

1 **From the River to the Sea, Using Large Language Models to Search**
2 **for Neolithic Vlaardingen Culture Sites in the Rhine-Meuse-Scheldt**
3 **Delta**

4 Van den Dikkenberg, Lasse^{1*}, Brandsen, Alex¹

5
6 ¹ Faculty of Archaeology, Leiden University – Leiden, The Netherlands

7
8 *Corresponding author

9 Correspondence: l.van.den.dikkenberg@arch.leidenuniv.nl

10
11
12 **ABSTRACT**

13 This paper presents a study on Vlaardingen Culture (3400-2500 BCE) sites in the
14 Rhine-Meuse-Scheldt delta using AGNES, an intelligent search engine for Dutch and
15 Flemish archaeological grey literature. The aims of this paper are twofold: 1) to provide
16 an up-to-date overview of Vlaardingen Culture sites; 2) to evaluate the performance
17 of AGNES in searching for period specific sites. Vlaardingen Culture (VLC) sites
18 usually consist of artefact scatters without clearly discernible house plans. These
19 scatters are often found amongst abundant remains from later periods. This type of
20 'by-catch' is usually not found in the metadata of archaeological reports. It can only be
21 recovered in full text searches. AGNES uses Large Language Models to allow
22 searches on archaeological concepts (in this case an archaeological culture) in full
23 texts extracted from three major repositories for Dutch (DANS and ARCHIS) and
24 Flemish (Onroerend Erfgoed) archaeology. This paper presents a search for VLC
25 sites, and a comparison of the retrieved information with a recently compiled overview
26 of VLC sites in the area. Using eight queries we retrieved 4532 hits, which were
27 subdivided into: relevant hits (n=430); semi-relevant hits (n=2133); and irrelevant hits
28 (n=1960). We recovered thirty previously unknown Vlaardingen Culture sites,
29 amounting to 19% of the total number of VLC sites (n=158). Not all sites could be
30 found in AGNES, older archaeological sites are often published in scientific and semi-
31 scientific journals, theses, or books. These publications are absent in the repositories
32 which can be accessed through AGNES, and by extent, they cannot be retrieved. As
33 such AGNES does not provide an alternative to traditional search methods.
34 Nevertheless, most of the newly found sites consist of sites which cannot be found by
35 searching the metadata of reports in DANS and ARCHIS. Therefore, AGNES proved
36 to be an essential and effective addition to traditional search methods.

37
38 **Keywords:** Grey literature; Information Retrieval; Vlaardingen Culture; Neolithic; Natural
39 language processing; Large Language Models

40

41

1. Introduction

42 The field of archaeology produces large amounts of textual data, from published books and
43 articles, to grey literature reports. In the Netherlands, a large proportion of this textual data is
44 produced by commercial archaeology units, who research and excavate in so-called development-
45 led archaeology. It estimated that over 5000 excavations and other investigations take place each
46 year (RCE 2023), each producing one or more publications. On top of that, the academic world
47 also produces books, papers and theses. The sheer size of all this data makes finding the right
48 information for synthesising research difficult, and existing search systems do not fully match the
49 requirements of the end users (Habermehl 2024). Using text mining and artificial intelligence to
50 better search through this literature has a long history (Amrani Abajian & Kodratoff 2008; Pajmans
51 & Brandsen 2009; Pajmans & Brandsen 2010; Richards, Tudhope & Vlachidis 2015; Vlachidis &
52 Tudhope 2012), but systems are often developed as proof-of-concepts, not evaluated and made
53 fit for purpose for users, and/or not maintained in the medium to long term. In the EXALT project,
54 we are working to create a search engine that uses Large Language Models (LLMs) to make
55 information retrieval in Dutch archaeology documents easier and more effective, in a user-centric
56 interface. The system is called AGNES (Archaeological Grey literature Named Entity Search), and
57 is described in more detail in section 2.1.

58

59 The aim of this study is twofold: 1) we aim to provide an up-to-date overview of Vlaardingen
60 Culture sites; 2) we aim to evaluate the performance of AGNES in searching for period specific
61 sites. As opposed to the earlier case-study, in which AGNES was used to find Merovingian
62 cremation graves in the Netherlands (Brandsen & Lippok, 2021), this present study is aimed at
63 finding sites attributed to the Vlaardingen Culture (3400-2500 BCE). The vast majority of
64 Vlaardingen Culture sites consist of artefact scatters without clear house plans (Van Gijn & Bakker
65 2005). Only a few sites contain clearly discernible house plans (Stokkel 2017; Van Beek 1990; Van
66 Kampen 2013; Van Zoolingen 2021; Verhart 1992). These artefact scatters are often found as 'by-
67 catch' on archaeological excavations. By-catch refers to "one or a few finds that are different from
68 the rest of the excavation" (Brandsen & Lippok 2021). Because AGNES allows for full text searches
69 it is well equipped for finding such by-catch, unlike metadata searches. Because of the nature of
70 Vlaardingen Culture sites they provide an ideal case to test the efficacy of AGNES.

71

72 The case-study on Merovingian cremations aimed to find something very specific (scarce
73 cremation graves) within Merovingian contexts, the current study has a broad aim, finding anything
74 that is attributed to the Vlaardingen Culture. In the study on Merovingian cremations it was
75 demonstrated that AGNES excelled at finding these very specific types of finds. Brandsen and
76 Lippok were able to recover a total of 23 previously unknown Merovingian cremation graves
77 (2021). In this paper we test the usefulness of AGNES in making period specific site overviews,
78 focussing on sites attributed to the Vlaardingen Culture.

79

80 The research questions are as follows:

81

- 82 • Compared to previously known sites, what unknown sites can we find with AGNES?
- 83 • What does this mean for the usefulness of AGNES?
- 84 • What do the newly rediscovered sites mean for the distribution of Vlaardingen Culture sites
85 in the Rhine-Meuse-Scheldt delta?

86 1.1 The Vlaardingen Culture dataset

87 Settlements from the Vlaardingen Culture are mainly located in the coastal area of the
88 Netherlands, notably in wetland areas along rivers and the coastal dunes. The most complete
89 overview of Vlaardingen Culture sites so far is presented in the distribution map by Verhart and de
90 Ridder, an overview containing ca. eighty sites (Verhart & de Ridder 2010). This is more than the
91 overview presented five years earlier in The Prehistory of the Netherlands, where it was mentioned

92 that there are about thirty Vlaardingen Culture sites in the Netherlands (Van Gijn & Bakker 2005).
93 Neither of these overviews present us with a list of sites. Furthermore, because the last overview
94 of sites was made in 2010 newly excavated sites are lacking from these existing overviews. Recent
95 discoveries have challenged some of the notions previously held about this period. The excavation
96 at Den Haag Wateringse Binnentuinen zone 3 and Den Haag Noordweg 76 revealed that
97 temporary settlements, previously thought to be a feature exclusively present on sites located on
98 river levees, also occur in the coastal dune area of the Netherlands (Bulten & Stokkel 2017;
99 Raemaekers 2003; Van Zoolingen & Rieffe 2021). Furthermore, the recent excavations at Den
100 Haag Steynhof and Den Haag Wateringse Binnentuinen were amongst the largest and best
101 documented Vlaardingen Culture settlements to date (Bulten & Stokkel 2017; Van Zoolingen &
102 Bulten 2021). The site Veldhoven Habraken demonstrated that permanent Vlaardingen Culture
103 sites can also be expected further inland on the sandy soils of Noord Brabant (Van Kampen 2013).
104 These new discoveries highlight the need for a renewed overview of Vlaardingen Culture sites. An
105 attempt to create such an overview was undertaken by Van den Dikkenberg as part of his current
106 PhD project, which is part of the Putting Life into Late Neolithic Houses project. Based on
107 previously published overview studies, known site reports, and queries in ARCHIS and DANS, an
108 overview was compiled of 129 known Vlaardingen Culture sites in the Netherlands and Belgium.
109 This dataset will be compared with the new results, as it is a fair reflection of what information can
110 be found with tools and databases besides AGNES.

111 1.2 AGNES dataset

112 The aim is to incorporate all open access documents about Dutch archaeology and
113 neighbouring countries, which is still underway. At the time of this study, just over 188,000
114 documents are included in AGNES, from the following sources:

- 115
- 116 • Documents from the DANS archive marked with the 'Archaeology' tag, up to December
117 2021
- 118 • Documents from the ARCHIS database, up to December 2021
- 119 • Documents from the Onroerend Erfgoed archive, up to June 2024
- 120

121 For all of these sources, we only harvested and indexed PDF files, and these contain a
122 multitude of document types. This includes excavation reports, coring reports, appendices,
123 database descriptions, personal daily reports, maps, find lists and sometimes even photographs
124 stored within PDFs.

125 2. Methods

126 Below we will first introduce the AGNES search system, next we will discuss our search
127 methodology for the case-study on Vlaardingen Culture sites.

128 2.1 AGNES and Large Language Models

129 As mentioned in the introduction, the original literature search for VLC sites in archives was
130 done by metadata search; searching in e.g. the title, description, and keywords. However, this
131 metadata can be incomplete and/or inaccurate, and is missing detailed information (Habermehl
132 2024). Think of an excavation of a Roman encampment; the metadata is not going to mention a
133 single Neolithic find (by-catch), but this find is only mentioned in the excavation report. To solve
134 this, we can apply full-text search, searching through all of the text instead of just the metadata.
135 This would be a significant improvement, and is something the DANS archive has since
136 implemented. However, archaeological discourse includes a lot of synonyms and homonyms;
137 multiple words with the same or similar meaning (e.g. mediaeval and Middle Ages), and words with
138 multiple meanings ('Flint' being both a material and a surname), respectively, which makes
139 searching more difficult (Brandsen 2021). Within AGNES, we try to solve the homonym problem

140 using Named Entity Recognition (NER), a natural language processing technique that finds and
 141 extracts certain entity types (Tjong Kim Sang 2002). In our research, we target artefacts, time
 142 periods, contexts, species, materials and locations. Once the word 'flint' has been identified as a
 143 material by NER, we know it's not the surname flint, and we solve the homonym problem. For time
 144 periods, we try to solve the synonym problem by taking detected time period entities, and
 145 translating them to a start and end year. This way we can search for e.g. '500 to 1500 AD' and this
 146 will return results for both 'mediaeval' and 'Middle Ages', as well as sub-periods and single years
 147 within this year range.

148
 149 The NER is done using BERT (Bidirectional Encoder Representations from Transformers), the
 150 OG Large Language Model (Devlin et al. 2019). Similar to the newer GPT (Generative Pre-trained
 151 Transformer) models, BERT uses large amounts of unlabelled text data to pre-train a model,
 152 gaining an understanding of words and their contexts. We took these generic models, and further
 153 pre-trained them with texts from the archaeology domain (Brandsen 2021; Brandsen 2023;
 154 Brandsen et al. 2022). This created archaeology specific BERT models for Dutch, English, and
 155 German, and finally we fine-tune these with labelled NER data to be able to predict and extract
 156 entities. Together with the full text of the documents, these entities are indexed in ElasticSearch,
 157 an open source search engine (Gormley & Tong, 2015). We built a frontend to query all this
 158 information, specifically designed and evaluated with the archaeologists' needs in mind (Brandsen
 159 et al. 2019; Brandsen et al. 2021).

160 2.2 Search methodology

161 In total eight queries were entered in AGNES resulting in a total of 4532 hits (see table 1; see
 162 supplementary table 1). These were exported to a single CSV file. Usually we did not use start and
 163 end dates for the queries, to expand the number of potential hits. Only for the open query
 164 "vlaardingen*" did we include dates to limit the number of hits. We opted to use a broader date
 165 range (3800-2000 BCE) than the traditional starting and end dates for the Vlaardingen Culture
 166 (3400-2500 BCE). This would increase the opportunity to find VLC sites which have date ranges
 167 expanding beyond the generally accepted starting and end dates for the VLC.

168 Table 1. List of queries entered in AGNES with number of hits.

Start date (BCE)	End date (BCE)	Free text query	English Translation	Number of hits
		"vlaardingen cultuur"	Vlaardingen Culture	834
3800	2000	"vlaardingen*"	Vlaardingen	2483
		"vlaardingen stein wartburg"	Vlaardingen-Stein-Wartburg	11
		"vlaardingen stein wartberg"	Vlaardingen-Stein-Wartburg	4
		"vlaardingen groep"	Vlaardingen Group	265
		"vlaardingen stein"	Vlaardingen-Stein / Vlaardingen Stein	98
		"vlaardingencultuur"	Vlaardingen Culture	712
		"vlaardingengroep"	Vlaardingen Group	125

169
 170 Three additional columns were added to the CSV file to include the query text, a ranking of how
 171 relevant the publication was (see table 2), and lastly a column in which for irrelevant hits the reason
 172 was stated for the irrelevance of the hit (see table 3). The listed categories are an adapted version
 173 of those applied by Brandsen and Lippok (Brandsen & Lippok 2021). Relevant hits consist of site

174 reports on previously unknown Vlaardingen Culture sites (1), previously known Vlaardingen
 175 Culture sites (2), as well as reports in which previously unknown Vlaardingen Culture sites are
 176 mentioned in the text (3) or the literature lists (4). In these last two instances the hits indirectly led
 177 to the discovery of new sites. In addition to these categories we included three semi-relevant
 178 categories. The Stein group and Vlaardingen group are closely related and as a result many
 179 reports on Stein sites also discuss the relationships between these groups. This was considered
 180 a semi-relevant hit (5) because AGNES correctly identified that the text discussed the Vlaardingen
 181 Culture, but the report itself was not about a VLC site. Similarly, other site reports also discussed
 182 the Vlaardingen Culture (6) or Vlaardingen Culture sites (7). **This was often the case with research**
 183 **plans.** In the previous study by Brandsen and Lippok hits in research plans were considered
 184 irrelevant hits (Brandsen & Lippok 2021). This makes sense for cremations as the expectation of
 185 finding cremation graves can refer to a plurality of periods. However, if research plans specifically
 186 mention the Vlaardingen Culture they do refer specifically to what we aimed to find in our queries.
 187 Therefore in this case we often considered these as semi-relevant hits, rather than irrelevant hits.
 188 These research plans also often mentioned nearby Vlaardingen Culture sites (7).

189 Table 2. Relevance of AGNES hits.

Nr.	Relevance
1	Relevant (report about a Vlaardingen Culture site) unknown
2	Relevant (report about a Vlaardingen Culture site) known
3	Relevant (previously unknown Vlaardingen site mentioned in the text)
4	Relevant (previously unknown Vlaardingen site mentioned in the literature list)
5	Semi-relevant (Stein site publication, mentioning Vlaardingen Culture in discussion)
6	Semi-relevant (Vlaardingen Culture mentioned in a discussion)
7	Semi-relevant (a different Vlaardingen Culture site mentioned in the text based on previous research)
8	Not relevant (not a report about a Vlaardingen Culture site)

190 Irrelevant hits were classified along the typology presented by Brandsen and Lippok (Brandsen & Lippok
 191 2021). We added three frequently occurring categories. Being a period designation, "Vlaardingen
 192 Culture" was frequently mentioned in documents containing a database structure. Therefore this was
 193 added as a separate category (10). The Vlaardingen Culture is named after the type-site Vlaardingen
 194 Arij Koplaan (Van Regteren Altena et al. 1962a; 1962b, 1962c). Because of this our hits frequently
 195 contained reports which only mentioned the place name "Vlaardingen" (12). Similarly, geographical
 196 maps which included the city of Vlaardingen were also considered as a separate category of irrelevant
 197 hits (11).

198
 199 Table 3. Types of irrelevant hits.

Number	Type of irrelevant document
1	Wrong time period
2	Page listing abbreviations
3	Page containing research plan (plan van aanpak)
4	Unknown time period
5	Page containing list of time periods
6	Negation ("no vlaardingen culture")
7	Other
8	Literature list (only)
9	Coring chart

10	Database structure
11	Vlaardingen as a location on a map
12	Vlaardingen as place name in text

200

201 2.2.1. Irrelevant or incorrect?

202 Above we presented different categories of relevant, semi-relevant, and irrelevant hits. This is
 203 based on the usefulness of the hits for the archaeological case-study. Such hits are not necessarily
 204 incorrect. When a report deals with a Stein site and it mentions the Vlaardingen Culture the hit for
 205 the query “vlaardingen culture” is correct, but it is irrelevant as the report does not deal with a
 206 Vlaardingen Culture site. Similarly, hits in which “Vlaardingen Culture” is only mentioned in the
 207 literature list are correct hits in the sense that AGNES correctly identified matching terms, but
 208 because it concerns a hit in the literature list it is deemed irrelevant for the case-study.

209

3. Results

210 The different queries yielded in total 439 relevant hits (see table 4). This means that 9,7% of
 211 the hits consisted of relevant hits. In addition to those 2133 hits were classified as semi-relevant
 212 and 1960 hits were classified as irrelevant. For the full data, see (supplementary table 1,
 213 doi.org/10.5281/zenodo.13283972). The relevance of reports depended on the specific queries
 214 which returned the hits. For example, in the report of the excavation of Hellevoetssluis-Ossenhoek
 215 the site is consistently referred to as a ‘Vlaardingen-groep’ site (Goossens 2009). The query
 216 ‘vlaardingen groep’ in this case thus yielded relevant hits. The query ‘vlaardingen culture’ however
 217 does not yield relevant hits for this publication, it does however yield two hits in the bibliography in
 218 which publications are cited which mention the term ‘vlaardingen-culture’ (Goossens 2009: 177-
 219 179). A single publication can thus return multiple hits for multiple queries.

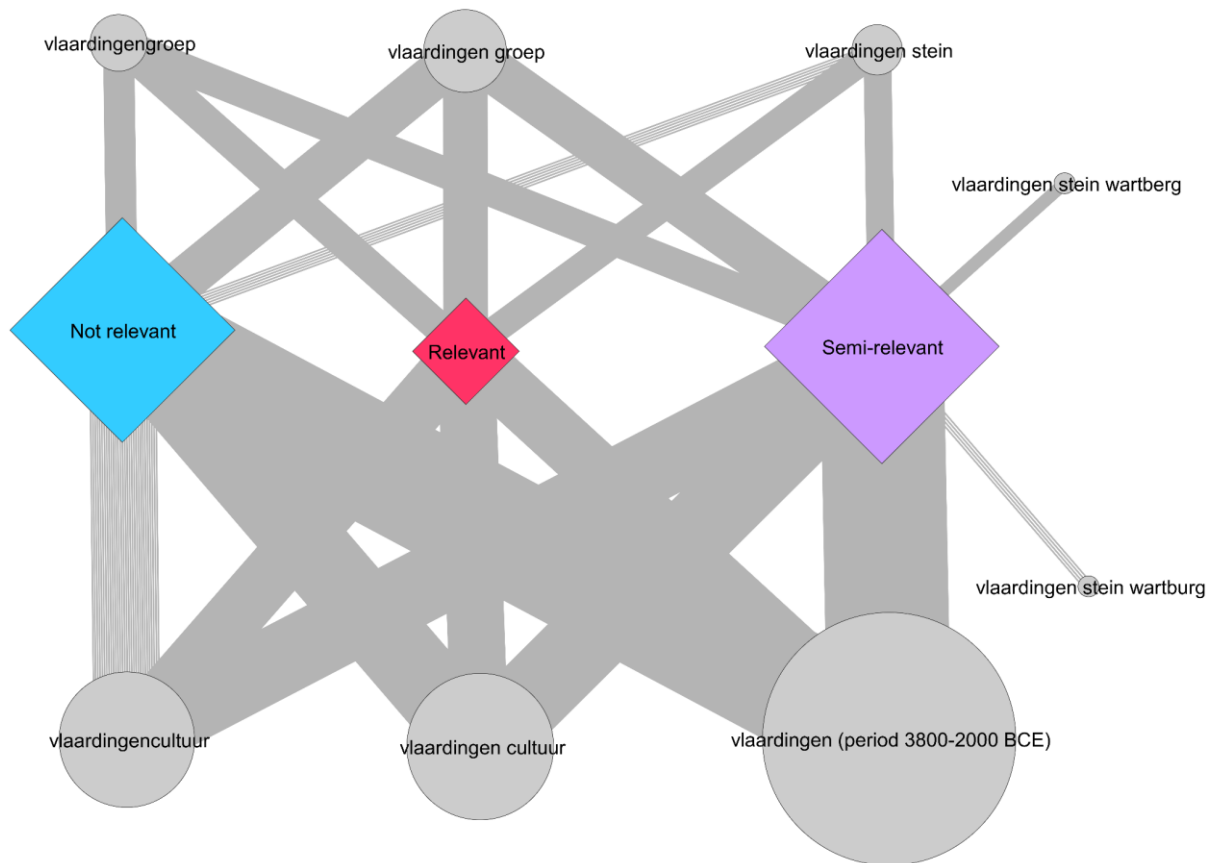
220

Table 4. Relevance of AGNES hits totals.

Relevance	Count
Relevant (report about a Vlaardingen Culture site) unknown	165
Relevant (report about a Vlaardingen Culture site) known	259
Relevant (previously unknown Vlaardingen site mentioned in the text)	9
Relevant (previously unknown Vlaardingen site mentioned in the literature list)	6
Semi-relevant (a different Vlaardingen Culture site mentioned in the text based on previous research)	1398
Semi-relevant (Stein site publication, mentioning Vlaardingen Culture in discussion)	65
Semi-relevant (Vlaardingen Culture mentioned in a discussion)	670
Not relevant (not a report about a Vlaardingen Culture site)	1960
Total	4532

221

222 The irrelevant hits were further subdivided into categories (see table 5), especially the query
 223 for “vlaardingen” yielded a large number of irrelevant hits (n=1506; see figure 1 and 2). These
 224 mostly related to the fact that Vlaardingen is a place name, therefore many reports on archaeology
 225 in the city of Vlaardingen were included in this query, as well as reports mentioning Vlaardingen
 226 as a place name in the text (n=897), or in a map (n=8). Frequently, hits included text only contained
 227 in the bibliography, a database structure, or lists of time periods. These hits were deemed irrelevant
 228 but they were generally not incorrect as AGNES correctly matched the search terms in the
 229 documents.



230

231 Figure 1. Network representation; two-mode network visualising the relevance of
 232 different queries, nodes are relatively scaled according to their degree centrality (%).
 233 Network visualising the different queries (grey) and relevant (red), irrelevant (blue),
 234 and semi-relevant hits (graph made in Visone).

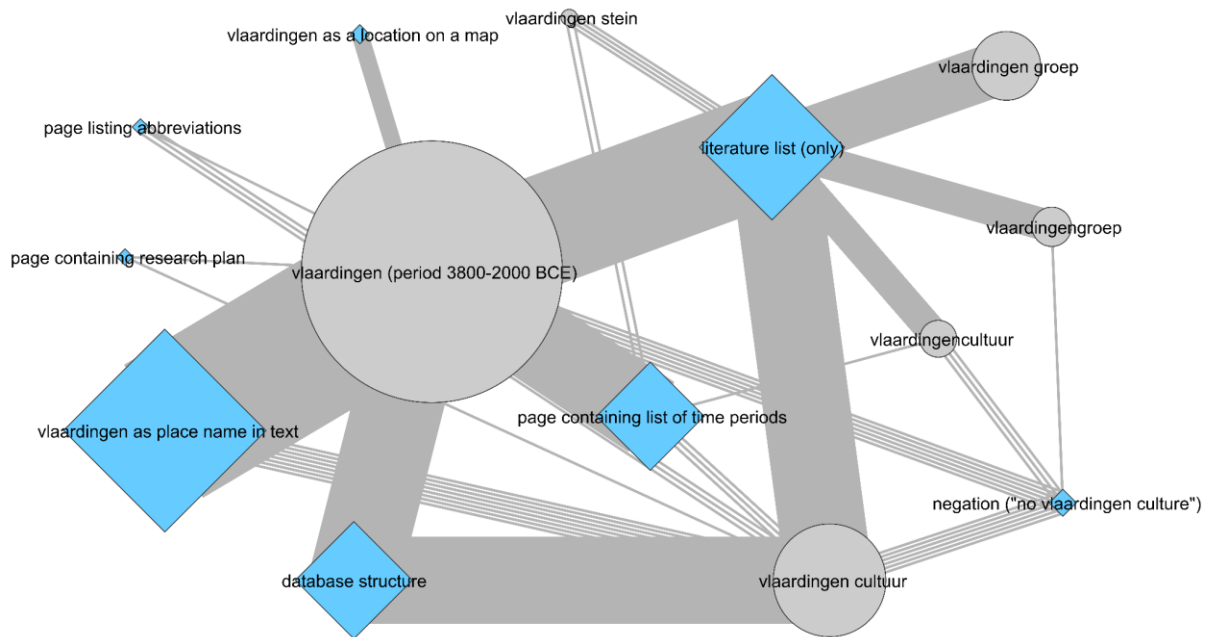
235 It is interesting that the queries for “vlaardingen stein wartburg” and “vlaardingen stein wartberg”
 236 only yielded semi-relevant hits (see figure 1). These terms were thus not used in reports about
 237 Vlaardingen Culture sites but they were used in discussions in other site reports.

238 Table 5. Reasons for irrelevant hits AGNES.

Irrelevance reason	Count
Page listing abbreviations	4
Page containing research plan (<i>plan van aanpak</i>)	2
Page containing list of time periods	257
Negation ("no vlaardingen culture")	16
Literature list (only)	465
Database structure	302
Vlaardingen as place name in text	905
Vlaardingen as place name on a map	8

239

240 In terms of irrelevant hits it is interesting to note that hits relating to database structures and
 241 lists of time periods never contained hits for “vlaardingengroep” or “vlaardingen groep”. Database
 242 structures systematically use the terms “vlaardingen” or “vlaardingen culture” (see figure 2).
 243 Standardised lists of time periods use the terms “vlaardingen”, “vlaardingen culture”, or
 244 “vlaardingen stein” (see figure 2). As such the irrelevant hits provide additional information relating
 245 to the terminology frequently employed in development-led archaeology.



247

248 Figure 2. Network representation; two-mode network visualising irrelevance types
 249 (blue) for different queries (grey), nodes are relatively scaled according to their
 250 degree centrality (%) (graph made in Visone).

251 3.1 Newly discovered sites

252 The queries yielded information on a total of thirty sites (19% of the total number of sites) which
 253 were not previously included in the overview (see table 6; see: supplementary table 2 and
 254 supplementary file 3 doi.org/10.5281/zenodo.13283972). In 27 instances this included hits on site
 255 reports of previously unknown Vlaardingen Culture sites. In three instances the hits consisted of
 256 indirect hits. These were site reports or research plans which mentioned previously unknown
 257 Vlaardingen Culture sites in their respective study areas.

258

259 For thirteen of the 89 sites which were not found in AGNES the publications were present in DANS
 260 or ARCHIS. In seven cases this concerned sites published in the monograph by Louwe Kooijmans
 261 (1974). These sites were found in the appendix where they were listed under the abbreviation "VL".
 262 As this abbreviation was not part of our queries these sites were not found in AGNES, even though
 263 the file in which they were listed was available. In one instance the report did not mention a cultural
 264 attribution, but the material could be attributed to the VLC based on the characteristics of the finds.
 265 For four sites the documents were present in DANS but these were not imported in AGNES due
 266 to errors during the PDF text extraction process (De Koning 2010; Eijskoot 2004; Eimermann 2008;
 267 Van den Broeke 1993). In one instance the BERT model did not recognise a date correctly (Hiddink
 268 2000: 11). The publication in this case mentions the term "Vlaardingen" along with several other
 269 Neolithic cultures and a date range (4000-2000 BCE). But, as the date range was not recognised
 270 as a time period, the query incorrectly did not match this page.

271 Table 6. Results per site, newly found sites, previously known sites and sites of
 272 which the reports are not in AGNES.

Result per site	Count
Found exclusively in AGNES	27
Found exclusively indirectly in AGNES	3
Found previously and in AGNES	39

Not found in AGNES queries (pdf not present in DANS or ARCHIS)	76
Not found in AGNES queries (pdf is present in DANS)	13
Total	158

273

274

275

276

277

278

279

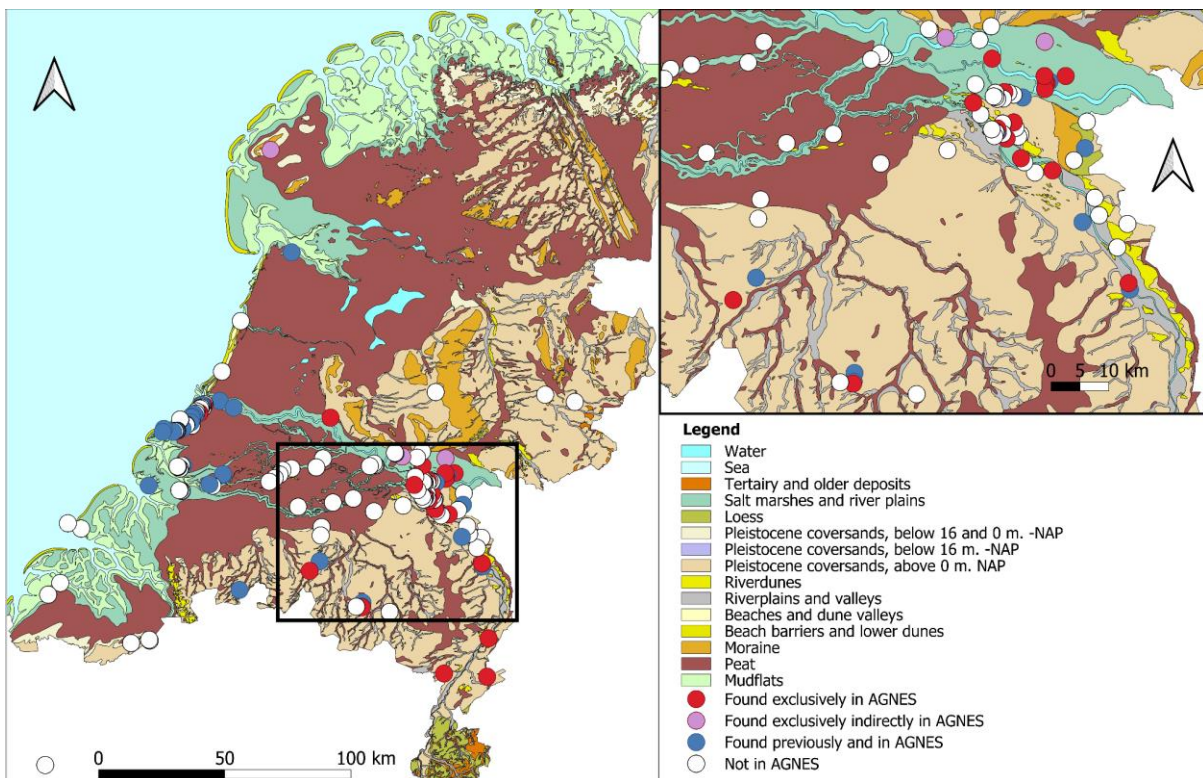
280

281

282

In one instance the unknown site was only mentioned in the bibliography of a report. The article concerned a previously unknown site on the island of Texel (Van Noort 1998). The site was discovered by an amateur archaeologist and it was published in a local historical journal, which explains why it remained unknown despite being published in 1998. This is an interesting find because the site is located much further north (see figure 3, the most northern purple dot on the map) than the most northern Vlaardingen Culture site known so far: Zandwerven (Van Gijn & Bakker 2005; Verhart & de Ridder 2010). It is located about twenty kilometres north of what is traditionally assumed to be the limit of the distribution of the Vlaardingen Culture (Van Gijn & Bakker 2005). As such the site presents a significant discovery.

283



284

285

286

287

288

Figure 3. Vlaardingen Culture sites plotted on the Palaeographic map of the Netherlands ca. 2750 BCE (Vos et al. 2020), for Belgium a palaeographic map for this period was not available, nevertheless the Belgian site Oudenaarde Donk is also plotted (see bottom-left corner).

289

290

291

292

293

294

295

296

297

298

Interestingly, many of the newly discovered sites are located in the eastern Netherlands in the area of Nijmegen (see figure 3). In the 2010 overview it was also noted that this area yielded a high number of Vlaardingen Culture sites. At the time this concentration could mostly be attributed to the tireless efforts of the local AWN (Association of Archaeology Volunteers) dependence. Between 1970 and 2000 this group discovered a great deal of sites in the Nijmegen and Wijchen area (Teubner & Tuijn 2010; Verhart & de Ridder 2010). The sites which were newly found in AGNES mostly consist of recent (post 2010) excavations in the area. Often these reports mention the previous studies conducted by local volunteers ('t Hart, Norde & Tuinstra 2019:12; Heirbaut 2010, 12; Janssen 1989; Janssen & Tuijn 1978). These studies thus led to a better formulation of archaeological expectations, and in turn to a better formulation of research plans. In recent years

299 this led to new excavations which in turn led to new discoveries. It is an excellent example of how
300 citizen science contributed to development-led archaeology.

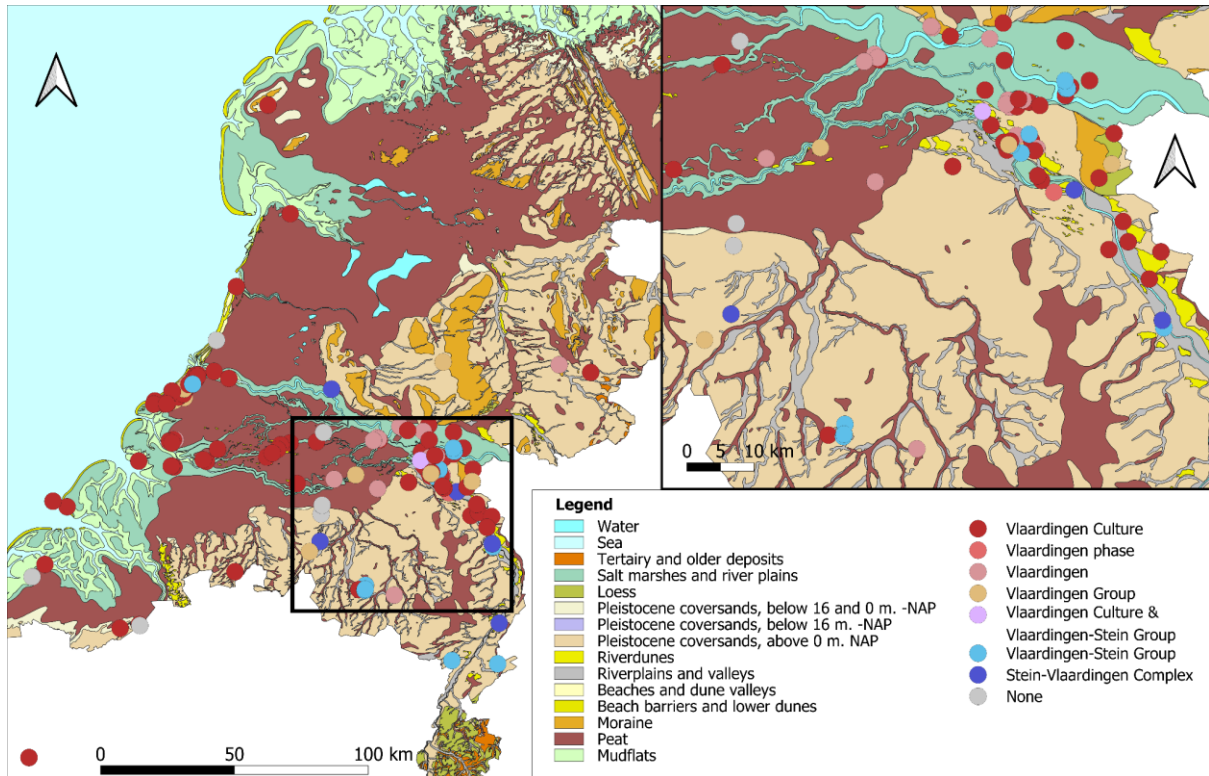
301

4. Discussion

302 Below we will discuss two main themes which popped up during our investigation. The first is
303 the plurality of terms used to describe Vlaardingen Culture sites. The second part will discuss the
304 data sources accessed in this study, with the aim of explaining on the one hand why we were able
305 to find previously unknown Vlaardingen Culture sites through AGNES, while we will also discuss
306 the types of sources which were missed during this study. This section will also provide
307 recommendations on how to make these sources more accessible for future studies.

308 4.1 Vlaardingen or Stein group?

309 It is not our aim to tackle the issue of defining what the Vlaardingen Culture is. We decided to
310 focus on the Vlaardingen Culture and not on the Stein group. Sites attributed to the latter are thus
311 excluded in our overview. The debate surrounding the distinction between these groups has been
312 the subject of several studies (Louwe Kooijmans 1983; Van den Dikkenberg 2024; Van Gijn &
313 Bakker 2005; Verhart 2010). In this paper we decided to adhere to the conclusions presented by
314 the excavators of these sites. In seven cases an exception was made. These consist of sites
315 without a cultural attribution but where the find material is dated to this period; sites where the
316 material is consistent with Vlaardingen Culture material; and sites which are geographically located
317 in the area of the Vlaardingen Culture. This is for example the case with the single find of a flint
318 oval axe in Elshout. Similar finds of single oval flint axes in the western Netherlands are
319 consistently attributed to the Vlaardingen Culture (Dorenbos & Koot 2010; Groenman-van
320 Waateringe & Van Regteren Altena 1966). For the other sites a plurality of cultural attributions was
321 used; Vlaardingen Culture, Vlaardingen group, Vlaardingen-Stein group (or Stein-Vlaardingen
322 group), and Stein-Vlaardingen Complex (see figure 4). It seems that different terms here are
323 generally not applied based on differing archaeological characteristics. Rather, it seems that terms
324 are regionally dependent. Sites in Zuid Holland (and more generally in the western Netherlands)
325 are usually referred to as Vlaardingen Culture sites. Sites in, and around, the border zone between
326 the Vlaardingen and Stein group are often referred to as “Vlaardingen-Stein group” or “Stein-
327 Vlaardingen Complex” (see figure 4).



329

330

331

332

333

334

335

Figure 4. Vlaardingen Culture sites plotted on the Palaeographic map of the Netherlands ca. 2750 BCE (Vos et al. 2020), for Belgium a palaeographic map for this period was not available, nevertheless the Belgian site Oudenaarde Donk is also plotted (see bottom-left corner). The sites are plotted according to their cultural attribution, on the left the supposed border area between the Vlaardingen and Stein group.

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

The fact that these cultural attributions in the literature depend more on the geographical location of these sites rather than the archaeological material is occasionally made explicit. The Stein site Schoolstraat in Thorn is for example attributed to the Stein group because “*the Vlaardingen Culture predominantly occurs in the coastal area of the western and southern Netherlands, as well as the riverine area in the central Netherlands, the site probably represent remains from the Stein group*”¹ (De Ridder 2011: 27). This problem was already envisioned in 1983 by Louwe Kooijmans when he first defined the characteristics of the Stein Group “The here discussed Late Neolithic find groups display a high degree of affinity with the Vlaardingen Culture, so much so that we believe that, had they been found in the delta region, they would have been, without much trouble, classified as Vlaardingen”² (Louwe Kooijmans 1983: 64). The supposed differences between these groups; the presence of Lousberg axes, blade technologies, and axe production for the Stein group vs. the presence of ceramic baking plates and pottery with perforations under the rim for the Vlaardingen Culture, might be real to some extent (Louwe Kooijmans 1983; Van den Dikkenberg 2024; Verhart 2010). Nevertheless, they are too often absent, making cultural attributions for sites discovered in development-led archaeology difficult.

1 “Omdat de Vlaardingencultuur voornamelijk in het holocene kustgebied van west en zuid-west Nederland, alsmede het midden-Nederlandse rivierengebied voorkomt, betreft het hier waarschijnlijk de culturele nalatenschap van de Stein-groep. Vindplaatsen hiervan kennen we uit Limburg (met een concentratie in Midden-Limburg aan weerszijden van de Maas), het aangrenzende Rijnland, Noord-Brabant en het oostelijk rivierengebied” (De Ridder 2011: 27).

2 “De hier besproken, laat-neolithische vondstgroepen tonen een grote verwantschap met die van de Vlaardingen-cultuur, zodanig zelfs, dat wij het idee hebben dat zij, waren ze in de delta gevonden, zonder veel moeite „VL” waren genoemd” (Louwe Kooijmans 1983: 64).

351 It is noteworthy that the term Vlaardingen-Stein-Wartberg complex is not applied in any of the
352 reports. Based on the irrelevant hits this term also seems to be avoided in lists of abbreviations,
353 periods, and in database structures (see figure 2). Such lists systematically employ either the term
354 Vlaardingen, or a variant of Vlaardingen Culture. Overall it seems that these are the dominant
355 terms used in development-led archaeology (see figure 2 and 3).

356 4.2 Datasets

357 There are clear regional differences in terms of which sites were found by AGNES. As
358 mentioned before a lot of sites found in AGNES were located in the Nijmegen area. Furthermore,
359 clearly a lot of previously known sites in Zuid Holland were found as well (see figure 3). Several
360 other areas are however largely missed. None of the sites in Zeeland were for example found in
361 AGNES. This is not entirely surprising as all but one of these sites were excavated in the twentieth
362 century. As such their reports were not deposited in DANS or Archis, and by extension they could
363 not be found in AGNES. Older sites are often published as articles in archaeological journals,
364 rather than as site reports deposited in DANS. This is for example the case with many of the key
365 Vlaardingen Culture sites including the type site Vlaardingen Arij Koplaan and the key sites of
366 Zandwerven, Haamstede Brabers, and Leidschendam Prinsenhof (Clason 1962; Glasbergen,
367 Groenman-van Waateringe & Hardenberg-Mulder 1967; Van Iterson Scholten 1988; Van Regteren
368 Altena 1958; Van Regteren Altena & Bakker 1961; Van Regteren Altena et al. 1962a; Verhart
369 1992). Similarly, archaeological sites discovered by volunteers are also generally absent in DANS.
370 They are often published in either the Westerheem (now Archeologie in Nederland) or in local
371 AWN reports. This is for example the case with many of the previously mentioned sites discovered
372 by volunteers in the Nijmegen area between 1970 and 2000 (De Jong 1986; 1988; Janssen 1976;
373 1980; 1989; 1993; Janssen & Tuijn 1978; Koolen 1976). A similar problem applies to many of the
374 sites found in the central part of the Netherlands and eastern parts of Zuid Holland. For example
375 a series of sites in the municipality Molenlanden were discovered by local AWN volunteers in the
376 1960's. These sites were published in the dissertation of Louwe Kooijmans, but they do not have
377 formal published excavation reports (Louwe Kooijmans 1974).

378
379 A core strength of AGNES is that it allows us to find 'by-catch' in archaeological reports. Many
380 of the previously unknown Vlaardingen Culture sites found in AGNES can be considered to be 'by-
381 catch'. This is for example the case with the site Nijmegen Park Waaijenstein. The excavation
382 focussed on a Roman period settlement in the area. The metadata of the report in DANS only
383 mention the Roman period settlement. The Vlaardingen Culture remains at the site consist of three
384 ceramic sherds and a few flint artefacts (Daniël 2018). The site Bergharen de Weem presents a
385 similar case, the report titled "On the edge of a mediaeval settlement"³ is focussed on mediaeval
386 finds. The metadata in DANS mention Neolithic remains but those are not specified. As such they
387 would not be found with the queries such as "Vlaardingen Culture". Only in the full text of the report
388 is it specified that these remains consist of flint and ceramics from the Vlaardingen Culture
389 (Diepeveen & Van Enckevort 2009). It is not surprising that many of these by-catch finds are
390 located in the region of Nijmegen. Nijmegen is the heart of the Roman Netherlands, it is the oldest
391 city in the country, and a major centre during the mediaeval period. Archaeological excavations
392 frequently yielded large quantities of finds from these periods. It is perhaps unsurprising that a
393 handful of Vlaardingen Culture sherds or flint artefacts on these excavations do not end up in the
394 metadata of these reports. This is also no longer problematic as we were now able to retrieve this
395 kind of information through AGNES.

396
397 Unfortunately we can not calculate the recall (and by extension, the F1 score), as the total
398 amount of relevant documents in the collection is unknown. However, the current case-study
399 yielded a precision of 9,7%, in terms of relevant hits. This is a much higher precision than that of
400 the previous case-study which dealt with Merovingian cremation graves. In this case-study only

3 "Aan de rand van een middeleeuwse nederzetting" (Diepeveen & Van Enckevort 2009).

401 2,1% of the hits consisted of relevant hits (Brandsen & Lippok 2021). This is a significant
402 improvement, which can partly be attributed to improvements following the recommendations
403 made in 2021, and partly due to the type of queries. As part of the EXALT project several other
404 case-studies will be carried out to further assess the efficiency of the system. It will be interesting
405 to see whether the improved precision is indeed a constant factor or whether this is largely case-
406 dependent.

407

408 The original overview which was compiled of Vlaardingen Culture sites took several months to
409 be completed and it has been further refined over the past years. Going through the AGNES
410 queries took about two to three weeks. It is clear that AGNES vastly speeds up the process of
411 compiling such overviews. This can mostly be attributed to the fact that AGNES provides direct
412 access to the relevant literature, and more specifically to the relevant pages contained within those
413 documents.

414

415 Finally, two main problems in AGNES were identified during the case study; PDFs not being
416 imported into AGNES due to PDF text extractions errors, and BERT missing a date range. The
417 PDF text extraction process has been updated after the case study to solve this problem, by using
418 a different tool less prone to errors (PyMuPDF). In a future version of AGNES, we will re-index all
419 the documents again, which means these missed documents will be available in the future.
420 Regarding the BERT error, the models currently have an F1 score of around 84% for detecting
421 time periods (Brandsen et al. 2022), meaning that around 16% of time periods are missed or
422 incorrectly classified. In future work, we want to improve on this performance by improving the
423 BERT model, potentially with more training data, or by using newer techniques such as GPT
424 models.

425

5. Conclusion

426 In the present study we aimed to test how well AGNES was equipped for generalised queries
427 aimed at finding period specific sites, in this case sites attributed to the Vlaardingen Culture. As
428 such our tests deviated from an earlier case-study which was aimed at finding highly specific
429 information, in this case on Merovingian cremation graves (Brandsen & Lippok 2021). We can
430 conclude that AGNES also greatly contributes to more general queries. Through AGNES we found
431 thirty previously unknown Vlaardingen Culture sites. This included the most northern Vlaardingen
432 Culture site ever found. As such, the study also contributed to our understanding of the
433 geographical spread of the Vlaardingen Culture phenomenon.

434

435 Newly discovered sites often consisted of by-catch on excavations in which the majority of finds
436 were dated to other (later) periods. Although AGNES greatly contributed to our overview the
437 program also missed a large number of known Vlaardingen Culture sites. This is mainly the case
438 because older sites are usually not published in site reports which have been deposited in DANS
439 or Archis. Often such sites are published in scientific and semi-scientific publications. Sites by
440 amateur archaeologists are usually published in local archaeological and historical journals. It is
441 recommended that such publications are digitised and deposited in DANS. This will increase the
442 visibility and usefulness of citizen science.

443

444 Another problem we observed is that there is no consensus amongst authors on which terms to
445 use to describe Vlaardingen Culture sites. They are referred to as: Vlaardingen Culture,
446 Vlaardingen group, Vlaardingen-Stein group (or Stein-Vlaardingen group), or Stein-Vlaardingen
447 Complex. It seems that the terms are applied rather arbitrarily, when sites are located in the border
448 area between the Stein group and Vlaardingen Culture they are often referred to as Stein-
449 Vlaardingen or Vlaardingen-Stein sites. In the western Netherlands the term Vlaardingen Culture
450 is more systematically applied. The problem seems to stem from the fact that a distinction between
451 the Stein group and Vlaardingen Culture is problematic, especially when dealing with a small

452 number of finds, discovered by chance, as is often the case in development-led archaeology. While
453 we were not able to provide a satisfactory solution to this problem, the study provided valuable
454 insights into the terminologies employed in development-led archaeology.

455
456 We can conclude that AGNES cannot be used as an alternative to established search methods
457 for creating thematic or temporal site overviews, because archaeological sites are not exclusively
458 published in excavation reports. Nevertheless, it provides an effective tool for finding
459 archaeological 'by-catch'. In this case 19% of the Vlaardingen Culture sites in our overview were
460 discovered exclusively through AGNES. As such, AGNES provided an essential tool for compiling
461 this overview. This type of 'by-catch' cannot be effectively found through other means, therefore it
462 is recommended that AGNES is used systematically in tandem with established search methods.
463

464 Acknowledgements

465 We like to thank Annelou van Gijn, PI of the Putting Life into Late Neolithic Houses project, for
466 supporting the collaboration between the Putting Life and the EXALT projects. We would like to
467 thank all the people who have provided access to difficult to find literature for the compilation of
468 the original overview.
469

470 Funding

471 This research was funded by an NWO grant for the project: EXALT: Excavating Archaeological
472 Literature - Enabling scientific discovery by applying multilingual text mining to archaeological
473 documents (grant number: AIB.19.011). The original overview of Vlaardingen Culture sites in the
474 Netherlands was created as part of the NWO project Putting Life into Late Neolithic Houses:
475 Investigating domestic craft and subsistence activities through experiments and material analysis
476 (grant number: AIB.19.020).

477 Conflict of interest disclosure

478 The funding source was not involved in study design; in the collection, analysis and
479 interpretation of data; in the writing of the report; and in the decision to submit the article for
480 publication. The authors declare that they comply with the PCI rule of having no financial conflicts
481 of interest in relation to the content of the article. The authors declare the following non-financial
482 conflict of interest: Alex Brandsen is a recommender for PCI Archaeology. The authors report there
483 are no other competing interests to declare.

484 Data, scripts, code, and supplementary information availability

485 The data used in this research is available in this Zenodo archive (
486 doi.org/10.5281/zenodo.13283972), together with the supplementary materials (1-3). The BERT
487 models and code are also available (Brandsen, 2023).

488 References

- 489 't Hart, R, Norde, EHL and Tuinstra, JWD. 2019. *Prehistorische bewoning langs een geul in*
490 *het plangebied Hekkeslag in Ewijk, gemeente Beuningen; archeologisch onderzoek: een*
491 *opgraving*. RAAP-Rapport 3679. Weesp.
- 492 Amrani, A, Abajian, V and Kodratoff, Y. 2008. A chain of text-mining to extract information in
493 archaeology. In: *Information and communication technologies: From theory to applications*,

- 494 ICTTA 2008, Damascus, Syria. Damascus. pp. 1-5.
495 DOI: <https://doi.org/10.1109/ICTTA.2008.4529905>
- 496 **Brandsen, A.** 2021. ArcheoBERTje - A Dutch BERT model for the Archaeology domain. *Zenodo*
497 *Repository*. DOI: <https://doi.org/10.5281/zenodo.4739063>
- 498 **Brandsen, A.** 2021. *Digging in Documents - Using Text Mining to Access the Hidden Knowledge*
499 *in Dutch Archaeological Excavation Reports*. PhD Thesis. Leiden University.
- 500 **Brandsen, A.** 2023. Archaeology specific BERT models for English, German, and Dutch.
501 Amsterdam. *Proceedings of the CAA2023 Conference*. DOI:
502 <https://doi.org/10.5281/zenodo.8296920>
- 503 **Brandsen, A, Lambers, K, Verberne, S and Wansleeben, M.** 2019. User Requirement
504 Solicitation for an Information Retrieval System Applied to Dutch Grey Literature in the
505 Archaeology Domain. *Journal of Computer Applications in Archaeology*, 2(1): 21-30. DOI:
506 <https://doi.org/10.5334/jcaa.33>
- 507 **Brandsen, A and Lippok, F.** 2021. A burning question – Using an intelligent grey literature
508 search engine to change our views on early medieval burial practices in the Netherlands.
509 *Journal of Archaeological Science*, 133: 105456.
510 DOI: <https://doi.org/https://doi.org/10.1016/j.jas.2021.105456>
- 511 **Brandsen, A, Verberne, S, Lambers, K and Wansleeben, M.** 2021. Usability Evaluation for
512 Online Professional Search in the Dutch Archaeology Domain. *arXiv*.
513 DOI: <https://doi.org/10.48550/arXiv.2103.04437>
- 514 **Brandsen, A, Verberne, S, Lambers, K and Wansleeben, M.** 2022. Can BERT Dig It? - Named
515 Entity Recognition for Information Retrieval in the Archaeology Domain. *Journal on*
516 *Computing and Cultural Heritage*, 15(3): 1-18. DOI: <https://doi.org/10.1145/3497842>
- 517 **Bulten, EEB and Stokkel, PJA.** 2017. Synthese. In: Stokkel, PJA and Bulten, EEB (eds.), *De*
518 *Wateringse Binnentuinen, Gemeente Den Haag: Een vlaardingennederzetting in het*
519 *Wateringse Veld*. Haagse Oudheidkundige Publicaties 20. Den Haag: Afdeling Archeologie &
520 Natuur- en Milieueducatie Dienst Stadsbeheer. pp. 283-305.
- 521 **Clason, AT.** 1962. Nieuwe opgravingen van de neolithische nederzetting te Zandwerven,
522 Gem.Opmeer, III. *Varia Bio-Archaeologica*, 18: 201-219.
- 523 **Daniël, AAWJ.** 2018. *Een inheems-Romeinse nederzetting in plangebied Park Waaijenstein te*
524 *Nijmegen Noord*. Proefsleuvenonderzoek project Npw2. Archeologische Berichten Nijmegen
525 83. Nijmegen.
- 526 **De Jong, JPM.** 1986. Op de Craayenberg. In: *Jaarverslag 1986 AWN afd. Nijmegen e.o.* pp. 24-
527 26.
- 528 **De Jong, JPM.** 1988. Op de Craayenberg II. In: *Jaarverslag 1988 AWN afd. Nijmegen e.o.* pp.
529 11-13.
- 530 **De Koning, MWA.** 2010. *Onder de rook van Wijchen : vondsten van de Vlaardingens-Groep, de*
531 *klokbekercultuur, sporen van ovens en resten van een mogelijk grafveldje uit de late bronstijd-*
532 *Romeinse tijd*. Archeologische Berichten Wijchen 11. Nijmegen.
- 533 **De Ridder, JAA.** 2011. *Middenneolithische vondsten aan de Schoolstraat en Molenweg te*
534 *Thorn, gemeente Maasgouw Een Inventariserend Veldonderzoek in de vorm van*
535 *proefsleuven*. ADC-Rapport 2908. Amersfoort.
- 536 **Devlin, J, Chang, MW, Lee, K and Toutanova, K.** 2019. BERT: Pre-training of deep
537 bidirectional transformers for language understanding. *arXiv*. DOI:
538 <https://doi.org/10.48550/arXiv.1810.04805>
- 539 **Diepeveen, MC and Van Enckevort, H.** 2009. *Aan de rand van een middeleeuwse nederzetting.*
540 *Archeologisch onderzoek in de bouwput van het Kulturhus in Berharen, gemeente Wijchen.*
541 *Archeologische Berichten Wijchen 9*. Nijmegen.
- 542 **Dorenbos, O and Koot, H.** 2010. Wonen bij de kust, Vlaardingens bewoners van de strandwal.
543 *Westerheem*, Special 2010: Vlaardingens-cultuur(2): 96-113.
- 544 **Eijskoot, YDRT.** 2004. *Freegolf 6.161. Basisverslag van de archeologische begeleiding van de*
545 *uitbreiding van het Freegolfterrein*. VLAK-verslag 18. Vlaardingens.

- 546 **Eimermann, E.** 2008. *Gemeente Hulst, Grote Bagijnestraat nr. 21-29. Sporen van*
547 *laatmiddeleeuwse activiteiten. Een Archeologische Begeleiding en Inventariserend*
548 *Veldonderzoek in de vorm van een proefsleuf.* ADC-Rapport 1225. Amersfoort.
- 549 **Glasbergen, W, Groenman-van Waateringe, W and Hardenberg-Mulder, GM.** 1967.
550 Settlements of the Vlaardingen culture at Voorschoten and Leidschendam (I). *Helinium* 7: 4-
551 31.
- 552 **Goossens, TA.** (Ed.) 2009. *Opgraving Hellevoetsluis-Ossenhoek. Een nederzetting van de*
553 *Vlaardingen-groep op een kwelderrug in de gemeente Hellevoetsluis.* Archol-Rapport 87.
554 Leiden.
- 555 **Gormley, C and Tong, Z.** 2015. *Elasticsearch: The Definitive Guide: A Distributed Real-Time*
556 *Search and Analytics Engine.* Cambridge: O'Reilly Media.
- 557 **Groenman-van Waateringe, W and Van Regteren Altena, JF.** 1966. Voorschoten. *Bulletin van*
558 *de Koninklijke Nederlandse Oudheidkundige Bond*, 65(5): 88-89.
- 559 **Habermehl, D.** 2024. *Over zaaien en oogsten II. Onderzoek naar de kwaliteit en bruikbaarheid*
560 *van archeologische rapporten voor synthetiserend onderzoek. Een aanvulling (2024) op*
561 *eerder onderzoek (2019).* Amersfoort: Rijksdienst voor Cultureel Erfgoed.
- 562 **Heirbaut, ENA.** 2010. Inleiding. In: Heirbaut, ENA, J. Hendriks, J. and Hermsen I (eds.), *Onder*
563 *een deken van zand. Archeologisch onderzoek op de Meshallen en de Kraanvogel te*
564 *Wijchen.* Archeologische Berichten Wijchen 10. Nijmegen: Gemeente Nijmegen, Bureau
565 Archeologie en Monumenten. pp. 1-18.
- 566 **Hiddink, HA.** 2000. *Groesbeek-Klein Amerika. Aanvullend archeologisch onderzoek van*
567 *bewoning uit de prehistorie en de Romeinse tijd en een mogelijk pre-Flavisch*
568 *openluchtheiligdom.* Rapportage Archeologische Monumentenzorg 74. Amersfoort:
569 Rijksdienst voor het Oudheidkundig Bodemonderzoek.
- 570 **Janssen, AJ.** 1976. Wychen, randweg oost. In: *Jaarverslag 1976 AWN afd. Nijmegen e.o.* pp.
571 16-18.
- 572 **Janssen, AJ.** 1980. Vlaardingen-kultuur op de Blankenburg te Beuningen. In: *Jaarverslag 1980*
573 *AWN afd. Nijmegen e.o.* pp. 28.
- 574 **Janssen, AJ.** 1989. Een neolithische nederzetting in het Ewijkse Veld. *Westerheem*, 38(3): 133-
575 145.
- 576 **Janssen, AJ.** 1993. Neolithische vondsten van het Woerdje aan de Ooigraaf te Beuningen.
577 *Westerheem* 42(6): 287-294.
- 578 **Janssen, AJ and Tuijn, WN.** 1978. De Homberg te Wijchen. *Westerheem* 27(4-5): 238-255.
- 579 **Koolen, M.** 1976. Vondsten op "Hoogbroek" te Wijchen. In: *Jaarverslag 1976 AWN afd.*
580 *Nijmegen e.o.* pp. 11-15.
- 581 **Louwe Kooijmans, LP.** 1974. The Rhine/Meuse delta. Four studies on its prehistoric occupation
582 and Holocene geology. *Analecta Praehistorica Leidensia* 7: 169-445.
- 583 **Louwe Kooijmans, LP.** 1983. Tussen SOM en TRB, enige gedachten over het laat-neolithicum
584 in Nederland en België. *Bulletin van de Koninklijke Musea voor Kunst en Geschiedenis* 54(1):
585 55-67.
- 586 **Paijmans, H and Brandsen, A.** 2009. What is in a Name: Recognizing Monument Names from
587 Free-Text Monument Descriptions. In: Van Erp, MGJ, Stehouwer, JH and Van Zaanen, M
588 (eds.), *Proceedings of the 18th Annual Belgian-Dutch Conference on Machine Learning*
589 *(Benelearn).* Tilburg: Tilburg Center for Creative Computing. pp. 2-6.
- 590 **Paijmans, H and Brandsen, A.** 2010. Searching in archaeological texts: Problems and solutions
591 using an artificial intelligence approach. *PalArch's Journal Of Archaeology Of*
592 *Egypt/Egyptology*, 7(2): 1-6.
- 593 **Raemaekers, D.** 2003. Cutting a long story short? The process of neolithization in the Dutch
594 delta re-examined. *Antiquity*, 77(298): pp. 740-748.
595 DOI: <https://doi.org/10.1017/S0003598X00061688>
- 596 **RCE.** 2023. *Erfgoedmonitor 2023 - Feiten en cijfers over ons erfgoed in 2023.* Amersfoort:
597 Rijksdienst voor Cultureel Erfgoed.

- 598 **Richards, J, Tudhope, D and Vlachidis, A.** 2015. Text Mining in Archaeology: Extracting
599 Information from Archaeological Reports. In: Barcelo, JA and Bogdanovic, I (Eds.),
600 *Mathematics and Archaeology*. Boca Raton: CRC Press pp. 240-254. DOI:
601 <https://doi.org/10.1201/b18530-15>
- 602 **Stokkel, P.** 2017. Sporen en structuren In: Stokkel, PJA and Bulten, EEB (eds.), *De Wateringse*
603 *Binnentuinen, Gemeente Den Haag: Een vlaardingennederzetting in het Wateringse Veld*.
604 Haagse Oudheidkundige Publicaties 20. Den Haag: Afdeling Archeologie & Natuur- en
605 Milieueducatie Dienst Stadsbeheer. pp. 53-100.
- 606 **Teubner, B and Tuijn, W.** 2010. Vlaardingens-cultuur in Wijchen en omgeving. *Westerheem*,
607 Special 2010: Vlaardingens-cultuur (2): 180-185.
- 608 **Tjong Kim Sang, EF.** 2002. Introduction to the CoNLL-2002 Shared Task: Language-
609 Independent Named Entity Recognition. *arXiv*. DOI:
610 <https://doi.org/10.48550/arXiv.cs/0209010>
- 611 **Van Beek, BL.** 1990. *Steentijd te Vlaardingens, Leidschendam en Voorschoten. De*
612 *vondstverspreiding in laat-neolithische nederzettingen in het hollandse kustgebied Universiteit*
613 *van Amsterdam*. Unpublished thesis (PhD). Universiteit van Amsterdam.
- 614 **Van den Broeke, PW.** 1993. A crowded peat area; observations in Vlaardingens-West and the
615 Iron Age habitation of southern Midden-Delfland. *Analecta Praehistorica Leidensia* 26: 59-82.
- 616 **Van den Dikkenberg, L.** 2024. Vlaardingens culture or Stein group? From object biographies to
617 communities of practice. In: Verbaas, A, Langejans, G, Little, A and Chan B (eds.), *Artefact*
618 *Biographies from Mesolithic and Neolithic Europe and Beyond. Papers in honour of Professor*
619 *Annelou van Gijn*. *Analecta Praehistorica Leidensia* 52. pp. 63-72.
- 620 **Van Gijn, AL, and Bakker, JA.** 2005. Megalith builders and sturgeon fishers. Middle Neolithic B:
621 Funnel Beaker culture and Vlaardingens group. In: Louwe Kooijmans, LP, Van den Broeke,
622 PW, Fokkens, H and Van Gijn AL (Eds.), *The Prehistory of the Netherlands*. Amsterdam:
623 Amsterdam University Press. pp. 281-306.
- 624 **Van Iterson Scholten, FR.** 1988. Inventory and Protection of the Archaeological Monument of
625 Zandwerven. *Helinium* 28(2): 232-243.
- 626 **Van Kampen, JCG.** 2013. Bewoning in het Neolithicum. In: Van Kampen, JCG and Van den
627 Brink, V (Eds.), *Archeologisch onderzoek op de Habraken te Veldhoven. Twee unike*
628 *nederzettingen uit het Laat Neolithicum en de Midden Bronstijd en een erf uit de Volle*
629 *Middeleeuwen*. Zuidnederlandse Archeologische Rapporten 52. Amsterdam: VUHbs
630 archeologie. pp. 39-61.
- 631 **Van Noort, G.** 1998. De Neolithische boeren van de Vlaardingens-cultuur op Texel. *Uitgave van de*
632 *Historische Vereniging Texel* 48: 3-7.
- 633 **Van Regteren Altena, JF.** 1958. Nieuwe opgravingen van de Neolithische nederzetting te
634 Zandwerven gemeente Zandbroek. *West Frieslands Oud en Nieuw* 25: 144-159.
- 635 **Van Regteren Altena, JF and Bakker, JA.** 1961. De neolithische woonplaats Zandwerven
636 (N.H.). In: Glasbergen W and Groenman-van Waateringe W (Eds.), *In het voetspoor van van*
637 *A.E. van Giffen*. Groningen: J.B. Wolters. pp. 33-40.
- 638 **Van Regteren Altena, JF, Bakker, JA, Clason, AT, Glasbergen, W, Groenman-van**
639 **Waateringe, W and Pons, LJ.** 1962a. The Vlaardingens culture I. *Helinium* 2: 3-35.
- 640 **Van Regteren Altena, JF, Bakker, JA, Clason, AT, Glasbergen, W, Groenman-van**
641 **Waateringe, W and Pons, LJ.** 1962b. The Vlaardingens culture II. *Helinium* 2: 97-103.
- 642 **Van Regteren Altena, JF, Bakker, JA, Clason, AT, Glasbergen, W, Groenman-van**
643 **Waateringe, W and Pons, LJ.** 1962c. The Vlaardingens culture III. *Helinium* 2: 215-243.
- 644 **Van Zoolingen, RJ.** 2021. Sporen en Structuren. In: Van Zoolingen, RJ and Bulten, EEB (Eds.),
645 *Steynhof (Anne Frankpark) Gemeente Den Haag Een Vlaardingennederzetting in het*
646 *Wateringse Veld*. Haagse Oudheidkundige Publicaties 23. Den Haag: Afdeling Archeologie &
647 Natuur- en Milieueducatie Dienst Stadsbeheer. pp. 45-82.
- 648 **Van Zoolingen, RJ and Bulten EEB.** 2021. Synthese. In: Van Zoolingen, RJ and Bulten, EEB
649 (Eds.), *Steynhof (Anne Frankpark) Gemeente Den Haag Een Vlaardingennederzetting in het*

650 *Wateringse Veld*. Haagse Oudheidkundige Publicaties 23. Den Haag: Afdeling Archeologie &
651 Natuur- en Milieueducatie Dienst Stadsbeheer. pp. 291-309.

652 **Van Zoolingen, RJ and Rieffe, C.** 2021. *Noordweg 76 Gemeente Den Haag Sporen uit het*
653 *neolithicum, de ijzertijd en Romeinse tijd in het Wateringse Veld*. Haagse Archeologische
654 Rapportage 2301. Den Haag: Afdeling Archeologie & Natuur- en Milieueducatie Dienst
655 Stadsbeheer.

656 **Verhart, LBM.** 1992. Settling or trekking? The late Neolithic house plans of Haamstede-Brabers
657 and their counterparts. *Oudheidkundige Mededelingen uit het Rijksmuseum van Oudheden te*
658 *Leiden* 72, 73-99.

659 **Verhart, LBM.** 2010. De zuidelijke tegenhangers van de Vlaardingencultuur: een schamel spoor
660 van scherven. *Westerheem*, Special 2010: Vlaardingencultuur(2): 207-222.

661 **Verhart, LBM and De Ridder, T.** 2010. Een waaier van Vlaardingencultuur. *Westerheem*, Special
662 2010: Vlaardingencultuur(2): 4-11.

663 **Vlachidis, A and Tudhope, D.** 2012. A pilot investigation of information extraction in the
664 semantic annotation of archaeological reports. *International Journal of Metadata, Semantics*
665 *and Ontologies* 7(3): 222-222. DOI: <https://doi.org/10.1504/IJMSO.2012.050183>

666 **Vos, P, Van der Meulen, M, Weerts, HJT and Bazelmans, J.** 2020. *Atlas of the Holocene*
667 *Netherlands, landscape and habitation since the last ice age*. Amsterdam: Amsterdam
668 University Press.