Please cite this as: Deckers, P. 2025 All the single ladies? Detector finds, dispersed data and suboptimal sources in the study of Viking Age metalwork, Internet Archaeology 68. https://doi.org/10.11141/ia.68.6

All the Single Ladies? Detector finds, dispersed data and suboptimal sources in the study of Viking Age metalwork

Pieterjan Deckers

In a growing number of countries, the activities of hobby detector users are lauded as a vital source of information about the past. The present paper argues that this statement is not unequivocal. An all too idealist view of the impact of detector finds, and of the recording schemes that capture and publish this data, obscures the laborious process that constitutes the heuristic stages of finds research. As a case study of select types of (pre-)Viking Age metalwork found across multiple northern European countries demonstrates, metalwork data are dispersed across multiple sources, often not conforming to academic standards. This situation complicates the retrieval, validation, processing and publication of information and hampers participation of the interested public. Through the case study analysis, the paper explores the factors contributing to the knowledge potential of detector finds data and the suboptimal sources containing such data. As a conclusion, it provides pointers forward for individual researchers and the discipline at large.

1. Introduction

In a growing number of countries in Europe, hobby metal detecting is not just tolerated through legislation and policy but perceived as a vital way to improve our knowledge of the past and to facilitate public participation in archaeology (Dobat *et al.* 2020). Northern Europe, the geographic region most directly discussed in this article, represents a range of attitudes and legislations towards hobby detecting (e.g. Axelsen 2022; Dobat 2013; Rundkvist 2008; Wessman *et al.* 2016). A number of studies have previously demonstrated the importance of detector finds within these jurisdictions (e.g. Bergstrøm 2023; Gilså 2021; Horsnæs 2021; Immonen 2023; Kaas 2024; Martens 2022; Oksanen and Wessman this issue). In particular, Denmark stands out due to its long collaboration between detectorists and archaeologists, in recent times through the successful DIME database.

This paper examines the reality of artefact heuristics across present-day national boundaries in this region. The aim is to reflect critically on the potential that detector finds hold for the generation of new knowledge and for the inclusion of non-professional audiences. Notably, this study highlights the challenges flowing from the diversity and dispersal of data sources. This leads to a number of secondary questions: how dispersed are different categories of finds data, and what effort is needed to optimally exploit these? Given the variable accessibility and quality of data sources, does this effort weigh up to the benefits i.e.

knowledge gain? And finally, what problems and opportunities does this current situation hold for the participation and inclusion of hobby practitioners?

The challenges alluded to above are most stringently illustrated by the contribution to research of the Portable Antiquities Scheme (PAS), the immensely successful voluntary recording scheme for England and Wales which, since 1997, has amassed a freely accessible online database of over 1.8 million artefact records. Highlighting its centrality to the debate, the PAS features as frequently in arguments for collaboration with non-professional detector users, as it does in criticism of permissive approaches (e.g. Brodie 2021; Hardy 2017; Lecroere 2016). In the past decade, the PAS has been repeatedly put forward as a prime example in discussions on the potential of 'Big Data' in archaeology, and has become so large and encompassing that it is employed as a proxy indicator for developments on grand temporal and spatial scales (e.g. Bevan 2012; Cool and Baxter 2016; Cooper and Green 2017; Oksanen and Lewis 2023). This promise of an unprecedented knowledge gain has been adopted with enthusiasm by proponents of a permissive approach to detecting, as evident in the number of PAS-inspired finds recording schemes that have sprung up in recent years, including the aforementioned DIME (Wessman *et al.* 2023).

However, this ideal is not always reflected in the reality of metalwork studies. While in some British artefact studies, the PAS delivers the vast majority of the relevant information (e.g. Webley 2020), in others this number is much lower. For example, recent comprehensive catalogues of Bronze Age axes found in Britain, and of Roman coins from England and Wales included, respectively, 22.6% and c. 35% records originating in the PAS (Bevan *et al.* 2024; Henry 2024, 11). In many cases, a diversity of sources needs to be consulted to create a complete inventory of relevant metalwork artefacts. For instance, for a study of early medieval equal-armed brooches, in addition to the PAS, the author consulted Historic Environment Records (HERs) maintained by local authorities, the archives of regional and national museums, the UK Detector Finds Database run by the detecting community, and a long list of publications, many of which needed to be trawled manually and cross-checked for inconsistencies and overlap (Weetch 2014, 26–35). This effort is not just a matter of compiling a complete catalogue, but also one of mitigating bias, as different object types can be represented in these sources to different degrees (e.g. Weetch 2014, 39; Henry 2024, 11–12).

If this routine heuristic work of collecting data on metalwork artefacts within England and Wales can rely so heavily on other sources besides PAS, both in quantitative and qualitative terms, the same is true to even greater extent in other countries where detector finds are presently not recorded and published systematically, or where detector finds recording initiatives have not reached the same level of hegemony. For instance, Webley (2020, 105) notes how only 34% of the non-English artefacts in his inventory are the outcome of metal detecting. In this respect, not much has changed since 1997 (the year PAS was launched), when “archaeologists (had) to manage with a complex set of overlapping collections of information gathered by different individuals and organisations at different times for different purposes” (Fraser 1997, 19). Evidently, the key challenge in metalwork studies, in Britain and elsewhere, remains better characterised as 'data dispersal' than as 'data deluge', the term used by Bevan (2015) to describe the unmanageable and increasing amount of data threatening to drown archaeologists in this digital era.

Andrew Bevan and Jeremy Huggett (both citing PAS as a key example) have justly highlighted the issues that spring from combining multiple archaeological datasets (Bevan 2012; Huggett 2015). However, information comes in much more dispersed and messy formats than the structured digital datasets these authors have in mind. This results in a 'tedious and frustrating labour' of retrieval and integration (Chapman and Wylie 2016, 95–97) or, more analytically put, the 'sociotechnical struggle' of data friction: making information travel between 'sites of use' (data formats, people, institutions, scholarly disciplines) incurs 'costs in time, energy, and attention required simply to collect, check, store, move, receive, and access data' (Edwards 2010, 84, also see Huggett 2022, 282).

A second, frequently lauded aspect of public finds recording schemes is the broad access to archaeological data it affords to members of the public, and the participation and inclusion this fosters (e.g. Beck and Neylon 2012, 490; Thomas and Pitblado 2020, 1063). This potential for democratisation is a stated aim of the PAS and of initiatives inspired by it, and constitutes an additional argument for the permissive policies and attitudes to metal detecting of which recording schemes are typically a central element (Dobat *et al.* 2019; 2020; Lewis 2016; Wessman *et al.* 2019). However, the impact of such benefits is partly determined by the accessibility and quality of the resources available to the public. In this respect too, the dispersed state of finds data raises the question to what extent such goals can be realised beyond the level of contributory citizen science, i.e. crowdsourcing (Wessman *et al.* 2023, 91).

In short, there exists an idealised 'PAS premise' of knowledge gain and inclusion driven by the collaborative construction of a coherent, representative and accessible dataset, which remains unrealised to a greater or lesser extent in the countries that have adopted this model. The key reason for this is the dispersed state of artefact data. Through a case study of the inventorisation of certain groups of Late Iron Age Scandinavian metalwork (outlined in Section 2), this paper aims to expose the seemingly mundane and often hidden heuristic work that is required, to explore its implications, and to consider how challenges can be mitigated or overcome. To that end, Section 3 provides a brief review of the data. Section 4 analyses the knowledge contribution from including detector finds and integrating dispersed source material, while Section 5 discusses the challenges and costs incurred in doing so. Finally, in light of this cost-benefit balance, Section 6 discusses strategies for individual researchers and proposes ways to address the 'PAS premise' for the discipline as a whole.

2. The case study

In order to address these questions, this paper analyses the dataset behind a previously published study of a set of casting moulds from the Viking-Age trading place Ribe (Deckers *et al*. 2021). As part of that study, metalwork finds belonging to the same classes as the impressions in the reconstituted mould fragments were catalogued across northern and western Europe. The cut-off date for this data collection was 2020. The data collection was limited to information gathering from the published record, online sources and direct contact with colleagues. To the best of the author's knowledge, no data were included on finds which could be considered illicit. Social media and hobbyist resources were consulted only for countries where metal detecting can be practiced legally.

The artefact types under consideration are represented in Figure 1 - brooches depicting horses with and without riders, pendants depicting hair-gripping figures, and so-called 'Valkyrie' pendants or mounts. Horse-shaped brooches (in some cases possibly serving as pendants) date partly to the pre-Viking period, but at least one subtype with realistically depicted horse-tack dates to the early 9th century CE (Kleingärtner 2003; Nancke-Krogh 1978). 'Valkyrie' pendants or mounts are largely 9th-century in date. Although having attracted scholarly attention before (e.g. Pentz 2018; Wicker 2020), Deckers *et al.* (2021) represented the first study aiming to collate a full corpus, including as many examples of this artefact class as could be retrieved from both academic publications and less traditional sources, especially taking advantage of information shared in online communities of hobby detectorists. Meanwhile, a new publication has further expanded the corpus of one class of

'Valkyrie' pendants/mounts (Gardeła *et al.* 2022), but these new finds and any other published after 2020 are not considered here.

Figure 1: Examples of the principal classes of artefacts considered here: 'Valkyrie' pendants/mounts class 1, single unarmed woman (a); class 2, single armed woman (b); pendants/mounts depicting hair gripping figures (c); horse-shaped brooches (d); class 3, scene with armed woman and rider (e); brooches depicting a horse and rider (f). (a) Klinta. Gabriel Hildebrand, Historiska museet/SHM; (b) Skodsebølle. National Museum of Denmark; (c), (d), (e) Tissø; National Museum of Denmark/Pia Brejnholt; (f) Birka bj. 825. Christer Åhlin, Historiska museet/SHM. All images CC-BY 4.0.

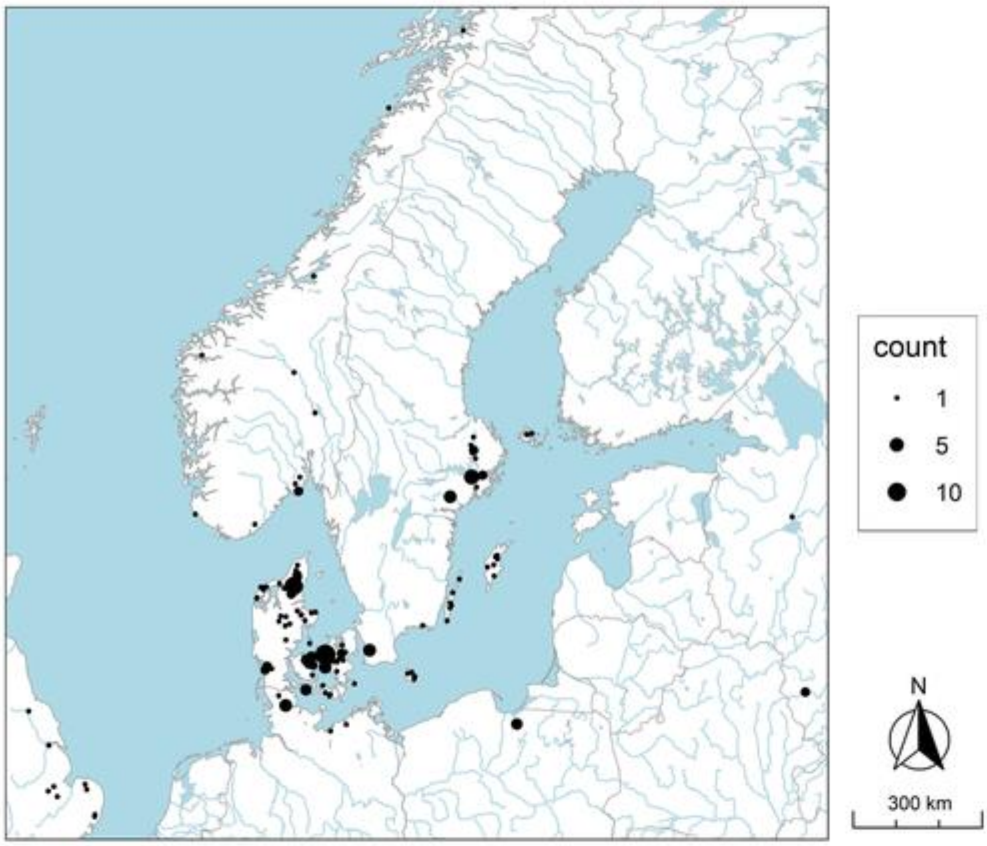

Figure 2: Distribution map of finds in the inventory, grouped by finds location (image by author). Bounding box coordinates: xmin -2.476932, ymin 51.652750, xmax 32.865994, ymax 69.281149

- The finds inventory with locations (.xlsx)
- Inventory description with references (.pdf)

This pre-existing artefact inventory is repurposed here for an analysis of the heuristic process. With this in mind, the primary dataset was enriched with paradata concerning the contexts and pathways of the information (Huvila 2022). This included the provenance of each find (whether it stems from excavation, antiquarian activity or hobby detecting); the type and year of the earliest retrievable publication in which it was mentioned; and the earliest academic or most accessible and reliable source type in which the artefact was published by the time of the review in 2020. The source types are ranked by perceived ease of use and scientific reliability in Table 1 and labelled 'optimal' and 'suboptimal' accordingly. In section 5, these characteristics will be explored and critically examined in terms of aspects adopted from the FAIR principles for open digital data – findability, accessibility, interoperability and reusability - and including a reflection on the situated nature of this assessment. The dataset was further expanded with information to serve the analysis on the knowledge value of finds. This includes a new typological classification, the criteria for which are explained in Section 4.

Table 1: Overview of source types consulted for this study

Source type	Description	No. of items (publication status 2020)	Designation in the analysis (see text)
Artefact databases	Online finds recording schemes (e.g. PAS, DIME) and online museum catalogues (e.g. British Museum, Unimus, SOL/Sydvestjyske Museer)	19	optimal
Academic literature	Monographs, contributions in edited volumes, journal papers, exhibition catalogues	110	optimal
Grey literature	Archaeological reports; dissertations; institutional databases (e.g. *Danefæ*)	32	suboptimal
Popular media	Popular archaeology magazines (e.g. *Skalk*), detectorist magazines (e.g. The Searcher), institutional webpages and blog posts	8	suboptimal
Hobbyist resources	Detectorist fora, community databases	8	suboptimal
Social media	Facebook pages and groups dedicated to hobby detecting	18	suboptimal
Unpublished information	Retrieved through personal contacts with archaeologists, museum staff and finders-collectors	9	suboptimal

In the context of this analysis, it is important to note that the dataset represents a thorough but not a completely exhaustive heuristic effort. Some obscure publications of artefacts may have been missed. Likewise, no observations from hands-on study are included in the analysis. Such an objective was not feasible in the context of COVID-19 lockdowns, and would also have been hindered by the dispersal of finds across many collections held by institutions and private individuals. However, it is improbable that this affects the points developed in this paper in any meaningful way. In fact, the principal arguments made here regarding the accessibility and usefulness of source material rely exactly on the fact that this represents a realistic case study. The case study thus stands as a real-world example of commonly employed 'distant' approaches to metalwork artefact corpora (as exemplified in the studies cited in the Introduction).

3. The nature of the data

To better grasp the composition of the dataset, a distinction was made between finds resulting from scientific or development-led excavation (n=44), from hobby metal detecting (n=123), and from antiquarian activities or other circumstances which cannot be readily deduced from the publications record (n=35) (Figure 3). Immediately striking in Figure 3 is how the numerical importance of detector finds is determined most strongly by the preponderance of those finds in the Danish corpus (106 detector finds vs. 18 from other provenances). As noted in the introduction, this highlights the impact of a permissive and cooperative approach to detecting, characterised by the long-standing *Danefæ* system and, more recently, the online finds recording scheme DIME.

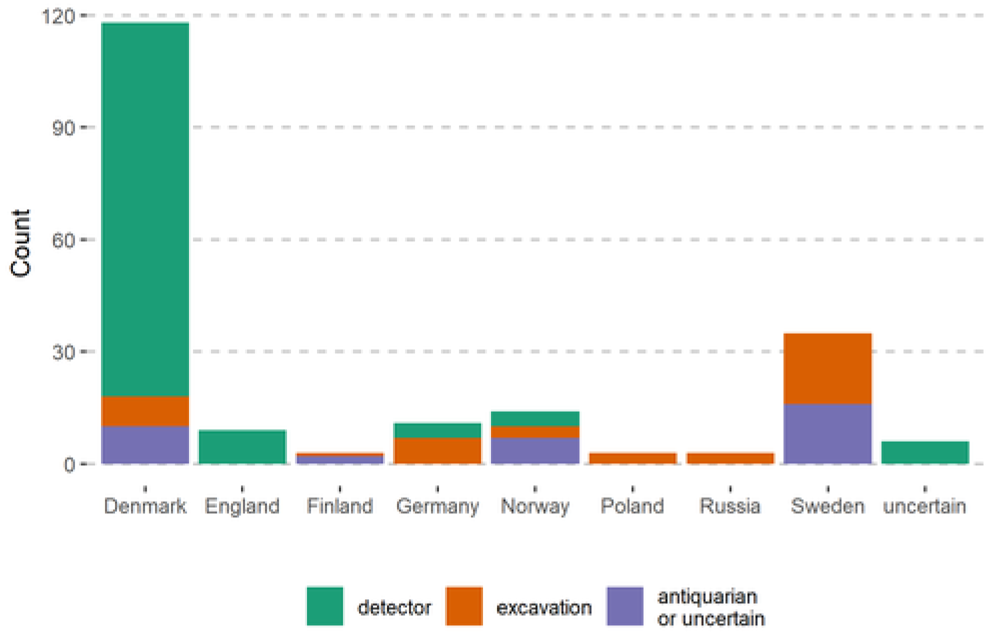

Figure 3: Provenance of finds of all classes considered in this paper, by country (image by author)

The data for England similarly illustrate the impact of permissive regulation and intensive public finds recording through the PAS. Here, the artefact classes under consideration are so rare that they have yet to turn up in regular excavation; all 9 relevant artefacts were recovered through hobby detecting. For Norway and Germany (Schleswig-Holstein), which also allow metal detecting but (as yet) have no comprehensive, publicly available finds recording scheme, the impact of detecting is smaller but nonetheless noticeable. This puts Sweden in stark perspective. Discounting hobby detecting, this country yields a corpus of relevant finds larger than Denmark's. Consequently, one may only surmise what detector users could turn up, and how this would affect the current state of knowledge.

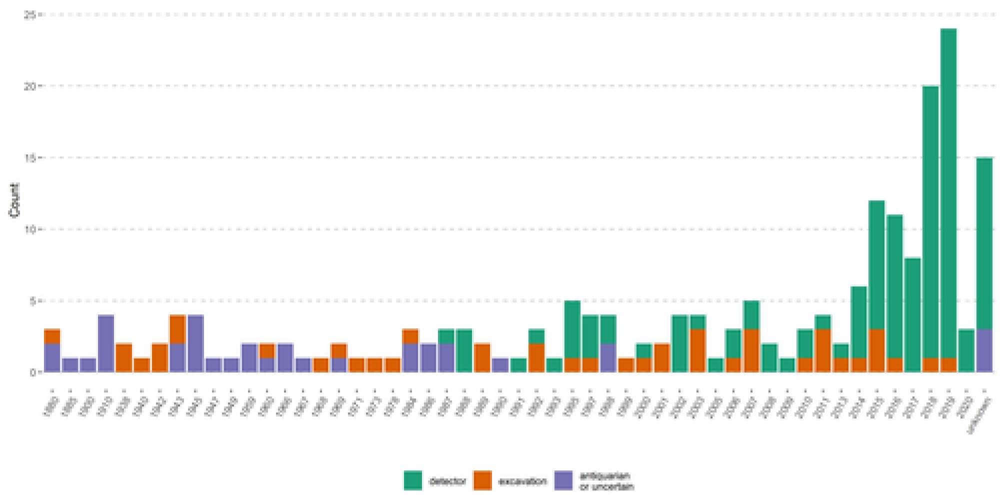

Figure 4: Provenance of finds of all classes considered in this paper, by year of first publication (image by author)

The impact of metal detecting is made all the more noticeable when considering the publication dates of detector finds (Figure 4). While the corpus of artefacts aggregated over a period of 140 years, hobby detector finds only appear in the 1980s. Certainly from the mid-2010s, they form an unprecedented addition, in numerical terms overshadowing by far the cumulative contribution of professional fieldwork.

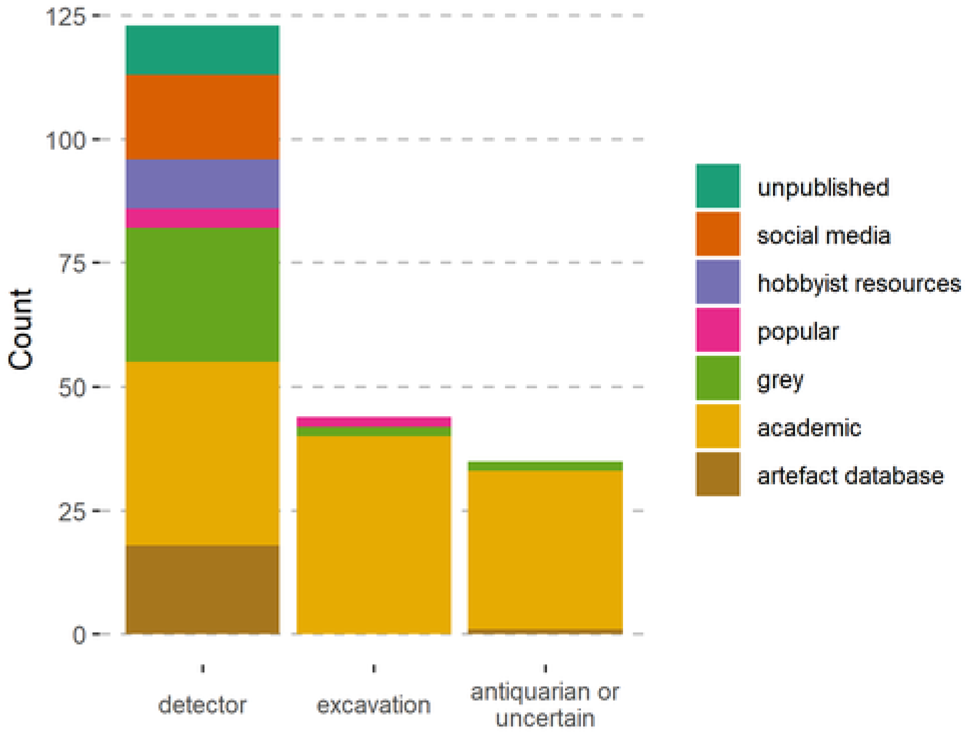

Figure 5: Publication status (situation 2020) of metalwork artefacts considered in the case study, by provenance (image by author)

A range of publication types was consulted during the course of collating the dataset (Table 1). As evident from Figure 5, the provenance of an artefact determines to a significant extent its publication trajectory. From the researcher's point of view, artefacts are preferably published in an easily accessible and reliable format. This can be in a scholarly publication or, as has recently become more common for detector finds in a number of countries, through the finds databases of recording schemes. Figure 5 shows, however, that many relevant artefacts for this study – detector finds especially – were found in a plethora of other source types.

Beyond this dispersal, what stands out is the suboptimal nature of many of these sources: far from academic sources such as journal papers or reasoned finds corpora, many detector finds are published in places that are ephemeral and/or difficult to access, including grey and popularising literature, social media and hobbyist websites. Other finds remain unpublished and can only be found through personal contacts with archaeologists, museum staff or finder-collectors themselves.

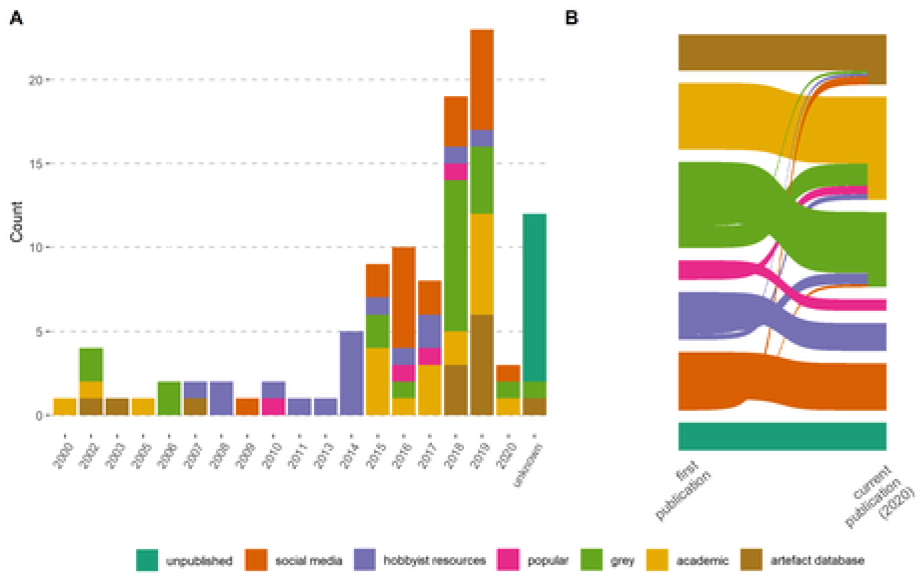

Figure 6a: Nature of the first publication of detector finds of all classes, first published between 2000-2020 or publication date unknown; 6b: Data journeys of all detector finds of all classes, between their first publication and their publication status in 2020 (image by author)

Figures 6a and 6b zoom in on the 'data journeys' (Huggett 2022) of detector finds from the last two decades under consideration in this study. As shown in Figure 6a, detectorists began publishing their finds on online fora and homegrown databases from the mid-2000s onwards. Around 2015, social media (notably, Facebook) supplanted such homegrown initiatives as the main outlet for sharing information on detector finds. Also visible are the contributions of artefact databases (notably, the Danish DIME since 2018) and a number of individual academic publications that added significantly to the dataset (e.g. Christiansen 2019).

A substantial amount of detector finds remains unpublished through more formal venues. Of the 63 detector finds for which an initial publication occurred through suboptimal sources between 2000 and 2020, only 8 had been published through academic publications or publicly accessible artefact databases by 2020. It would be a mistake to dismiss such material as 'floating culture' (Daubney 2017), circulating indefinitely, detached from its contextual information and hence of limited use for research. In many cases, it is just a matter of time for finds to be studied and published in academic venues. In Denmark, public finds typically take some years to percolate through the *Danefæ* system. This entails that they are first collected at regional museums, before being selected and further processed at the National Museum. Nonetheless, more generally the slow and partial progress of these data journeys is notable (Figure 6b), posing a significant hurdle to the exploitation of this material.

4. Quantifying knowledge gain

The aim of spending considerable time and effort surveying both academic and less traditional sources for metalwork finds relevant to the case study was, obviously, to maximise knowledge gain. In a sense, the simple quantitative analysis above can stand as a first indication of this knowledge gain. However, the notion of marginal returns should be considered: what is the return on investment of collecting information on detector finds in particular, given their complicated publication history?

In this section, knowledge gain is explored by adopting a standard analytical approach to metalwork artefacts: the development of a classification based on morphological criteria, followed by a consideration of the geographical distribution of the identified types or classes. The new information this adds in terms of morphological diversity and spatial patterning will then be assessed, differentiating between the contribution, firstly, of hobby detector finds, and secondly, of suboptimal source types regardless of provenance. This part of the analysis will focus on a portion (n=140) of the dataset considered so far, notably those belonging to the horse-shaped brooches and to classes 1 and 2 of the 'Valkyrie' pendants/mounts: single-standing female figures, respectively unarmed and armed.

4.1. Class richness

Moving beyond the basic division of 'Valkyrie' pendants/mounts into (long-recognised) classes, a new branching classification is offered here that demonstrates the potential for subgroupings within the material (i.e. class richness) and illustrates the increasing diversity of the corpus as it grew through time (Table 2; also see supplementary materials for the class attributions of individual artefacts). Subclasses are expediently split off based on obvious morphological differences and commonalities, with an added criterion that a subclass must contain at least two members.

Table 2. A classification of 'Valkyrie' pendants/mounts class 1 and 2 and horse-shaped brooches. For 'Valkyries' class 2, reference is made to the variants defined by Gardeła *et al*. (2022) where possible. The provenance of the find that allows the definition of the subclass and its source type are indicated if they are detector finds and/or from a suboptimal source

(Sub)class	(Sub)class description	Earliest possible year of definition with detector finds	Earliest possible year of definition without detector finds	Identifiable in 2020 with only academic publications and artefact databases?	Spatial information added through detector finds and suboptimal sources
'Valkyrie' pendants/mounts		1938	1938	Yes	
1	single, unarmed woman	1993	2011	Yes	Expansion (reliant on detector finds)
1.1	carrying drinking horn	1938	1938	Yes	Expansion (reliant on detector finds and suboptimal sources)
1.1.1	head tilted backward, hairknot on top of head, small horn/drinking cup held outwards by short arm stretched out straight	1992	NOT	Yes	
1.2	hand raised to neck, gripping shawl, necklace, strand of hair?	1940	1940	Yes	Expansion (reliant on detector finds)
1.2.1	simple facial details, pretzel-like hairknot, characteristic dress design	1995	2003	Yes	
1.3	no arms depicted (folded under shawl?)	1938	1938	Yes	Expansion (reliant on detector finds and suboptimal sources)
1.3.1	slender figure with stylized dot-pattern decoration	2015	2015	Yes	
2	single, armed woman	1997	2019	Yes	Expansion (reliant on detector finds)
2.1	armed with sword and shield	2007	NOT	Yes	Expansion (reliant on detector finds and suboptimal sources)
2.1.1	conical helmet, shield-arm behind shield	2017	NOT	No	Regionality (reliant on detector finds and suboptimal sources)

2.1.2	shield carried under-arm and inside out, with shield grip held horizontally (Gardela et al. 2020, variant 3)	2014	NOT	Yes	Only known from detector finds
2.1.3	sword diagonally upwards and sheath diagonally downwards behind shield	2020	NOT	No	Expansion (reliant on detector finds and suboptimal sources)
2.2	armed with sword, shield and horn (Gardela et al. 2020, variant 5)	2019	NOT	Yes	Regionality (reliant on detector finds and suboptimal sources)
2.3	armed with shield and holding horn	2020	NOT	No	Regionality (reliant on detector finds and suboptimal sources)
2.4	armed with shield and spear (Gardela et al. 2020, variant 6)	2007	NOT	Yes	Regionality (reliant on detector finds)
Horse-shaped brooches		1900	1900	Yes	
1	complex animal-style design	1973	1973	Yes	
2	simple design	1910	1910	Yes	Expansion (reliant on detector finds)
2.1	elongated, stylized profile	1942	1942	Yes	Expansion (reliant on detector finds)
2.1.1	decoration and shape referring to animal styles	1949	1949	Yes	
2.1.2	simple geometric grooved decoration including vertical lines between shoulder and hip	1986	1986	Yes	Expansion and regionality (reliant on detector finds)
2.1.3	baseline, little to no decoration or marked details	1942	1942	Yes	Expansion (reliant on detector finds)
2.2	high, more naturalistic profile with long muzzle dipping down in front of the body	1943	1943	Yes	Expansion (reliant on detector finds and suboptimal sources)
2.2.1	almond-shaped marking on shoulder and/or hip	unknown	NOT	No	Expansion (reliant on detector finds and suboptimal sources)
2.2.2	no decoration except for some naturalistic detailing of the head	1945	1945	Yes	Expansion (reliant on detector finds)
2.2.3	neck bent downwards, geometric decor including ring-dot motifs and lines on the body	2015	NOT	No	Only known from detector finds, suboptimal sources
2.2.4	no detailing, squat neck	2013	2013	No	
3	curved neck, head bent downwards, only two legs showing, support bars between head or tail and body	1910	1910	Yes	Expansion (reliant on detector finds and suboptimal sources)
4	Lying position	1984	1984	Yes	
5	realistic depiction of riding tack	1995	2014	Yes	

Table 2 brings together the results of the classification and additional information to assess knowledge gain. The table shows the year in which each subclass could have been first defined from the corpus of published examples by that year, including or excluding detector finds. In this way, the contribution of detector finds to class richness is made clear. The identification of 10 out of 31 subclasses (32%) relies exclusively on detector finds. These finds were all published after 1990: out of the 16 new subclasses identifiable since then, only 6 could have been made without detector finds.

The epistemic importance of detector finds is particularly evident for the Valkyrie pendants/mounts class 2, of which only two examples were known from excavation by 2020. Without detector finds, we would hardly have known anything about this group, apart from the existence of a small group of variant pendants depicting armed women. With detector finds, this group could possibly have been identified 22 years earlier. In the other two classes, the marginal knowledge gain from detector finds concerns mostly small, more detailed levels of subclasses.

We also gain an insight into the impact of including suboptimal sources in the finds survey. If the survey were limited only to the optimal source types - open artefact databases and academic publications - 6 out of 31 (sub)classes (19%) would have remained unidentified. For the most part, here too, the subdivisions that would be missed are those lower in the hierarchy.

4.2. Spatial distribution

Spatial knowledge gain can be assessed in relation to the (sub)classes described above and conceptualised in two ways. First, by expansion: the presence of one or more items in a region where the type was previously unattested. The regions used in this analysis correspond to countries, with the exception of Denmark (Western Denmark/Jutland vs Eastern Denmark/the islands) and Sweden (southern vs central). Second, by advancing knowledge of the regionality of subclasses: the distribution of at least two items in an area that is significantly constricted in comparison with the wider distribution of their parent class.

The gain from hobby detecting is obvious. Discounting the two subclasses known solely through such finds, 19 out of 29 classes gain meaningful spatial information thanks to detector finds data. Ten classes gain this information due to the inclusion of suboptimal sources. Overall, the 52% of detector finds refine our knowledge on the spatial distribution of 66% of the (sub)classes. Amongst the unarmed 'Valkyries', the return is even greater: detector finds represent 20% of the finds yet yield new spatial knowledge on 57% of classes. Amongst both groups of 'Valkyrie' pendants/mounts, the inclusion of suboptimal sources (regardless of provenance) also yielded a substantial return. This pattern is even more pronounced when considering solely hobbyist resources and social media (resp. 5% and 14% of finds vs. new spatial data on 14% and 25% of classes).

Figure 7. Distribution map of regional subclasses of 'Valkyrie' pendants/mounts class 2 (image by author). Bounding box coordinates: xmin -1.631865, ymin 51.652750, xmax 13.160615, ymax 57.861915

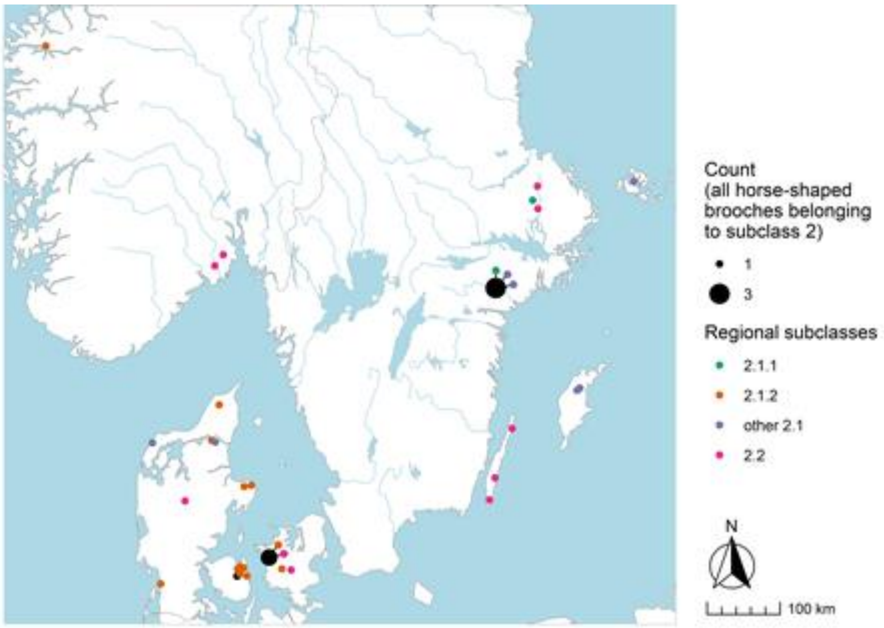

Figure 8. Distribution map of regional subclasses of horse-shaped brooches (image by author). Bounding box coordinates: xmin 4.963231, ymin 54.755547, xmax 20.908514, ymax 62.414102

While earlier publications observed the morphological diversity and broad distribution of 'armed Valkyries' across Denmark, this new analysis highlights the regionality of a number of subclasses (Figures 7 and 8). The two known examples of 'armed Valkyrie' subclass 2.4 were found in close proximity in eastern England at a distance of c.20km as the crow flies, and were likely a local product associated with the persistent Scandinavian presence in that region since the mid-9th century CE. Also notable is the presence of a number of subclasses in western Denmark (subclasses 2.1.1, 2.2) or on the islands (2.3). Horse-shaped brooches also display some geographic patterning. Compared to subclass 2.2, subclass 2.1 is much more common in Denmark than in Sweden (at a ratio of 2.67 to 1 in Denmark vs. 1.2 to 1 in Sweden), and within subclass 2.1, subclasses 2.1.1 and 2.1.2 display the same regionality, appearing to be restricted to central Sweden and Denmark respectively.

These patterns possibly reflect regional networks of craft traditions and social interaction also noted in other aspects of Viking-Age material culture (Sindbæk 2008). Some caution is warranted, however e.g. the apparent regional restriction of 'Valkyrie' pendants/mounts type 2.1.3 in the western Baltic is belied by a casting mould of the same subclass found in Ribe in SW Jutland, perhaps indicating long-distance trading connections (Blair *et al*. forthcoming).

4.3. Factors determining knowledge gain

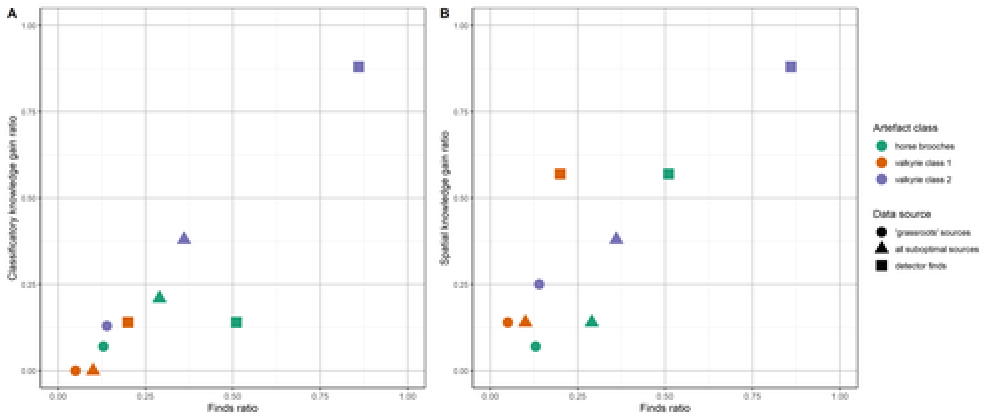

Figure 9: Scatter plot demonstrating the knowledge gained by including detector finds and suboptimal sources within the inventory, differentiated across artefact classes and data sources. 'Grassroots' sources are those built and maintained by the hobby community: social media groups and hobbyist resources such as online fora and community databases. 9a: morphological knowledge (Y-axis reflecting the additional subclasses identified divided by the overall number of subclasses); 9b: spatial (Y-axis reflecting the number of subclasses for which new spatial information was gained, divided by the overall number of subclasses) (image by author)

It stands to reason that, as a finds corpus grows, so does the knowledge that can be gained from it (for class richness e.g. Beck and Jones 1989; Colwell and Chao 2022). This marginal return can be conceived as the rate of knowledge gained for a marginal increase of finds. Here, it is visually expressed as the proportion of subclasses identified thanks to the use of detector finds or suboptimal sources (Figure 9a), and, likewise, the proportion of subclasses receiving new typological or spatial information (Figure 9b), plotted for different artefact classes and data sources. Combined, the data points from both graphs yield an overall Pearson correlation coefficient of 0.85, demonstrating the strong positive correlation of including detector finds and exploiting suboptimal sources with knowledge gain. Nonetheless, notable differences exist across the different artefact groups in this dataset. Although the specific nature of the case study and the small size of the artefact (sub)classes do not warrant a full statistical analysis, placed in context these quantitative results do point towards a number of factors determining the knowledge value of detector finds and suboptimal sources.

The size of the 'traditional' corpus (i.e. non-detector finds from optimal sources) presumably plays a role. Amongst the 'Valkyrie pendants' class 2 (86% detector finds, 36% from suboptimal sources), the knowledge gain correlation is practically 1-to-1. Marginal returns from detector finds and suboptimal sources likely diminish as more knowledge is available through non-detector finds and from optimal sources, both in absolute and relative terms. Amongst the horse-shaped brooch corpus, the largest group under consideration, 51% of detector finds only result in 14% of the new subdivisions. Likewise, the inclusion of suboptimal sources, which amount to about 28% of the horse-shaped brooches, has yielded only one new subclass (out of 14, or 7%), and contributes new spatial information to only one subclass.

The nature of the material also appears to have an impact. Knowledge value, as defined here, is dependent on the potential for hierarchical classification. Some groups of material are difficult to capture in such a scheme, due to their very high or low internal diversity. It is for these reasons that some artefact groups from the larger dataset were excluded from the analysis to begin with: on the one hand, the 'Valkyrie' pendants/mounts class 3 (n=45, of which 36 detector finds) that are for the most part highly similar; on the other, a disparate and selective group of items (n=17) that are iconographically or stylistically relevant but not necessarily susceptible to typological classification.

Fragmentation grade and the quality of the record may also limit the potential for classification. For example, 8 out of 10 horse brooches of type 2.1.2 are detector finds recorded after 2008. Both their apparent lack of diversity or decoration, but also their fragmented state and the relatively low quality of many photographs self-recorded by detectorists on the DIME platform (Dobat *et al.* 2019, 8) precluded further subdivision.

Overall, based on the photographs available, detector finds are fragmented or significantly damaged more often than other finds in the dataset (49.6% vs. 34.2%). This could represent the greater vulnerability of metal in the topsoil to plough damage and other forms of degradation (e.g. Christiansen 2019, 24–28; Haldenby and Richards 2010; Oksanen and Wessman this issue). Alternatively, it might reflect certain biases in antiquarian, professional and academic practices of recovery, recording and publication. Such limitations could partly be overcome by detailed, hands-on study and recording, but this is not always feasible in the type of larger-scale, transnational metalwork studies exemplified here.

The present analysis also likely underestimates the information potential of groups of highly diverse or individually unique items. Typological approaches are less suited for such material, as exemplified in the horse-shaped brooches of subclass 5 (n=12, of which 9 detector finds) or the diverse horse-shaped brooches that could not be attributed to subclasses at all (n=32, of which 17 detector finds). In these cases, the finds data, including photographs and locations, certainly hold potential for knowledge gain through more appropriate research approaches than the one employed here, e.g. from comparative iconographical or stylistic angles.

Finally, contextual factors play a role, notably the nature of the archaeological record and the policies and attitudes towards metal detecting. The unarmed 'Valkyrie' pendants/mounts are a case in point. This group presents a varied picture of knowledge gain from detector finds and suboptimal sources, notably a strong contribution to spatial knowledge, and conversely, a very limited contribution to morphological diversity. A plausible reason for this is that the morphological diversity of the Swedish corpus (n=11), mostly excavated finds from Viking-Age burials, largely determines the overall diversity of the class. The smaller number of (partly) suboptimally published detector finds from Denmark (n=4), where the Viking-Age burial record is much sparser, mainly add new knowledge regarding the geographic distribution of the established morphological classes, but at the same time reinforcing the general image that class 1 'Valkyries' are predominantly a Swedish type.

Despite some limitations and nuances, it is fair to conclude that making efforts to include finds from hobby detecting can be of significant value, even if it requires integrating data sources that are difficult to access and process. This leads on to the next question: what costs and challenges are associated with these benefits?

5. Shades of grey: challenges associated with data from suboptimal sources

5.1. Understanding the suboptimal

Long before the digital revolution, accessibility and reuse of data was considered a major concern for archaeology and other disciplines, encapsulated by problems such as the 'information explosion' and the accessibility of 'grey literature' (e.g. Aitchison 2010; Cunliffe 1983; Frere 1975; Thomas 1991). In the typological scheme of sources outlined in Table 1, the term 'grey literature' has been used in its strict definition of unpublished institutional reports and repositories, but more broadly the term represents 'an assortment of more ephemeral (re)sources unlikely to be systematically collected' (both narrow and broad definitions from Adams *et al.* 2017, 434). The latter definition conforms well with what has been described in this paper as 'suboptimal sources'.

Researchers have noted the heterogeneity of such sources, and the associated uncertainties related to accessibility, reliability, precision, completeness and processes of quality assurance, as well as the time and effort required to retrieve, validate, analyse, preserve and cite them (Adams *et al.* 2017; Kepes *et al.* 2012; Schöpfel 2018). Building on this literature and based on first-hand experience with the source material, this section approaches these characteristics through the lens of the FAIR principles, which allows for more nuanced comparison of the types of source material encountered in the case study analysis.

Table 3: The FAIR principles as originally defined for use with digital datasets, and as adapted to use here in the evaluation of diverse types of archaeological source material

Principle	**Original criteria (from https://www.go-fair.org/fair-principles/)**	**Generalised criteria applied in Table 4**
Findability	Metadata and data should be easy to find for both humans and computers.	How easy is it to find the source, and to locate relevant information within it (i.e. artefact data needed for the case study)?
Accessibility	Once the user finds the required data, she/he/they need to know how they can be accessed, possibly including authentication and authorisation.	What obstacles exist to gain access to the source?
Interoperability	The data usually need to be integrated with other data. In addition, the data need to interoperate with applications or workflows for analysis, storage, and processing.	How easy is it to extract, transfer and integrate information into a new (structured, digital) dataset ready for subsequent analysis?
Reusability	To achieve (the optimal reuse of data), metadata and data should be well-described so that they can be replicated and/or combined in different settings.	How easy is it to convey this information to others, and to lead others to the original source?

While emerging from digital data management in the natural and medical sciences (Wilkinson *et al.* 2016), the FAIR principles are readily adapted to reflect the useability of a broader range of source types, digital and analogue, in the context of archaeological research (Table 3). Inspired by this framework, the data friction generated by the use of such sources can be disentangled into four major facets, each a cluster of criteria representing the advantages and limitations to scientific research of sources for archaeological finds data.

Table 4: FAIR principles as applied to archaeological artefact data. To aid visual interpretation for some readers, the expedient green-red colour scheme offers an impression of higher or lower scores within each of the FAIR criteria (columns)

Source type	**Findable**	**Accessible**	**Interoperable**	**Reusable**	
Finds recording schemes	Centralized, structured, individual finds easy to find (depending on quality and validation of input)	Online, no barriers for most purposes	Digital, structured, completeness and reliabilty depending on data standards and validation	Easily citable, often CC licence (quality of image depending on user or validation)	
Academic literature	Dispersed but strongly referenced	Depending on (physical and online) access and availability	(Mostly) analogue and unstructured; mostly complete, reliable	Easily citable, images may be copyrighted	Increasing effort Decreasing usability ↓
Grey literature	Dispersed; findabiliity depending on references, online findability or personal connections	Depending on (physical and online) availability or personal connections	May be analogue and unstructured; mostly complete, reliable	Easily citable, images may be copyrighted	

Popularizing sources

Popular print media	Dispersed, typically not online, few references	Depending on library access and availability	Analog, unstructured; may be incomplete or unreliable	Usually citable, images may be copyrighted
Institutional websites and blogs	Dispersed, depending on search engine hits; unstructured data	No barriers	Digital, unstructured, may be incomplete	Difficult to cite, not persistent, image licensing not always clear
Hobbyist resources	May be somewhat centralized and structured, with search functionalities (esp. community databases)	Dependent on user access	Digital, not always well-structured, may be incomplete or unreliable	Difficult to cite, uncertain persistence, no clear image licensing
Social media	Challenging (finding dispersed groups, varied terminology and language, finds data dispersed across posts and comments, …); search functionality dependent on the availability of commonly used terminology	Dependent on user access (e.g. closed groups on Facebook)	Digital, unstructured, often incomplete, may be unreliable	Difficult to cite or even irretrievable (due to limited persistence and accessibility); privacy issues, restrictive terms of use
Unpublished information	Depending on personal connections	Depending on personal connections	Unstructured; complete and reliable (depending on personal connections and trust)	Difficult or impossible to retrieve by reader

Table 4 shows that any sources other than the large finds recording schemes databases and, for the most part the academic literature, come with significant limitations to the collection and processing of finds data. They are thus rightly described as 'suboptimal', often in more ways than one.

In any study of a particular class of finds, it is laborious to compile all relevant examples. One aspect is to identify the relevant sources; findability is thus affected by the particular heterogeneity and dispersal of suboptimal sources, as well as by the presence or absence of bibliographic referencing. A second aspect is the ability to locate relevant finds data within a source. Some digital repositories, including some hobbyist resources, store information in structured formats and feature search facilities. In those cases, findability is dependent on the quality of the metadata. Not even finds recording schemes are always optimal in this regard; for instance, classificatory information in the PAS is structured only to a limited degree. On the other hand, finds information is very hard to locate in sources missing both structure and referencing, such as popularizing publications and social media; even if the search capabilities of digital media may mitigate this problem to some extent.

In this case study, a saving grace was that it concerns some of the most iconic artefact classes from the Viking Age, easily recognised as depictions of 'Valkyries' by professionals and hobbyists alike. Consequently, finds belonging to these classes are deemed significant and therefore attractive to publish, even individually, and simple search terms were available that could be expected to yield a majority of relevant items even from obscure sources. In researching artefact groups that are less universally recognisable, the quality of recording, the motivation for publishing and hence, findability, will decrease especially in suboptimal sources. As a result, traditional academic sources will be relatively more important for studying those groups.

Obstacles to accessibility exist in the digital realm (user restrictions in databases, subscriptions and paywalls for digital publications, closed membership on online fora, private groups on social media) as much as in the physical (access to libraries and the availability of publications). The access to some information may even rely on personal communication, for instance with regards to unpublished items in museum stores or private collections.

Interpersonal connections may play a role in overcoming some other barriers too, for instance in gaining access to private online communities.

The transferability of the data - interoperability - presented relatively few problems across different source types in this relatively small-scale case study. Issues here relate to the availability of structured, digital records that can easily be transferred, and to the ease with which visual documentation can be stored and analysed. In these respects, only finds recording schemes with well-structured metadata and export capability are truly optimal. Even so, the lack of digital structured data from most other source types strongly affects only the compilation and processing of larger datasets. As the number of relevant items to be catalogued increases, so does the impact of data friction. In these cases, a growing reliance on more optimal sources can be expected, not just because of the greater cost efficiency in finding, accessing and validating relevant items, but also due to a better interoperability. While reliability and precision may become less of an issue as the larger number of data points compensates to some degree for small proportions of fuzzy or erroneous data, optimal interoperability - more precisely, the availability of structured digital data - becomes a *conditio sine qua non* for the most encompassing studies as the burden of integrating suboptimal sources - even dispersed academic publications - becomes insurmountable. The 'big data' studies cited in the introduction, relying largely or exclusively on the PAS dataset, exemplify this point.

A second, more impactful aspect of interoperability is data quality. Ideally, an artefact record includes a description, photographs and a findspot. Many sources provide this information, although the completeness and reliability may vary for multiple reasons. Uncertainty over author expertise, an issue flagged by several of the authors on grey literature cited above**,** played less of a role in this case study. Conceivably, author expertise may affect the identification - and hence, findability - of a record. However, the expertise and mutual support existing within the hobby community usually lead to a (broadly) correct identification, at least for the artefact groups under consideration. Furthermore, hobbyist sources such as homegrown databases and social media posts typically include photographs - if not always from sufficient or suitable angles - that allow for a necessary validation of the identification and cross-checking to eliminate duplicate records. The twelve records in the dataset without images are all of professional origin, mostly finds listed but not pictured in academic literature.

Beyond the notion of expertise, Ta more useful divide is the motivation of data providers. The context, purpose and intended audience of the publication typically determine which information is shared or withheld. Some of the information may thus be omitted, fuzzy or misleading, an issue that does not just affect suboptimal sources but possibly also the PAS (Brodie 2021, 89–90). In particular, (exact) finds locations may be omitted or generalised because they are considered irrelevant for the purposes of the message (e.g. in academic publications discussing typology), in order to prevent unwelcome visits to finds-rich sites by competing detectorists, or with an eye on protecting the vulnerable archaeological record.

Within the dataset, 20 finds have a location that is known only at country or region level. Three are antiquarian finds (out of 35 finds of antiquarian and uncertain provenance). The remainder are detector finds (17 out of 123), mostly Danish, for which it is likely that a more exact finds location will eventually be recorded by the responsible regional museum. For six of these finds, known through Danish detector groups on social media, no findspot at all is reported.

Reusability, finally, is undoubtedly one of the principal concerns in scientific research, as this has been identified as a fundamental attribute of scientific data: “Making data travel is (...) a necessary (...) condition for their prospective use as evidence” (Leonelli 2015, 7). Criteria here are whether information is uniquely identifiable and citable, and whether it can be reproduced in further research (e.g. with regards to copyright of artefact photographs). An underlying concern, especially for digital sources, is data persistence. Social media posts may be deleted, user accounts closed, and platforms shut down. Many online sources must be considered ephemeral in the mid- to long term, in terms of the persistence of unique identifiers (such as URLs), changes to conditions of use, and the bare availability of data (e.g. Zubiaga 2018). The analysis of reusability reveals a very clear distinction between commonly used scientific resources, and sources of information here labelled suboptimal.

5.2 FAIR, to whom?

The argument in this section so far represents the particular perspective of the trained and research-oriented archaeologist. However, the challenges encountered when exploiting sources of different types are not absolute; they are determined as much by the purpose of the source as by the position and intent of those consulting it. Many sources are not produced to convey scientific data. Reports and institutional databases principally serve to manage cultural resources and collections. Popularizing literature such as archaeology magazines or museum blogs is intended to educate or entertain the public, using artefacts mostly as illustrations in a broader narrative about the past or about archaeological practice. On social media and detector fora, authors share their latest discoveries and ask peers for artefact identification, while hobbyist databases are first and foremost a showcase of finds for the benefit of the hobby community.

For a professional researcher affiliated with a university or museum, access to institutional libraries and paywalled journals is typically not an obstacle, but for independent researchers and hobbyists, it most certainly is. In the databases of finds recording schemes, some levels of data access are typically reserved for vetted researchers (Wessman *et al.* 2023, 98). Competencies play a role as well: for non-professionals, consulting academic sources to retrieve information or using academic conventions to ensure reusability in publication, is not self-evident. We must therefore also pose the question which sources are FAIR to whom, and, ultimately, whether the participative approaches and citizen science philosophies associated with permissive standpoints on metal detecting are as effective as we think.

6. Strategies and future directions in addressing data dispersal

Detector finds can only generate knowledge in particular conditions, notably where a pathway exists connecting finders and researchers: a legal, ethical and infrastructural framework through which detector finds information can be recorded, shared, processed and published, most successfully via finds recording schemes. Even if such pathways exist, detected artefacts and artefact groups must lend themselves to recovery and recording or handing over to authorities by hobbyists. That is to say: by merit of their material, design or socio-cultural associations, they must be identified as items holding archaeological significance and research potential.

Secondly, the information drawn from hobby detector finds is better suited to some types of research than others. As Brodie (2021, 88) points out, many studies involving detector finds are comparable to the case study driving this paper: not site-level studies, which require a holistic approach to multiple material categories and the availability of contextual information; nor in-depth studies on specific (small groups of) artefacts which necessitate hands-on access for examination and scientific analysis. Rather, detector finds shine in the analysis of the morphological diversity and spatial distribution of classes or types of artefacts. Additionally, and importantly, their significance is most keenly realised not in isolation, but through analysis in conjunction with excavated finds that complete the corpus and provide context, including chronology.

Accepting these limitations of detector finds data and, as with any dataset, keeping in mind potential sources of bias (Robbins 2013), the knowledge value of detector finds is evident; the problems rather reside in the sources through which this data is made available. This section identifies some pointers on how to meet these challenges, both for individual researchers weighing the costs and benefits of including suboptimal source types in a finds review, and for an archaeological discipline confronted with the shortcomings inherent in its current approaches to the data- and knowledge-generating activities of the wider public.

6.1. Integrating suboptimal sources in research

As established in this paper, a significant proportion of detector finds data resides in suboptimal formats: dispersed, ephemeral, difficult to access and cite, and sometimes containing imprecise or unreliable information. The analysis shows that it may be worthwhile to include suboptimal sources, but this is not a given. The cost-benefit ratio of including this data may be relatively low, because of the time and effort required, due to the nature of the material, because the data is problematic in some way, or because the knowledge gain is too small relative to the information gathered from more optimal sources. Consequently, the inclusion of suboptimal sources as a heuristic strategy merits closer consideration.

Several criteria for this decision have emerged from this analysis, in part inspired by recommendations made by Adams *et al.* (2017) in their study of the use of grey literature in scientific reviews.

1. If there is a sufficiently large and representative body of data from academic sources, the consultation of suboptimal sources might not be needed.
2. Are suboptimal sources likely to provide information relevant to the research question? Expanding the dataset with information from suboptimal sources may shed light on some aspects of knowledge, e.g. it might shed light on the spatial distribution of a typologically well-described artefact class, or it could help demonstrate the presence of an artefact class in parts of the landscape less likely to yield examples due to biases in the archaeological record. On the other hand, in some cases, pursuing suboptimal sources might not be worthwhile e.g. if exact find locations or high-quality photographs are required.
3. Is the retrieval and use of information from suboptimal sources ethically acceptable? The decision to include evidence from such grassroots sources should be carefully considered within legal and deontological contexts. Beyond the ethics of reusing data from social media (discussion of which currently focuses on privacy e.g. Kessler *et al*. 2023; Lathan *et al*. 2023), ethical concerns about the scholarly use of information reported on hobby fora or social media may be widely held within the professional community, especially in countries where hobby detecting is illegal or strongly restricted (e.g. Lecroere 2016).
4. How salient is the study subject in suboptimal sources? The artefacts studied here, especially 'Valkyrie' pendants/mounts, grab scholarly interest as much as the public imagination due to their evocative imagery and vivid associations. This is not the case for all artefact types, and this may affect the recognition of finds as items of interest, the recording of associated information, and the motivation to publish.
5. Non-academic sources can be overwhelming in their diversity, degree of dispersal and the amount of data they contain. Conversely, it may be uncertain whether any useful information might emerge at all from a review. In those situations, it is useful to explore or sample the information held in suboptimal sources through a pilot study or a limited case study, which may help with delimiting and targeting the heuristic effort.

Several of these criteria require knowledge of local legislation and professional ethics, the local language, the online platforms and other publication venues commonly used, and the type of information likely to be published there. In some cases, personal connections might be needed to gain access, e.g. to unpublished data and closed groups on social media. This local familiarity is an important obstacle associated with suboptimal sources, not just for non-local researchers aiming to study the material from within any country or jurisdiction, but also for research projects with a larger, transnational scope. In this study, for instance, a thorough exploration of hobbyist resources and social media in Poland, the Baltic countries and Russia was out of scope for this reason.

6.2. Addressing suboptimal sources as a discipline

Beyond the heuristic considerations for individual researchers, the discipline as a whole would be served by taking steps towards an improved integration of the prolific suboptimal data relating to finds made by the public.

It is clear that with an eye on findability, accessibility, interoperability and reusability of metalwork finds data, recording schemes such as the PAS outcompete other academic source types. At present, however, for more of the national finds recording schemes and for a broader range of research questions, significant hurdles exist to truly realise the status that PAS data already holds as a big data proxy for certain overarching questions about the past (see Introduction).

In order to achieve this, pathways for unpublished and suboptimally published data towards full recording need to be expanded to improve the representative claims of finds recording schemes. Important hurdles exist here too. One is that a digital recording scheme cannot exist solely as digital infrastructure, but requires flanking measures, not least a network of locally accessible heritage professionals and the concomitant time and effort required to process the flood of finds that the detecting hobby generates (e.g. Axelsen and Fredriksen 2024, 375). A second hurdle is the simple observation that grassroots platforms and social media often suit hobbyists' needs much better for sharing information than academic formats.

In view of these problems, it is worth reconsidering the role of social media in archaeology. As recently pointed out (Schofield *et al.* 2021, 440), social media has mostly been used in archaeology as a medium for outreach rather than a source of data. It will be useful to build guidelines and procedures towards valorisation of suboptimal data by responding to the 'data friction' issues already invoked above, from efficient methods for data capturing to the legality of reproduction. An even more fruitful way forward would be to further enhance and integrate opportunities for public participation.

6.3. Data dispersal and participation

Given the considerable public interest it evidently generates, metalwork finds data holds significant, yet unfulfilled potential for public participation in science. A blossoming ecosystem of autonomous research and interaction testifies to the depth of expertise and the desire for knowledge shared within hobby communities. Often led by community-acknowledged experts, these activities take place on online fora, homegrown databases, blogs and social media, in addition to specialised hobby magazines, both small-scale and with national distribution, such as *The Searcher* in the UK or *Fund&Fortid* in Denmark. Only a small part of this output reaches the academic record, including occasional contributions to peer-reviewed scholarly publications (the Danish brooch table [Abramsson 2014], widely used by amateurs and professionals alike, being a notable example; also see Bondesson and Bondesson 2021; Haldenby 2012).

One view on such autonomous citizen science is to question it as “a strategic maneuver to shape the narrative surrounding (detectorists') actions” (Mroczek 2024, 347). However, more productive would be to explore how to better accommodate such expressions of non-professional interest and improve their value to the discipline as a whole. As it stands, two out of three conditions for 'independent amateur research' identified by Mahr and Dickel (2019, 12–13) are fulfilled in the existence of a shared research object - metal(-detected) archaeological artefacts - and the availability of a peer community with social as well as scientific motivations. It is the third condition, infrastructure for networking (i.e. for making data and knowledge travel) that is currently lacking. To help resolve this, detector finds recording schemes could further expand their role as prominent hubs for participation in archaeology by facilitating data exploration and enrichment. Mapping of search results is a standard functionality (albeit often limited by the restrictions imposed on exact finds locations), but recent developments such as CoinSampo (a dataset of medieval and Early Modern coin finds from Finland) allow the public to filter and visualise data in various innovative formats (Oksanen *et al.* 2023). In this way, the platform encourages the exploration of patterns and the discovery of data-driven narratives.

A further step could be to enable flexible, multivocal enrichment of finds data for research purposes. Some platforms with user-driven recording (notably, DIME) already allow for structured artefact classification by finders. Further expansion of this functionality would not only solve issues of transnational harmonisation of datasets, but also cater for the research interests of various disciplines and interest groups, including those belonging to the wider public (Deckers 2021). In turn, this would further bolster the potential of the explorative functionalities highlighted above.

By becoming platforms for exploration and enrichment, finds recording schemes could create a new virtuous cycle, on top of that driven by crowdsourcing, professional research and (often delayed and selective) public-oriented feedback. By supporting members of the public in their own knowledge discovery and in sharing the outcomes, such developments - within reach of current technology (e.g. semantic computing) and relatively easy to implement at least partially in existing infrastructure - could help bridge some of the gap between academic research and the often-valuable independent research of hobby communities (Lewis *et al.* forthcoming).

For true participatory science to be successful, a further hurdle has to be overcome. As noted before, detector finds recording schemes, even the PAS, do not contain all relevant data needed to develop valid research on metalwork artefacts, and for hobby researchers, it is the academic source material that can be challenging to consult. Information on some artefacts can be gathered from the individual online catalogues of many museums and other heritage instances, but large-scale, open-access digital exposition of data at the level of individual artefacts from excavated assemblages is much less common at present (e.g. Deckers 2023; Feugère *et al.* 2021; Pesonen *et al.* 2024). Consequently, research efforts of non-professionals are hampered by the findability, accessibility, interoperability and reusability of data from dispersed sources, arguably even more than those of professional researchers.

Whereas some restrictions to data access are unavoidable, the problem of the dispersal and public accessibility of data is a vital one to address, and not just from the idealistic perspective of Open Science (e.g. Beck and Neylon 2012; Pétursdóttir 2020). Empowering citizen archaeology to become truly autonomous and integrating the outcomes into mainstream archaeology would benefit the quality of citizen contributions and significantly raise the capacity of the discipline to deal with the 'data deluge'. This viewpoint is no more than an extension of recognizing the unparalleled and unpaid contribution to the discovery of artefacts and sites by the fieldwork of the detecting community. After all, '(t)he world is large and there are few archaeologists in it' (Rundkvist 2009); why not also foster 'data-work' by the interested public?

7. Conclusion

Placed within a broader context, this paper responds to the challenge put forward by Jeremy Huggett (2022, 289), who stated that, in contrast to the nature of the archaeological record, the nature of archaeological data, its journeys and frictions, is sorely understudied. The case study presented here highlights the dispersed and messy state of metalwork finds data in particular, and the numerous forms of friction that any heuristic exercise in artefact studies encounters.

This state of dispersal is one aspect undercutting what has been labelled here the 'PAS premise': the partly implicit aspiration of and/or expectation for detector finds recording schemes such as the PAS to be representative, practically encompassing datasets of archaeological metalwork artefacts. Certainly, in the Nordic Countries, the reality is far different. Their knowledge potential notwithstanding, detector finds have complex and diverse publication histories involving a range of offline and online media driven by professional, institutional and hobbyist efforts, even in countries that are at the forefront of collaborative approaches. Beyond dispersal, the suboptimal character of many sources for (metalwork) finds data adds significantly to the data friction researchers are confronted with, and necessitates well-considered heuristic strategies.

The second aspect of the 'PAS premise' is the partially fulfilled aim of public inclusion. The dispersed state of metalwork finds data throws up obvious barriers for participation and autonomous citizen science. In terms of retrieving and processing information, the problems posed by suboptimal sources faced by non-professional researchers are to some extent the mirror image of those met by academics and other professionals. Nonetheless, the participative potential of finds recording schemes can provide a way forward, if future developments are sufficiently multivocal to foster autonomous citizen science.

Acknowledgements and data availability statement

This paper finds its origin in a presentation given at the workshop 'Current opportunities in public archaeology research - Uses and Users of the Data' organised on 6-7 October 2022 at the University Museum in Bergen, Norway, as part of the workshop series 'From Treasure Hunters to Citizen Scientists' generously funded by NOS-HS grant no. 335213. In addition to the rich discussions taking place during those workshops, the ideas expressed here greatly benefited from the innovative and collaborative environment fostered by the European Public Finds Recording Network. I extend my thanks to all colleagues and friends within these groups. Furthermore, I am particularly grateful to Eljas Oksanen, Anna Wessman and the external reviewer, Raimund Karl, for their thorough review and helpful comments which significantly improved this paper.

- The finds inventory with locations (.xlsx)
- Inventory description with references (.pdf)
- Also available at: https://rdr.kuleuven.be/dataset.xhtml?persistentId=doi:10.48804/2VKN2P

The finds inventory underpinning this paper was compiled as part of the Northern Emporium research project, led by Søren Sindbæk at Aarhus University and funded by the Carlsberg Foundation Semper Ardens grant CF16-0008: Northern Emporium. Additional information was kindly provided during the preparation of this paper by Rikke Søgaard (National Museum in Copenhagen), Andres Dobat (Aarhus University), Claus Feveile (Museum VEST) and Anna Wessman (University Museum of Bergen). Nonetheless, I take full responsibility for any errors and misjudgements in this paper.

Data availability statement: the full artefact inventory used in the case study is available via this article as well as via KU Leuven repository. https://rdr.kuleuven.be/dataset.xhtml?persistentId=doi:10.48804/2VKN2P

Bibliography

Abramsson, G. 2014 'Kronologisk Fibeltavle (inkl. stangnåle, dobbeltknapper og stangknapper) Udg. 1.0'. https://harja.dk/vare/fibeltavle/ [Last accessed 19 April 2025]

Adams, R.J., Smart, P., and Huff, A.S. 2017 'Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies', *International Journal of Management Reviews* **19**(4), 432–454. https://doi.org/10.1111/ijmr.12102

Aitchison, K. 2010 'Grey Literature, Academic Engagement, and Preservation by Understanding', *Archaeologies* **6**(2), 289–300. https://doi.org/10.1007/s11759-010-9145-5

Axelsen, I. 2022 'Collaboration and Communication between Hobby Metal Detectorists and Archaeologists in Norway', *Advances in Archaeological Practice* **10**(3), 295–310. https://doi.org/10.1017/aap.2022.14

Axelsen, I. and Fredriksen, C. 2024 'Organically Grown Archaeological Databases and their “Messiness”: Hobby Metal Detecting in Norway', *European Journal of Archaeology* **27**(3), 372–392. https://doi.org/10.1017/eaa.2024.10

Beck, A. and Neylon, C. 2012 'A vision for Open Archaeology', *World Archaeology* **44**(4), 479–497. https://doi.org/10.1080/00438243.2012.737581

Beck, C. and Jones, G.T. 1989 'Bias and Archaeological Classification', *American Antiquity* **54**(2), 244–262. https://doi.org/10.2307/281706

Bergstrøm, I.I. 2023 'This brooch became high fashion during the late Viking and Middle Age'. https://web.archive.org/web/20241220084803/https://www.sciencenorway.no/archaeology-medieval-history-viking-age/this-brooch-became-high-fashion-during-the-late-viking-and-middle-age/2189994 [Last accessed: 29 December 2024].

Bevan, A. 2012 'Spatial methods for analysing large-scale artefact inventories', *Antiquity* **86**, 492–506. https://doi.org/10.1017/S0003598X0006289X

Bevan, A. 2015 'The data deluge', *Antiquity* **89**, 1473–1484. https://doi.org/10.15184/aqy.2015.102

Bevan, A., Northover, P., Bray, P., Bonacchi, C., Colledge, S., Crellin, R., Gwilt, A., Hamilton, H., Hart, P., Kaleta, R., Keinan-Schoonbaert, A., Knight, M., Laws, K., Lodwick, M., Martinón-Torres, M., Needham, S., O'Connor, B., Perucchetti, L., Pett, D., Wexler, J., and Wilkin, N. 2024 'A Catalogue of British Bronze Age Axes, Including Basic Typology, Compositional Analyses and Associated Radiocarbon Dates', *Journal of Open Archaeology Data* **12**. https://doi.org/10.5334/joad.119.

Blair, J., Bogucki, M., Deckers, P., Jöns, H., Hilberg, V., Hedenstierna-Jonson, C., Leroy, I., Messal, S., Sindbæk, S.M., Verslype, L., Wade, K. (forthcoming) 'The network of early medieval emporia' in S. Croix, P. Deckers, and S.M. Sindbaek (eds) *Ribe Reflections. New Light on Northern Emporia*.

Bondesson, T. and Bondesson, L. 2021 'Update on brooches with Balt traits in Norway – the map redrawn by metal detecting', *Fornvännen* **116**, 145–150.

Brodie, N. 2021 'What is the thing called the PAS? Metal-detecting entanglements in England and Wales', *Revista d'Arqueologia de Ponent* **30**, 85–100. https://doi.org/10.21001/rap.2020.30.4

Chapman, R. and Wylie, A. 2016 *Evidential Reasoning in Archaeology*, London-New York: Bloomsbury.

Christiansen, T.T. 2019 'Metal-detected Late Iron Age and Early Medieval Brooches from the Limfjord Region, Northern Jutland: Production, Use and Loss', *Journal of Archaeology and Ancient History* **24**, 3–134.

Colwell, R.K. and Chao, A. 2022 'Measuring and Comparing Class Diversity in Archaeological Assemblages: A Brief Guide to the History and State-of-the-Art in Diversity Statistics' in M.I. Eren and B. Buchanan (eds) *Defining and Measuring Diversity in Archaeology: Another Step Toward an Evolutionary Synthesis of Culture*, Berghan Books. 263–294. https://doi.org/10.3167/9781800734296

Cool, H.E.M. and Baxter, M.J. 2016 'Brooches and Britannia', *Britannia* **47**, 71–98. https://doi.org/10.1017/S0068113X16000039

Cooper, A. and Green, C. 2017 'Big questions for large, complex datasets: approaching time and space using composite object assemblages', *Internet Archaeology* **45**. https://doi.org/10.11141/ia.45.1

Cunliffe, B. 1983 *The Publication of Archaeological Excavations*, Report of a joint working party of the Council for British Archaeology and the Department of the Environment.

Daubney, A. 2017 'Floating culture: the unrecorded antiquities of England and Wales', *International Journal of Heritage Studies* **23**(9), 785–799. https://doi.org/10.1080/13527258.2017.1325770

Deckers, P. 2021 'LOD typology: land of opportunity?', *Germania* **98**(2020), 222–225. https://doi.org/10.11588/ger.2020.1

Deckers, P. 2023 'Masses of Medieval Metal: A Quantitative Approach to Metalwork from Medieval Cities in Flanders (AD 1000–1600)' in J. Sawicki, M. Lewis, and M. Vargha (eds) *A United Europe of Things: Portable Material Culture across Medieval Europe*, Cham: Springer International Publishing. 115–124. https://doi.org/10.1007/978-3-031-48336-3_12

Deckers, P., Croix, S., and Sindbæk, S.M. 2021 'Assembling the Full Cast: Ritual Performance, Gender Transgression and Iconographic Innovation in Viking-Age Ribe', *Medieval Archaeology* **65**, 30–65. https://doi.org/10.1080/00766097.2021.1923893

Dobat, A.S. 2013 'Between rescue and research: An evaluation after 30 years of liberal metal detecting in archaeological research and heritage practice in Denmark', *European Journal of Archaeology* **16**, 704–725. https://doi.org/10.1179/1461957113Y.0000000041

Dobat, A.S., Christiansen, T.T., Jessen, M.D., Henriksen, M.B., Jensen, P., Laursen, S.V., Ruhe, R., Holst, M.K., and Arntsen, F. 2019 'The DIME project. Background, status and future perspectives of a user driven recording scheme for metal detector finds as an example of participatory heritage', *Danish Journal of Archaeology* **8**, 1–15. https://doi.org/10.7146/dja.v8i0.111422

Dobat, A.S., Deckers, P., Heeren, S., Lewis, M., Thomas, S., and Wessman, A. 2020 'Towards a Cooperative Approach to Hobby Metal Detecting: The European Public Finds Recording Network (EPFRN) Vision Statement', *European Journal of Archaeology* **23**, 272–292. https://doi.org/10.1017/eaa.2020.1

Edwards, P.N. 2010 *A Vast Machine: Computer models, climate data, and the politics of global warming*, Cambridge, Mass.: MIT Press.

Feugère, M., Sueur, Q., and Vigier, E. 2021 'Typologie 2.0 – Datenbanken in der Archäologie: das europäische Projekt Artefacts.mom.fr', *Germania* **98**(2020), 193–210. https://doi.org/10.11588/ger.2020.1

Fraser, D. 1997 'The British Archaeological Database' in J. Hunter and I. Ralston (eds) *Archaeological Resource Management in the UK: An Introduction*, Stroud, Gloucestershire: Sutton Pub.: Institute of Field Archaeologists. 19–29.

Frere, S.S. 1975 *Principles of Publication in Rescue Archaeology*, London: Department of the Environment.

Gardeła, L., Pentz, P., and Price, N. 2022 'Revisiting the "Valkyries": Armed Females in Viking Age Figurative Metalwork', *Current Swedish Archaeology* **30**, 95–151. https://doi.org/10.37718/CSA.2022.10

Gilså, N. 2021 *Et spørgsmål om stil: Urnesspænder i gennembrudt arbejde fra tidlig middelalder*, Højbjerg: Middelalderarkæologisk Forum.

Haldenby, D. 2012 *Early-Medieval 'Collared' Pins*, FRG Datasheet 44, The Finds Research Group AD 700-1700. https://findsresearchgroup.com/FRG44-datasheet.pdf

Haldenby, D. and Richards, J.D. 2010 'Charting the effects of plough damage using metal-detected assemblages', *Antiquity* **84**, 1151–1162. https://doi.org/10.1017/S0003598X00067144

Hardy, S.A. 2017 'Quantitative analysis of open-source data on metal detecting for cultural property: Estimation of the scale and intensity of metal detecting and the quantity of metal-detected cultural goods', *Cogent Social Sciences* **3**(1), 1298397. https://doi.org/10.1080/23311886.2017.1298397

Henry, R. 2024 'A New Corpus of Roman Coins from England and Wales. An Overview of the Evidence and Analysis of the Data', *Britannia* **55**, 1–35. https://doi.org/10.1017/S0068113X2400028X.

Horsnæs, H.W. 2021 'Finds of ancient coins from the Viking Age to modern periods in the territories of Denmark/Znaleziska monet antycznych od okresu wikinskiego do czasów współczesnych z terenow Danii' in M. Bogucki, A. Dymowski, and G. Snieżko (eds) *Slivers of Antiquity. The use of Ancient Coins in Central, Eastern and Northern Europe in the Medieval and Modern Periods/Okruchy starożytności. Użytkowanie monet antycznych w Europie Środkowej, Wschodniej i Północnej w średniowieczu i okresie nowożytnym*, Warsaw: Warsaw University Press. 177–200. https://doi.org/10.31338/uw.9788323547051.pp.177-200

Huggett, J. 2015 'Digital Haystacks: Open Data and the Transformation of Archaeological Knowledge' in A. Wilson and B. Edwards (eds) *Open Source Archaeology: Ethics and Practice*, Berlin: De Gruyter Open. 6–29. https://doi.org/10.17613/yfss-zt74

Huggett, J. 2022 'Data Legacies, Epistemic Anxieties, and Digital Imaginaries in Archaeology', *Digital* **2**(2), 267–295. https://doi.org/10.3390/digital2020016

Huvila, I. 2022 'Improving the usefulness of research data with better paradata', *Open Information Science* **6**(1), 28–48. https://doi.org/10.1515/opis-2022-0129

Immonen, V. 2023 'Exploring diversity through archaeology – medieval seal stamps and metal detecting in Finland' in M. Karlsson (ed) *Sigill i Norden: Bidrag från en nordisk konferens om sigill: Riksarkivet i Stockholm, 28–29 oktober 2021 / Seals in the Nordics: Proceedings from a Nordic Conference on Seals: The National Archives, Stockholm, 28–29 October 2021*, Riksarkivet. 51–60.

Kaas, M.H. 2024 *Brooches and Beyond: A New Corpus of Viking-Age Equal-Armed Brooches in a Nordic Context with an Analytical Focus on Typology, Networks, and Supra-regionality*, Unpublished MA dissertation, Aarhus University, Aarhus.

Kepes, S., Banks, G.C., McDaniel, M. and Whetzel, D.L. 2012 'Publication Bias in the Organizational Sciences', *Organizational Research Methods* **15**(4), 624–662. https://doi.org/10.1177/1094428112452760

Kessler, M., Marino, F. and Liska, D. 2023 'Netnographic research ethics in applied linguistics: A systematic review of data collection and reporting practices', *Research Methods in Applied Linguistics* **2**(3), 100082. https://doi.org/10.1016/j.rmal.2023.100082

Kleingärtner, S. 2003 'Eine pferdegestaltige Fibel aus Uppåkra Stilistische und kulturhistorische Betrachtungen' in B. Hårdh (ed) *Fler fynd i centrum. Materialstudier i och kring Uppåkra*, Stockholm. 123–135.

Lathan, H.S., Kwan, A., Takats, C., Tanner, J.P., Wormer, R., Romero, D., and Jones, H.E. 2023 'Ethical considerations and methodological uses of Facebook data in public health research: A systematic review', *Social Science & Medicine* **322**, 115807. https://doi.org/10.1016/j.socscimed.2023.115807

Lecroere, T. 2016 '"There Is None So Blind as Those Who Won't See": Metal Detecting and Archaeology in France', *Open Archaeology* **2**(1), 182–193. https://doi.org/10.1515/opar-2016-0014

Leonelli, S. 2015 'What Counts as Scientific Data? A Relational Framework', *Philosophy of Science* **82**(5), 810–821. https://doi.org/10.1086/684083

Lewis, M. 2016 'A Detectorist's Utopia? Archaeology and Metal-Detecting in England and Wales', *Open Archaeology* **2**(1), 127–139. https://doi.org/10.1515/opar-2016-0009

Lewis, M., Ehrnsten, F., Kurisoo, T., Oksanen, E., and Rohiola, V. forthcoming 'Improving Systems for Processing Public Finds: digital technology and citizen science', *European Journal of Archaeology*.

Mahr, D. and Dickel, S. 2019 'Citizen science beyond invited participation: nineteenth century amateur naturalists, epistemic autonomy, and big data approaches avant la lettre', *History and Philosophy of the Life Sciences* **41**(4), 41. https://doi.org/10.1007/s40656-019-0280-z

Marginal return 2025 *Wikipedia*. https://en.wikipedia.org/wiki/Marginal_return [Last accessed 19 April 2025]

Martens, J. 2022 'Ball Brooches in the Age of Citizen Science', *PLURAL. History, Culture, Society* **10**(1), 95–114. https://doi.org/10.37710/plural.v10i1_5

Mroczek, D. 2024 'Archaeological Heritage at Risk: Poland's Problem with Treasure Hunters' in T.Ø. Kuldova, J. Østbø, and C. Shore (eds) *Compliance, Defiance, and 'Dirty' Luxury: New Perspectives on Anti-Corruption in Elite Contexts*, Cham: Springer Nature Switzerland. 321–352. https://doi.org/10.1007/978-3-031-57140-4_11

Nancke-Krogh, S. 1978 'Ribehesten og dens slægtninge', *Kuml* **1978**, 179–191. https://doi.org/10.7146/kuml.v27i27.106894

Oksanen, E., Ehrnsten, F., Rantala, H., and Hyvönen, E. 2023 'Semantic Solutions for Democratizing Archaeological and Numismatic Data Analysis', *Journal on Computing and Cultural Heritage* **16**(4), 1–18. https://doi.org/10.1145/3625302

Oksanen, E. and Lewis, M. 2023 'Evaluating Transformations in Small Metal Finds Following the Black Death', *Medieval Archaeology* **67**(1), 159–186. https://doi.org/10.1080/00766097.2023.2204727

Oksanen, E. and Wessman, A. 2025 'New horizons in understanding Finnish Iron Age material culture through metal-detected finds', *Internet Archaeology* **68**. https://doi.org/10.11141/ia.68.4

Pentz, P. 2018 'Viking art, Snorri Sturluson and some recent metal detector finds', *Fornvännen* **113**, 17–33.

Pesonen, P., Moilanen, U., Roose, M., Saipio, J., Tiilikkala, J., Sanwal, U., Immonen, V., Vesakoski, O., and Onkamo, P. 2024 'Archaeological Artefact Database of Finland (AADA)', *Scientific Data* **11**(1), 815. https://doi.org/10.1038/s41597-024-03602-8.

Pétursdóttir, Þ. 2020 'Collaborative, Public, Participatory: Pros and Cons of an Open Archaeology', *Norwegian Archaeological Review* **53**(1), 1–4. https://doi.org/10.1080/00293652.2020.1782463

Robbins, K.J. 2013 'Balancing the scales: Exploring the variable effects of collection bias on data collected by the Portable Antiquities Scheme', *Landscapes* **14**, 54–72. https://doi.org/10.1179/1466203513Z.0000000006

Rundkvist, M. 2008 'För en liberalisering av de svenska metallsökarreglerna', *Fornvännen* **103**, 118–122.

Rundkvist, M. 2009 'Review of S. Thomas & P.G. Stone (eds) 2009 Metal Detecting and Archaeology, Newcastle University: Woodbridge', *Fornvännen* **104**, 234–235.

Schofield, J., Praet, E., Townsend, K.A., and Vince, J. 2021 '"COVID waste" and social media as method: an archaeology of personal protective equipment and its contribution to policy', *Antiquity* **95**(380), 435–449. https://doi.org/10.15184/aqy.2021.18

Schöpfel, J. 2018 'Grey Literature and Professional Knowledge Making' in L. Börjesson and I. Huvila (eds) *Research Outside The Academy. Professional Knowledge-Making in the Digital Age*, Palgrave Macmillan. 137–153. https://doi.org/10.1007/978-3-319-94177-6_8

Sindbæk, S.M. 2008 'The Lands of Denemearce: Cultural Differences and Social Networks of the Viking Age in South Scandinavia', *Viking and Medieval Scandinavia* **4**, 169–208. https://doi.org/10.1484/J.VMS.1.100310

Thomas, R. 1991 'Drowning in data? - publication and rescue archaeology in the 1990s', *Antiquity* **65**(249), 822–828. https://doi.org/10.1017/S0003598X00080546

Thomas, S. and Pitblado, B.L. 2020 'The dangers of conflating responsible and responsive artefact stewardship with illicit and illegal collecting', *Antiquity* **94**(376), 1060–1067. https://doi.org/10.15184/aqy.2019.201

Webley, R. 2020 *Conquests and Continuity: Portable metalwork in Late Anglo-Saxon and Anglo-Norman England, c. AD 1000-1200*, PhD dissertation, University of York. https://etheses.whiterose.ac.uk/id/eprint/30872/

Weetch, R. 2014 *Brooches in Late Anglo-Saxon England within a North West European Context. A study of social identities between the eighth and eleventh centuries*, PhD dissertation, University of Reading.

Wessman, A., Koivisto, L., and Thomas, S. 2016 'Metal Detecting in Finland - An Ongoing Debate', *Open Archaeology* **2**(1). https://doi.org/10.1515/opar-2016-0006

Wessman, A., Thomas, S., Deckers, P., Dobat, A.S., Heeren, S., and Lewis, M. 2023 'Hobby Metal-detecting as Citizen Science. Background, Challenges and Opportunities of Collaborative Archeological Finds Recording Schemes', *Heritage & Society* **16**(2), 89–108. https://doi.org/10.1080/2159032X.2022.2098654

Wessman, A., Thomas, S., and Rohiola, V. 2019 'Digital Archaeology and Citizen Science: Introducing the goals of FindSampo and the SuALT project', *SKAS* **1**, 2–19.

Wicker, N.L. 2020 'Humans and Animals: The Changing Corpus of Danish Viking Art' in A. Pedersen and S.M. Sindbæk (eds) *Viking Encounters: Proceedings of the 18th Viking Congress, Denmark, August 6–12, 2017*, Aarhus: Aarhus University Press. 413–425.

Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B. 2016 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific Data* **3**(1), 1–9. https://doi.org/10.1038/sdata.2016.18

Zubiaga, A. 2018 'A longitudinal assessment of the persistence of Twitter datasets', *Journal of the Association for Information Science and Technology* **69**(8), 974–984. https://doi.org/10.1002/asi.24026