Arch-I-Scan Colloquium and Workshop –
Artificial Intelligence and Pottery Identification and Analyses
College Court, Leicester April 20-21
Willow Room

Each paper c. 20 mins with 10 mins discussion

Wednesday evening (19th April)
19.00: Dinner at The Cradock Arms (201 Knighton Road)

Thursday morning (20th April)
9.00 – 9.20: Arrival and refreshments (College Court coffee lounge)

Introduction
9.30 – 9.35: Pim Allison – Introduction to Arch-I-Scan Project

Artificial Intelligence and mathematical perspectives (discussion led by Daan van Helden)
9.35 – 10.05: Ivan Tyukin – The story of Arch-I-Scan: the power and limitations of data-driven AI
10.05 – 10.30: Zedong Zheng – Creating a diverse and representative sherds dataset for AI training: Segmentation and simulations

10.30 – 11.00: Refreshments (College Court coffee lounge)

11.00 – 11.30: Tatiana Tyukina – Automated recording of rim diameters
11.30 – 12.00: Qinghua Zhou – Basic machine and deep learning for the classification of ceramic fabrics at Olynthos
12.00 – 12.30: Danai Kafetzaki – Typology and Chronology Algorithmically: Case studies in Sagalassos Red Slip Ware

12.30 – 13.30: Lunch (College Court Restaurant)

Thursday afternoon
Archaeological perspectives
Arch-I-Scan's recording processes and a spin off (discussion led by Pim Allison with Tatiana Tyukina)
13.30 – 14.00: Daan van Helden – Development and refinement of Arch-I-Scan's recording processes
14.00 – 14.30: Fiona Seeley – Through another lens: observations from a Roman pottery specialist on their involvement with the Arch-I-Scan project
14.30 – 15.00: Lisa Nevett and David Stone – A pilot AI-based analysis of ceramic fabrics at Olynthos
15.00 – 15.30: Discussion involving Nick Cooper

15.30 – 16.00: Refreshments (College Court coffee lounge)

Developing approaches to ceramic analyses (discussion led by Fiona Seeley)
16.00 – 16.30: Alasdair Gilmour – Visualising pottery *en masse* through principal component analysis of shape descriptors derived from typological images
16.30 – 17.00: Pim Allison and Daan van Helden – Approaches to analysing inter-site vessel use distribution
17.00 – 17.30: Jesús Bermejo Tirado – Ceramic variability and use-wear analysis as an archaeological index of wealth inequality in Roman Spain: Some methodological reflections
17.30 – 18.00: Round-up discussion

Thursday evening
18.00 – 19.00: Drinks (Garden Lounge)
19.00 – 21.00: Dinner (College Court Restaurant)

Friday morning (21st April)
9.00 – 9.20: Arrival and refreshments (College Court coffee lounge)

<u>Future plans and prospects for Artificial Intelligence in ceramic studies</u> (discussion led by Pim Allison)
9.30 – 10.00: Holly Wright – The ArchAIDE Project: Lessons learned and future work
10.00 – 10.30: Discussion on Arch-I-Scan's data and code preservation and dissemination

10.30 – 11.00: Refreshments (College Court coffee lounge)

11.00 – 13.00: Discussion on future plans and prospects for AI and ceramics

13.00 – 14.00: Lunch (College Court Restaurant)

Arch-I-Scan Colloquium and Workshop – Abstracts

The story of Arch-I-Scan: The power and limitations of data-driven AI

Ivan Tyukin, King's College London - ivan.tyukin@kcl.ac.uk

In this talk we will present the computational concept of Arch-I-Scan and discuss the outcomes of the project achieved so far. We will also discuss a number of technical challenges which the team had to overcome and discuss approaches we adopted in the project to overcome these challenges.

Creating a Diverse and Representative sherds Dataset for AI training: Segmentation and Simulations

Zedong Zheng, University of Leicester - zz288@leicester.ac.uk

The identification and classification of pottery sherds is one of the most important and challenging tasks in the Arch-I-Scan Project. However, the lack of sufficient training data makes it difficult to develop accurate and reliable machine learning models. With this in mind, the project is creating a diverse and representative dataset based on two methods: Segmentation, to isolate the clay pot fragments from the background by using the real photos and simulation, to provide a multi-angle sherd generation method based on a new view of rendering. The proposed way and the resulting dataset will support researchers and practitioners to develop an accurate and comprehensive learning model with more potential, ultimately contributing to a better understanding and preservation of cultural heritage.

Automated recording of rim diameters

Tatiana Tyukina, University of Leicester - tt51@leicester.ac.uk

Measuring the diameter of pottery fragments is one of the key elements of the documenting process in archaeology. If the fragment shows the rim, the measurement is usually obtained via a rim chart. In our work, we will showcase our attempt in creating a tool for the automated recording of rim diameters. We discuss various approaches to measuring a pot diameter off the basis of the photo of a sherd. We consider some issues of the capturing process that can lead to incorrect results and present our recommendations.

Basic machine and deep learning for the classification of ceramic fabrics at Olynthos

Qinghua Zhou, University of Leicester - qz105@leicester.ac.uk

In this paper, I detail the exploration of machine and deep learning approaches for classifying ceramic fabrics carried out on a small sample of ceramic fabric images as a preliminary study in Olynthos. We produced a fabric image pre-processing pipeline that contains both conventional image processing techniques and an automated in-focus regions segmentation neural network. In this study, we implemented (1) basic feature extraction with machine learning classifiers and (2) deep neural networks with ensemble/transfer learning for the classification of ceramic wares.

Typology and Chronology Algorithmically: Case studies in Sagalassos Red Slip Ware

Danai Kafetzaki, Katholieke Universiteit Leuven - danai.kafetzaki@kuleuven.be

In archaeological databases there is currently plenty of digital output in pottery photos, scans, 3D models and drawings. To support archaeological research, automated methods are increasingly utilized in sustaining classification models. Yet, there is potential for advancement in creating, rethinking, documenting and updating typological arrangements by means of label-driven or data-driven algorithmic approaches. Specifically, the definition of a pottery class and the mental image of a class, in its morphology and textual description, show promise in digital archaeological research. In this talk, I add to the tools to construct and maintain a pottery class, incorporating information vagueness using fuzzy logic and statistical techniques. The methodological aspects in class (re-)definition are implemented on two-dimensional technical drawings of Roman finewares produced in Sagalassos and refer to knowledge elicitation, shape quantification, data modelling and interpretation. The results of class (re-)definition are assessed for their relevance in chronological inference, using extrinsic and intrinsic information of the assemblages that the profiles are found. This talk highlights quantitative and qualitative aspects of typological arrangement and provides new perspectives of working with the pottery archaeological record.

Development and refinement of Arch-I-Scan's recording processes

Daan van Helden, University of Leicester - dpvh2@leicester.ac.uk

In this paper I will go over the different iterations of the photography process and the reasoning behind decisions and changes that were made. The process has changed in response to local conditions and intervening pandemics as well as developing insights of both the archaeologists and mathematicians on the team. As we gained a more detailed appreciation of the possibilities of the technology and responded to the difficulties of the task in hand and the emerging limitations of the technology, the photography process had to be adapted and became a more integral part of the AI research than we had perhaps expected at the outset. Covid's initial implacability and time's inexorable progress has meant that some of these changes got bunched up at the latter end of the data gathering period but they have given us useful insights into the extent to which AI technology can deal with the realities of archaeological data.

Through another lens: observations from a Roman pottery specialist on their involvement with the Arch-I-Scan project

Fiona Seeley, freelance heritage professional - fiseeley@tiscali.co.uk

For the London scanning sessions, the Arch-I-Scan project photographed the samian pottery from four of the largest Roman sites excavated in London during the last ten years by Museum of London Archaeology (MOLA), as well as 389 complete vessels held in the collections of the Museum of London. Tens of thousands of sherds were retrieved and photographed from over 1900 boxes of pottery by groups of Arch-I-Scan project members and volunteers for c. 60 days. This paper will include observations on this process from the perspective of a Roman pottery specialist including the benefits and limitations for the project objectives and for those involved, in particular the volunteers and ceramicists.

Fabric Recognition with AI: An Offshoot of the Arch-I-Scan Project

Lisa Nevett, University of Michigan - lcnevett@umich.edu; David Stone, University of Michigan - dlstone@umich.edu

The paper presents the preliminary results of a collaboration between the Arch-I-Scan Project and the University of Michigan to explore fabric recognition with AI at the site of Ancient Olynthos (Greece). In June 2022 we took 2000 close-up photographs of fifth- and fourth-century BCE pottery collected during the fieldwork of the Olynthos Project using a modified version of the Arch-I-Scan Project methodology. The photographs derived from a variety of vessels including those for storage, transport, cooking, and dining. The photographs were first processed so that the computer learned which parts of the image were in focus. Next, we classify these images using supervised and unsupervised methods. The results were promising. A classification using supervised learning showed an accuracy of c.45% among the six wares represented, and an accuracy of c.75% in the top two wares, against a result by chance of c.17%. The results were higher than 80% when the wares were more distinct. The results of unsupervised learning were inconclusive at present, suggesting that a larger and more balanced dataset would be desirable. Our goal is now to raise funding to continue the study.

Visualising pottery *en masse* through principal component analysis of shape descriptors derived from typological images.

Alasdair Gilmour, University of Exeter - acg218@exeter.ac.uk

Typological drawings of ceramic vessels fill Roman archaeological excavation reports and conspectuses, but these detailed images are all too infrequently recognised and utilised as the rich source of archaeological data that they are. Making use of evidence from across the North-western provinces, this paper aims to demonstrate the potential that these images have for providing fresh insights into pottery and its relationship with people in the Roman world. By deriving a select series of unitless shape measures from typological pottery images and visualising them through techniques like PCA (principal component analysis), it is possible to identify and explore variation in ceramic forms within and between type series, archaeological assemblages, and wider inter-artefactual domains in a way that goes beyond etic and regionally variable modern categorisations of pottery forms.

Approaches to analysing inter-site vessel use distribution

Pim Allison, University of Leicester - pma9@le.ac.uk; Daan van Helden, University of Leicester - dpvh2@leicester.ac.uk

The fundamental rationale behind the Arch-I-Scan project is to be able to collate ceramic data, and particularly for *terra sigillata*, from excavations more comprehensively and more efficiently, so that these data are closer to representing a true record of the likely assemblages of such vessels used at these sites than are the current sampling approaches. More comprehensively recorded collections are more useful for consumption-oriented analyses. This paper explores quantitative approaches to *terra sigillata* data from the Vindolanda Trust, Museum of London, and University of Leicester Archaeology Service collections for investigating similarities and differences among the assemblages from these collections.

Ceramic variability and use-wear analysis as an archaeological index of wealth inequality in Roman Spain: Some methodological reflections

Jesús Bermejo Tirado, Universidad Carlos III de Madrid - jbtirado@hum.uc3m.es

In recent years, in the framework of different research projects, we have had the opportunity to analyze domestic assemblages from different sites from Roman Spain. Beyond the traditional value that this type of finds has had in terms of chronological markers, our work has been oriented to the analysis of the use and consumption practices related to these products in the framework of different communities. To this end, we have developed a methodology of analysis based on three fundamental points: the typological analysis of the pieces, the analysis of the production traces related to the operational chains of production, and the use-wear analysis.

The main objective of this paper is to explain our sampling procedure and the different formats of documentation we are generating. To do so, we will detail several examples chosen from a series of cases, such as the Roman villa of Fuente Álamo (Córdoba, Spain), several Roman peasant farms from the Madrilian region, and other urban area contexts from New Carthage (Cartagena, Spain). Finally, we will also discuss the quantitative patterns inferred in this way and the analytical possibilities that this type of methodology offers us in the framework of studies on consumption patterns in the Roman world.

The ArchAIDE Project: Lessons learned and future work

Holly Wright, Archaeology Data Service, University of York - holly.wright@york.ac.uk

The ArchAIDE project worked to optimise and economise identification processes, developing a new system that streamlines the practice of pottery recognition in archaeology, using automated image recognition technology. At the same time, ArchAIDE worked to ensure archaeologists remained at the heart of the decision-making process within the identification workflow, and focused on optimising tasks that were repetitive and time consuming. Specifically, ArchAIDE worked to support the essential classification and interpretation work of archaeologists (during both fieldwork and post-excavation analysis) with an innovative app for tablets and smartphones. The ArchAIDE project was completed in 2019, but there were useful lessons learned, and plans for future work are underway. This presentation will discuss these lessons and the focus of the pending COST Action application: Managing Artificial Intelligence in Archaeology (MAIA).