

An exploration of NLP and NER for enhanced search in osteoarchaeological and palaeopathological textual resources

Can this be the future of bioarchaeological data reuse?

Search

Archaeology Data Service | University of York

NLP - PROCESSING OF TEXTUAL DOCUMENTS

NER - RECOGNITION AND CLASSIFYING OF TERMS



National archive for archaeological data

Over 1 million files 25 years



Freely disseminate digital resources made by research



42km new tunnels >200 archaeologists

30-40m deep >10,000 artefacts

Closed example of documents

Previous works:



ArchaeoTools
(2007)



STAR
(2007)



STELLAR
(2010)



SENESCHAL
(2013)

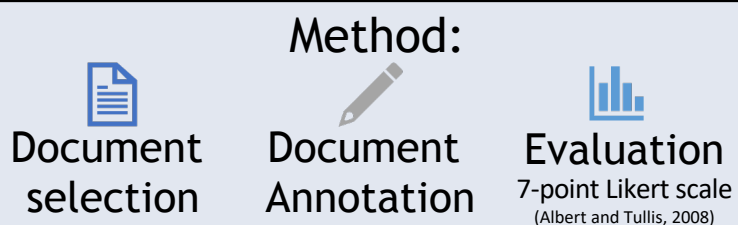


ZooArch
(2017)



OsteoArch
(2019)

(Talboom 2017; Talks 2019)



Osteoarchaeological Entity Search

Filter: Sort By (Alphabetical (A to Z), Alphabetical (Z to A)), LCSH (Scotland--history (3), Archaeology--periodicals (5), Archaeology (51), Fieldwork (20), Leprosy--hospitals (2), Leprosy (2), Excavations (archaeology)--england (2), Grey Literature (35)), General Subject (General (18), Chronological (6), Building (6), Artefact (6), Maritime (1)), Apply Filters

Search results: Showing 159 matching documents

- 0759-01 (0759-01.pdf) Molar ... suggesting that this was an undisturbed primary deposit of an immature pig (the third mandibular molar had not yet erupted). No evidence of butchery was noted. A number of bones showed signs of ...
- Untitled report (arcsmh97_evaluation_report.pdf) Molar ... distal radius (a measurable item) from an adult individual (from [7]) and a horse maxillary molar from a medium pony sized animal. All of the remaining hand recovered bones were shaft pieces ...
- Untitled report (BHB_PXAssessment_text.pdf) Molar ... AD 43-100 ? Skull vault white ? 44 Fill of pot c. AD 43-100 1361 Skull vault ,mandible, third molars, mastoid, rib, radius, ulna, fibula white 2 ? (male and female ?) 45 Fill of pot LIA- AD ...
- Untitled report (Pepper_Hill_human_remains.pdf) Molar

Result: Key:

		Reliable?	Time saving?	Accessible?	Use again?	Useful?
Less than 5						
More than 5						
Value	##					
Expert	26	4.1	3.9	5.1	3.5	4.4
Student	42	4.9	4.9	5.8	5.2	5.3
Layman	15	5.5	5.7	6.3	4.3	6.3
All	83	4.8	4.7	5.7	4.5	5.2

Overall successful, particularly for public and students. Such technologies should be used again.

Thanks to Colleen Morgan, Holly Wright, The Archaeology Data Service, Malin Holst, Carrie Wright and Jordan Hart

Albert, William, and Thomas Tullis. 2013. *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Burlington: Morgan Kaufmann.

Meghini, Carlo, Franco Niccolucci, Achille Felicetti, Paola Ronzino, Federico Nurra, Christos Papatheodorou, Dimitris Gavrilis *et al.*. 2017. "ARIADNE: A Research Infrastructure for Archaeology." *Journal on Computing and Cultural Heritage* 10, no. 3: 1-27. <https://doi.org/10.1145/3064527>.

Richards, Julian, Stuart Jeffrey, Stewart Waller, Fabio Giravegna, Sam Chapman, and Ziqi Zhang. 2011. "The Archaeology Data Service and the Archaeotools Project: Faceted Classification and Natural Language Processing." In *Archaeology 2.0 new approaches to communication & collaboration*, edited by Eric Kansa, Sarah Whitcher Kansa and Ethan Watrall, 31-56. California: Cotsen Digital Archaeology.

Talboom, Leontien. 2017. "Improving the discoverability of zooarchaeological data with the help of Natural Language Processing." *Archaeological Information Systems*, University of York.

Talks, Alphaeus. 2019. "An exploration of NLP and NER for enhanced search in osteoarchaeological and palaeopathological textual resources." *Digital Archaeology*, University of York.

Want help? Please let me know. For a copy of the report or help/support in similar projects contact me at: AlfieTalks@live.co.uk www.AlfieTalks.com/