Reusing 3D Measurement Data of Lithic Artifacts to Develop Analytical Methods

Florian Linsel, Jan Philipp Bullenkamp & Hubert Mara

Abstract – This article focuses on how we reused a 3D data publication of lithic artifacts from the Upper Palaeolithic site of Grotta di Fumane and how it influenced our work creating workflows and new software solutions. These models are annotated recreating already published drawings and enriched by the scar position in an operational sequence. The annotated 3D models are then used as base dataset for developing new software solutions. The first is an automatic segmentation using Morse Theory, in which their performance was tested against the dataset. And in the second, it was used to build graph models and create a parameter-based predict algorithm to automatically predict the relation between adjacent scars.

Key words – archaeology; research software; Open data; computer science; Upper Palaeolithic; operational sequence; 3D models

Titel – Nachnutzung von 3D-Messdaten zur Methodenentwicklung für die Analyse von lithischen Artefakten

Zusammenfassung – Der Fokus dieses Artikels liegt auf der Wiederverwendung einer 3D-Datenpublikation von lithischen Artefakten aus der jungpaläolithischen Fundstelle von Grotta di Fumane und der Frage, wie diese unsere Arbeit bei der Entwicklung von Arbeitsabläufen und neuen Softwarelösungen positiv beeinflusst hat. Die Modelle werden nach dem Vorbild bereits veröffentlichter Zeichnungen annotiert und mit der zeitlichen Einordnung der Schlagnarben innerhalb einer Chaîne Opératoire angereichert. Die annotierten 3D-Modelle dienen als Grunddatensatz für die Methodenentwicklung. Die erste Methode ist eine automatische Segmentierung basierend auf der Morse-Theorie, deren Leistung anhand des Datensatzes evaluiert wurde. Die zweite Methode wurde verwendet, um Graphenmodelle zu erstellen und einen parameterbasierten Vorhersagealgorithmus zu entwickeln, der automatisch die Beziehung zwischen benachbarten Narben bestimmt.

Schlüsselwörter – Archäologie; Archäoinformatik; Forschungssoftware; Informatik; Jungpaläolithikum; chaîne opératoire; 3D-Modelle

Introduction

Motivated by the challenge of attributing leaf points, a Palaeolithic artifact category, to *anatomically modern humans*, Neanderthals or both (Flas, 2011; Jöris et al., 2022; Swainston, 1999), we are developing digital tools to analyse stone artifacts and their processing steps (**Fig. 1**). So far, leaf points have only been found at one site in the context of human remains (Mylopotamitaki et al.,

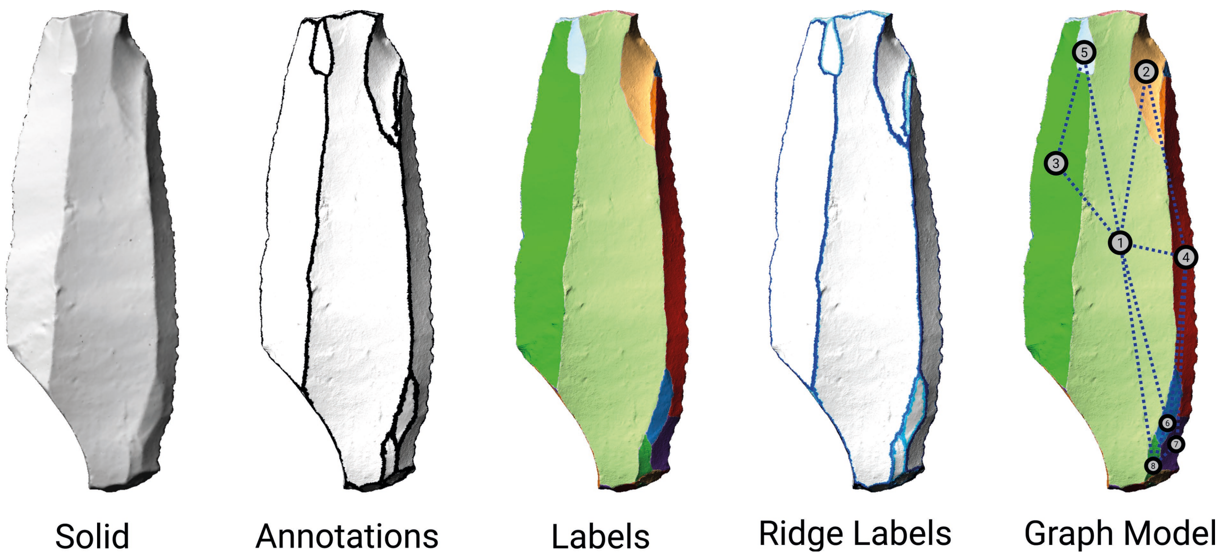

Fig. 1 Renderings of meshes along the annotation process: solid rendering (Solid), a colour-coded drawing of the ridges done using Blender (Blender Online Community, 2018) (black:ridge; white=surface feature), the labels using GigaMesh (colour: labels) and the graph model of the scar-ridge patterns.

Received: 6 Jan. 2025
accepted: 10 Feb. 2025
published online: 14 Feb. 2025

NWVA 2024 Bochum: Digitale Archäologie

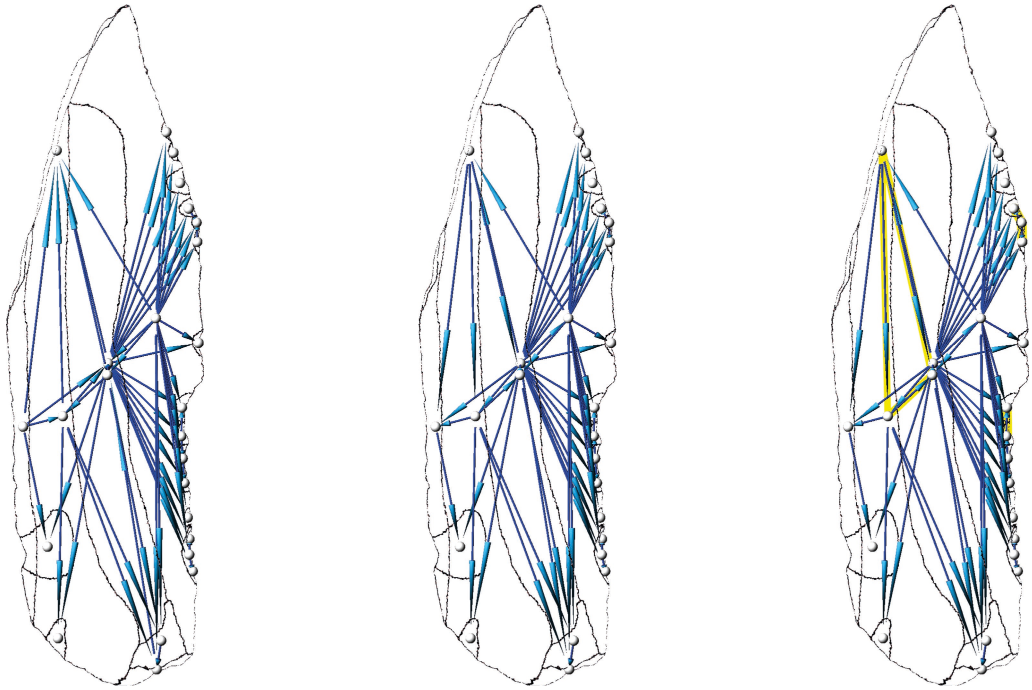

Fig. 2 3D visualisation of directed graph model of GdF b-207; (a) Directed by manual operational sequence; (b) Directed by network parameter (Degree); (c) Directed by Degree with highlighted differences to manual operational sequence.

2024). As most of the sites where leaf points were found are old excavations and the original find context is therefore difficult to reconstruct, the only way to determine the species of their inhabitants is on the basis of how lithic artifacts were made (Flas, 2011; Jöris et al., 2022). In the line of technological investigation, we wanted to analyse lithic artifacts using 3D models. In order to obtain the maximum amount of information, we used high-resolution 3D measurement data, which reproduces the geometric properties of the surface precisely and without interpretation.

The search for 3D measurement data revealed that many datasets are published with restrictive licenses, are unpublished or only contain empty promises in the accompanying publications. Only the dataset from Falcucci and Peresani (2023) was usable for our numerical experiments in terms of scope, quality of the 3D measurement data and licencing, even if it did not fall within the period of the original research question. This data was then cleaned, filled and orientated using the *GigaMesh* preprocessing routine (Mara et al., 2010). The MSII curvature values were also calculated for the detection of the concave and convex scars (Mara, 2012). An important question was how to process these pre-processed data in order to evaluate computer-aided, newly developed methods against archaeological interpretations. This evaluation data includes manual segmentations and graph models of the temporal ordered traces in the *operational sequence*.

The manual segmentation was based on the archaeological drawings. The manual segmentation is based on colouring in the edges of the surface features, the ridges, with Blender (Blender Online Community, 2018), for which there is also a separate workflow (Linsel et al., 2024a). This data can be used for the development of segmentation algorithms (Bullenkamp, Kaiser et al., 2024; Bullenkamp, Linsel et al., 2022) and for the creation of *operational sequence* (*OS*) graph models.

The graph models are based on the manual segmentations and represent the neighborhood of the scars. However, the temporal sequence cannot be determined from the segmentations alone. In archaeology, graphs (Harris matrices) or drawings are usually created to visualize the chronological sequence of the scars. Such drawings were also created for some of the scanned artifacts (Falcucci, Conard et al., 2017; Falcucci & Peresani, 2018). Based on this information, annotations were created that contain the chronological information. This data can also be used to verify graph-based methods (Linsel et al., 2024b) (**Fig. 2**). The proposed workflow enables the creation and evaluation of an automatic segmentation algorithm and it allows the algorithmic investigation of workflows in the production of stone artifacts by combining graph data and 3D measurement data (Linsel et al., 2024a). Inspired by the initial problems, the following questions are investigated in this study: How can already published 3D measurement data and their derivatives answer archaeological questions and how can reproducible workflows be created and existing ones optimised so that the resulting data can be sustainably structured and openly published?

Related Work

Even though data collections are becoming more often published, the availability of data and in particular of 3D models of cultural heritage is still in its infancy. But the landscape is changing with promising national data initiatives for cultural heritage like the CND3D in France (Quantin et al., 2023) and the National Research Data Infrastructure (ger. *Nationale Forschungsdateninfrastruktur*) (*NFDI*) in Germany (NFDI4objects: Bibby et al., 2023; NFDI4culture: Altenhöner et al., 2019).

In the field of lithic analysis with 213 studies published using 3D models until 2022, it is in a better situation than most other archaeological fields but it is still not common practice (Wyatt-Spratt, 2022). Even though it was promised to be "revolutionary" for lithic analysis (Grosman, 2016), 3D modelling is still considered rather a fancy commodity then a necessity for creating reproducible research, reflected in 61 studies (28.6 %) in which 3D models are only used for visualisation purposes (Wyatt-Spratt, 2022). In the last years the number of digitized lithic artifact collections increased (Di Maida & Hageneuer, 2022; Harmand et al., 2015) but most of them require either a registration or a bulk download is not implemented. But the total number of working groups publishing their datasets as open access data publication is growing daily (e.g. Falcucci & Moroni, 2024; Lombao, 2019; Porter et al., 2019). One benefit is that most of these are published at one platform, Zenodo, but they are still only findable with expert knowledge.

A common trait of an archaeological investigation is that it often ends with publishing the data while reusing the data is not a standard but rather an exemption for specific data-types relating to e.g. chronometric data, e.g. radiocarbon dates of the Iberian Peninsula between 45-30 ka BP (Díaz-Rodríguez et al., 2023), drawings, images or outlines often reused in *GMM* (Matzig et al., 2023). Currently, the reuse of 3D data is not a very common, but reusing 3D models is at least a subject worth a keynote presentation at the SUMAC '24 (Moitinho de Almeida, 2024). Unfortunately, in their summary, Moitinho de Almeida (2024) didn't elaborated the specific implementations and workflows applied for doing so.

In lithic analysis and archaeology in general, many researchers still rely heavily for illustrating artifacts on manual drawings derived from caliper measurements and/or photographs (Dryer & Mazierski, 2009; Magnani, 2014). An emerging alternative involves creating these illustrations directly on 3D models by encoding features such as scars and ridges as vertex colours. These colour-coded annotations can be transformed into specific labels, such as scars (Linsel et al., 2023) or structural features of buildings (Buldo et al., 2024). Presently, the software available for artifact annotation is limited, with tools often tailored to specific artifact categories, such as cuneiform tablets (Homburg et al., 2022) or lithic artifacts (Bullenkamp et al., 2024). Annotated 3D models also enable the development of segmentation algorithms capable of partitioning 3D models into regions that resemble manual segmentations (Pulla et al., 2001; Richardson et al., 2014). These manual segmentations, while subject to variation depending on researcher expertise and model resolution, remain critical for evaluating the precision of segmentation algorithms, as visual inspection is otherwise the primary evaluation method. Ideally, an automatically labelled dataset that includes reduction sequences information would enhance reproducibility and automation, but such a dataset is not yet available.

Beside drawings and 3D models, the third category, graph models, are often used to simplify the stratigraphic relation of layers (Herzog & Scollar, 1991) or social network analysis (Brughmans, 2013). In some cases, graph models are used to display the technical procedures, known as *operational sequence* (*OS*) or *Chaîne Opératoire*, leading to a lithic artifact. These graph models are either used to model the relation between artifacts by refitting, or the relation between scars (Kot et al., 2024; Soressi & Geneste, 2011) or working stages on one artifact using a scar-pattern, also called diacritic, analysis (Dauvois, 1976; Pastoors et al., 2015; Richter, 2004; Tafelmaier et al., 2022), which are based on ranked surface attributes.

Richardson et al. (2014) presented a segmentation algorithm for a scar detection on 3D models, which lead ultimately to creating an adjacency graph model of these segments. In a later study, Grosman (2016) underlined the potential of graph models and the possibilities of enriching these

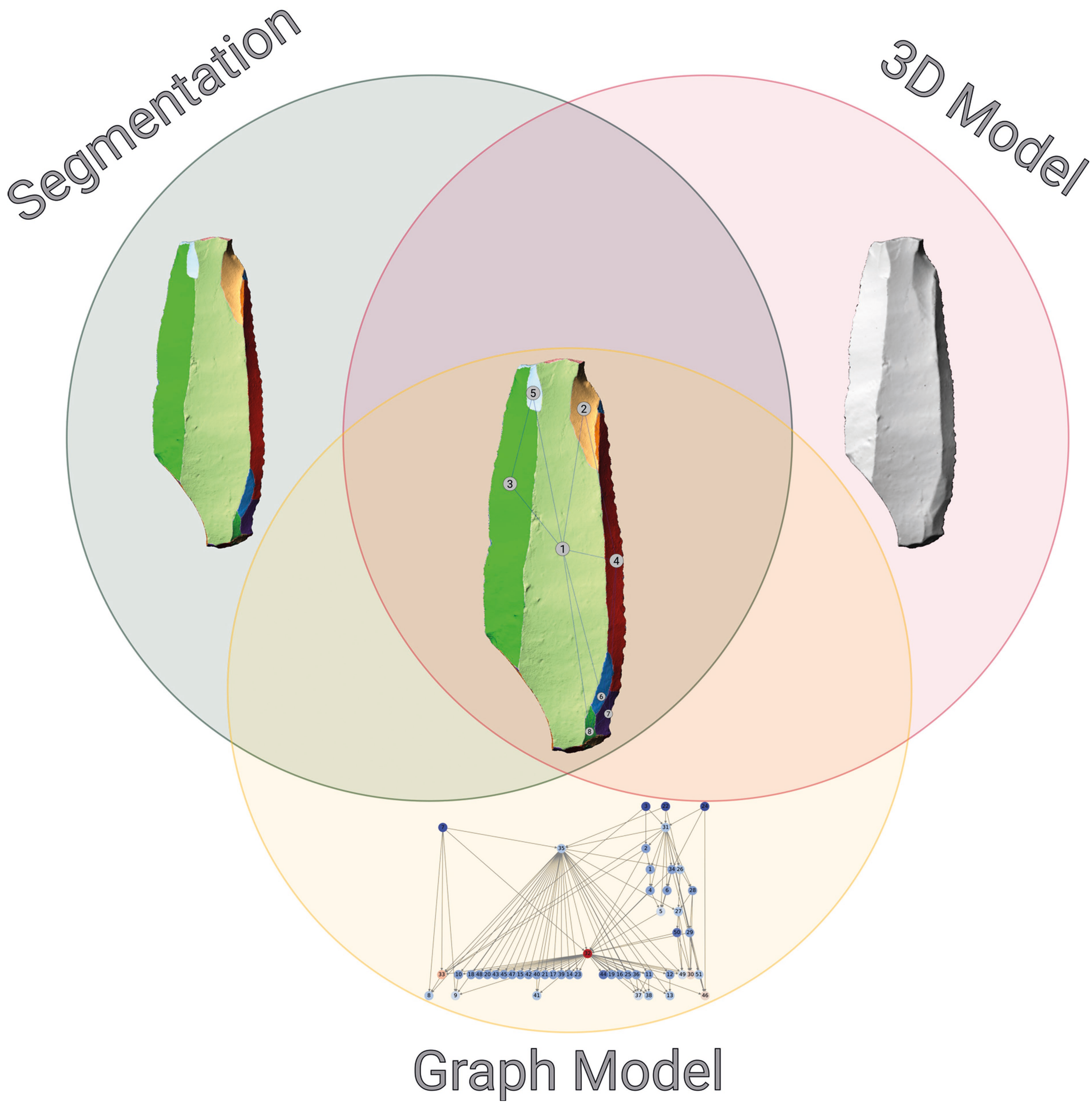

Fig. 3 An overview of all 3 manual data types necessary for this study: a 3D model, a vertex segmentation of scars and a graph model of the operational sequence.

with "inaccessible data" like the distribution of scar counts, their areas, shapes, and mean concavity leading to a "more precise description and analysis of the lithic artifact surface and the underlying production technology". Grosman (2016) even referenced a manuscript "L. Grosman, E. Richardson, U. Smilansky, manuscript in preparation", which should include graph based analysis with parameter enriched graphs, which to the authors' knowledge has never been published. Similar to the proposed approach, in our previous studies (Linsel et al., 2024a,b), we used manual annotations to build adjacency graph models, directed these with manual interpretations of the temporal position of its scars and created a visualisation of the relation between scars creating a 3D model of the graph (Linsel et al., 2024b) (**Fig. 2**). We also explored different parameters either were archaeologically derived parameters like *Integral Invariants of Polylines* (*IIoP*) or the curvature along the ridges, the scar area or network parameters like degree or betweenness centrality to predict the temporal relation of adjacent scars (Linsel et al., 2024a).

The Ideal Dataset and the Real Dataset of *Grotta di Fumane*

This study focuses on the reuse of data and how three distinct documentation record types can be combined. If considering the perfect dataset for a graph based analysis to find the best parameter to approximate an *operational sequence* digitally it consists of 3 types of data (**Fig. 3**):

1. 3D models published under a non-restrictive license (CC BY 4.0 or a more open option) preferably with high resolution,
2. scar segmentations dividing the surface in distinct segments of faces and vertices, and
3. the temporal relation between these scars either assigning a position in the *operational sequence* to each scar or a file containing the scar relations from the older (source) to the younger scar (target).

At the beginning of this research in 2022 there existed no datasets which fulfilled all 3 data requirements. Even 3D models of leaf points were either not published or non existent. For lithic artifacts in general the picture was a bit better regarding 3D data publications but also no annotated data was available for direct usage. After widening the search to include artifacts from a similar period between 45-40 ka cal BP, for one site, the northern Italian site of *Grotta di Fumane* (*GdF*), its publications contained enough non-restrictively published data to be used to create an annotated 3D dataset. The publications included 3D models (Falcucci & Peresani, 2023) and drawings (Falcucci, Conard, et al., 2017; Falcucci & Peresani, 2018) of multiple artifacts with the temporal scar position within the *operational sequence* by applying a diacritic analysis (Dauvois, 1976).

The artifacts included in the 3D data publication are mainly blades and bladelets but also cores and flakes from its Upper Palaeolithic sequence (41,000-33,000 cal BP) (Higham et al., 2009). It consists of 948 scans created using structured light or micro-CT scanners with resolutions of 40,000-400,000 vertices (Falcucci & Peresani, 2023). To test the automatic segmentation 62, artifacts were first annotated (Linsel et al., 2023), but for none of them *operational sequence*s were identified. However, *operational sequence*s were available for 44 other artifacts. An artifact (*GdF* b-207) has also been added.

Even though the dataset was not yet the ideal dataset, it enabled us to develop the following workflow for creating manual segmentations and graph models as well as develop methods to automatically segment surface features and predict the temporal relation between adjacent surface features. This workflow has 16 steps belonging to 6 categories (**Fig. 5**).

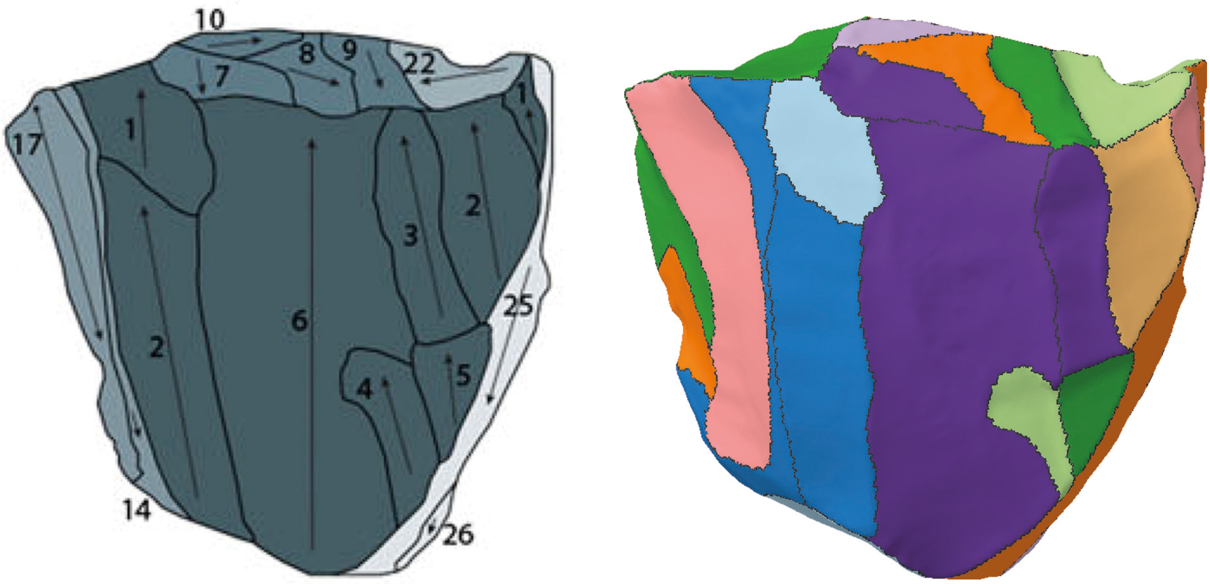

Fig. 4 Depictions of an artifact (GdF c-49): (a) Drawing with temporally ordered scars (drawing: A. Falcucci; Falcucci, Conard, et al., 2017); (b) 3D Model with annotations using Blender (Blender Online Community, 2018).

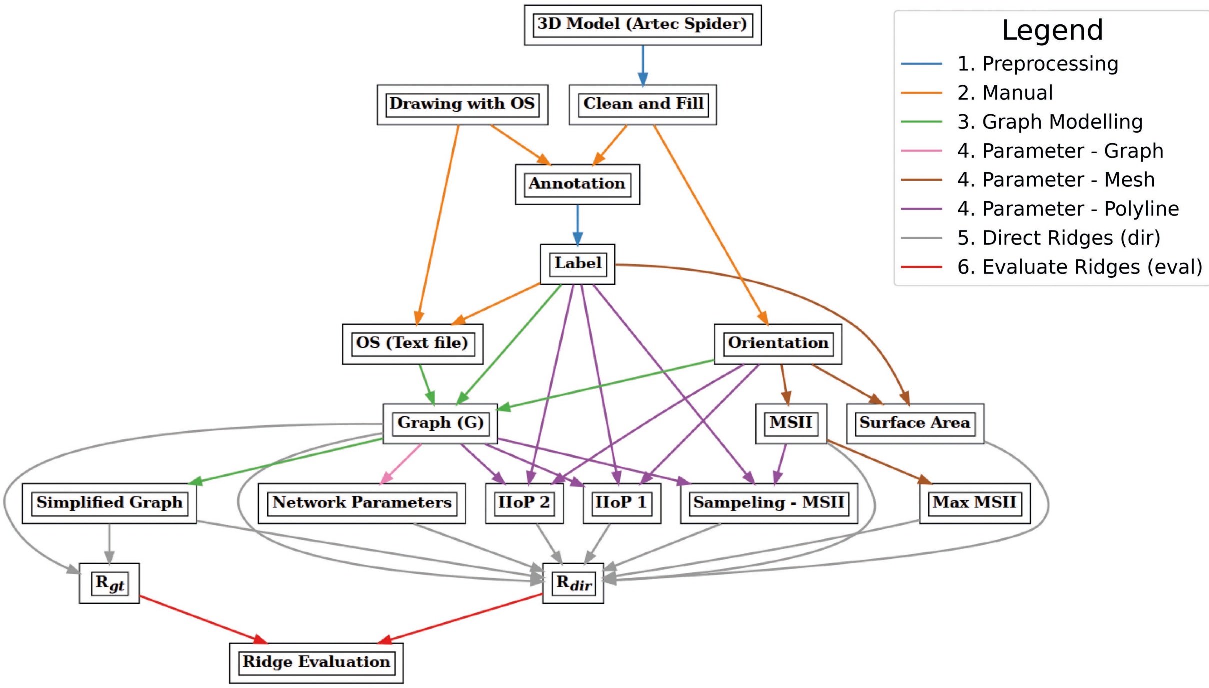

Fig. 5 Workflow of all manual and automatic steps for conducting Linsel et al. (2024a). The edge colour is determined by the category of the target of the edge.

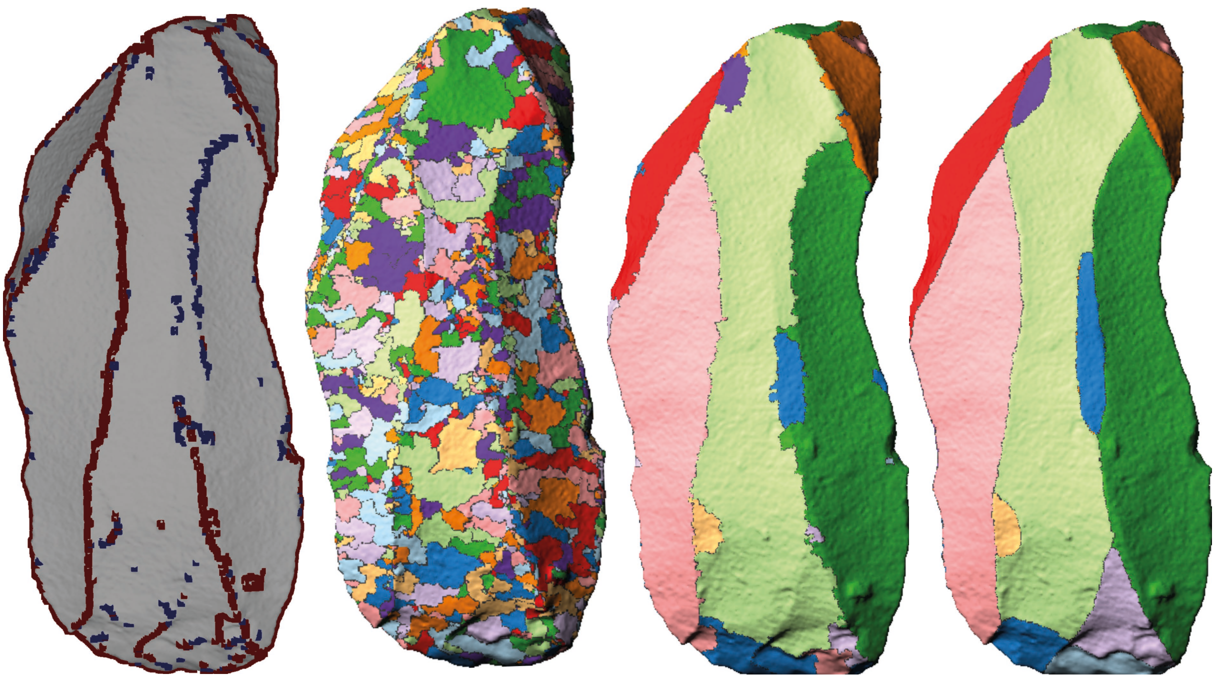

Fig. 6 (a) strong (red) and weak (blue) ridges; (b) oversegmentation; (c) merged result; (d) original labels.

Preprocessing and Manual Segmentation

As a first step, the 3D models were cleaned and filled using *GigaMesh* to ensure that the 3D model has specific surface properties (Mara et al., 2010) enabling to calculate certain surface parameters. The model should be watertight and without mesh errors like non-manifold edges or self-intersecting triangles. The resulting meshes are then orientated using *GigaMesh* and in a next step, the *Multi-Scale Integral Invariants* (*MSII*) values get calculated to approximate the maximum surface curvature (Mara et al., 2010). The MSII values can be used for determining whether the surface is convex, flat or concave.

The cleaned and filled mesh are also used to create the manual segmentation as a digital version of the manual drawing. The manual segmentation reproducing the drawings with *operational sequence* (*OS*) was done using Blender (Blender Online Community, 2018) by colouring in the vertices black and the surface features white (**Fig. 1**: Annotation). To complete reproduce the drawings while considering the available 3D models, all surface features, scars, cortex and other alterations need to be marked as distinct segments on the surface of the 3D model (**Fig. 4**). If surface features are not noticeable or need adjustments, the annotation was altered accordingly. After finishing the annotation, these colour values are then converted in numeric labels, referencing each label as an identifier of each distinct surface features to each vertex (**Fig. 1**: Labels; **Fig. 5**: Label).

Automatic Segmentation using Morse Theory

The automatic segmentation is often based on models like k-Means or in this case Morse theory, which is a concept from the field of topology. For a more detailed analysis compare Bullenkamp et al. (2024) and Bullenkamp et al. (2022). It can be used for segmenting meshes like the surface of 3D models using values assigned to the vertices of the mesh. In this study, *MSII* curvature values were taken as function on the vertices, such that high curvature values identify peaks and ridges, and lower or no curvature corresponds to the flatter scar areas on the model's surface. A skeleton of the mesh can be determined by finding local maxima, minima and saddle points and connecting them by following ridges and valleys (following the steepest gradients). This separates the mesh into smaller enclosed patches.

Then strong and weak ridges can be filtered with two thresholds from the lines of the skeleton, adding weak ridges only if they are connected to strong ridges. These smaller patches were used as an oversegmentation and by merging smaller patches of the oversegmentation the final labelling can be obtained. This was done by defining a threshold at which there is not enough ridge detected along the boundary between two adjacent patches. This ensures that each segmented area is distinct and separated by significant ridges, providing a clear and detailed analysis of the model's surface. The steps of this process are visualized in **Fig. 6**.

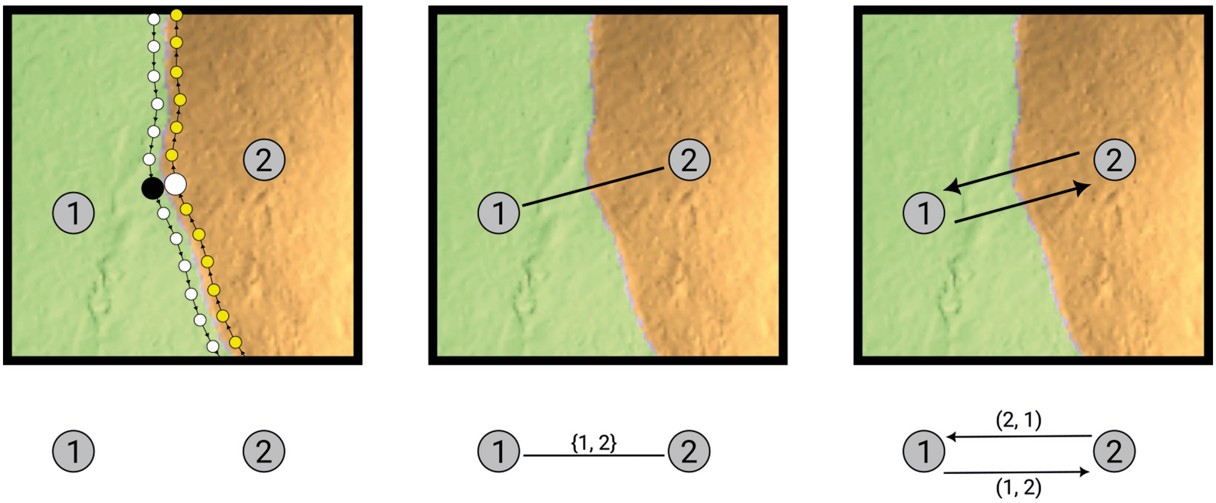

Fig. 7 Connecting neighboring scar labels according to their bordering polylines.

Graph Models

Either automatically created labels and the manual labels can be used to create the graph models but assuming that the manual labels better represent the surface features, the graph models will be based on the manual labels. The next step is assigning the second information of the drawings, the *operational sequence*, to the label ids in a text file. Strictly speaking the labels represent not only scars labels but also labels of all surface features, therefore, the *operational sequence* needs to be adjusted accordingly. For instance, cortex, which could also be considered 0, was due to technical definition that labels are only allowed to be positive integers, is defined as 1 and all others needed to be adjusted by adding one to the position. When the labels were adjusted to reflect the new procedural positions, slight adjustments to the original dataset have been done.

The label ids are also used as nodes in the adjacency graph (G), in which adjacent surface feature are connected with a link (**Fig 7**). All vertices belonging to the label are then referenced as a property to the label node making the part accessible inside the graph. The outline of the label area is a closed polyline, each vertex belonging to these lines is connected to the vertices of a neighboring polyline. The ridge between two adjacent labels is then the part of both polylines directly connected to each other, both stored as attribute of the link of the graph model. By doing so, both polylines can be compared according to their relative properties. This is relevant because many archaeologically derived attributes rely on ridge properties and hence their digital twin, these two polylines.

One additional idea was to simplify the graph because the edge retouches were easily visible in the graph due to their isolated position between dorsal and ventral side and their unique property to be mostly isolated only linking to two neighbours. Based on that observation, all labels with 2 or less neighbours were deleted resulting in a simplified graph model (**Fig. 5**, Simplified Graph). Parameters derived from the Mesh, the Polyline or the Graph

The adjacency graph models are not yet representing the complete *operational sequence* because the relation between scars are not yet determined. For determining these, similar studies of scar-pattern analysis used each a set of 5 attributes, their list position indicating their importance (Pastoors et al., 2015; Richter, 2004; Tafelmaier et al., 2022). In a recent study (Linsel et al., 2024a), we approximated 4 attributes and added 5 parameters, which were either derived from the mesh, the polyline or the graph model (**Tab. 1**). The objective was to approximate the most common archaeological attributes and try to recreate these computationally and add parameters common-

Data Source	Parameter	Archaeological Attribute
Mesh	Surface Area	-
	MAX MSII Curvature	Younger scar is more convex than older scar (RSP-2)
	MSII Curvature	Younger scar is more convex than older scar (RSP-2)
Polyline	Length of Polyline (IIoP 1)	Younger scar ridge follows older one (RRP-2).
		Younger scar ridge cuts across older scar (RRP-3).
	Angle of Polyline (IIoP 2)	Younger scar ridge follows older one (RRP-2).
		Younger scar ridge cuts across older scar (RRP-3).
	Curvature along Polylines ($samp_{in}$ – MSII)	Younger scar is more concave along ridges than older scar (RRP-1)
Graph	Degree	-
	Betweenness Centrality	-
	Degree Centrality	-

Tab. 1 List of parameters used to approximate the archaeologically determined scar properties.

ly used for meshes like the surface area and the MSII as curvature approximation and graphs like Degree, Betweenness Centrality and Degree Centrality (Hagberg et al., 2008).

The parameters for the polylines were not yet available and hence were implemented, for more details see (Linsel et al., 2024a). Similar to the *MSII* being a curvature approximation of a mesh, the *Integral Invariants of Polylines* (*IIoP*) are curvature approximations of a polyline. For approximating the curvature along the polylines, for each vertex of a polyline the surrounding vertices of the same label within a defined radius were sampled and the mean MSII curvature was determined ($samp_{in}$ – MSII).

Direct Ridge Links

In ridge direction prediction, as in archaeological practice, two adjacent labels are compared on the basis of individual parameters. If one label has a higher value, the link points towards it or vice versa. As *Integral Invariants of Polyliness* and $samp_{in}$ – MSII contain all vertices of the polylines and *MSII* even all vertices of each label, the mean of all parameters will be used for the comparison between the scars. In the workflow (**Fig. 5**), the ridges are noted as R, while R_{dir} represent all parameter-directed and R_{GT} the manual *operational sequence* directed ridges. To evaluate the parameter-directed predictions, they are then compared to the original *operational sequence* ridge data.

Results

For both areas, the automatic segmentation and the analysis of graph models, the results should be considered as preliminary because the goal was show how segmentation and working with graph models can open new avenues of research. All evaluations rely on manual data and hence will contain inaccuracies and interpretational bias but due to the lack of unbiased data, it will be used as *ground truth* (*GT*) dataset. To evaluate the automatic segmentation the created labels will be compared with the manually segmented labels. Similarly, the parameter-directed predictions will be compared with the manual *operational sequence* data. Also, due to the lack of third party annotations or reliable approaches, which evaluated results beyond a visual inspection, the results can only be discussed on a surface level.

Results – Segmentation

In previous studies, the accuracy of the Morse theory based segmentation reached an average

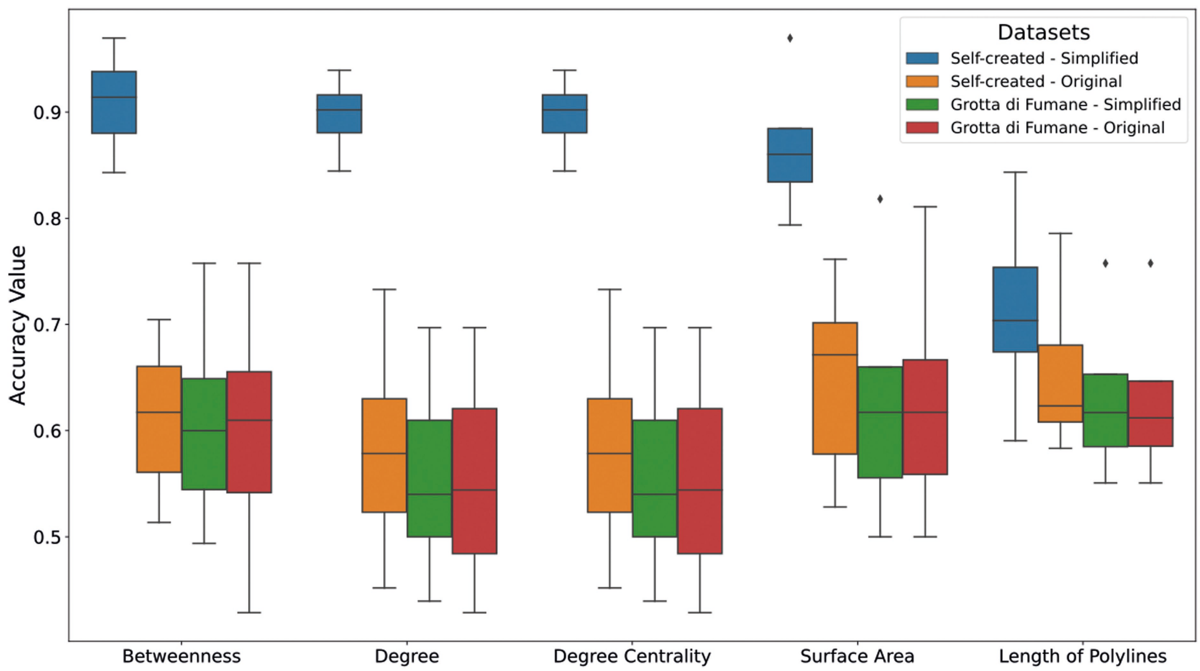

Fig. 8 Top 5 performing parameters based on median accuracy.

of around 91.26 % ranging from 81.2 % to 97.2 % when testing a whole range of parameters and taking the best combination for each artifact.

Results – Graph Model
The results of the *Grotta di Fumane* dataset was compared to a smaller sample of 15 artifacts scanned using a GOM Scan 1 100 (point distance: 0.037 mm), annotated and interpreted by our working group. For the *Grotta di Fumane* dataset, the results are not very convincing. The best performing parameter was one of the archaeological derived parameters, *IIoP* 1, with 62.73 % mean accuracy and the overall prediction of the original and the simplified Graphs only reached a mean accuracy between 55.14 % and 62.73 %. For the self-created 3D models, the results are more promising with slightly elevated mean accuracies of the original graph model (58.11 % and 64.95 %), and a noticeable improvement for the simplified graphs (58.43 % and 90.84 %). The best reperforming parameter is the betweenness centrality of the simplified graph model with a mean accuracy of 90.84 %, ranging between 84.30 % and 96.97 %.For a better comparison, a list of the 5 top performing parameters was created based on the best median performance across the datasets, placing the three network parameters in the Top 3 (betweenness, degree, degree centrality), than the surface area and last the length of polyline (*IIoP* 1) (**Fig. 8**).

Resume of the reused Dataset

Even though the results for the *GdF* dataset vary between the methods, achieving in the automatic segmentation a mean accuracy of 91.26 % and not being as precise as those derived from the artifacts scanned with GOM Scan 1 100 in the direction prediction, they enabled the research in the first place. Without having the 3D models as test dataset, the development of all methods and algorithms applied to the second dataset would not have been in the same state. Further having these meshes with lower resolution at hand opened up completely new questions like: Is there a specific resolution needed to apply specific methods? Should all 3D datasets be published with its measurement file? Which processing and post-processing steps are required to use and which need to be documented to interpret 3D models accurately?

Benefits and Restrictions of working with Annotated 3D Models

In this study, 3 manual data sources were combined to build graph models to make operational sequences analysable not only visually but computationally. All approaches, either the methods of the digitisation process or the methods developed have their restrictions and benefits. On a digitisation level, scanning, manual segmentation and creating *operational sequence* are all individual crafts. Many issues are common for the methods and arise from the experience of the person, who is performing the step, the software used, by surface properties (roughness, reflectiveness), the material (transparency) and the overall shape of the artifact. The scanning can also be influenced by the lighting, the general setting of the scanner and by the device itself (**Fig. 9**).

3D Modelling
By no means, creating 3D models is often quite costly and if done with a *structured light scanner* a time-consuming endeavour often taking over 30 minutes per artifact. Further documenting artifacts using 3D scanning is always connected to numeric variance. An object scanned twice will vary and hence even at the beginning of working with 3D models the method is not completely reproducible. However, with the rising availability of micro-CT scans, already applied to digitize lithics (Göldner et al., 2022) and bones (O'Neill et al., 2024) on a larger scale in a time (35 seconds of human time and about 2 minutes total per object) and in a cost-effective way (0.98-1.40 USD per 3D model) resulting in high resolution scans (0.0418 and 0.056 mm) (O'Neill et al., 2024) beyond level of drawings with ±1mm accuracy, these arguments loses its validity over time.

One of the biggest issues for 3D modelling is its structuring and long term storage. One high resolution scan with the GOM Scan 1 for example ranges between 100 MB to 2 GB for a standard ply file. While working with 3D models after nearly each step the data gets reproduced, resulting in bigger 3D models in the process and larger needs for storage. This is one of the reasons why researchers in archaeology only publish a selection of their 3D data. It is often one step too much to publish the data, especially if it is done correctly, a challenging endeavour.

Segmentation
Whether creating manual drawings or segmentations, both are similar in their time-consumption

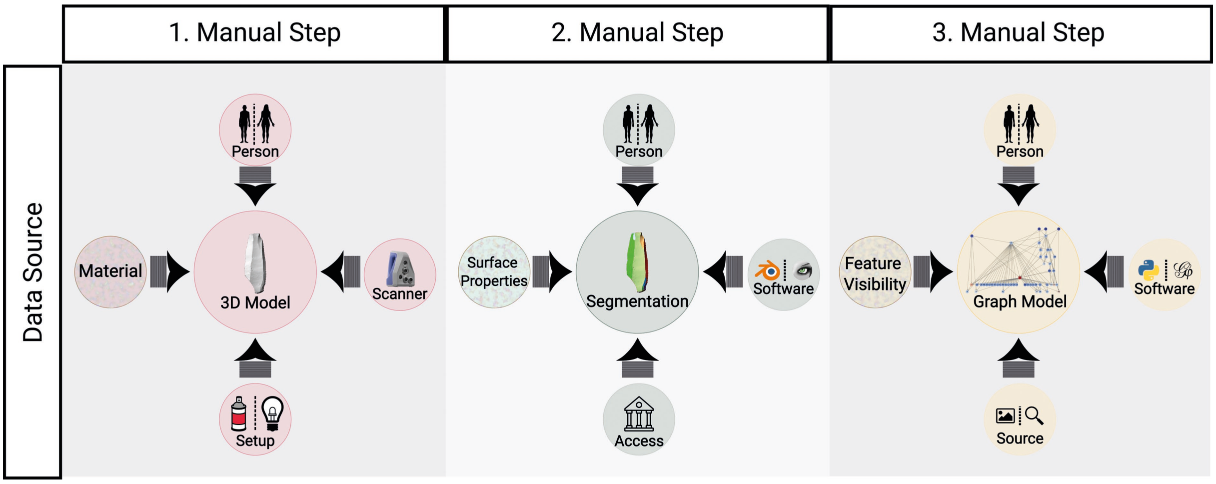

Fig. 9 Manual steps and the what can influence their accuracy.

having the benefit of directly creating an undirected graph model of the segmented area. Which can then be used to help creating the operational sequence. But technical knowledge like how to use special software like MeshLab or Blender and data structuring are still imperative for performing these tasks, because easy to use workflows and guides are still not available.

On a bigger picture, standing in front of the *Replication Crisis* in all academic discipline, reproducibility can also be a central problem in today's archaeological discourse but unfortunately many studies don't incorporate or discuss the issues luring behind antiquated technics like manual record keeping or creating drawings (Farahani, 2024). Even when drawings are recreated directly as manual segmentations on the 3D models they will differ, due to the interpretational biases in the drawings as well as the manual segmentations. The problem will not disappear, and they only should be used as long as more sophisticated experimental data isn't available.

Additionally, manual record keeping and drawings are also not easily reusable and transformable. As a rule of thumb 3D data can be simplified and transformed in 2D, categorial data or to simple parameter but the other way, creating 3D data out of 2D drawings is not reliably possible (yet) (NVIDIA et al., 2024). Hence the best advice is to create data in the best resolution possible.

Operational Sequences

As Kot et al. (2024) already pointed out, the scars and surface features visible on an artifact will not allow to model the complete operational sequence and hence only show a certain state in the operational sequence. However, our approaches investigate the relation between scars and if assumed the attributes for interpreting the temporal relation between adjacent scars are correct, it results in a distinct partly ordered graph, which are comparable to others in a similar state. Further, without even discussing the interpretational bias in the refitting process, the archaeological record always has gaps, what applies also to the manual or automatic refitting, resulting also in incomplete and partly ordered graphs. Hence, some frequencies will be missing in a refitting.

Even having exceptionally precise records at hand, including videos, drawings and keeping a written record like in shown by Kot et al. (2024) lead to inherent problems of reliability and reproducibility. Only based on the article and its appendix, its data cannot be reliably reproduced and interpreted. The major benefit of the automatic segmentation and the direction prediction are that anybody who has the code, the datasets and a computer can recreate the results. Additionally, if new methods are developed to extract new parameters or a better function is developed for existing ones, they can simply be added while the older implementation is removed from the workflow, making this approach scalable.

Outlook

This study showed on the example of the *Grotta di Fumane* dataset what can be done with already published 3D models and how it can impact research in real-time. That should be a friendly reminder that we as researchers will never know what other people or even generations of researchers can do with openly published data. Even if the data is not exemplarily structured and doesn't fulfil the highest standards. It would be ideal if every data publication is published according to the guidelines FAIR data (Wilkinson et al., 2016) and *Linked Open Data* (*LOD*) (Berners-Lee, 2006; Thiery and Mees, 2023), and with the highest standard available to every researcher with a minimal restrictive copyright but even if it is not ideal it can give crucial insights of how to improve workflows and adjust software solutions, which can ultimately lead to a higher data standard. The collection of publications used in this study (Falcucci, Conard, et al., 2017; Falcucci & Peresani, 2018; 2023) is a good example for exactly this problem. Collectively, it is one of the best datasets, where 3D models and their drawings with *operational sequence* were published but it was not the perfect dataset. The 3D models were either simplified (Göldner et al., 2022) or their precision doesn't fulfill the accuracy requirements to analyse lithic artifact on the scale, we need to conduct our research. But for many other studies like their own research of their morphometrical differences (Falcucci & Peresani, 2022), the precision is more than enough. But it gave us the opportunity to develop and test our methods. And it allowed a comparison with a high-resolution dataset, which led to questions like which resolution is best for creating 3D models of lithic artifacts? Do we need to publish the unsimplified meshes and how can we teach the knowledge about these standards? Do we need to publish the scan project and how can we ensure the readability, while scan-software is in many cases proprietary?

And this is not the end of the journey because the *Grotta di Fumane* dataset will be reused again to develop methods for graph comparisons.

Acknowledgements

We thank Jürgen Richter for his help in creating the *chaîne opératoires* as well as Armando Falcucci for his help handling the 3D models used from *Grotta di Fumane* (*GdF*). We want to thank Rebecca Gnau and Naira Khachatryan for their help in creating the annotations and graph models of the artifacts. We also thank the anonymous reviewers for their time and suggestions!

Fundings

This research was funded by the Martin-Luther University Halle-Wittenberg and conducted at the Institute of Computer Science with contributions from the Department of History and Cultural Studies at the FU Berlin.

References

Altenhöner, R. u.a. (2019). *Fokusthemen und Aufgabenbereiche* für *eine Forschungsdateninfrastruktur zu materiellen und immateriellen Kulturgütern.* (Living Document der NFDI-Initiative NFDI4Culture). https://doi.org/10.5281/zenodo.2763575.

Berners-Lee, T. (2006). *Linked Data*. https://www.w3.org/DesignIssues/LinkedData.html [14.8.2018].

Bibby, D. et al. (2023). *NFDI4Objects – Proposal.* https://doi.org/10.5281/zenodo.10409227.

Blender Online Community (2018). *Blender – a 3D modelling and rendering package*. Amsterdam: Blender Foundation.

Brughmans, T. (2013). Thinking Through Networks: A Review of Formal Network Methods in Archaeology. *Journal of Archaeological Method and Theory,* 20, 623-662. https://doi.org/10.1007/s10816-012-9133-8.

Buldo, M., Agustín-Hernández, L. & Verdoscia, C. (2024). Semantic Enrichment of Architectural Heritage Point Clouds Using Artificial Intelligence: The Palacio de Sástago in Zaragoza, Spain. *Heritage,* 7, 6938-6965. ISSN: 2571-9408. https://doi.org/10.3390/heritage7120321.

Bullenkamp, J. P. et al. (2024). Discrete Morse theory segmentation on high-resolution 3D lithic artifacts. *it – Information Technology*. https://doi.org/10.1515/itit-2023-0027.

Bullenkamp, J. P., Linsel, F. & Mara, H. (2022). Lithic Feature Identification in 3D based on Discrete Morse Theory. In F. Ponchio & R. Pintus (eds), *Eurographics Workshop on Graphics and Cultural Heritage.* (pp. 55-58). Delft, Netherlands: The Eurographics Association. https://doi.org/10.2312/gch.20221224.

Dauvois, M. (1976). *Précis de dessin dynamique et structural des industries lithiques préhistoriques*. Périgueux: Pierre Fanlac.

Di Maida, G. & Hageneuer, S. (2022). The DISAPALE Project: a new digital repository of lithic and bone artefacts. *Lithic Technology, 47(4), 283-295.* https://doi.org/10.1080/01977261.2022.2048511.

Díaz-Rodríguez, M., Nielsen, T. K., Maier, A. & Riede, F. (2023). An Annotated Compilation of Chronometric Dates for the Middle-Upper Palaeolithic Transition (45-30 ka BP) in Northern Iberia (Spain). *Journal of Open Archaeology Data,* 11, 10. https://doi.org/10.5334/joad.113.

Dryer, M. & Mazierski, D. (2009). Illustrating Artifacts Using Digital Rendering Techniques. In A. Falcucci, N. J. Conard & M. Peresani (2017). A critical assessment of the Protoaurignacian lithic technology at Fumane Cave and its implications for the definition of the earliest Aurignacian. *PLOS ONE,* 12, 1-43. https://doi.org/10.1371/journal.pone.0189241.

Falcucci, A. & Moroni, A. (2024). *The Open Aurignacian Project. Volume 2: Grotta di Castelcivita in southern Italy*. Zenodo. https://doi.org/10.5281/zenodo.10631389.

Falcucci, A. & Peresani, M. (2018). Protoaurignacian Core Reduction Procedures: Blade and Bladelet Technologies at Fumane Cave. *Lithic Technology,* 43, 125-140. https://doi.org/10.1080/01977261.2018.1439681.

Falcucci, A. & Peresani, M. (2022). The contribution of integrated 3D model analysis to Protoaurignacian stone tool design. *PLOS ONE,* 17, 1-29. https://doi.org/10.1371/journal.pone.0268539.

Falcucci, A. & Peresani, M. (2023). *The Open Aurignacian Project. Volume 1: Fumane Cave in northeastern Italy*. Version 2.1.1. Zenodo. https://doi.org/10.5281/zenodo.7664308.

Farahani, A. (2024). Reproducibility and Archaeological Practice in the Journal of Field Archaeology. *Journal of Field Archaeology, 49*, 391-394. https://doi.org/10.1080/00934690.2024.2391623.

Flas, D. (2011). The Middle to Upper Paleolithic transition in Northern Europe: the Lincombian-Ranisian-Jerzmanowician and the issue of acculturation of the last Neanderthals. *World Archaeology,* 43, 605-627. https://doi.org/10.1080/00438243.2011.624725.

Göldner, D., Karakostis, F. A. & Falcucci, A. (2022). Practical and technical aspects for the 3D scanning of lithic artefacts using micro-computed tomography techniques and laser light scanners for subsequent geometric morphometric analysis. Introducing the StyroStone protocol. *PLOS ONE,* 17, 1-12. https://doi.org/10.1371/journal.pone.0267163.

Grosman, L. (2016). Reaching the Point of No Return: The Computational Revolution in Archaeology. *Annual Review of Anthropology, 45,* 129-145. ISSN: 1545-4290. https://doi.org/https://doi.org/10.1146/annurev-anthro-102215-095946.

Hagberg, A. A., Schult, D. A. & Swart, P. J. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. In G. Varoquaux, T. Vaught & and J. Millman (eds), *Proceedings of the 7th Python in Science Conference*. (pp. 11–15). Pasadena, CA USA: SciPy.

Harmand, S. et al. (2015). 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature,* 521, 310-315. https://doi.org/10.1038/nature14464.

Herzog, I. & Scollar, I. (1991). A New Graph Theoretic Oriented Program for Harris Matrix Analysis. In S. Rahtz & K. Lockyear (eds), *CAA90. Computer Applications and Quantitative Methods in Archaeology 1990.* (BAR International Series 565). (pp. 52-59). Oxford: Tempus Reparatum.

Higham, T. et al. (2009). Problems with radiocarbon dating the Middle to Upper Palaeolithic transition in Italy. *Quaternary Science Reviews, 28,* 1257-1267. https://doi.org/10.1016/j.quascirev.2008.12.018.

Homburg, T., Zwick, R., Mara, H. & Bruhn, K. Chr. (2022). Annotated 3D-Models of Cuneiform Tablets. *Journal of Open Archaeology Data, 10,* 4. https://doi.org/10.5334/joad.92.

Jöris, O., Neruda, P., Wiśniewski, A. & Weiss, M. (2022). The Late and Final Middle Palaeolithic of Central Europe and Its Contributions to the Formation of the Regional Upper Palaeolithic: a Review and a Synthesis. *Journal of Paleolithic Archaeology,* 5, 17. https://doi.org/10.1007/s41982-022-00126-8.

Kot, M., Tyszkiewicz, J. & Gryczewska, N. (2024). Can we read stones? Quantifying the information loss in flintknapping. *Journal of Archaeological Science, 161,* 105905. https://doi.org/10.1016/j.jas.2023.105905.

Linsel, F., Bullenkamp, J. P. & Mara, H. (2023). *3D Data Derivatives of Grotta di Fumane: GigaMesh-processed, Annotations and Segmentations*. Version 1.0.0. Zenodo. https://doi.org/10.5281/zenodo.8288967.

Linsel, F., Bullenkamp, J. P. & Mara, H. (2024a). From Scar to Scar: Reconstructing Operational Sequences of Lithic Artifacts using Scar-Ridge Pattern-based Graph Models. https://doi.org/10.5281/zenodo.10969327.

Linsel, F., Bullenkamp, J. P. & Mara, H. (2024b). Linking Scars: Topology-based Scar Detection and Graph Modeling of Paleolithic Artifacts in 3D. https://doi.org/10.5281/zenodo.10477448.

Lombao, D. (2019). *VRM experiment raw data*. Zenodo. https://doi.org/10.5281/zenodo.3368658.

Magnani, M. (2014). Three-Dimensional Alternatives to Lithic Illustration. *Advances in Archaeological Practice, 2,* 285-297. https://doi.org/10.7183/2326-3768.2.4.285.

Mara, H. (2012). *Multi-scale integral invariants for robust character extraction from irregular polygon mesh data*. https://doi.org/10.11588/heidok.00013890.

Mara, H., Krömker, S., Jakob, S. & Breuckmann, B. (2010). GigaMesh and Gilgamesh 3D Multiscale Integral Invariant Cuneiform Character Extraction. In A. Artusi et al. (eds), *VAST: International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage*. (pp. 131-138). Paris, France: The Eurographics Association. https://doi.org/10.2312/VAST/VAST10/131-138.

Matzig, D. N. et al. (2023). Research compendium for 'Benchmarking methods and data for the whole-outline geometric morphometric analysis of lithic tools'. https://doi.org/10.5281/zenodo.7757171.

Moitinho de Almeida, V. (2024). God or the Devil are in the Details Too: Reusing 3D Digital Resources for Cultural Heritage Research. In *Proceedings of the 6th workshop on the analysis, understanding and promotion of heritage contents*. (SUMAC '24). (p. 33). Melbourne VIC, Australia: Association for Computing Machinery. https://doi.org/10.1145/3689094.3689474.

Mylopotamitaki, D. et al. (2024). Homo sapiens reached the higher latitudes of Europe by 45,000 years ago. *Nature,* 626, 341-346. https://doi.org/10.1038/s41586-023-06923-7.

NVIDIA et al. (2024). Edify 3D: Scalable High-Quality 3D Asset Generation. https://doi.org/10.48550/arXiv.2411.07135.

O'Neill, R. C. W., Yezzi-Woodley, K., Calder, J. & Olver, P. J. (2024). En masse scanning and automated surfacing of small objects using Micro-CT. https://doi.org/10.48550/arXiv.2410.07385.

Pastoors, A., Tafelmaier, Y. & Weniger, G. Chr. (2015). Quantification of late Pleistocene core configurations: Application of the Working Stage Analysis as estimation method for technological behavioural efficiency. *Quartär, 62*, 62-84. https://doi.org/10.7485/QU62_3.

Porter, S., Roussel, M. & Soressi, M. (2019). A comparison of Châtelperronian and Protoaurignacian core technology using data derived from 3D models. *Journal of Computer Applications in Archaeology, 2,* 41-55. https://doi.org/10.5334/jcaa.17.

Pulla, S., Razdan, A. & Farin, G. (2001). Improved curvature estimation for watershed segmentation of 3-dimensional meshes. *IEEE Transactions on Visualization and Computer Graphics, 5,* 308-321.

Quantin, M. et al. (2023). Combining FAIR principles and long-term archival of 3D data. In *ACM Web3D conference*. (pp. 16:1-6). San Sebastian: Association for Computing Machinery. https://doi.org/10.1145/3611314.3615913.

Richardson, E., Grosman, L., Smilansky, U. & Werman, M. (2014). Extracting Scar and Ridge Features from 3D-scanned Lithic Artifacts. In Earl, G. et al. (eds), *Archaeology in the Digital Era: 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA)*. (Computer Applications and Quantitative Methods in Archaeology) (pp. 83-92). Amsterdam: Amsterdam University Press. https://doi.org/10.1515/9789048519590-010.

Richter, J. (2004). Copies of flakes: operational sequences of foliate pieces from Buran-Kaya III level B1. *The Middle Paleolithic and Early Upper Paleolithic of Eastern Crimea, 3,* 233-247.

Seales, W. B. et al. (2016). From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi. *Science Advances,* 2, e1601247. https://doi.org/10.1126/sciadv.1601247.

Soressi, M. & Geneste, J. M. (2011). The history and the efficacy of the chaîne opératoire approach to lithic analysis. *PaleoAnthropology, 2011,* 334-350. https://doi.org/10.4207/PA.2011.ART63.

Swainston, S. (1999). Unlocking the Inhospitable. In: *Dorothy Garrod and the Progress of the Palaeolithic*. (pp. 41-56). Oxford: Oxbow Books. https://doi.org/10.2307/j.ctvh1dppc.9.

Tafelmaier, Y. et al. (2022). *Methods for the Analysis of Stone Artefacts*. Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-39091-4.

Thiery, F. & Mees, A. (2023). Taming Ambiguity – Dealing with doubts in archaeological datasets using LOD. In *Human History and Digital Future: Proceedings of the 46th Annual Conference on Computer Applications and Quantitative Methods in Archaeology*. (p. 744). Tübingen: Tübingen University Press. https://doi.org/10.15496/publikation-87762.

Wilkinson, M. D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3,* 160018. https://doi.org/10.1038/sdata.2016.18.

Wyatt-Spratt, S. (2022). After the Revolution: A Review of 3D Modelling as a Tool for Stone Artefact Analysis. *Journal of Computer Applications in Archaeology, 5(1), 215-237*. https://doi.org/10.5334/jcaa.103.

Abbreviations

GMM	*Geometric Morphometric Methods*
GM	*GigaMesh*
GT	*Ground truth*
GdF	*Grotta di Fumane*
IIoP	*Integral Invariants of Polylines*
LOD	*Linked Open Data*
MSII	*Multi-Scale Integral Invariants*
NFDI	National Research Data Infrastructure (ger. *Nationale Forschungsdateninfrastruktur*)
OS	*Operational sequence*

Authors' Contributions
Linsel, Florian: Author, Coding; Bullenkamp, Jan Philipp: Editor; Mara, Hubert: Supervisor.

About the authors

FLORIAN LINSEL holds a master's degree in archaeology from the University of Cologne and is currently working in the eHumanities group Informatics and Cultural Heritage at the Martin Luther University Halle-Wittenberg. As a PhD student in computer science, his work focuses on the analysis of lithic artifacts with statistical methods to determine the operational sequences using 3D models.

JAN PHILIPP BULLENKAMP received a master's in mathematics from Heidelberg University. Currently he works on segmentation methods on 3D meshes as a PhD student in computer science in the eHumanities group Informatics and Cultural Heritage at the Martin Luther University Halle-Wittenberg.

PROF. DR. HUBERT MARA received a master's in computer science from Technical University Vienna, Austria, and his PhD in 2012 from Heidelberg University, Germany. As part of his doctoral thesis, he developed the GigaMesh Software Framework, which calculates and visualizes the local mathematical curvature of surfaces. He then headed the Forensic Computational Geometry Laboratory in Heidelberg until 2020 and moved to Mainz University of Applied Sciences, as managing director of the Mainz Center for Digitality in the Humanities and Cultural Sciences (mainzed). He then was between 2021 and 2024 head of the eHumanities group Informatics and Cultural Heritage at Martin-Luther University Halle-Wittenberg. Since 2024 he is Professor for Archaeoinformatics in the Institute for Computational Ancient Studies (CompAS) at the Free University of Berlin.

Florian Linsel, Jan Philipp Bullenkamp
Institute of Computer Science
Martin-Luther University of Halle-Wittenberg
Halle, Germany
florian.linsel@informatik.uni-halle.de

Hubert Mara
Institute for Computational Ancient Studies (CompAS)
Free University of Berlin
Berlin, Germany

https://orcid.org/0000-0001-8855-9332
https://orcid.org/0000-0001-6580-8885
https://orcid.org/0000-0002-2004-4153