

WG4 Exploratory Workshop, 31 March, 2020

Holly Wright | Archaeology Data Service | University of York







# The 'R' in FAIR

TO BE RE-USABLE: meta(data) have a plurality of accurate and relevant attributes.

- R1.1. (meta)data are released with a clear and accessible data usage license.
- R1.2. (meta)data are associated with their provenance.
- R1.3. (meta)data meet domain-relevant community standards.





#### R1.3. (meta)data meet domain-relevant community standards.

- Easier to reuse data sets if they are similar: same type of data, data organised in a standardised way, well-established and sustainable file formats, documentation (metadata) following a common template and using common vocabulary.
- A submitter may have valid and specified reasons to divert from the standard good practice for the type of data to be submitted.
- The data's reliability lies in the eye of the beholder and depends on the intended application.

https://www.go-fair.org/fair-principles/r1-3-metadata-meet-domain-relevant-community-standards/





# The data's reliability lies in the eye of the beholder and depends on the intended application.

- If reliability is in the eye of the beholder, then how do we know we can trust data enough to re-use it?!
- What does reliability mean within archaeology?
- How much do we need to know about data before we trust it enough to re-use? Traditionally very subjective:
  - Do I trust the methodology used to create the data?
  - Do I trust the person who created the data as a researcher?





# The Evolution of Rome's Maritime Façade: archaeology and geomorphology at Castelporziano (Data Paper) •

Like

▼ Tweet

Altmetric 0
Citations ?

#### Amanda Claridge<sup>1</sup> and Helen Rendell<sup>2</sup>

- 1. Department of Classics, Royal Holloway, University of London, UK. Email: A.Claridge@rhul.ac.uk
- 2. Department of Geography, University of Loughborough, UK. Email: H.M.Rendell@lboro.ac.uk

Cite this as: Claridge, A. and Rendell, H. 2013 The Evolution of Rome's Maritime Façade: archaeology and geomorphology at Castelporziano (Data Paper), Internet Archaeology 35. https://doi.org/10.11141/ia.35.11

#### **Dataset Location**

This dataset has been deposited with the Archaeology Data Service. doi: 10.5284/1000127

#### Referee

Referee statement by Charles French

#### **Dataset Content**

The dataset (Claridge and Rendell 2010) comprises both original digital and digitised (scanned) archival records relating to archaeological, topographical and geophysical fieldwork carried out by British teams at Castelporziano from 1984 to 2009 and geomorphological and environmental fieldwork in the area during 2006-

2000 (Bandall at al. 2007) Biolist at al. 2000). The data are organized in 9 costions and E2 cylesostics:





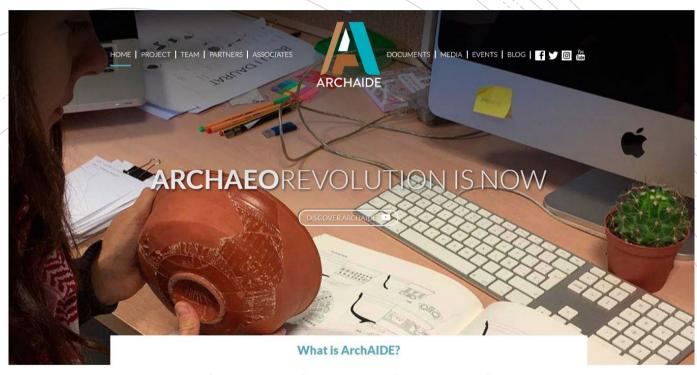


# The data's reliability lies in the eye of the beholder and depends on the intended application.

- Our understanding of how digital archaeological data is created is also becoming more opaque:
  - Digital workflows becoming more complex with more variables
  - Archaeologists are increasingly using AI (image recognition, NLP) but algorithms are not transparent
- Are we trusting data in ways that we should not because we don't understand their biases and complexities?

















archaide.eu









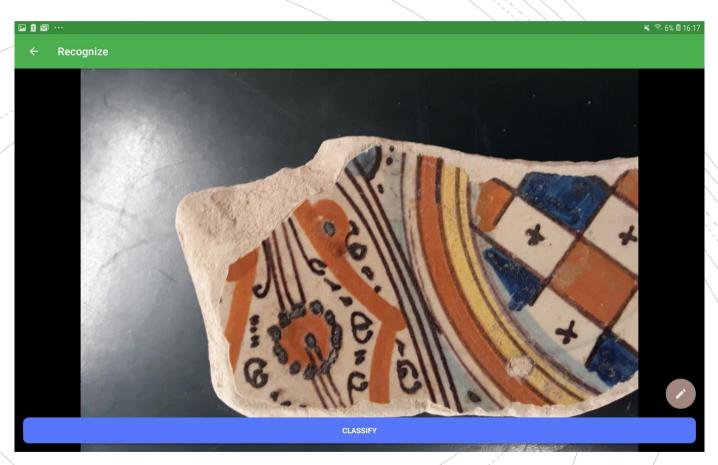


#### **ArchAIDE: Finding a Balance**

- Identifying and classifying archaeological pottery is often a very tedious and time-consuming task; urge to automate as much as possible was strong
- Virtual assemblages: if the (re)user doesn't trust the data created by the app, they will not use it
- Needed to build trust into the workflow: How do we automate mundane tasks while building in key decision points allowing users to trust the identification is correct?

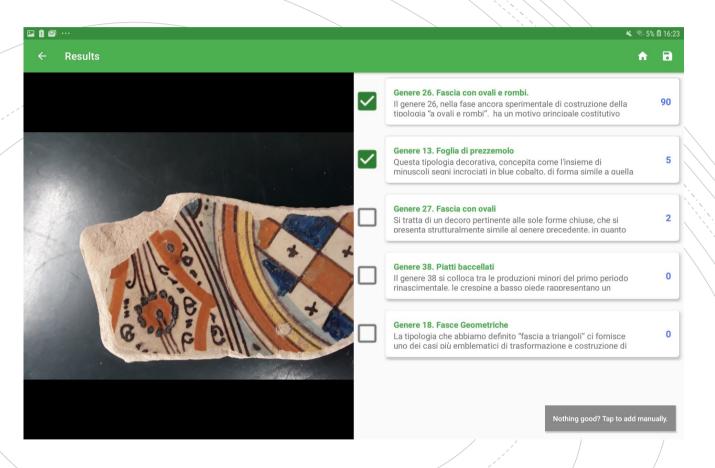








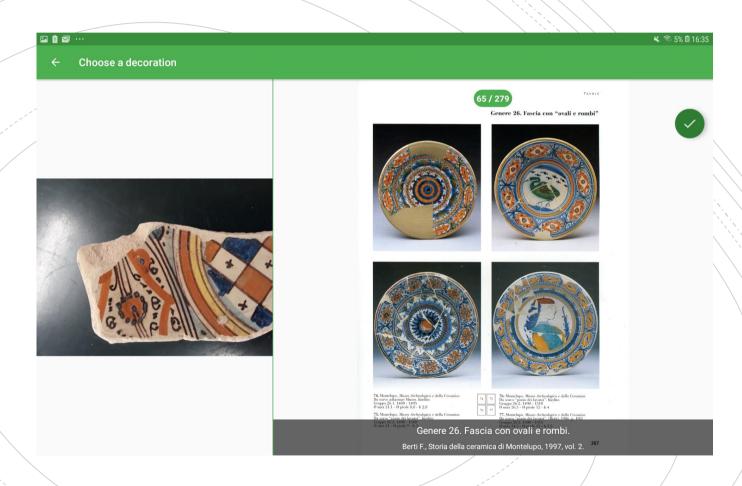




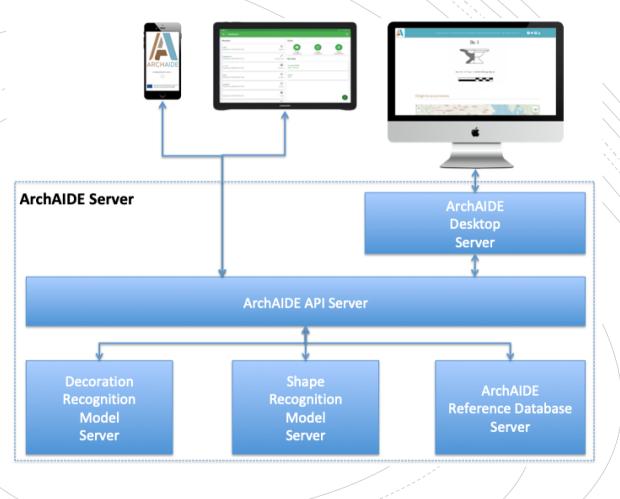




EUROPEAN COOPERATION
IN SCIENCE AND TECHNOLOGY









#### Things to think about for SEADDA

- What is required for archaeologists to trust a resource enough to re-use it? How do we create best practice when we use so many different workflows?
- How do we build in key decision points into opaque workflows to ensure the data can be trusted?
- When we think about re-use, should we be starting with trust?





# Thank you!!

# **Holly Wright**

holly.wright@york.ac.uk

#### **Archaeology Data Service**

http://archaeologydataservice.ac.uk

#### **SEADDA COST Action**

http://seadda.eu

@seadda\_cost

