Arts and Humanities Research Council

Collaborative Doctoral Partnership

Historic England

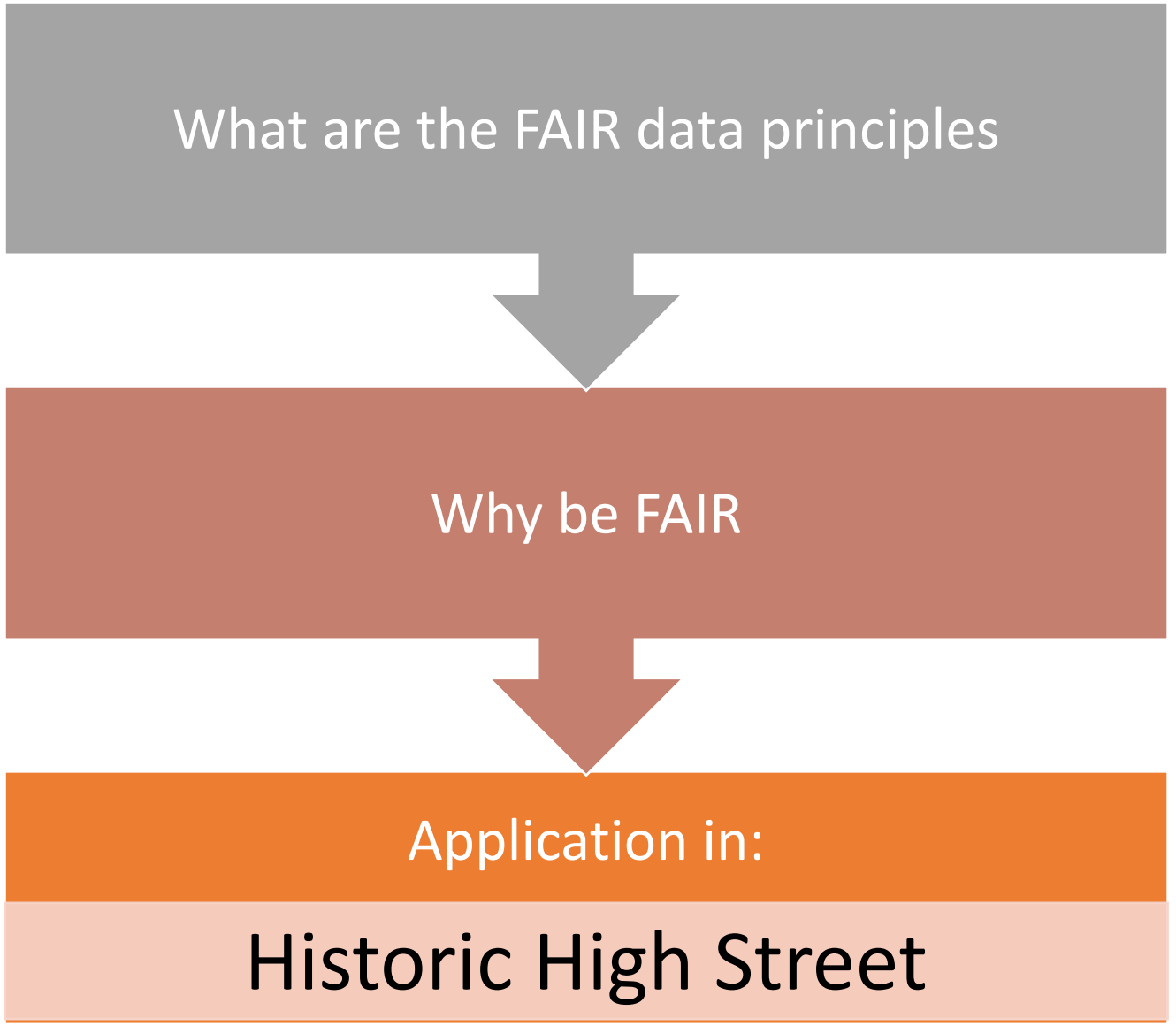
How FAIR is the Historic High Street:

and why this matters

University of York

ads ARCHAEOLOGY DATA SERVICE

Introduction



The FAIR data principles

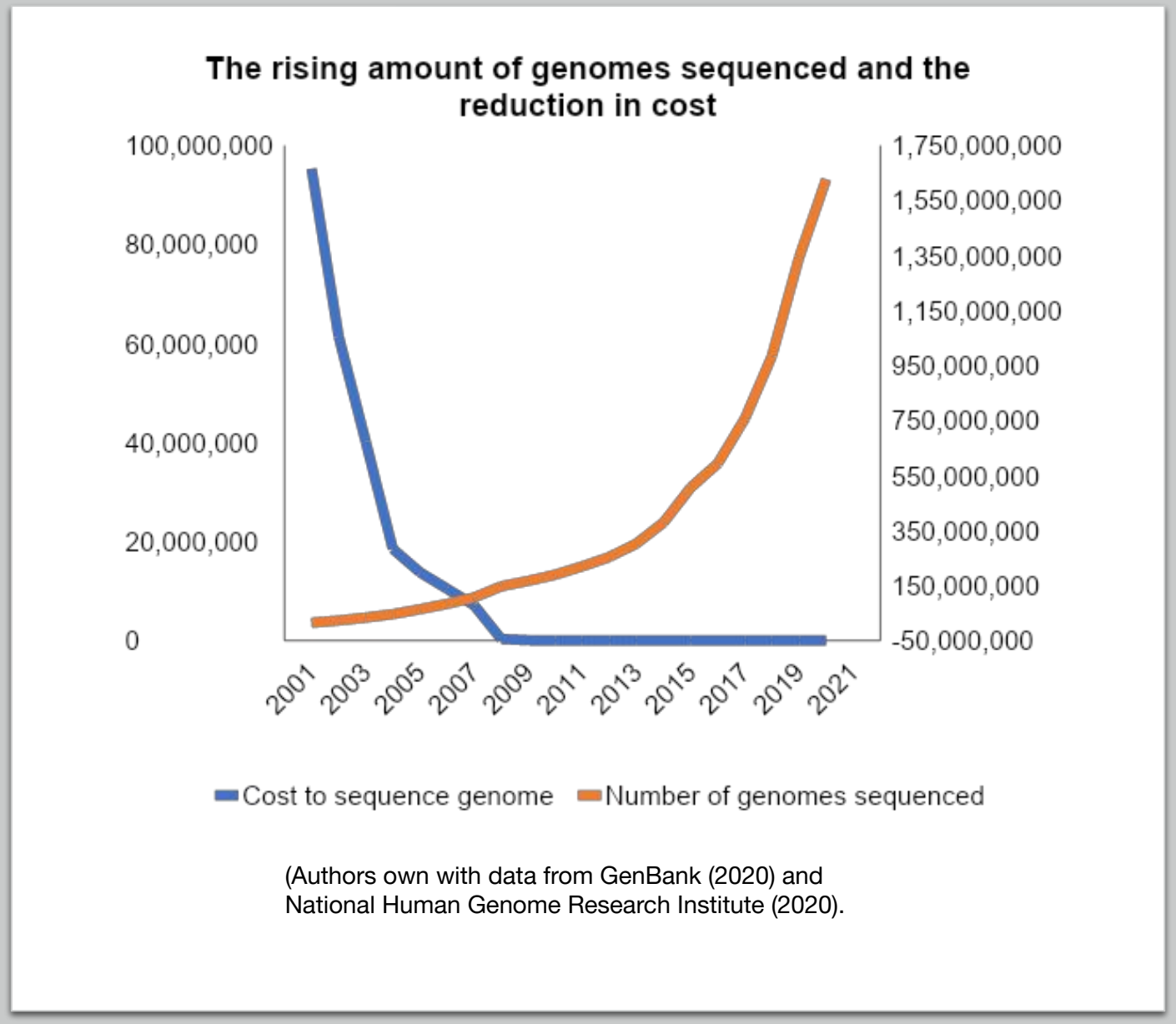
(Wright and Richards, 2020)

Findable	Persistent IDs	Metadata schemas	PIDs in metadata	
Accessible	Communication protocols	Harvestable metadata and Endpoints	Open Access	Repositories
Interoperable	Metadata models	Standardised file formats	Ontologies	Controlled vocabulary
Reusable	Systematic documentation	Community standards	Detailed metadata	Usage licence

(Authors own)

Why be FAIR

- Archaeology is a destructive process (Oakley, 2005, 171; Pálsdóttir, 2019, 2)
- More and more data created (Green *et al.*, 2017, 180).
- Increase in misuse of PDF format (Evans and Moore, 2014, p. 124; Kansa et al. 2020, p. 45; Sobotkova, 2018, p. 121).
- To have access to it in the future



(Authors own with data from GenBank (2020) and National Human Genome Research Institute (2020).

Findable

Persistant identifiers

Metadata Schemas

PIs in metadata

Accessible

Communication protocols

Harvesting

Open Access

Repositories

Interoperable

Metadata models

Standardised file formats

Ontologies

Controlled vocabulary

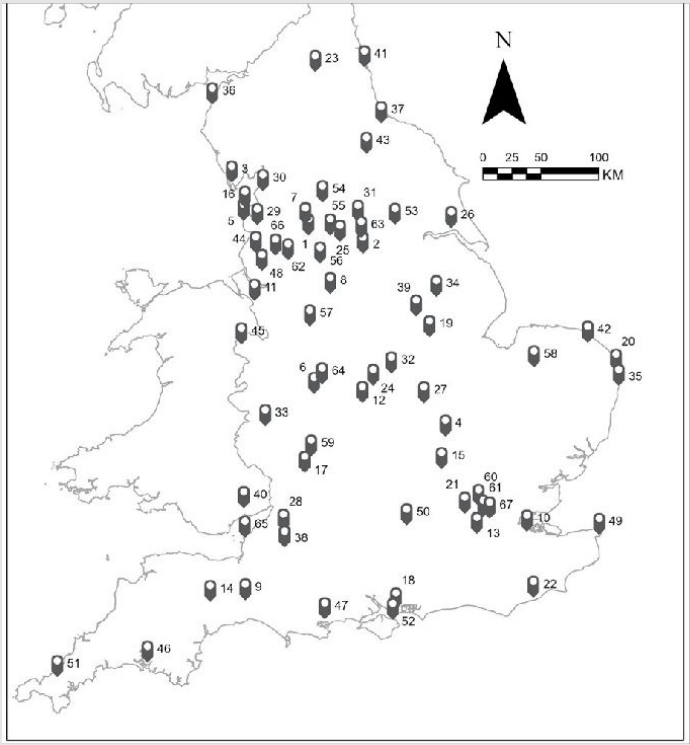
Findable

Systematic documentation

Community Standards

Detailed metadata

Usage license



1 Bacup, Rossendale
2 Barnsley
3 Barrow in Furness
4 Bedford
5 Blackpool
6 Brierley Hill
7 Burnley
8 Buxton
9 Chard
10 Chatham
11 Chester
12 Coventry
13 Croydon
14 Cullompton
15 Dunstable
16 Fleetwood
17 Gloucester
18 Gosport
19 Grantham
20 Great Yarmouth
21 Harlesden
22 Hastings
23 Hexham
24 Hinckley
25 Huddersfield
26 Hull
27 Kettering
28 Keynsham
29 Kirkham
30 Lancaster
31 Leeds
32 Leicester
33 Leominster
34 Lincoln
35 Lowestoft
36 Maryport, Cumbria
37 Middlesbrough
38 Midsomer Norton
39 Newark-on-Trent
40 Newport
41 North Shields
42 North Walsham
43 Northallerton
44 Ormskirk
45 Oswestry
46 Plymouth
47 Poole
48 Prescot
49 Ramsgate
50 Reading
51 Redruth
52 Ryde
53 Selby
54 Skipton
55 Sowerby Bridge
56 Stalybridge
57 Stoke on Trent
58 Swaffham
59 Tewkesbury
60 Tottenham
61 Tower Hamlets
62 Tyldesley, Greater Manchester
63 Wakefield
64 Wednesbury
65 Weston-super-Mare
66 Wigan
67 Woolwich

(Authors own)

Historic High Street

- Ensuring the accessibility and reuse of data created from the High Street
- HAZ – stakeholders for past
- HSHAZ - economic, social and cultural recovery
- Many datatypes

Previous studies

- Historic Town Atlas – interoperability of datasets between cities
- EUS and HLC – how characterisation assists FAIR
- Mapping Medieval Chester – the relationships between datasets
- City Witness – how interoperability helps with lack of contemporary
- Know Your Place – inclusion of community datasets
- Layers of London – how to access community datasets with iteration
- CHARTEX – how to access textual documents using NLP

Methodology

Needs Analysis

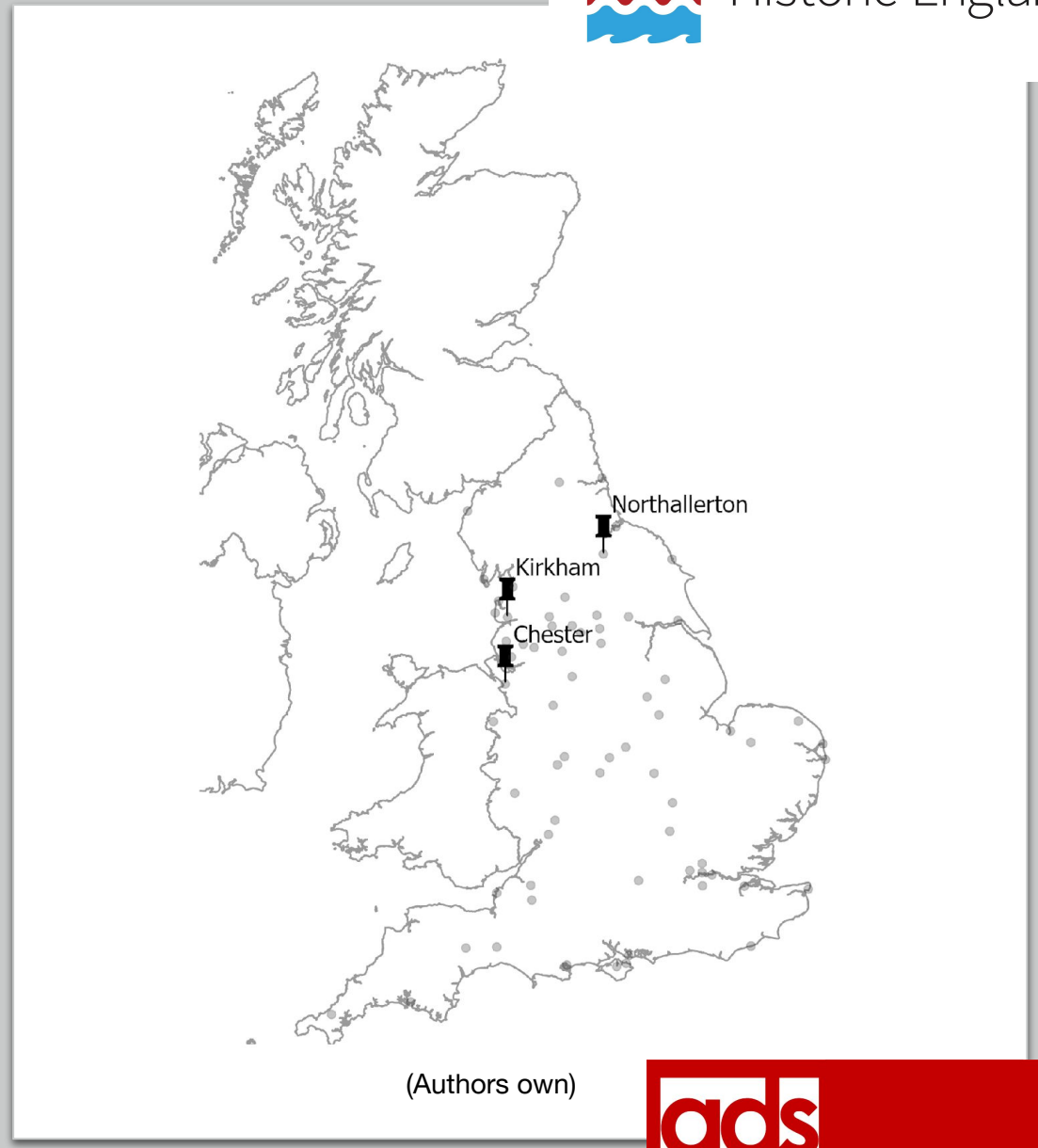
Ensure the long-term preservation and reusability of data to researchers and public

Iterate strategies of FAIR data

4 case studies

Case studies

1. Chester – “complete” dataset, for data capture and management practices
2. Northallerton – what data is being reused
3. Kirkham – beginning of HSHAZ work
4. Fourth?



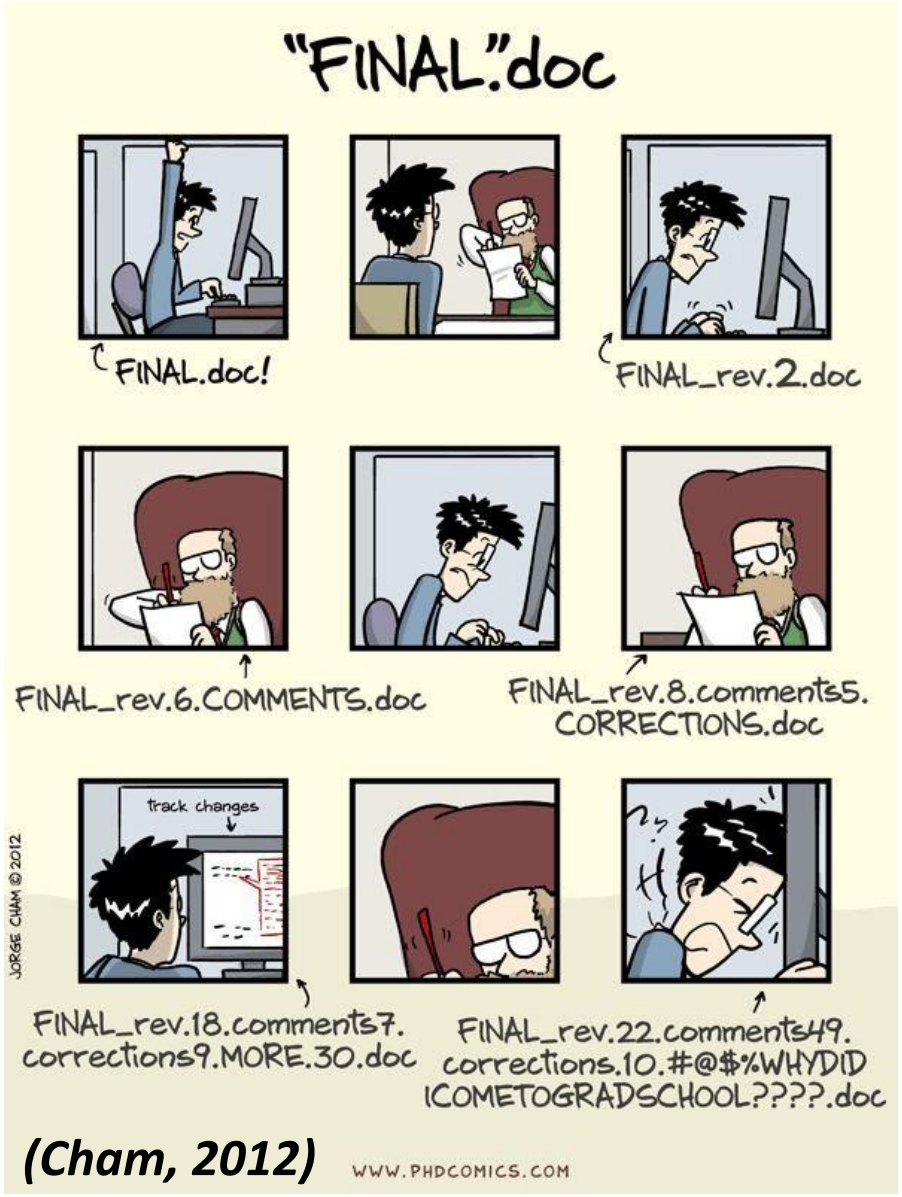
(Authors own)

Historic England

ads ARCHAEOLOGY DATA SERVICE

Take home lessons: Top 10 tips for data management

1. Create an ORCiD account
2. Consider archiving your datasets
3. Make a data management plan and update it
4. Display a clear usage license
5. Have contact details to enable reuse
6. Consider your datafiles – what's their format
7. Systematically document – make it clearer where things are
8. Set up clear data sequencing
9. Consider creating metadata
10. Back up frequently



(Cham, 2012)

Conclusion

WHAT IS THE FAIR DATA PRINCIPLES

WHY USE THEM

HOW TO ACCESS DATA INSIDE PDFS

HOW THEY ASSIST WITH THE HISTORIC HIGH STREET

References

- Wright, H. and Richards, J. (2020). D.5.3 Data Curation Policy. E-RIHS. [Online]. Available at: https://doi.org/10.1080/00934690.2018.1511960 (accessed 13 May 2020).
- Oakley, K. (2005). Forensic archaeology and anthropology: An Australian perspective. *Forensic science, medicine, and pathology*, 1(3), 169–172.
- Pálsdóttir, A. H., Bläuer, A., Rannamäe, E., Boessenkool, S. and Hallsson, J. H. (2019). Not a limitless resource: ethics and guidelines for destructive sampling of archaeofaunal remains. *Royal society open science*, 6(10), 191059.
- Green, E. D., Rubin, E. M., and Olson, M. V. (2017). The future of DNA sequencing. *Nature News*, 550(7675), 179.
- Evans, T. N. and Moore, R. H. (2014). The use of PDF/A in digital archives: a case study from archaeology. *International journal of digital curation*, 9(2), 123-138.
- Kansa, S. W., Atici, L., Kansa, E. C. and Meadow, R. H. (2020). Archaeological analysis in the information age: guidelines for maximizing the reach, comprehensiveness, and longevity of data. *Advances in archaeological practice*, 8(1), 40–52.
- Sobotkova, A. (2018). Sociotechnical obstacles to archaeological data reuse. *Advances in archaeological practice*, 6(2), 117-124.
- Brandsen, A, Verberne, S, Wansleeben, M and Lambers, K 2020 Creating a Dataset for Named Entity Recognition in the Archaeology Domain. In: Calzolari, N, Bechet, F, Blache, P, Choukri, K, Cieri, C, Declerck, T, Goggi, S, Isahara, H, Maegaard, B, Mariani, J and Mazo, H (eds.). *Proceedings of the 12th Language Resources and Evaluation Conference Marseille, 11 to 16 May 2020.* Marseille: The European Language Resources Association, pp. 4573-4577.
- Richards, J., Jeffrey, S., Waller, S., Ciravegna, F., Chapman, S. and Zhang, Z. (2011). The Archaeology Data Service and the Archaeotools Project: Faceted Classification and Natural Language Processing. In: E. Kansa, S. Whitcher Kansa and E. Watrall (eds). *Archaeology 2.0 new approaches to communication & collaboration*. California: Cotsen Digital Archaeology, pp. 31-56.

- Tudhope, D., Binding, C., Jeffrey, S., May, K. and Vlachidis, A. (2011). A STELLAR role for knowledge organization systems in digital archaeology. *Bulletin of the Association for Information Science and Technology*, 37(4), 15-18.
- May, K., Binding, C. and Tudhope, D. (2015). Barriers and opportunities for linked open data use in archaeology and cultural heritage. *Archäologische informationen*, 38(1), 173-184.
- Binding, C. and Tudhope, D. (2016). Improving interoperability using vocabulary linked data. *International journal on digital libraries*, 17(1), 5-21.
- Talboom, L. (2017). *Improving the discoverability of zooarchaeological data with the help of Natural Language Processing*. Unpublished: University of York. MSc Archaeological Information Systems.
- Richards, J. D. (1997). Preservation and re-use of digital data: the role of the Archaeology Data Service. *Antiquity*, 71(274), 1057-1059.
- Holly Wright, Archaeology Data service, pers. comm. February 2019
- Keily, J. (2017). Tunnel: The archaeology of Crossrail. London: Museum of London Docklands.

Arts and Humanities Research Council

Collaborative Doctoral Partnership

Historic England

Any questions?

Alfie Talks: AT1186@york.ac.uk

UNIVERSITY of York

ads ARCHAEOLOGY DATA SERVICE