMATA include academics, museum professionals, heritage supervisory bodies, students, and professional archaeologists. Each group will interact with the project's outputs in ways tailored to their needs. Academics will use the data for research, education, and outreach, while the system also holds significant potential for teaching purposes, particularly in demonstrating how movable instruments can generate valuable insights into cultural heritage. Ministries and heritage bodies will utilise the data for heritage protection, preservation initiatives, and public communication, supporting responsible cultural heritage stewardship. In museums, the digitised data and instruments will enable curators, archaeologists and researchers to improve artefact processing, inventory management, analysis, publication, and educational outreach.

The system, designed to be mobile, will be deployed in museum settings, research centres, and storage facilities, allowing professionals to capture high-quality data directly on-site. By facilitating broader and easier access, AUTOMATA's approach enhances the usability of digitised data and tools across a wider array of settings and user groups than previously possible.

2.1 User requirements

2.1.1 Contextual or institutional constraints

Common user requirements can be easily identified, as they are linked to widespread research questions shared by many archaeologists, providing a solid foundation for collaboration. However, there are also contextual and/or institutional constraints that must be addressed, which introduce additional requirements.

- In the museum research context, these constraints can be listed:
  - Selection of material: the process of defining the need for obtaining specific additional data, its purpose and its usability;
  - Storage and accessibility of original material;
  - Digitisation: labour and time consumption;
  - Data input and linking (as a process of connecting digital items/models with textual/tabular data/metadata);
  - Storage and management of digital data (quantity and size of data, adequate database, its maintenance and updates);
  - Long-term accessibility, preservation and searchability/retrieval of data (question of all FAIR data principles).
- In the preventive archaeological research context, we can underline mainly the following constraints:
  - Digitisation time;
  - Portability of the system;
  - Management of the mass of documentation produced.
- In laboratory research context:
  - Need for shortening digitation time (not only digitation, the entire process is time-consuming);
  - High cost of archaeometric analysis;
  - Democratisation of these types of analysis and the knowledge produced by them;
  - Difficult access to European archaeological material (for extra-Europe users).

2.1.2 Artefacts accessibility

Accessibility to archaeological artefacts for digitisation varies among institutions, with some facing bureaucratic and logistical challenges. In some cases, moving artefacts requires specific authorisations and certain objects cannot be relocated. This limitation highlights the need for a highly movable digitisation system that can be deployed directly within storage facilities, where Internet connectivity may also be limited. A flexible, standalone system capable of operating offline would address these accessibility challenges, ensuring that digitisation can occur without moving sensitive artefacts or relying on continuous online access.

Generally, few legal or political barriers to artefact access as long as digitisation occurs on their premises. This ease of access allows smoother workflows in institutions where artefacts are readily available and internet connectivity is stable. However, to accommodate varying degrees of connectivity and regulatory constraints, the project emphasises the need for a system that can be transported across different storage locations and operated in diverse institutional environments, both within and outside the EU. An effective approach is to adopt an "agile" working method, starting with specific sensors to ensure consistent data acquisition during the initial stages. However, by designing the system to be scalable and adaptable, it becomes possible to integrate additional or alternative sensors as needed over the course of the project. This flexibility aligns with the project's long duration and evolving requirements, ensuring that the digitisation process remains accessible and adaptable to various artefact storage conditions and institutional regulations while supporting efficient data collection in diverse environments.

2.1.2 Legal, political barriers and ethical considerations

Legal and ethical considerations in the AUTOMATA project play a crucial role in guiding how artefacts are accessed, digitised, and shared, ensuring both compliance with regulatory frameworks and adherence to responsible digitisation practices. Legal barriers related to Cultural Heritage digitisation can vary significantly by region. In some cases, restrictions may apply to the dissemination of cultural heritage images, requiring adherence to specific licensing conditions. For example, images may be shared for research and non-commercial purposes under a CC-BY-ND-NC licence when offered as open access. In other instances, the sharing of 3D models may require prior authorisation from national agencies and be limited to a CC-BY-NC licence.

Archaeometric data generated during the project will be owned by AUTOMATA and openly shared under a CC-BY licence. However, the project may need to consider whether archaeometric data embedded within 3D models should be handled separately to ensure compliance with regulatory and contextual requirements for specific uses.

Ethically, the digitisation process encompasses several stages — from selection and evaluation to metadata collection and digital distribution — each of which can raise unique challenges. Potential ethical issues include the manipulation of digital content, biases in selection and interpretation, and concerns over privacy, access, and authenticity. To address these risks, the AUTOMATA project will align its design with ethical principles such as autonomy, non-maleficence, beneficence, justice, and explicability, as outlined by the High-Level Expert Group on Artificial Intelligence (HLEG, 2019). A careful, principled approach is essential to ensure the integrity, transparency, and fairness of the digitisation process across all regions and cultural contexts. By proactively addressing these legal and ethical dimensions, the AUTOMATA project aims to foster an environment where digitised data on Cultural Heritage can be shared responsibly, respecting both local regulatory standards and broader ethical considerations to maintain trust and accountability in the digitisation of cultural heritage. It is also important to reflect on the potential licensing options and the implications for data reuse, particularly given the risk of data extractivism approach to data. While datasets may remain open, the meta-value they generate — such as the models created — can become private property, raising questions about equitable access and the balance between openness and commercialisation.

2.2 System requirements

2.2.1 System portability and operational autonomy

The AUTOMATA project's digitisation system will be designed with a nuanced approach to portability, responding to stakeholders' needs for artefact accessibility across various storage and study environments. Rather than emphasising traditional portability, the system will be developed to be movable, allowing it to be transported between buildings, storage facilities, and archaeological centres to enable on-site digitisation of artefacts, with the goal of being used even in institutions where moving artefacts requires formal authorisation or is restricted. The system should be compact enough to fit through standard doors and elevators and will likely be mounted on a wheeled platform to facilitate movement within controlled indoor spaces. Its dimensions and mobility will be determined during the design and prototyping phases to ensure alignment with user needs and functional stability.

Environmental adaptability will also be a key design consideration. The system will need to function effectively in moderately dusty environments typical of storage facilities and archaeological centres. It will be built to operate within a temperature range of 10°C to 50°C and 30% to 70% relative humidity and will feature a built-in buffer battery to protect the electronic devices from fluctuations and brief interruptions of the power supply. Additional design features, such as shielding to protect operators from radiogenic sources and integrated environmental monitoring sensors, will enhance its suitability for diverse working conditions. However, the system will not be intended for use in active excavation sites, as outdoor environments present challenges such as variable lighting, higher dust exposure, and weather instability. Instead, the design will prioritise a flexible and resilient mobile digitisation cell optimised for indoor archaeological and storage settings. This approach will support efficient and accessible digitisation across museums, research centres, and other controlled environments, with consideration for regulatory compliance across various regions, including both EU and non-EU countries.

The AUTOMATA digitisation system will be designed to operate with a level of partial autonomy that requires active, but not continuous, oversight by trained personnel. While the system can automate significant portions of the digitisation workflow, it will necessitate an operator with appropriate archaeological training, who can prepare artefacts for scanning, monitor system functions, insert metadata, and address any alerts or malfunctions. The operator will play a crucial role in setting up and organising artefacts, ensuring they are clean and properly arranged, and entering key data points such as artefact inventory numbers and excavation context (where, by who, how), minimal characteristics (SU n°, size, weight, material, etc.). This active human oversight will support system efficiency and data integrity, as the operator can intervene in the workflow to manage any procedural adjustments or data validation needs that arise. Furthermore, the system will require a structured protocol for operator interaction, allowing manual input where necessary and minimising direct sensor contact to protect equipment. Additionally, the human operator’s involvement will be essential for ongoing AI learning, contributing expertise in both archaeology and technology to guide the system’s adaptation to artefact variability. This human-in-the-loop approach ensures that the system can meet high standards of precision, reliability, and adaptability while enhancing the digitisation process’s overall effectiveness.

2.2.2 Artefacts’ handling and system type

The AUTOMATA system will need to handle a wide range of archaeological artefacts, primarily focusing on ceramics and lithics. These objects vary significantly in size, weight, and state of preservation. Most artefacts are expected to fall within manageable parameters — which will be better defined during the prototype testing phase, but for now, weight limits of a maximum of 1 kg and size constraints approximately between 10 and 120 mm will generally be sufficient. Archaeological artefacts could vary from intact objects to fragmented pieces, each requiring specific handling approaches. For example, ceramic pieces may range from just a few centimetres to several tens of centimetres, with weights spanning from a few grams to multiple kilograms. Typical excavation contexts can yield hundreds of artefacts packed in 20 to 40-litre containers, emphasising the need for a system capable of efficiently processing numerous items. The handling system will also need to accommodate extremely small artefacts, depending on the desired resolution and analytical targets. The system must be gentle enough to avoid damaging non-durable or perishable artefacts while capturing high-resolution details that support archaeological analysis. This versatility will allow the AUTOMATA system to handle thousands of artefacts for AI training and analysis, with extensive data collection facilitated offline through photographs, spectra, and other non-contact methods. The handling and movement of artefacts during the digitisation process is a critical concern for the AUTOMATA project, as it requires a balance between precision and safety to prevent any risk of damage. The digitisation system's robotic arm must be carefully designed to minimise contact pressure and avoid scratching or breaking delicate artefacts. This will involve tailoring the system’s gripping strength and positioning based on the specific characteristics of each artefact type, with particular attention to fragile materials such as pottery, which can be susceptible to edge damage, and small, sensitive lithics. Additionally, since the arm should accommodate various artefact sizes, from standard pottery dimensions to smaller flint pieces, it may necessitate dynamic gripping positions and flexible rotational axes to capture all visual angles without compromising artefact stability.

To further enhance artefact safety, the project will implement standard precautionary measures, such as pedestals or support wedges, to stabilise items during handling. While manual handling may be preferred initially for particularly fragile items, the system will include a restraint setup to protect artefacts as they are positioned for scanning. The handling system will also address practical concerns, such as dust management, to maintain equipment and sensor functionality. Although engineering limitations mean it may be impossible to guarantee absolute artefact safety, a rigorous assessment of acceptable handling standards will be established, ensuring that potential risks are minimised and limited to a very low mishandling rate.

In designing the AUTOMATA system, two primary options for artefact handling have been proposed: a conveyor belt and a robotic arm. Each system offers distinct advantages, with the conveyor belt providing efficient, linear movement and the robotic arm offering flexible, precise handling.

The conveyor belt system is designed to transport artefacts sequentially between stations, where sensors or cameras capture data. Artefacts placed on the belt would move along a fixed path, passing through different stations as needed for digitisation. This setup would allow for rapid processing, which is especially useful when handling a large volume of artefacts in a structured sequence. However, one limitation of the conveyor system is its inability to reorient artefacts dynamically; objects would generally stay in a fixed position as they move, requiring sensors or cameras to remain stationary. This lack of flexibility could pose challenges for capturing complex shapes or surfaces from all necessary angles.

The robotic arm system, on the other hand, is designed for a more adaptable approach to artefact handling. A robotic arm would pick up each artefact individually, positioning it within an acquisition area equipped with a transparent surface. Cameras placed above, below, and on the sides would capture a comprehensive view of the artefact without needing to move it to different stations. The robotic arm can also reorient artefacts as needed, ensuring that every surface and angle is thoroughly scanned. After initial imaging, the arm could move the artefact to additional areas for further analyses and return it to the original container once digitisation is complete. This flexible positioning and reorientation capability makes the robotic arm more suitable for capturing intricate details and handling artefacts of various shapes and sizes.

While both conveyor belts and robotic arms have been proposed, testing will prioritise a robotic arm-centric approach due to its flexibility, precision, and suitability for handling a diverse range of artefacts. This choice aligns with the project's goals of high-quality, comprehensive data capture, ensuring that the system can effectively accommodate various artefact types and digitisation requirements.

A hybrid configuration combining a conveyor belt with robotic arms at specific stations has also been considered. This setup would allow the conveyor belt to transport artefacts between stations, where robotic arms could reorient and position each item for optimal data capture. However, given the additional complexity and space requirements, the consortium views a primarily robotic arm-based approach as more feasible and adaptable.

Another focus during testing will be the computational complexity associated with the robotic arm's reorientation capabilities. Ensuring that the system can efficiently handle the required computations for precise artefact positioning and imaging will be critical. System testing will assess the arm's ability to provide accurate positioning across different artefacts and contexts, weighing these benefits against the computational resources involved.

The design of the arm and the gripping device are also functional for interaction with other sensors involved in the digitisation process. For example, when specifying the integration of 3D modelling into the AUTOMATA robotic system, an important consideration is that all methods require a full, all-around view of the object. Regardless of whether the artefact or the sensor is mobile, the gripping device or placement surface will inevitably obscure a portion of the artefact at any given time. This necessitates at least one change in grip or position during the scanning process. Using a transparent placement surface or enclosure to capture 3D data has been proposed, but this approach requires careful evaluation. While it may provide unobstructed views, it could introduce distortions in the captured data. To address these challenges, the software must incorporate robust segmentation capabilities, accurately distinguishing between artefact and non-artefact elements (such as the gripping device or surface) and automatically removing undesired parts from the final model. These considerations are critical for ensuring the precision and reliability of the digitisation process and will be thoroughly analysed in D2.3 (*System Specification*).

2.2.3 Sensors specifications and requirements

The AUTOMATA project enhances digitisation by integrating 3D modelling with portable, non-destructive, and non-invasive archaeometric instruments. This approach is designed to study, digitise, and preserve ceramic and lithic artefacts while ensuring minimal physical interaction with fragile and valuable objects, safeguarding their integrity during in-depth analysis. Recent advancements in miniaturisation make these tools practical for on-site use, enabling immediate results and reducing the need for artefact movement. The system prioritises adaptability over achieving maximum accuracy in every scenario, tailoring methods to archaeologists' specific needs and requirements. It also supports integrating emerging sensor technologies like hyperspectral imaging (HSI) with smartphone lenses (Stuart M.B. *et al.* 2021). While these solutions may offer lower accuracy compared to advanced systems, they are well-suited for initial screenings, providing preliminary assessments of artefacts' visible and physico-chemical characteristics and facilitating the efficient collection of large datasets. By creating comprehensive yet scalable digital records, the system enables detailed analyses to be conducted selectively on artefacts of particular interest. This approach not only meets preservation and research objectives but also supports the efficient digitisation of diverse archaeological collections. The project's flexible methodology ensures the system can adapt to a wide range of contexts, laying a solid foundation for both current needs and future advancements in archaeological research.

The following Table provides an overview of the sensors available to the project, highlighting their raw data output formats and interfaces (physical and software) and demonstrating how they can be integrated into the system to effectively achieve the project's digitisation objectives.

Sensor	Producer	Team	Raw data output format	Interface (physical and software)
iPhone LiDAR and photo scanning 	Apple	UNIPI; HUJ	- Point Cloud Data: captures spatial information as a dense collection of points representing the 3D structure of an object or environment. - Typically exported in standard 3D formats such as .obj, .usd, or .ply, depending on the app used.	- Phone screen: phone's touch interface to control scanning and adjust settings. - Reality Composer App: Apple's app for creating, visualising, and exporting 3D data. - Compatibility with third-party apps like Polycam, 3D Scanner App, or Scaniverse for enhanced functionality and diverse output options.
Polymetric PT-M4 3D scanner	**Polymetric**	HUJ	Compatible software: QTSculpture. It produces a visual 3D	The scanner comprises a projector and camera lenses mounted on a rod,

			model and point-cloud coordinate data. Exports geometry as .wrl/.vrml, .stl, .obj and .ply files.	a tripod, a PT-M controller unit, and various cables. It also has a separate turntable. It connects to an external standard computer
IQ camera for VIS Hyperspectral imaging 	Specim	UBM	- Reflectance spectra that can be processed into conventional images (RGB or Infrared False-Color) or more advanced statistical treatments like Chromatic Derivative (ChromaDI), True-Color Infrared, or Multi-illumination Hyperspectral eXtraction (MHX). These images are typically large files (spectral data in .csv, .tiff and .hdr) that require further processing through software or custom routines for data analysis.	- Camera screen. - Specim IQ studio software (basic tools for spectral exploration). - Third-party software like Evince (for multivariate analysis) or Spectral Python (for custom programming). - Python workflows for custom image processing (e.g., shadow correction, automatic detection of spectral signals).
p-XRF Vanta 	Olympus	UNIPI; UBM	- Common formats include .csv, .xls, and .txt, suitable for analysis in various software. -2 spectra (1 for each beam). - Semiquantitative with an in-built calibration system.	- Handheld device with an integrated screen for real-time data visualisation and control. - Compatible with software like Excel, SPSS, MATLAB, R, and Past4 for data analysis. - Olympus Vanta software on laptop.
Raman Bravo duoLaserTM (785 nm, 852 nm)	Bruker	UNIPI	- Raman spectra in text file (.dpt), .dat or opus format (.0).	- Integrated touchscreen or buttons for basic operation and control. - Uses Bruker OPUS or similar proprietary software for spectral

				acquisition, processing, and analysis. - Data can be exported to external systems for further analysis, compatible with third-party tools (e.g. Origin or Spectragryph) or custom workflows.
i-Raman Plus 785H 	Metrohm - BWTek	UBM	- Spectral Data: Text files (e.g., .txt) compatible with BWIQ® software for multivariate analysis and BWID® software for material identification. - Raw spectral data with options for advanced processing via third-party software or programming routines.	- Touchscreen or control panel; onboard display. - BWIQ® Raman data analysis software. - Compatible with generic data analysis platforms like MATLAB, Origin, or Python libraries for custom workflows and spectral processing. - Data exportable in standard formats for use in multivariate statistical analysis software like R or SPSS.

Table 2. Overview of the sensors to be used in the AUTOMATA project, detailing their producers, associated teams, data output formats, and interfaces.

The project must carefully consider which sensor types to integrate, balancing versatility with the system's technical limits. Structured light and LiDAR are essential to high-resolution 3D model creation (see fig. 1, 2, 3), with digital cameras also playing a key role in photogrammetry-based scanning. While for material identification, initial artefacts' mapping and surface composition analysis, the HSI camera can be used efficiently to capture detailed spatial and spectral data. The quality of the hyperspectral data interpretation (images and spectra) depends on the sensor setup and the quality of the reflectance spectra database. In any case, despite algorithms for the post processing of data, HSI analysis may need confirmation by other analytical methods such as X-ray fluorescence (EDXRF) and Raman spectroscopy. Elemental screening using p-XRF is highly valued for its non-invasive capability to conduct provenance studies on ceramics and lithics, offering a preliminary dataset to guide more focused or destructive analyses. This screening function allows the system to gather data which can be used to create clusters of artefacts based on molecular/chemical composition, enabling a deeper examination of selected pieces using complementary techniques. Similarly, Raman spectroscopy is advantageous for studying pigments and decorative elements, providing detailed insights into specific features.

Establishing clear specifications is crucial to ensuring that each sensor is compatible with the robotic system's technical requirements while meeting the diverse analytical demands of archaeological artefacts. This careful balance enables the development of a well-calibrated and efficient digitisation setup. The AUTOMATA project's sensor requirements underscore the importance of precise handling and tailored environmental configurations to achieve accurate and consistent data capture across a wide range of archaeological artefacts.

The specifications of each sensor type introduce key considerations for weight, adaptability, and compatibility within the robotic system. For instance, the Bravo Raman spectrometer weighs 2-3 kg, which is hardly manageable for a robot, whereas lighter devices like iPhones with LiDAR enhance integration ease. Sensor weights and dimensions vary significantly across technologies such as photogrammetry, lasergrammetry, structured light, and hyperspectral imaging, with some devices being bulkier. Techniques like Hyperspectral Imaging (RTI) require precise light source control (see fig. 10 and paragraph 1.4.1.1.1.), while photogrammetry and lasergrammetry may require filters and controlled lighting to handle reflective surfaces.

Each technology's unique requirements for resolution, sensor-to-artefact distance, cost, and spatial demands further influence integration within the robotic cell. Battery life is critical for sensors like the Raman Bravo or the Vanta p-XRF, which rely on a portable power source. Indoor settings, generally, provide stable power sources to mitigate this need. In almost all cases, an external power supply is preferred to avoid interruptions and support consistent data capture. However, portable batteries could be beneficial for potential off-the-grid functionalities.

Contact-based versus contactless data acquisition is a central consideration. While contactless methods are generally preferred for preservation and ease, sensors like p-XRF and Raman spectrometers require close contact with the artefact to collect precise data, meaning that they must remain stationary during measurement to avoid data interference or inaccuracies caused by movement. In contrast, HSI can operate without direct contact, as well as photogrammetry, providing a more comprehensive overview of the artefact.
These differences necessitate a careful assessment of the target accuracy for each sensor, as well as spatial and spectrometric resolutions, depending on the artefact's characteristics, such as the smallest measurement area that can be analysed, the material matrix or decorative elements, accounting for possible surface alterations. Optimal sensor placement and integration configuration, including distance settings, lighting control, and contact preferences, are critical for achieving consistent, high-quality data capture. Moving artefacts during the digitisation process, rather than sensors, has practical benefits by enabling a stable, multi-sensor setup, reducing the need for recalibrating sensor positions, and ensuring consistent data collection. This approach also simplifies the workflow, as a single camera or sensor can track the artefact's position through various stages, ensuring accurate and consistent data capture. For example, placing artefacts on a transparent plane under a fixed camera array could allow multiple-angle capture simultaneously, provided the artefact is clean and stable. Sensor mobility should depend on the type of acquisition: 2D or zonal acquisitions might benefit from stationary sensors with multi-angle coverage, whereas 3D acquisitions offer three potential configurations. These include rotating the artefact in front of a fixed sensor, moving the sensor around the artefact, or using multiple sensors (such as in photogrammetry) to capture the full 3D structure. For 3D scanning, moving the artefact might be preferable, as it ensures all surfaces are captured accurately and avoids the technical complexities of sensor reorientation.

Alternatively, a robotic arm equipped with a wrist-mounted camera could grasp artefacts from the initial container, position them at acquisition stations, and monitor their alignment, simplifying the workflow. Despite the flexibility of moving sensors, this approach introduces complexity that may not align with the project's focus on accuracy and consistency. A stable, artefact-focused system better supports the AUTOMATA project's goals of streamlining and adapting to various artefact types.

Sensor interfaces vary depending on the manufacturer and type of data out. Spectral data formats (e.g. from p-XRF or Raman spectrometers) lack a single standard and require specific postprocessing protocols in order to be transformed into .csv files. Many instruments require proprietary software to visualise and process the raw data (see Table 2 and fig. 9).

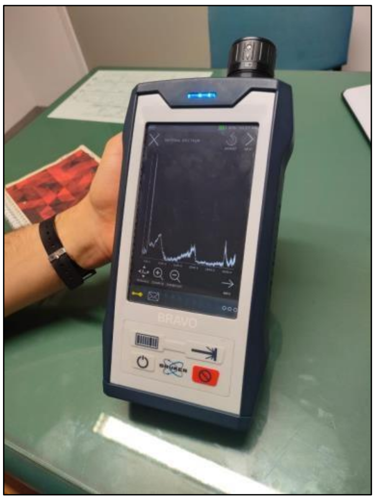

Fig. 9. Portable Raman Bravo spectrometer in use at the LAD (Archaeometry and Diagnostic Laboratory) of the University of Pisa, displaying the resulting spectrum from a lithic fragment.

Most sensors, such as the IQ hyperspectral machine (fig. 10), the Vanta p-XRF (fig. 11) and the Raman Bravo spectrometer (fig. 9), are designed for autonomous data acquisition without the need for a computer connection. These handheld devices can be operated completely autonomously and directly from the devices' screens. However, in a static setup, they can be connected to a computer via cable and managed remotely through software, allowing real-time data visualisation.

Fig. 10. On the left, IQ camera analysing the Aubusson tapestry, Cité Internationale de la tapisserie (France). On the right, IQ camera at Zaragoza Museum (Spain). As described in paragraph 1.4.1.1.1., artefacts can be placed on a surface with two halogen lamps positioned at a 45-degree angle and the IQ camera placed centrally. Ideally, images should be captured from approximately 50 cm away from the target (although a shorter distance may be used), and a Spectralon tile should always be positioned near the objects.

Fig. 11. Portable XRF Olympus Vanta C Series available at Archéosciences Bordeaux (UBM).

2.3 Data collection and storage

Once data from sensors has been collected, they must be effectively integrated into a unified system for analysis, interrogation, and visualisation. The AUTOMATA project envisions a solution based on the capability to integrate 3D models with detailed analytical data, allowing researchers to explore and query data spatially, within a 3D interface. This solution will be based on the RIS3D (Referenced Information System in 3D) concept (Dutailly B. *et al.* 2023), which provides a centralised platform for managing diverse datasets in a 3D environment. This approach is particularly well-suited for archaeological artefacts, as it combines geometric, textural, and chemical data into a single, cohesive system, facilitating advanced analysis and comparative studies.

Automating the integration of these datasets is critical to streamlining workflows and reducing manual input. This involves defining compatible data structures, ensuring alignment of sensor outputs with the 3D coordinate system, and supporting direct input from sensor software. Through this approach, RIS3D becomes a robust framework for managing and visualising archaeological datasets in a dynamic and interactive 3D environment, aligning with the AUTOMATA project's goals of enhancing artefact digitisation and analysis workflows.

2.3.1 Automating the Referenced Information System in 3D (RIS3D) process and software requirements

The RIS3D serves as a centralised platform for visualising 3D objects and environments alongside their associated digital data, seamlessly integrating diverse datasets, including hyperspectral, XRF, Raman, and other archaeometric analyses, into a single database. It supports both polygonal meshes and point clouds.

Data are stored in a PostgreSQL relational database, with a NodeJS-based web server managing database administration, user accounts, and the storage of large files.

A 3D viewer, built with Unity, displays the 3D data and associated information. This viewer relies on an API provided by the web server for database queries and access.

The data are stored in the database in JSON format, which offers several major advantages:

1. Flexibility in data storage: JSON accommodates a wide range of data types, including text annotations (author, comments), measurements (distance, volume), external links (URIs), images, files (photos, spreadsheets, PDFs), dates (ISO text or timestamps), and booleans. More complex records can be

described using nested JSON objects, such as: camera positions (to save viewpoints), numerical series (e.g., temperature curves, category counts, spectra), 3D voxel volumes (e.g., CT scans, ground-penetrating radar), 3D images (e.g., X-ray fluorescence, hyperspectral imaging), projected images (camera + image data).

2. Tree-structured organisation: JSON's hierarchical structure eliminates the need for multiple tables and joins. This allows users to easily define their data structure without requiring relational database expertise.
3. User-defined schema: RIS3D does not predefine the structure of the data to be stored, allowing users to design schemas tailored to specific project needs.
4. Field-level accessibility: each JSON field is accessible via queries. This makes it possible to retrieve specific values through SQL queries while benefiting from indexing for efficiency.

To integrate archaeometric and 3D data, RIS3D supports three types of anchors:

1. 3D Points: simple X, Y, Z coordinates.
2. Surface Outlines: used to define areas larger than a single point.
3. 3D Volumes: typically 3D elements derived from geometric modeling, commonly used for architectural reconstructions.

To automate the integration of 3D models and analytical data, the following steps are required:

1. Define a JSON structure to accommodate all acquired data, ensuring compatibility with RIS3D.
2. Specify a compatible 3D model file format, such as GlTF.
3. Locate each analysis within the 3D coordinate system of the acquired object. This is the most critical step, as the position and orientation of each sensor must be known at the time of acquisition.

Manual data entry should not be required. Scripts can help populate the database using files generated by sensor software and place 3D models in the correct directory for the RIS3D to process.

For spectrometric analyses (to be associated as meshes in 3D models), sensor parameters must be known to simulate the projection of the image into 3D correctly. This will require calibration and an iterative approach tailored to each piece of equipment.

2.3.2 Data quality assurance and control

Data management presents a significant challenge, particularly due to the data-intensive nature of 3D models enriched with archaeometric information. While chemical data is relatively lightweight, handling 3D data requires an eco-responsible approach. This means digitising what is necessary to address the specific scientific question rather than uniformly capturing an entire object at the same resolution and precision. Instead, the 3D digitisation process should be adapted to focus on different parts of the object based on the relevance of the information or analyses those areas can provide. This also entails exploring data compression techniques for 3D models to manage storage without sacrificing accuracy, supporting easy retrieval and processing. Moreover, the system should include a digital ID system to ensure each artefact can be easily found post-digitisation, facilitating further analysis or re-evaluation.

Ensuring data quality in the AUTOMATA project is essential for creating reliable, reusable digital models that meet both current project goals and potential future uses. Quality assurance for digitised artefacts will involve multiple criteria, including metrological accuracy, visual fidelity, and the comprehensiveness of metadata. Metrological criteria emphasise the importance of minimising uncertainties in the 3D model's mesh vertices and surface reflectance (SV-BRDF) to capture optical properties accurately. Visual criteria will require a comparison between a photograph and the 3D model under identical lighting and viewing conditions, ensuring the digital model faithfully represents the artefact's appearance.

To support the long-term usability of the data, the metadata accompanying each model should be as detailed as possible, capturing not only what is necessary for integration with the RICHeS (Research Infrastructure for Conservation and Heritage Science, UK Research and Innovation, 2024) infrastructure but also additional information that may be valuable for future research contexts. This aligns with the "archive once, reuse everywhere" principle, ensuring the models are versatile and compatible with a range of applications. Input from 3D domain experts is critical to define these standards comprehensively, and consultations will be conducted to ensure all reuse needs are anticipated.

To maintain high data quality, calibration will be conducted before each major acquisition series, with a test acquisition on a standard artefact to validate sensor setup and ensure consistency. Acquisition thresholds and spatial resolutions will require manual validation to align with each artefact's specific characteristics. This human-in-the-loop approach will ensure a human expert assesses and monitors model quality, allowing adjustments to be made based on real-time evaluation. Furthermore, the system should enable precise zoning for high-detail surface analysis, especially for areas like painted decorations, where pigments or specific textures must be analysed in detail. The system should flag noisy or unreliable data, setting a quality threshold for spectra to ensure only high-quality measurements are used in analyses. Additionally, data from physico-chemical analyses, such as XRF, Raman, or hyperspectral imaging, should adhere to sampling protocols that consider the surface area and zoning to capture reliable results for different types of artefacts. Lastly, non-destructive analysis is prioritised; however, when destructive or micro-destructive analysis is critical, sampling may be conducted on representative objects in a controlled laboratory setting as a separate task. The data obtained from these destructive analyses can later be manually linked to the digital model, providing detailed insights into the material composition. This multifaceted approach to data quality assurance and minimum quality requirements will support the creation of accurate, detailed, and adaptable 3D models that meet diverse research and preservation needs and will be thoroughly explained in D2.3 (*System Specification*).

Furthermore, all information that helps estimate the quality of the resulting data, under the Heritage Digital Twin (HDT) framework and the FAIR (Findable, Accessible, Interoperable, Reusable) principles must be provided. They will be the interpretation referential for future use and training for more automatic processes.

The AUTOMATA project will build on established metadata frameworks to develop interoperable solutions aligned with European infrastructures like ARIADNE, Europeana, EOSC and ECCCH. Practitioners will use structured guidelines to ensure metadata captures critical details across the entire workflow. Key elements include:

- **Overview/Planning**
    - Purpose: Why is the data being created? Are there limitations to the approach?
    - Expected reuse/intended audience: Are there restrictions or specific outputs?
    - Needs: Are there requirements for data dissemination or preservation?
- **Preparation**
    - Documentation of location, data collection procedures, and equipment testing.
- **Collection/Creation**
    - Identifying raw data and any initial cleaning.
    - Selecting appropriate data formats (native or open, standards-compliant).
    - Documenting equipment settings and protocols used during collection.
- **Processing (Post-Acquisition)**
    - Documenting intermediary datasets and final outputs for different purposes.
    - Including protocols and policies to assist future users in data understanding.
- **Long-term Curation**
    - Ensuring formats are suitable for preservation and clear file relationships.
    - Structuring datasets meaningfully and including relevant documentation.

The AUTOMATA project will focus on the 3D referencing of acquired data using annotations, integrating manual and robotic acquisitions with detailed descriptions of hardware configurations and precise measurement positioning. Existing standards, such as the Web Annotation Data Model developed by W3C (used in IIIF for annotating images and volumes), will guide the development of a new 3D annotation proposal. Practitioners will also adopt and extend common vocabularies like PeriodO, VAIF, Geonames, and PACTOLS, ensuring metadata interoperability. Demonstrated frameworks, including Quantin M. *et al.* (2023), highlight the compatibility of these vocabularies with broader European infrastructures. This approach aligns with ongoing efforts in ECCCH calls and ECHOES projects, addressing interoperability, harvesting, and reuse. Ultimately, AUTOMATA's metadata strategy will facilitate advanced digitisation workflows and long-term preservation while ensuring seamless data sharing and integration.

3 Optimising analytical workflows

3.1 Step-by-step process

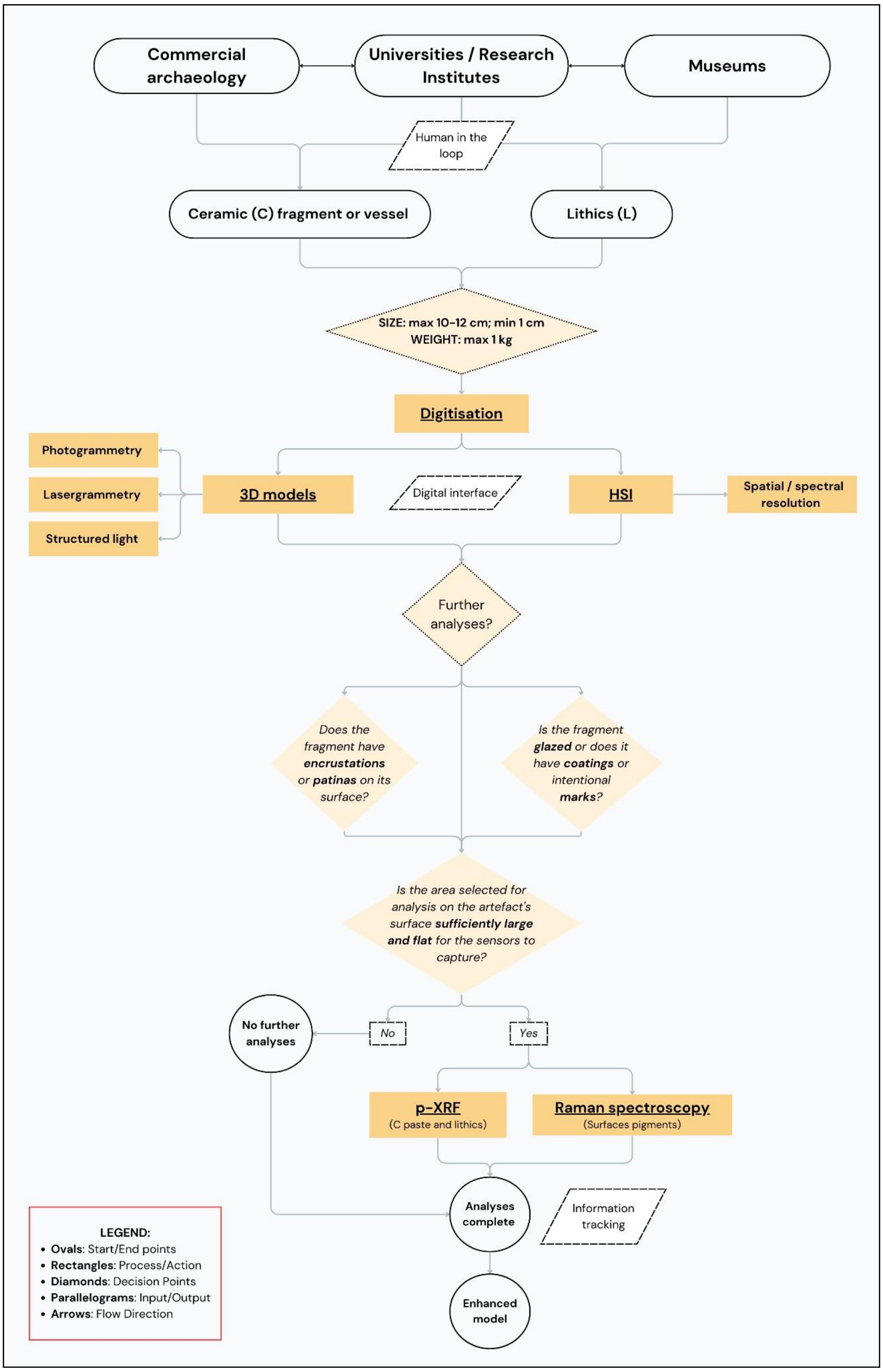

Fig. 12. Workflow illustrating the hypothesised analytical process for ceramic and lithic fragments. Starting with sample selection based on size and weight, the process moves through digitisation (3D modelling and hyperspectral imaging) and decision points for further analyses. Depending on sample surface conditions and suitability, the workflow progresses to either p-XRF for material composition or Raman spectroscopy for surface pigment analysis, or both.

The workflow shown in the diagram (fig. 12) outlines the journey of the archaeological artefacts through the AUTOMATA system to address specific research questions and meet the needs of commercial archaeologists, universities, research institutes, and museums.

Preliminary Setup and launching

The acquisition environment, including light conditions, dust, or other potential disturbances must be considered. Ideally, if possible, the system should be connected to the Internet so that data can be synchronised with the AUTOMATA cloud repository in real-time. In offline scenarios, data will be synchronised as soon as a connection is available. The operator prepares a batch of sherds — typically up to 50 pieces for small objects — for scanning, enters metadata and monitors the system's functions. Although much of the process is automated, human intervention remains critical at this stage to handle artefacts, address any alerts or malfunctions, and mitigate challenges posed by surface alterations that interfere with the recognition and characterisation of the artefacts. After this initial setup, the operator can leave the workstation and let the system operate autonomously. Light and acoustic signals, as well as notifications via email or mobile alerts, will provide updates on process start, completion, malfunctions or errors. The system aims to process 30-50 artefacts within 2-4 hours, aligning with the target of approximately 5 minutes per artefact. The operator must also determine the required accuracy level for the digitisation. For small fragments without distinctive features (e.g. rims or bases), a low-accuracy model may suffice. However, for pieces with distinctive elements, high-accuracy 3D modelling should be selected, always considering the interests, the needs and the requirements of the user. The current design focuses on fragments between 1-12 cm, excluding complete forms, covering the majority of archaeological fragments typically found in archaeological investigations.

Automated digitisation phases

Phase 1: 3D Modelling

3D modelling serves as the foundational requirement, ensuring reliable digital model generation. The system will include a 3D sensor to meet this requirement. While the specific methodology will be detailed in Deliverable 2.3 he project duration allows for ongoing experimentation to refine techniques. Current options under consideration include photogrammetry, laser scanning, and structured light (Table 1).

Phase 2: Hyperspectral Imaging (HSI) and sensors integration

The second phase involves the application of hyperspectral imaging (HSI), either integrating miniaturised sensors (Stuart M.B. *et al.* 2021) onto the robotic arm or using commercially available devices, such as the IQ camera (Table 2). The outcomes of these two initial phases not only contribute to the creation of an enhanced model but also determine whether additional analyses using other sensors are required. At this stage, a decision is made regarding whether the artefact should undergo further analyses. This determination primarily depends on the artefact's surface characteristics — such as the presence of encrustations, patinas, glazes, coatings, or intentional marks —, whether the selected area for analysis is sufficiently large and flat to ensure accurate sensor readings as well as the number and type of sensors available at the institution operating the system. If these conditions are met, advanced analyses may proceed. For example:

- portable X-Ray fluorescence (p-XRF) (refer to paragraph 1.4.1.2) to examine the ceramic paste or lithic composition,
- Raman spectroscopy (refer to paragraph 1.4.1.3) to investigate surface pigments and coatings.

Should the artefact fail to meet the necessary requirements, it exits the workflow without undergoing additional analyses.

Tracking and Decision Support

The system guarantees comprehensive tracking of each artefact through a dedicated interface, linking scanned data, analytical results, and physical storage location. This enhanced model serves not only as a repository of essential information but also as a decision-support tool, guiding operators through each stage of the process and enabling informed decisions based on sensor integration and workflow outcomes.

3.2 Digitisation time

The digitisation time is a crucial factor in ensuring the full functionality and sustainability of the AUTOMATA system. Excessively long processing times risk undermining the tool's usability, while overly rapid workflows can compromise the quality of the data collected. It is therefore essential to strike a balance that accommodates both operational and archaeological requirements, ensuring an efficient workflow and scientifically reliable outcomes. The time needed for digitising each object can vary significantly depending on the research questions, the resolution required, the nature and complexity of the objects and the analytical methods applied. Artefacts may require different amounts of time for analysis and digitisation, depending on their characteristics and nature. Some may need minimal processing, while others require multiple techniques, significantly extending the duration of the analytical process. Additionally, the time required for data acquisition and processing can vary significantly depending on the instruments used and the artefact's conditions. For instance, painted or glazed objects may require multiple analyses involving different sensors. This variability highlights the importance of tailoring workflows to the specific features of each artefact to ensure accurate and comprehensive results.

Method	Average Time	Key Factors
3D Model Production	180 - 600 sec (3 - 10 min)	Object size, model resolution, method used (photogrammetry -number of photos-, scanner resolution), processing speed
HSI Imaging	30 - 180 sec (0 - 3 min)	Sensor resolution and spectral range, surface irregularities, lighting conditions
p-XRF Analysis	60 - 120 sec (1 - 2 min) per measurement	Surface homogeneity, object positioning, number of measurements
Raman Spectroscopy	60 - 300 sec (1 - 5 min) per measurement	Fluorescence issues, material complexity, number of measurements

Table 3: Average time required for the analytical methods considered for AUTOMATA project, along with the key factors influencing these durations, already discussed in paragraphs 1.3.1 and 1.4.1.

Considering only the operational timing of the sensors, it is possible to roughly estimate the time required to produce an enhanced digital model of a single fragment. These estimates include a margin to account for the grasping and movement operations performed by the robotic arm. However, the time required for data transmission, reception, and potential on-the-fly processing has not been included at this stage. These factors will be addressed in detail in the deliverable D2.3 focusing on the system's technical aspects. The following scenarios outline the estimated time required for processing a single artefact under different workflows:

- **Baseline Workflow (3D modelling and hyperspectral imaging)**: the system completes the 3D modelling of the artefact, enriched with hyperspectral imaging (HSI) data. The process takes approximately **4 minutes per artefact**, with the following timings:

  - 3D Modeling: 180 seconds
  - HSI: 30 – 60 seconds

- **Multi-sensor Workflow (including p-XRF and Raman spectroscopy)**: this workflow incorporates additional analytical techniques, such as portable X-Ray fluorescence (p-XRF) and Raman spectroscopy, alongside 3D modelling and HSI.

  - Minimum time: approximately **5 minutes**
    - 3D Modelling: 180 sec
    - HSI: 30-60 sec
    - p-XRF: 60 sec
    - Raman spectroscopy: 60 sec

  - Average acquisition times: approximately **8–10 minutes**
    - 3D Modelling: 180 sec
    - HSI: 150 sec
    - p-XRF: 90 sec
    - Raman spectroscopy: 180 sec

  - Maximum time: could exceed **15 minutes**
    - 3D Modelling: 300 sec
    - HSI: 300 sec
    - p-XRF: 120 sec
    - Raman spectroscopy: 300 sec

AUTOMATA aims to offer a system that is able to complete the automation process within a minimum average time of five minutes per artefact. This timeframe is considered optimal for ensuring an operational workflow combined with massive digitisation. Achieving this minimum timeframe is feasible where time is optimised and certain compromises are made, depending on the research questions, as well as the conditions and the nature of the artefacts being studied.

This may involve:

- Optimising time: process multiple fragments simultaneously, allowing fragments to be subjected to different sensors concurrently, streamlining the workflows when multiple analyses are required. This solution helps streamlining the process also for high-resolution data capture given the possibility to reduce the digitisation time even during maximum time multi-sensor workflow.
- Using low-resolution scans 3D modelling, which significantly reduces the time needed (see paragraphs 1.3. and 1.4.);
- Limiting p-XRF and Raman analyses to fewer points or specific areas of the artefacts, focusing only on the most relevant features or zones that will provide the essential data;
- Skipping or reducing the depth of certain analyses. This option depends on the nature of the material. For example, well-preserved and relatively homogeneous materials may not need detailed Raman spectroscopy or hyperspectral imaging on every point or layer.

When compromises cannot be made and more comprehensive analyses, as well as high-resolution and high-accuracy data are essential, the five-minute timeframe solution is unlikely to be sufficient. Full, detailed analyses — especially those involving high-resolution 3D scanning or multiple measurements — will generally exceed this short timeframe. In particular, complex artefacts, such as those with intricate textures or features, require more time to capture accurate data. Additionally, issues like fluorescence interference in Raman spectroscopy can slow down the analysis, as they require adjustments to the setup or extended acquisition times.

Such a targeted approach based on the combination of the above solutions allows for flexibility and customisation to meet diverse requirements. On one hand, the system supports rapid and efficient digitisation with less data depth, suitable for general screening, which can be completed within an average of five minutes per artefact. On the other hand, it also provides the capability to collect data with greater depth and accuracy, accommodating more detailed analyses by extending processing times. This adaptability ensures that the system can cater to both quick overviews and comprehensive investigations, depending on the specific research needs.

4 Conclusions

In conclusion, this deliverable establishes the foundation for the AUTOMATA project by outlining the key steps and methodologies that will drive the automated and enriched digitisation of ceramic and lithic artefacts. This document should be read in conjunction with subsequent deliverables, including D2.2 *State of the art on enhanced digitisation*, D2.3 *System specification*, and D2.4 *Ethical guidelines for trustworthy AI*, which together provide a comprehensive framework for the development and implementation of the AUTOMATA system. The proposed solutions have been carefully designed to address stakeholders' needs, navigate technical constraints, and, most importantly, respond to specific research questions. These considerations aim to optimise data acquisition processes while taking into account the constraints of the robotic system and the timing required to complete all workflow steps. By addressing the challenges inherent in traditional archaeological workflows, this approach offers innovative solutions to enhance efficiency, accuracy, and accessibility in artefact documentation and analysis.

A critical advantage of the AUTOMATA project is its portability, as well as its emphasis on non-destructive and non-invasive portable analytical instruments. The development of a movable and autonomous system ensures that digitisation can occur directly in museums or storage facilities. This flexibility broadens the potential for collaboration and data sharing between institutions, making high-quality digitisation more accessible to a wider range of heritage professionals. The use of non-destructive and non-invasive instruments allows for in situ analyses, addressing the challenges associated with the movement and handling of fragile artefacts. This approach is particularly valuable in contexts where artefacts cannot be easily transported due to their condition, size, or legal restrictions.

One of the core principles emerging is the need for adaptability in digitisation processes. Rather than aiming for maximum precision in all cases, the project prioritises flexibility, tailoring methods and tools to the specific needs of each scenario. For example, the use of lightweight and low-resolution instruments provides a rapid and efficient means for initial assessments. Although these tools may not offer the same precision as advanced systems, they allow for the quick screening of large datasets, helping to identify artefacts that require more detailed investigation. This strategy ensures that digitisation remains scalable and responsive to the constraints faced by archaeologists, such as time limitations during excavations or resource availability in different institutions.

The workflow for digitising artefacts should therefore be flexible but maintain a fixed general structure. The process moves through clear stages: artefact selection, 3D modelling, and physico-chemical analyses. This workflow allows for the efficient management of large collections, which may require different levels of accuracy in digitisation, ensuring that each step is documented and reproducible. By employing 3D scanning techniques, the AUTOMATA system captures detailed information about the geometry, texture, and visual appearance of artefacts. The model can be integrated with physico-chemical information about the materials from non-invasive techniques like HSI, portable X-ray fluorescence (p-XRF), and Raman spectroscopy. The result is an enriched digital model that combines visual and compositional data, enhancing both documentation and future research potential. This workflow is designed to optimise data resolution and time efficiency at every stage. Traditional archaeological workflows can be time-consuming, requiring significant manual effort for artefact documentation and analysis. The AUTOMATA system aims to drastically reduce the time required to document and analyse artefacts by streamlining the process while integrating several methods in a unique workflow. For example, using low-resolution 3D models and rapid HSI scans allows for quick initial assessments, while more detailed analyses can be applied selectively when required. This approach balances speed and accuracy, ensuring that the process is efficient without compromising the quality of the results.

Compliance with ECCCH

Ultimately, the AUTOMATA project places significant attention on the creation of comprehensive metadata for enhanced digital models. By adhering to established standards and frameworks such as the FAIR principles and CIDOC-CRM, the project ensures that digitised artefacts are well-documented and interoperable across different systems. This is particularly critical in aligning with the ECCCH (European Collaborative Cloud for Cultural Heritage) model, which provides a structured framework for data sharing and collaboration across institutions. Detailed metadata captures not only the artefact's physical and contextual information but also the digitisation process, including details on the methods, instruments, and calibration settings employed. The inclusion of paradata — information on the decision-making and processing steps during digitisation — further enhances transparency and usability, enabling other researchers to evaluate or replicate the work. Through its metadata scheme, AUTOMATA supports long-term preservation, ensuring that digitised artefacts remain accessible and reusable for future research and educational purposes. By integrating metadata into a centralised system, the project also facilitates data interoperability and alignment with broader European infrastructures, such as ARIADNE and Europeana, fostering a collaborative ecosystem for the advancement of cultural heritage studies. This approach reinforces the ECCCH's objectives of promoting shared access, responsible data management, and innovation in cultural heritage preservation.

References

ARIADNE Infastructure, Horizon 2020, Grant agreement ID: 823914, 2019. DOI:10.5281/zenodo.7528754 [Online]. https://ariadne-infrastructure.eu/

Alaoui M'Darhri A. *et al.* (2019), *Share - Publish - Store - Preserve. Methodologies, Tools and Challenges for 3D Use in Social Sciences and Humanities*, European Commission ; Horizon H2020 Projects, Tech. rep. May 2019. DOI:10.5281/zenodo.3258316.

Apollonio F. I. (2024), *The Three-Dimensional Model as a 'Scientific Fact': The Scientific Methodology in Hypothetical Reconstruction*, Heritage, vol. 7, no. 10, pp. 5413-5427. DOI:2571-9408/7/10/255.

Behmann J., Acebron K., Emin D., Bennertz S., Matsubara S., Thomas S., Bohnenkamp D. *et al*. (2018), *Specim IQ: evaluation of a new, miniaturized handheld hyperspectral camera and its application for plant phenotyping and disease detection*, Sensors (Basel, Switzerland), 18(2) art. 441. DOI: https://doi.org/10.3390/s18020441

Belfiore C.M., Mazzoleni P., Manenti A.M., Mastelloni M.A., Corsale V., Barone, G. (2021), *Non-destructive XRF analysis of aegyptiaca from sicilian archaeological sites*, Mediterranean Archaeology and Archaeometry 21(1): p. 37–69.

Chatzistamatis S., Tsekouras E. G., Anagnostopoulos C.-N. (2023), *The Quality in 3D Acquisition of Cultural Heritage Assets: Challenges and Risks*, in 3D Research Challenges in Cultural Heritage III: Complexity and Quality in Digitisation.: Springer International Publishing, 2023, pp. 65-76. DOI:10.1007/978-3-031-35593-6_4.

Colomban P., Tournié A., Bellot-Gurlet L. (2006), *Raman identification of glassy silicates used in ceramics, glass and jewellery: a tentative differentiation guide*, Journal of Raman Spectroscopy, 37, pp. 841-852. Ffhal-00120401

Crocombe R., Leary P., Kammrath B. (2021), *Portable Spectroscopy and Spectrometry*, John Wiley & Sons, Ltd.

Cuomo di Caprio N. (2007), *Ceramica in archeologia 2. Antiche tecniche di lavorazione e moderni metodi di indagine*, Roma: "L'erma" di Bretschneider.

DGCNCT, *Study on quality in 3D digitisation of tangible cultural heritage : mapping parameters, formats, standards, benchmarks, methodologies, and guidelines*, Final study report: Publications Office of the European Union, 2022. DOI:10.2759/471776.

Dorn R. I. (2024), *Rock varnish revisited*, Progress in Physical Geography: Earth and Environment, 48(3), 389-419. https://doi.org/10.1177/03091333241248038

Dubinsky L., Grosman L. (2024), *Techné of Rock Engravings—the Timna Case Study*, J Archaeol Method Theory 31, 1917–1953. https://doi.org/10.1007/s10816-024-09658-5.

Dutailly B., Portais J. C., Granier X. (2023), *RIS3D: A Referenced Information System in 3D*, ACM Journal on Computing and Cultural Heritage, 15(4), 1-20.

Fihn Marberg J. *et al.* (2022), *D15.2 - Final report on ARIADNEplus services.* DOI:10.5281/zenodo.7528755.

Frahm E., Adler D. S., Gasparyan B., Luo B., Mallol C., Pajović, Tostevin G. B., Yeritsyan, Monnier G. (2020), *Every contact leaves a trace: Documenting contamination in lithic residue studies at the Middle*

Palaeolithic sites of Lusakert Cave 1 (Armenia) and Crvena Stijena (Montenegro), PLoS ONE 17(4): e0266362. https://doi.org/10.1371/journal.pone.0266362

Frahm E. (2024), *Protocols, pitfalls, and publishing for pXRF analyses: From "know how" to "best practices"*, Journal of Archaeological Science Reports, 60 (104831).

Galantucci L.M., Guerra M.G., Lavecchia F. (2018), *Photogrammetry Applied to Small and Micro Scaled Objects: A Review*, in Ni J., Majstorovic V., Djurdjanovic D. (eds), Proceedings of 3rd International Conference on the Industry 4.0 Model for Advanced Manufacturing. AMP 2018. Lecture Notes in Mechanical Engineering. Springer, Cham, https://doi.org/10.1007/978-3-319-89563-5_4

Galluzzi F., Manca R., Puntin M., Raneri S., Sciuto C., Benvenuti M., Chapoulie R. (2024), *Maiolica seen by Vis–NIR hyperspectral imaging spectroscopy: the application of an ultraportable camera at the Museo Nazionale del Bargello*, Eur. Phys. J. Plus, 139:629.

Grifoni E., Campanella B., Legnaioli S., Lorenzetti G., Marras L., Pagnotta S., Palleschi V., Poggialini F., Salerno E., Tonazzini A. (2019), *A New Infrared True-Color Approach for Visible-Infrared Multispectral Image Analysis*, Journal of Computing and Cultural Heritage, 12(2) Art. 8 1-11.

Grosman L., Smikt O., Smilansky U. (2008), *On the application of 3-D scanning technology for the documentation and typology of lithic artifacts*, Journal of Archaeological Science 35, 3101–3110.

Grosman L., Muller A., Dag I., Goldgeier H., Harush O., Herzlinger G., Nebenhaus K., Valletta F., Yashuv T., Dick N. (2022), *Artifact3-D: New Software for Accurate, Objective and Efficient 3D Analysis and Documentation of Archaeological Artifacts*, PLoS ONE 17(6):e0268401. https://doi.org/10.1371/journal.pone.0268401

Harush O., Glauber N., Zoran A., Grosman L. (2019), *On Quantifying and Visualizing the Potter's Personal Style*, Journal of Archaeological Science 108:104973. https://doi.org/10.1016/j.jas.2019.104973

Harush O., Grosman L. (2021), *Toward the identification of social signatures in ceramic production – An archaeological case study*. PLoS ONE 16(7): e0254766. https://doi.org/10.1371/journal.pone.0254766

Harush O., Roux V., Karasik A., Grosman L. (2020), *Social signatures in standardized ceramic production – A 3-D approach to ethnographic data*, Journal of Anthropological Archaeology 60, 101208.

Hernández V. *et al.* 2012, *Raman spectroscopy analysis of Palaeolithic industry from Guadalteba terrace river, Campillos (Guadalteba county, Southern of Iberian Peninsula)*, Journal of Raman Spectroscopy, Vol. 43, Issue 11. https://doi.org/10.1002/jrs.4104

High-Level Expert Group on Artificial Intelligence (HLEG) (2019), *Ethics guidelines for trustworthy AI. European Commission*. Retrieved from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Jehlička J., Culka A. (2022), *Critical evaluation of portable Raman spectrometers: From rock outcrops and planetary analogs to cultural heritage - A review*, Analytica Chimica Acta 1209, 339027.

Jouan P., Hallot P. (2020), *Digital Twin: Research Framework to Support Preventive Conservation Policies*, ISPRS International Journal of Geo-Information, vol. 9, no. 4, p. 228. DOI:10.3390/ijgi9040228.

Karasik A., Smilansky U. (2008), *3D scanning technology as a standard archaeological tool for pottery analysis: practice and theory*, Journal of Archaeological Science 35, 1148e1168.

Karasik A., Smilansky U. (2011), *Computerized morphological classification of ceramics*, Journal of Archaeological Science 38, 2644 - 2657.

Karasik A. (2010), *A Complete, Automatic Procedure for Pottery Documentation and Analysis*, Conference: Computer Vision and Pattern Recognition Workshops (CVPRW), DOI: 10.1109/CVPRW.2010.5543563

Legnaioli S., Lorenzetti G., Cavalcanti G. H. , Grifoni E., Marras L., Tonazzini A., Salerno E., Pallecchi P., G. Giachi, Palleschi V. (2013), *Recovery of archaeological wall paintings using novel multispectral imaging approaches*, Heritage Science 1 (1): 1-9.

Liu L., Miteva T., Delnevo G., Mirri S., Walter P., de Viguerie L., Pouyet E. (2023), *Neural networks for hyperspectral imaging of historical paintings: a practical review*. Sensors 23, 2419. https://doi.org/10.3390/s23052419

Maritan L. (2020), *Ceramic abandonment. How to recognise post-depositional transformations*. Archaeological and Anthropological Sciences, 12: 199.

Muller A., Barsky D., Sala-Ramos R., Sharon G., Titton S., Vergès J.M., Grosman L. (2023), *The limestone spheroids of 'Ubeidiya: intentional imposition of symmetric geometry by early hominins?,* Royal Society Open Science, 10:230671, http://doi.org/10.1098/rsos.230671

Muller A., Barkai R., Shemer M., Grosman L. (2022), *3D Morphology of Handaxes from Late Acheulean Jaljulia: a Flexible Reduction Strategy in the Lower Paleolithic Levant*, Archaeological and Anthropological Sciences, 14 (206):1-20. https://doi.org/10.1007/s12520-022-01671-7

Niccolucci F., Felicetti A., Hermon S. (2022), *Populating the Data Space for Cultural Heritage with Heritage Digital Twins*, Data, vol. 7, no. 8, p. 105. DOI:10.3390/data7080105.

Niccolucci F., Markhoff B., Theodoridou M., Felicetti A., Hermon S. (2023), *The Heritage Digital Twin: a bicycle made for two. The integration of digital methodologies into cultural heritage research*, Open Research Europe, vol. 3, no. 64. DOI:10.12688/openreseurope.15496.1.

Ogburn D., Sillar B., Sierra J. C. (2013), *Evaluating effects of chemical weathering and surface contamination on the in situ provenance analysis of building stones in the Cuzco region of Peru with portable XRF*, Journal of Archaeological Science 40 1823e1837.

Orton C., Hughes M. (2013) (2nd), *Pottery in Archaeology*. New York: Cambridge University Press.

Pasikowska-Schnas M., Lim Y.-S. (2023), *Artificial intelligence in the context of cultural heritage and museums - Complex challenges and new opportunities*, EPRS | European Parliamentary Research Service, European Parliamentary. [Online]. https://t.ly/tn7i9

Puntin M., Gattiglia G., Sciuto C. (2023), *Archeologie dell'invisibile. Tecniche archeometriche non distruttive per la caratterizzazione dei laterizi del sito romano di Massaciuccoli*, in "Rivista di Archeologia", 46.1, pp. 125-146.

Quantin M., Tournon S., Grimaud V., Laroche F., Granier X. (2023), *Combining FAIR principles and long-term archival of 3D data*, in Proceedings of the International ACM Conference on 3D Web Technology, pp. 16:1-6. DOI:10.1145/3611314.3615913.

Reneau S. L., Raymond R., Harrington C. D. (1992), *Elemental relationships in rock varnish stratigraphic layers, Cima volcanic field, California; implications for varnish development and the interpretation of varnish chemistry*, American Journal of Science, 292(9), 684-723.

Richardson E., Grosman L., Smilansky U., Werman M. (2014), *Extracting Scar and Ridge Features from 3D-scanned Lithic Artifacts*, in Earl G., Sly T., Wheatley D. *et al.* (eds.) *Archaeology in the Digital Era: Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26-29 March 2012*, 83-92, Amsterdam University Press.

Rousaki A., Vandenabeele P. (2021), In situ *Raman spectroscopy for cultural heritage studies*, Journal of Raman Spectroscopy, 52, 2178–2189.

Sciuto C., Cantini F., Chapoulie R., Cou C., De La Codre H., Gattiglia G., Granier X., Mounier A., Palleschi V., Sorrentino G., *et al.* (2022), *What Lies Beyond Sight? Applications of Ultraportable Hyperspectral Imaging (VIS-NIR) for Archaeological Fieldwork*. Journal of Field Archaeology, 47 (8), pp. 522-535.

Sciuto C., Geladi P., La Rosa L., Linderholm J., Thyrel M. (2018), *Hyperspectral Imaging for Characterization of Lithic Raw Materials: The Case of a Mesolithic Dwelling in Northern Sweden*, Lithic Technology, 1–14.

Statham N. (2019), *Scientific rigour of online platforms for 3D visualization of heritage*, Virtual Archaeology Review, vol. 10, no. 20, pp. 1-16. DOI:10.4995/var.2019.9715.

Stuart M.B. *et al.* (2021), *Low-Cost Hyperspectral Imaging with a Smartphone*, Journal of Imaging, 7, 136, https://doi.org/10.3390/jimaging7080136

Tournon S. *et al.* (2021), *The French National 3D Data Repository for Humanities: Features, Feedback and Open Questions*, in Computer Applications and Quantitative Methods in Archaeology (CAA) 2021, vol. Digital Crossroads. DOI:10.48550/arXiv.2211.04094. [Online]. https://3d.humanities.science/

Triolo P., Marras L., Adinolfi G, Carmagnola R., Legnaioli S., Raneri S., Palleschi V. (2020), *Imaging for Cultural Heritage and Archaeology*, 2020 IMEKO TC-4 International Conference on Metrology for Archaeology and Cultural Heritage: 512-516.

UK Research and Innovation (UKRI) (2024), Research Infrastructure for Conservation and Heritage Science (RICHeS). Retrieved from https://www.ukri.org/what-we-do/browse-our-areas-of-investment-and-support/research-infrastructure-for-conservation-and-heritage-science-riches/

Valletta F., Dag I., Grosman L. (2021), *Identifying Local Learning Communities During the Terminal Palaeolithic in the Southern Levant: Multi-scale 3-D Analysis of Flint Cores*, Journal of Computer Applications in Archaeology 4 (1):145–168. https://doi.org/10.5334/jcaa.74

Vandenabeele P. (2013), *Practical Raman Spectroscopy: An Introduction*, Wiley.

Vandenabeele P., Weis T.L., Grant E.R. *et al.* (2004), *A new instrument adapted to in situ Raman analysis of objects of art*, Anal Bioanal Chem 379, 137–142. https://doi.org/10.1007/s00216-004-2551-z

Wilkinson M. D. *et al.* (2016), *The FAIR Guiding Principles for scientific data management and stewardship*, Scientific Data, vol. 3, pp. 160018:1-9. DOI:10.1038/sdata.2016.18.

Yashuv T., Grosman L. (2024), *12,000-year-old spindle whorls and the innovation of wheeled rotational technologies*, *PLoS ONE* 19(11): e0312007. https://doi.org/10.1371/journal.pone.0312007