



**Negation Detection and Word Sense Disambiguation in
Digital Archaeology Reports for the purposes of Semantic
Annotation**

Journal:	<i>Program</i>
Manuscript ID:	PROG-10-2014-0076
Manuscript Type:	Article
Keywords:	Negation Detection, Word Sense Disambiguation, Digital Humanities, Natural Language Processing, CIDOC-CRM, Semantic Annotation

SCHOLARONE™
Manuscripts

View Only

Negation Detection and Word Sense Disambiguation in Digital Archaeology Reports for the purposes of Semantic Annotation

Purpose

The paper presents the role and contribution of Natural Language Processing Techniques, in particular Negation Detection and Word Sense Disambiguation in the process of Semantic Annotation of Archaeological Grey Literature. Archaeological reports contain a great deal of information that conveys facts and findings in different ways. This kind of information is highly relevant to the research and analysis of archaeological evidence but at the same time can be a hindrance for the accurate indexing of documents with respect to positive assertions

Design/methodology/approach

The paper presents a method for adapting the biomedicine oriented negation algorithm NegEx to the context of archaeology and discusses the evaluation results of the new modified negation detection module. A particular form of polysemy, which is inflicted by the definition of ontology classes and concerning the semantics of small finds in archaeology, is addressed by a domain specific word-sense disambiguation module.

Findings

The performance of the Negation Detection module is compared against a “Gold Standard” that consists of 300 manually annotated pages of archaeological excavation and evaluation reports. The evaluation results are encouraging, delivering overall 89% Precision, 80% Recall and 83% F-Measure scores. The paper addresses limitations and future improvements of the current work and highlights the need for ontological modelling to accommodate negative assertions.

Keywords: Negation Detection, Digital Humanities, Word Sense Disambiguation, CIDOC-CRM, Semantic Annotation, Natural Language Processing.

1 Introduction

The latest advances of semantic technologies have opened new innovative ways in which scholars can act and elaborate information via search and browsing software applications that process the “meaning” of words beyond the level of a simple and dry string matching process (O'Hara et al., 2010). Semantic metadata practices enrich information with conceptualisations that enable sophisticated methods for data publishing and pave new ways for information analysis and data integration (Bizer et al., 2009).

1
2
3
4
5
6
7
8
9
10 Named Entity Recognition (NER), also sometimes referred to as Named Entity
11 Recognition and Classification (NERC), is a particular Natural Language Processing
12 (NLP) task aimed at the recognition and classification of units of information to
13 predefined categories, such as names of person, location, organisation, expressions of
14 time, money, percentage etc. (Nadeau & Sekine 2007). NER can lead the delivery of
15 specific metadata in form of Semantic Annotations, which are usually generated with
16 respect to a given ontology and are aimed to automate identification of concepts and
17 their relationships in documents (Uren et al., 2006). Linguistic phenomena, such
18 polysemy and negation present a challenging ground for the application of such
19 semantic metadata technologies that ought to be explored and investigated for the
20 delivery of scholarly research solutions that could significantly enhance and advance
21 the ways in which the search of facts and findings is conducted.

22 The following sections of this paper discuss the method and evaluation results of the
23 negation detection and word sense disambiguation (WSD) modules of the OPTIMA
24 pipeline (Vlachidis, 2012) for the semantic annotation of archaeological grey
25 literature with respect to the CIDOC CRM ontology. CIDOC CRM is an ISO
26 Standard (ISO 21127:2006) comprehensive semantic framework that makes available
27 semantic definitions and clarifications that promote shared understanding of cultural
28 heritage information (Crofts et al., 2009).

29 OPTIMA contributed to the Semantic Technologies for Archaeological Research
30 (STAR) project (Tudhope et al., 2011), which explored the potential of semantic
31 technologies in query and integration of archaeological digital resources. The output
32 of the pipeline is delivered in the form of semantic indices, expressed as RDF triples
33 that enable semantically defined information retrieval and cross-searching over
34 disparate archaeological digital resources i.e. grey literature and datasets.

35
36 The paper highlights the essential role of the negation detection and WSD modules
37 for the aims of the pipeline and reveals the contribution of the NegEx algorithm
38 (Chapman et al., 2001) in the applied method. The necessary modifications of NegEx
39 are also discussed which led to the adaptation of a biomedicine oriented algorithm to
40 the negation requirements of the archaeology domain. The evaluation results of the
41 negation module of OPTIMA are encouraging, delivering high Precision (89%) and
42 Recall (80%) scores. The performance of the negation module is evaluated with the
43 Gold Standard method of humanly defined annotations. The contribution of the WSD
44 module is evaluated against the overall NER outcome of the OPTIMA pipeline and
45 example cases are discussed. The paper also discusses the issue of the limited support
46 of ontologies, in particular CIDOC-CRM, in modelling and representation of negated
47 findings and concludes with known limitations and future improvements of the work.
48
49
50
51
52
53
54
55
56
57
58
59
60

2 Background and Relevant Work

2.1 Negation Detection

Negation is an integral part of any natural language system. It is a linguistic, cognitive and intellectual phenomenon, which enables the users of a language system to communicate erroneous messages, the truth value of a proposition, contradictions, irony and sarcasm (Horn, 1989). Philosophers, from Plato to Spencer Brown have independently approached negation as a case of heteron (not-being) described as a “positive assertion of the existence of a relevant difference” (Westbury, 2000). Whereas, there is a plethora of studies and theories addressed to the complexity of negation and its characteristics from a philosophical and linguistic point of view, research on automatic detection of negation and representation of its semantics has not been extensive (Blanco and Moldovan, 2011)

In recent years, Natural Language Processing (NLP) applications have mainly drawn attention to the automatic detection of negation in biomedical text (Chapman et al., 2001) and in opinion mining and sentiment analysis (Maynard and Funk, 2011). The techniques and approaches that are employed to address the issue of negation within NLP vary and cover a wide spectrum of machine learning and ruled-based (regular expressions and syntactic processing) applications. Both machine learning and rule-based approaches have been reported as capable of addressing the task of automatic detection, with the rule-based approaches having an edge and being competitive in the biomedicine domain (Goryachev et al. 2006). However, there is little evidence of research aimed at the study of negation detection in the domain of archaeology, albeit some strong parallels can be drawn between archaeological and biomedicine research questions particularly when dealing with facts and findings (Vlachidis and Tudhope 2012).

In the scientific research of the biomedicine and the archaeology domain, appreciation and understanding of negated facts is as equally important as the interpretation of positive findings. For example the medical phrase “*The chest X-ray showed no infiltrates*” reveals a significant negated finding which can lead to a correct diagnosis of a cardiac condition. In archaeology “negative results are essential when providing an assessment of the archaeological potential of a specific site” (Falkingham, 2005), for example the phrase “*No traces of a Roman settlement have been discovered in the area*” can lead to specific conclusions with regards to settlement activity during the Roman period in a particular area. Being able to distinguish such negative assertions in context is highly desirable for the research and analysis of facts and findings especially when those activities are supported by information retrieval systems.

2.2 The NegEx Algorithm

NegEx (Chapman et al., 2001) is a specific algorithm targeted at the identification of negated findings in medical documents. The algorithm determines whether Unified Medical Language System (UMLS) terms of findings and diseases are negated in the context of medical reports. NegEx is particularly relevant to the scope of the

1
2
3
4
5
6
7
8
9 OPTIMA negation module, due to its rule-based design, the use of pattern matching
10 mechanism and the employment of vocabulary listings.

11
12 The design of the algorithm is based on the use of offset patterns that utilise a
13 negation related vocabulary. The vocabulary contains terms and phrases that denote
14 negation, which are invoked by a set of rules. The algorithm makes use of two
15 specific patterns; The first pattern [Pre-Neg] identifies negated UMLS terms in
16 phrases which commence with a negation phrase followed by a window of up to 5
17 tokens before matching an UMLS term, i.e. <negation phrase> * <UMLS Term>.
18 The second pattern [Post-Neg] is a reversed version of the above, which matches
19 negated UMLS terms that are up to five tokens prior to a negation phrase, i.e.
20 <UMLS Term> * <negation phrase>.

21
22 There are two main parallels for archaeological reports which support the adoption of
23 the NegEx approach by the OPTIMA negation mechanism. Firstly, the use of pattern
24 matching rules and vocabulary terms allows a smooth integration of the algorithm
25 within the requirements and scope of the OPTIMA pipeline for semantic annotation
26 via based rule-based techniques that are supported by knowledge organisation
27 resources (i.e. thesauri and glossaries). Secondly, the good performance of the
28 algorithm in detecting negations about findings in biomedicine context. In
29 archaeological reports, as in medical reports, authors frequently negate facts about
30 findings (Falkingham, 2005).

31 **2.3 Word Sense Disambiguation**

32 Word Sense Disambiguation (WSD) refers to the computational ability to identify the
33 different meanings of a word that has multiple meanings (Navigli, 2009). Early
34 attempts to answer the problem of polysemy via computational means originate back
35 to the 1950's. Initially the attempts were focused in limited domains or over small
36 vocabularies. From the 1980's improvements in the scalability of WSD systems were
37 made due to the advances of the available computational means and the progress of
38 Machine Learning (ML) techniques, enabling disambiguation over larger
39 heterogeneous resources.

40 WSD applications can be rule-based or Machine Learning. Machine learning
41 approaches can be supervised, which require a training set for determining the
42 disambiguation results, or they can be unsupervised. When rule-based, a WSD task
43 invokes hand-crafted rules which utilise contextual for determining the
44 disambiguation results. Rules can utilise templates which state that an ambiguous
45 word has a certain sense when a particular word(s) appears in a specific location
46 relative to that word. Knowledge based resources, such as dictionaries, glossaries,
47 thesauri etc., can also be used by a WSD task for supporting inference of word senses
48 in context.

49
50 Voorhees (1993) devised a WSD system based on WordNet exploiting synsets of
51 nouns for defining what she called hoods which were used to determine the sense of
52 ambiguous terms. Based on the assumption that a set of words that occur together in a
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9 context determine appropriate senses for one other, the system populated diverse
10 hoods with words from different synsets for an ambiguous word in a given context.
11 The hood with the largest number of occurrences determined the sense of an
12 ambiguous word. The approach signified the importance of contextual evidence in the
13 disambiguation process, in line with (Resnik, 1997) that linked the disambiguation
14 process with Selectional Preference. Selectional Preference is the tendency of words
15 to co-occur with words that belong to specific semantic groups.

16
17 WSD can be viewed as an automatic classification task that makes use of contextual
18 evidence and external knowledge resources for applying an appropriate class (word
19 sense) to ambiguous terms. The task of disambiguation can be focused on a particular
20 set of words thus to be “targeted”, or it can be applied to the vast range of all words in
21 document. Typically a WSD task is configured as an intermediate task of a larger
22 NLP application, either set up to execute as a standalone module or integrated within
23 the system architecture. Although, use of ML approaches can improve the
24 generalisability of a disambiguation method, still many WSD systems have inherited
25 limitations in terms of their performance and generalisation, especially when fine-
26 grained sense distinctions are employed by the disambiguation task (Navigli 2009).

27 **2.4 Ontological Polysemy**

28 Polysemy is defined as the condition where that same word can carry multiple
29 meanings (senses) e.g. *ring* (wedding vs. boxing) and should not be confused with
30 homonymy where two unrelated words share the same spelling (homographs) and
31 pronunciation (homophones) e.g. *saw* (see vs. tool).
32

33 The adoption of the CIDOC CRM ontology for driving the NER task of OPTIMA
34 brought a specific form of polysemy, which is inflicted by the definition of ontology
35 classes. The CRM-EH is concerned to model the processes involved in archaeological
36 excavation recording and analysis. Thus, a particular pottery fragment becomes a
37 ‘find’, as a result of being extracted from a context and recorded separately on site.
38 For example “pottery” in CRM-EH can be classified as Physical Object or as Material
39 with only one of the two classes considered to be correct in a given context. Usually
40 the distinction reflects the focus of scholarly inquiry rather than any absolute
41 semantics. An entity may be treated as either a material or as an object, depending on
42 the archaeological objective. Consider the term “pottery” in the phrases “...ditch
43 containing pottery and coins...” and “ditch containing pottery fragments”. In the first
44 case “pottery” refers to a physical object found in a ditch, while in the second case the
45 term refers to the material of fragments.

46 The volume of ambiguous terms (physical object vs. material) concerning the WSD
47 task is defined by a thorough examination of the domain specific vocabulary
48 resources contributing to the NER task. The NER phase of the OPTIMA pipeline
49 utilises a range of (English Heritage) thesauri and recording manual glossary
50 resources, such as the MDA Object Types thesaurus, Main Building Material
51 thesaurus, Small Finds Form glossary and Box Index Material Form glossary. A term
52 overlap study of the contributing vocabulary resources revealed the volume of
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
ambiguous terms addressed by the WSD phase (Vlachidis 2012). For example, there is an extensive overlap between the Small Finds Form glossary (aligned to Physical Object) and the Box Index Material Form glossary (aligned to Material). The overlap covers more than half of the glossary terms, such as "Animal bone", "Bone", "Fired clay", "Flint", "Glass", "Human bone", "Plaster", "Pottery", "Shell", "Slag", "Stone", "Wood".

15
16
17
18
19
20
21
22
23
This particular form of ambiguity that concerns the semantics of small finds in archaeology is also verified in Dutch and could be possibly evidenced in other European languages. A pilot study for the purposes of the FP7 Infrastructures Project ARIADNE (Advanced Research Infrastructure for Archaeological Dataset Networking in Europe) revealed a series of Dutch small find concepts, such as aardewerk (pottery), bot (bone), houtskool (charcoal), vuursteen (flint) etc. that carry polysemy characteristics analogous to the English small finds terms.

24 25 26 27 28 29 30 31 32 33 34 **3 Method**

35
36
37
38
39
40
41
42
43
44
The negation detection module of the OPTIMA pipeline is primarily developed to support the task of Named Entity Recognition (NER) with respect to the CIDOC-CRM entities *E19.Physical Object*, *E49.Time Appellation*, *E53.Place* and *E57.Material*. NER is a particular subtask of Information Extraction aimed at the recognition and classification of units of information to predefined categories (Nadeau and Sekine, 2007). On the other hand, the WSD module of the OPTIMA pipeline addresses a particular and domain specific type of polysemy inflicted by the definition of ontology classes. It is a form of polysemy that is dictated by the conceptual definitions of an ontological model mainly concerning the fine distinction between small find objects and materials affecting the overall precision of the NER task.

45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
Since the aim of the semantic annotation pipeline (OPTIMA) is to deliver semantic indices of archaeological grey literature, it is important to be able a) to exclude from indexing those occurrences of CRM entities that are negated b) to minimise the cases of incorrect classification of the recognised entities concerning the fine distinction between small find objects and materials. Thus, the fundamental aim of the negation detection and WSD modules is to strengthen the precision performance of the pipeline by discarding matches that could harm the validity of results at information retrieval level.

61 62 63 64 65 66 67 68 69 70 **3.1 Adapting the NegEx Algorithm in the Domain of Archaeology**

71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
The process of adaptation of the NegEx in the archaeological domain addressed a range of modifications relating to the coverage and use of negation glossaries as well as adaptations to the scope and application of the negation rules themselves. The main aim of the adaptation exercise was to apply the NegEx approach to the identification of negation phrases involving the four CRM entities (Physical Object, Time Appellation, Place and Material) which are targeted by the NER phase of the OPTIMA pipeline. Modification of the original pattern matching rules to aim at CRM entities instead of UMLS terms is a straightforward task. However, a range of

1
2
3
4
5
6
7
8
9 additional adaptation issues required further examination before porting the original
10 algorithm and glossaries in the archaeology domain.

11
12 The adaptation strategy considered the following issues which potentially affect
13 application of NegEx to a new domain: i) the size of the negation window which
14 originally had been set to a span of a maximum five word-tokens to fit particularly to
15 the writing style of medical text, ii) coverage and re-usability capacity of existing
16 negation glossaries to support the negation detection task in a new domain, iii)
17 usefulness of the *pseudo-negation* glossary list for limiting the scope of a negation
18 phrase and iv) review on the relevancy of the assumption that medical narrative is
19 “lexically less ambiguous than unrestricted documents” (Ruch et al., 2001) in the
20 context of archaeological reports. In addition, the reported limitation of NegEx at
21 targeting cases of conjunct negated terms was addressed during the adaptation task as
22 discussed in the next section.

23 *3.1.1 Corpus Analysis to Inform the Task of Adaptation*

24 The main aim of the bottom up corpus analysis was to reveal additional vocabulary
25 evidence which could be used by the negation detection mechanism in order to
26 improve adaptation of the algorithm to the context and the writing style of
27 archaeological reports. Therefore, it was decided that a negation window expanding
28 beyond the window limit of five tokens could be exercised for surfacing larger spans.

29
30 The first stage of corpus analysis extracted from a volume of 2460 archaeological
31 reports, phrases of a maximum of 10 tokens which contained negation moderators and
32 CRM entity matches. Using the existing NegEx [Pre-Neg] and [Post Neg] glossary
33 listings, the following two separate matching grammars were constructed:

```
34 ((Token.string!=". "}) [0, 5] {PreNeg} ({Token.string!=". "}) [0, 5]  
35 ((Token.string!=". "}) [0, 5] {PostNeg} ({Token.string!=". "}) [0, 5]
```

36 The grammars are almost identical; they only differ on the listing type which they
37 invoke (i.e. PreNeg or PostNeg). The rules translate as: match a span which expands 5
38 tokens before a glossary match and 5 tokens after a glossary match excluding full
39 stops (to prevent the rule expanding beyond the limits of a potential sentence). A
40 succeeding matching grammar was invoked for filtering out those phrases that did not
41 contain any of the four CRM entities failing within the scope of OPTIMA.

42 The second stage implemented a separate pipeline which post-processed the negation
43 phrases delivered by the first stage. The aim of this particular pipeline was to reveal
44 the most commonly occurring Noun and Verb phrases of the negated phrases output.
45 Such commonly appearing noun and verb phrases were then analysed to inform the
46 process of enhancement and adaptation of existing glossaries and negation grammars
47 to the context of archaeology.

48 In total, 29040 noun phrases and 14794 verb phrases were identified. From them
49 14686 were unique noun phrases and 2564 were unique verb phrases. Examining the
50 list of the most frequent noun phrases and comparing it with the list of the NegEx lists
51 it became apparent that some of the existing entries were not applicable to the
52 archaeology domain and returned no matches. Such entries are rather particular to the
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

medical domain, for example “suspicious”, “decline”, “deny” and “unremarkable for”. Moreover, frequently occurring negation classifiers of archaeological narrative, such as “unknown”, “unclear” and “undated” were not part of the initial NegEx lists. The adaptation exercise created new versions of the [PreNeg] and [PostNeg] lists adapted to the archaeology domain by removing the entries that are particular to the medical domain and by including new entries that are relevant to the archaeology domain.

The analysis of the verb phrases result revealed some very interesting vocabulary patterns. Examining the most commonly occurring verb phrases, a pattern emerged relating to use of passive voice utterances. For example the phrase “should not be considered” occurred 134 times, the phrase “was not excavated” 67 times, the phrase “were not encountered” 39 times, etc. Although, NegEx covered some cases of backward matching via the [Post-Neg] list for phrases where a negation classifier is found at the end of a phrase, the algorithm did not consider extensively the use of passive voice expressions apart from the case “being ruled out”.

The intellectual examination of the list of the frequently occurring verb phrases isolated a set of passive voice verbs that could be used to enhance the operation of the negation algorithm. The list of verbs constitutes a specialised vocabulary of 31 entries such as “appear”, “associate”, “compose”, “discover”, “encounter”, etc., which were composed under a new glossary listing named *Negation-verbs*. The glossary is used by the pattern matching rules discussed in the section below, for identifying negation in phrases, such as “*deposits were not encountered at the machined level*”.

An integral part of the NegEx algorithm is the [Pseudo-Negation] list which is responsible for limiting the scope of a match by identifying false negation triggers. Due to the elaborate and unrestricted report style of archaeological grey literature documents, it was decided to expand the negation window of the algorithm to larger phrases containing a maximum of 10 word tokens instead of 5 that are originally set by NegEx. Thus, the inclusion of the [Pseudo-Negation] operation seemed highly relevant for avoiding matches of positive entity assertions that adjoin with negation phrases.

The general principle of the [Pseudo-Negation] operation was adopted as a means to narrow the scope of a negation window. A new list [Stop-Neg] was created that contained 38 new entries originating from the empirical use of English when separating different clauses in a sentence. The lexical resource Wordnet (Miller, 1995) was employed in the construction of a list containing a range of entries such as “but”, “nonetheless”, “than”, “though” and relevant synonyms from the available synset hierarchies of the Wordnet.

The operation of the [Stop-Neg] list prevents matching beyond the scope of a negation phrase and does not exclude identification of conjunct entities. The original NegEx algorithm reported limitations on accurate matching of long lists of conjunct UMLS terms that expand beyond the word limit (5 tokens) of the negation window (Chapman et al., 2001). The OPTIMA pipeline is equipped with an Entity

Conjunction module which delivers matches of the same CRM entities conjunct with “and”, “or”, “commas” and other forms of hyphenation. Hence, the negation module can exploit conjunct matches, in order to deliver negation phrases that include a list of entities, as for example the phrase “*no evidence of archaeological features or deposits dating to the Neolithic or Bronze Ages*”. The inclusion of the [Stop-Neg] list in the negation grammars prevents the match of “*post-medieval spread*” as a negated case in the phrase “*absence of evidence after the Roman period, with the exception of the post-medieval spread*” while the first clause of the phrase is identified as a negated match.

3.2 The Negation Detection Module

The negation detection module of the OPTIMA pipeline incorporates the four glossary listings, [Pre-Neg], [Post-Neg], [Stop-Neg] and [Verb-Neg] with a set of information extraction grammars. A set of three different pattern matching rules is deployed for each of the four different CRM entity types that fall within the scope of the negation module. The arrangement of the negation rules avoids multiple annotation of the same phrase, even if more than one CRM entities are mentioned in a phrase. The description of grammars given below refers to a unified form of a CRM entity which encompasses all four different CRM types for simplicity.

The grammars deliver a single annotation span, which covers all CRM entities involved in a phrase. For example the phrase “*no evidence of Roman pottery*” delivers a single annotation spanning the whole phrase rather than two separate annotations for “*Roman*” and “*pottery*”. Similarly when conjunction of entities is present, the negation span covers all conjunct entities under a single annotation span

The following grammar is targeted at matching cases of negation which commence with a match from the [Pre-Neg] list and end in a CRM entity or a CRM conjunct entity match, for example “*absence of any datable small finds or artefacts*”.

```
{PreNeg}({Token, !StopNeg}) [0, 10] ({CRM} | {CRM_Conjunction})
```

The following grammar matches cases of negation which commence with a CRM entity or a CRM conjunct entity and end with a match from the [Post-Neg] list, for example “*wares such as tea bowl are particularly unlikely to exist*”.

```
({CRM} | {CRM_Conjunction}) ({Token, !StopNeg}) [0, 10] {PostNeg}
```

The following grammar is targeted at matching cases of negation which commence with a CRM entity or a CRM conjunct entity and end with a match from the [Verb-Neg] list, for example “*pottery and tile remains were not observed*”.

```
({CRM} | {CRM_Conjunction}) ({Token, !StopNeg}) [0, 10]
{Token.string=="not"} ({Token})? {VerbNeg}
```

3.3 Word Sense Disambiguation Module

The disambiguation module implements a set of contextual collocation templates expressed as hand-crafted (JAPE) information extraction rules. Contextual collocation refers to the location of ambiguous terms in relative location to non-ambiguous terms.

1
2
3
4
5
6
7
8
9 The module utilises three groups of annotation types: (i) non-ambiguous (single
10 sense) annotations of Physical Object, Time Appellation and Material, (ii) ambiguous
11 (Multisense) annotations of Physical Object and Material, and (iii) Token annotations
12 containing part of speech features.

13
14 The disambiguation module resolves the appropriate terminological (SKOS) reference
15 to ambiguous terms. For example when the term “brick” is disambiguated as material,
16 the terminological reference “97777” originating from the Main Building Material
17 thesaurus is assigned to the annotation. When the same term is resolved as physical
18 object, the terminological reference “96010” originating from the MDA Object Type
19 thesaurus is assigned instead.

20 *3.3.1 Annotation of Ambiguous Concepts*

21 The first phase of the WSD module is to annotate all ambiguous terms. Consider the
22 term “pottery” which is ambiguous and can refer either to a Physical Object or to a
23 Material. The term is annotated as ambiguous by the matching conditions
24 {Lookup.skosConcept ==“96010”}, where “96010” corresponds to the Physical
25 Object sense of pottery and {Lookup.skosConcept ==“97777”}, where “97777”
26 corresponds to its Material sense. This rule generates annotations of the type Mention,
27 any textual instance of an ambiguous term is assigned two annotations that have a
28 multisense property equal to “true” but one has type “Physical_Object” and the other
29 “Material” for maintaining a distinct terminological reference. The competing
30 Mention annotations are used as input by the disambiguation phase, which resolves
31 ambiguity and assigns the final sense, ontological and terminological reference, to a
32 textual instance.

33 *3.3.2 Rules for Resolving Ontological Polysemy*

34 The OPTIMA WSD module implements 15 different cases of contextual templates for
35 the automatic disambiguation of physical object and material instances in context.
36 The templates are grouped into three categories a) word pair cases b) concatenate
37 structures and c) syntactical patterns. The list of rules is not exhaustive but covers
38 common (empirically selected) lexical patterns that can be invoked by the
39 disambiguation process. A dual sense annotation (i.e. material and physical object) is
40 assigned to those entity cases that fail disambiguation. This particular choice favours
41 Recall rather than Precision resulting in a half-correct annotation of ambiguous terms
42 since only one of the two applied senses can be correct in a given context. On the
43 other hand, it ensures that annotations are not discarded due to their ambiguity but are
44 still revealed by the NER process.
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

3.3.3 Word Pair Cases

The word-pair rules define simple templates which examine the location of ambiguous terms in pair relation to other ambiguous and non-ambiguous terms¹. In total, three separate rules (grammars) of word pair cases contribute to the module. The rules follow a JAPE grammar expression which examines the annotation type of the pair entities. The following grammar targets cases of an ambiguous term followed by another ambiguous term e.g. “brick tile”. It resolves the left part of the pair as Material and the right part as Physical Object based on the empirical use of English where usually the material of an object is stated first.

```
{(Mention.type=="Material", Mention.multisense=="true"  
{Mention.type=="Physical_Object",Mention.multisense=="true"})}
```

Two additional variations of the above grammar target the cases of i) an ambiguous terms followed by a non-ambiguous term e.g. “pottery fragment” ii) a non-ambiguous terms followed by an ambiguous term e.g. “plaster tile”. The grammars always resolve the left part of the pair as Material and the right side as Physical object.

3.3.4 Concatenate Pattern Rules

The patterns of this category target cases of concatenation between ambiguous and non-ambiguous terms. The grammars are based on the assumption that co-ordinating concatenations join terms of the same kind. The concatenation operators employed by the grammars are comma “,”, forward slash “/”, the word “and” and the word “or”.

The following grammar resolves ambiguous terms as Materials when such terms concatenate with non-ambiguous terms of material sense e.g “brick and plaster” (“brick” is resolved as material due to plaster regarded as a non-ambiguous material term).

```
{(Mention.type=="Material",Mention.multisense=="true"):match  
{(Token.string=="and")|(Token.string=="or")|  
{Token.category=="",}|{Token.category=="/"})  
{E57}}
```

An inverted version of the above grammar addresses cases of a non-ambiguous term concatenating with an ambiguous term e.g. “plaster and brick”. Similarly, two additional versions of the above grammar addressing concatenation between ambiguous terms and non-ambiguous terms of physical object sense, for example “coin and brick” or “brick and coin”. The rules resolve the ambiguous term to a physical object sense.

¹ The ambiguous terms targeted by the disambiguation rules are described in JAPE grammar as *Mention* annotations, while the non-ambiguous terms are described with their CIDOC CRM annotation type i.e. E19 (Physical Object), E49 (Time Appellation), E53 (Place), E57 (Material)

3.3.5 Syntactical Pattern Rules

A range of elaborate syntactical patterns are employed by the module for addressing polysemy between physical object and material terms. The list of patterns is not exhaustive but is representative of the kind of templates that can be employed for tackling the particular form of polysemy addressed by the WSD module. The templates were derived empirically by examining archaeology grey literature documents and abstracting patterns from phrases which carry clues for disambiguation. The rules use input form ambiguous, non-ambiguous terms and tokens which are parameterised with part of speech features.

The disambiguation grammars generalise distinct syntactical patterns for resolving the polysemy behaviour of material and physical object terms. Based on the assumption that material descriptions precede noun descriptions, a set of rules exploits the part of speech input phrases for resolving the material sense to ambiguous terms. The following rule resolves the material sense in phrases, such as "...animal bone assemblage..." and "...brick pieces...", where "assemblage" and "pieces" are general noun descriptions.

```
{(Mention.type=="Material", Mention.multisense=="true"):match
({Token.category ==NN}|{Token.category ==NNS}|
{Token.category ==NNP}|{Token.category ==NNPS})
```

A second set of rules exploits particular syntactical clues, such as a determiner preceding an ambiguous term or use of the "made of" clause. The following rule resolve the physical object sense in the phrase "...artefacts made of wood..."

```
{(Mention.type=="Physical_Object",
Mention.multisense=="true")|{E19})
({Token.string=="of"}|({Token.string=="made"}
{Token.string=="of"}))
{Mention.type=="Material", Mention.multisense=="true"}|
{E57}}
```

A third set of rules, which is more particular to the domain of archaeology, examines the location of temporal descriptions in phrases such as "...sherds of Iron Age pottery..." and "...6th century pottery, at Puddlehill ..." for resolving the appropriate material or physical object sense to ambiguous terms. In the first phrase "pottery" is resolved as material while in the second is resolved as physical object.

```
{(Mention.type=="Physical_Object",
Mention.multisense=="true")|{E19})
({Token.category == IN}{E49})
({Mention.type=="Material", Mention.multisense=="true")|{E57})

({E49})
({Mention.type=="Physical_Object",
Mention.multisense=="true"):match
({Token.category !=NN}|{Token.category !=NNS}|
{Token.category !=NNP}|{Token.category !=NNPS})
```

4 Evaluation

For the purposes of evaluation a dedicated task addressed the performance of the Negation Detection module whilst the performance of the WSD has been addressed by the main evaluation task of the OPTIMA pipeline as discussed in Vlachidis (2012).

Typically the performance of Information Extraction systems is measured in Recall, Precision and F-Measure scores as established by the second Machine Understanding Conference, MUC 2 (Grishman and Sundheim 1996). The F-Measure score is the harmonious mean of Precision and Recall used to provide a comprehensive view of system's performance. Attempts to improve Recall will usually cause Precision to drop and vice versa. High scoring of F-Measure is desirable since it can be used to benchmark the overall system's accuracy (Maynard et al., 2006).

Partial matches are those having different annotation boundaries than the correct matches, either matching only a part or expanding beyond the limits of a correct match. Partial matches can be weighted with decimal values ranging from 0 to 1 depending on the importance of such matches in the system's accuracy. The set of the correct responses participating in system's evaluation were delivered by the method of manual annotation also known as "Gold Standard" definition, which is typically employed for comparison against system produced annotations.

4.1 Evaluation of the Negation Detection Module

The evaluation pursued the definition of a gold standard corpus tailored to serve the purposes of the evaluation task of the Negation Detection Modules. Such manual definitions are usually built by domain experts but their availability is often scarce. In the case of archaeological reports, there was no available gold standard of semantically annotated documents with respect to negated CIDOC-CRM entities.

The evaluation task treated partial matches of negated phrases as fully correct matches based on the flexible user-centred approach followed during the definition of the correct responses, which delivered negated phrases that were syntactically complete from a user's point of view. For example, an correct response might be "*No traces of a Roman settlement have been discovered in the area*". However, the negation algorithm is programmed to extract phrases that commence or end with a CRM entity, in this case only the first part is extracted ("*No traces of a Roman settlement*") delivering a partial match. Hence, the match can be treated as fully correct since the correct response is not defined with the algorithm in mind but with what is useful from an end-user point of view.

In total 10 grey literature documents of archaeological excavation and evaluation reports contributed to the gold standard definition. In archaeology, grey literature reports reflect the different stages of a fieldwork project worth recording and disseminating information about. They contain comprehensive explanations, diagrams, summaries and statistics that deliver in depth analysis and discussion

usually not possible to be accommodated by traditional publication. The evaluation corpus contained a set of archaeological excavation and evaluation reports, which typically contain rich discussion about the findings and excavation phases over other types of archaeological reports, such as watching briefs and observation reports. In addition, the selection process included reports from a range of different UK archaeological units aiming to cover different reporting styles and practices. The gold standard overall consisted from 300 pages which contained 144 cases of negation.

4.1.1 Results

Among the 10 documents that participated in the evaluation task, the negation detection module delivered an overall Recall score 80%, Precision 89% and F-Measure 83% (table 1). The Recall score of individual documents presents a fluctuation ranging from 50% to 100% while fluctuation of Precision scores is smaller ranging from 64% to 100%. The good precision performance of the module is also reflected by the standard deviation score which is 0.11(or 11%) with only one document scoring under 80%. On the other hand, the standard deviation of Recall scores is slightly higher 0.15 (or 15%) with half of the documents scoring under 80%.

Table 1. Performance of Negation Detection Module.

Document	Recall	Precision	F-Measure
Aocarcha1-11167	0.74	0.94	0.83
Birmingh2-28160	0.77	1.00	0.87
Essexcou1-10460	0.83	1.00	0.91
Essexcou1-5166	0.76	0.85	0.80
Foundati1-5205	0.87	1.00	0.93
Heritage1-10767	0.50	1.00	0.67
Heritage1-11948	1.00	0.83	0.91
Suffolkc-6115	0.85	0.89	0.87
Wessexar1-25626	0.70	0.64	0.67
Wessexar1-5680	1.00	0.83	0.91
Average	0.80	0.89	0.83

The negation detection module has overall delivered 114 correct and partially correct matches, 14 false positive (falsely identified) matches, while it missed 30 negation answers of the gold standard definition. The number of total false positive matches is half of those matches being missed. This significant difference between the two is also reflected in the Precision and Recall scores where missed matches directly affect recall and false positives precision. Overall, the negation module delivers better precision than recall, indicative of the module's capacity to accurately identify cases of negation while being challenged by the variety in which negation can be expressed in natural language.

4.2 Evaluation of the Word Sense Disambiguation Module

The main evaluation phase of the OPTIMA pipeline was based on an iterative process of Gold Standard definition via a pilot evaluation. The manual annotation instructions

were written to reflect the end-user aims of the evaluation (supporting retrieval and research of archaeological reports), hiding complex and sometimes unnecessary ontological details. Annotators were instructed to annotate at the level of archaeological concepts rather than identifying more abstract ontological entities in context. The instructions directed the task of manual annotation at the concepts of archaeological place, archaeological find, material of archaeological finds and time appellation, thus annotating textual instances that have a value from an archaeological point of view.

The particular task of the main evaluation of OPTIMA aimed to evaluate the contribution of the various NLP modules contributing to the NER phase. The pipeline was stripped of all modules that were used by the NER pipeline to improve accuracy of performance, such as the WSD module. A basic configuration (Basic) was executed and the results were used as indicator of the system performance, without the use of accuracy techniques. The contribution of the WSD module was then evaluated by adding the module to the Basic configuration and comparing the results.

The Basic configuration Recall rate is 89%, however, the Precision score is low (55%) which affects the overall F-measure score (65%) of the OPTIMA pipeline. The WSD module has a positive contribution increasing the Precision of the pipeline by 6%. When all NER supporting modules are combined in the pipeline the overall precision score improves by 23%.

The figure 1 below presents a document section in which the disambiguation of material-physical object sense has been achieved for the concepts of “brick” and “pottery”.

been related to the external red brick walls on the same alignment. It measured 0.80m+ by 0.30m by 0.15m high. Associated with this wall and external wall [1/008], [6/004] was a brick floor [1/028]. The brick from the floor was dated to the late 18th century. In this cluster accounts for the vast majority of finds recovered from the entire site. The cluster consists of seventeen pits, all of which had very similar form and characteristics. The pits were uniformly similar, with dense quantities of charcoal containing hazel nutshell, burnt bone fragments and Beaker pottery sherds. This suggests that the pits were all open and filled simultaneously from a common source. Eleven of these seventeen pits contained Early Bronze Age pottery. Fifteen also contained worked flint that appears to be contemporary with the pottery. A further five

Figure 2. Word Sense Disambiguation examples in the context of archaeological excavation reports.

The above example demonstrates the fact that the terms “brick” and “pottery” can have a material or a physical object sense depending on contextual arrangements. Thus, in the cases “red brick wall” and “brick wall” the sense is material, while in the case “The brick from the floor was dated to the late 18th century”, the sense is physical object. Similarly, in the case “Beaker pottery sherds”, “pottery” has a material sense while in the case “pits contained Early Bronze Age pottery”, the sense

1
2
3
4
5
6
7
8
9 is physical object. Different senses (i.e. CRM entities) are highlighted in different
10 colours; orange is used for physical object and purple for materials.
11

12 **5 Discussion**

13
14 The evaluation results revealed an encouraging performance of the negation detection
15 module which delivered Recall and Precision scores over 80%. Although, the
16 evaluation task had a limited scope and was based on the use of a small scale gold
17 standard definition, it suggests that negation in archaeological text can be addressed
18 with information extraction techniques that use a small set of domain oriented pattern-
19 matching rules. Our results agree with research findings from biomedical text
20 negation (Chapman et al., 2001) reporting that negation phrases typically comply with
21 the Zipf's law regarding the frequency distribution of words in human languages,
22 where a few very common negation patterns can capture a large portion of pertinent
23 negation cases.

24 According to the evaluation results, use of frequently occurring negation patterns in
25 extraction rules supports the system's precision. The vast majority of automatically
26 identified negation phrases (approximately 9 out of 10) delivered by the negation
27 module were correct. Incorrect cases (false positives) are primarily the result of
28 limitation in the vocabulary used to support the operation of extraction patterns and
29 not due to the incapacity of extraction patterns themselves. For example the phrase "*It*
30 *is not unusual to find solitary prehistoric cremations*" has a positive meaning which is
31 falsely identified as a negation case.

32 The OPTIMA algorithm, similarly to the NegEx algorithm (Chapman et al., 2001),
33 employs a specialised vocabulary [Stop-Neg], which limits the scope of negation. The
34 original NegEx glossary of pseudo-negation phrases is enhanced with additional
35 terms (drawn from ordinary use of English), through a WordNet Synset expansion
36 technique (section 2.2.1) to include a range of entries such as, "but", "nonetheless",
37 "though" and their synonyms. The expanded glossary failed to address fully all the
38 cases of double negatives as for example "not unusual", which has a fairly positive
39 assertion. Possibly use of double negatives is avoided in the restricted context of
40 narrative reports of medical records but in the context of archaeological reports such
41 double negated narratives may occur. Fewer false positive cases relate to the
42 operation of matching patterns as for example the phrase "*non-intrusive survey had*
43 *accurately predicted the ridge and furrow*" where "non" applies only to the
44 immediate noun that follows.

45 The recall performance of the negation module is reasonable (approximately 8 out
46 of 10) but not as high as precision. The capacity of the algorithm to identify all correct
47 cases of negation in text is challenged by the sometimes creative and indirect writing
48 style of archaeological reports. For example the phrase "*The low quantity and quality*
49 *of the remains encountered on the site suggests that there is only a minor*
50 *archaeological implication*" clearly suggests that findings do not have an
51 archaeological interest. However, it is formulated in an indirect style, which does not
52 invoke any negation triggers that could be matched by the module.
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Other cases of missed examples concern use of passive voice utterances that do not employ clear negation classifiers but verbs which are loaded with negative sense. For example the phrase “*both these deposits were largely absent*” is missed due to the definition of matching patterns that expect a negation classifier at the beginning of a phrase for example “*Absence of deposits*”, or a negation formation at the end of a phrase for example “*deposits were not largely present*”. Fewer examples of non-identified cases concern limitations of the NER vocabulary itself. For example the phrase “*there was virtually no artefactual evidence recovered*” is missed because “*artefactual evidence*” is not recognised as a CRM entity.

5.1 Method Limitations

The development approach of constructing a negation detection mechanism based on shallow parsing delivered results that suite the aims of semantic indexing. Shallow parsing analyses a sentence or a phrase to its constituent parts but without analysing their role or their internal structure. The negation detection module has managed to identify with reasonable success a vast range of phrases containing one or more CRM entities which were negated via a lexical classifier. Upon successful identification of a negation phrase, all CRM entities of the phrase were discarded from further indexing. This approach did not harm the quality of the indexing due to the vast amount of CRM entities being delivered by the NER phase. On the other hand, this particular approach might be considered as a blanket practice that does not support the aims of a detailed and meticulous text mining effort.

Looking closer at the following example “*No artefacts were retrieved from this deposit*” it is clear that there is absence of artefacts. However, the same absence does not apply to the deposit itself which does exist but under the current configuration is excluded from indexing. Adoption of deep parsing techniques, which analyse the role and structure of the constitute parts of a sentence, could be sufficient to address such cases of detailed negation assignment on the level of subject clause. However, the semantic annotation of such negated cases with respect to ontology classes may prove a challenging task as discussed below.

Consider the above example “*No artefacts were retrieved from this deposit*”. Assignment of the *E19.Physical_Object* class to “*artefacts*” instance assumes a positive assertion. Similarly an ontological model may define a relationship property between place and physical object. Again this kind of property assumes a positive assertion which does not cover the cases where an object is not in place.

A specific project which addressed the issue of factual argumentation using the CIDOC-CRM ontology is the Integrated Argumentation Model (IAM) (Doerr et al. 2011). Although, factual argumentation is a broader epistemological issue that concerns falsification or verification of arguments, the aim of IAM to connect such epistemological aspects with instances of a formal ontology could be potentially useful and applicable to the semantics of negation assertion. The project presented benefits to archaeological reasoning for a particular case (the natural mummy Oetzi) but its applicability in the context of semantic annotation of archaeological text remains untested.

1
2
3
4
5
6
7
8
9
10 Providing a semantic annotation i.e. assigning classes or properties, to textual
11 instances that are negated is not always viable within the scope of an ontology that
12 assumes only positive assignment. A possible answer to this limitation might be
13 addressed with the introduction of a property for declaring the sense of an instance for
14 example *has_sense*, positive or negative. However, introduction of such property will
15 significantly increase the chain of triples defined by a SPARQL query on an
16 application level even for the simplest queries.

17
18 An alternative approach could be the introduction of negative print of all ontological
19 classes and properties in order to accommodate negative assertions. Thus, every class
20 or property of an ontological model would have its equivalent contrasting class in the
21 sense of “matter, anti-matter”. This particular approach though, would double the size
22 of an ontological model which could lead to issues relating to the maintenance and
23 version control of an ontological model.

24 **6 Conclusions**

25 The paper presented the results of a negation detection module targeted at identifying
26 negated cases of four CIDOC-CRM entities in the context of semantic indexing of
27 archaeological grey literature for information retrieval. The evaluation results
28 demonstrate the capacity of rule-based information extraction techniques to accurately
29 detect a large portion of negation phrases. The employment of three small scale
30 glossaries that support the operation of a few simple pattern matching expressions has
31 proved sufficient to deliver high Recall (80%) and Precision (89%) scores. Current
32 limitations of the method relate to the employment of shallow parsing techniques that
33 do not support deeper analysis of negation phrases, the capacity of glossaries to cover
34 all possible vocabulary variations, and the adequacy of pattern matching rules to
35 address every single case of negation which can be expressed by a creative and
36 sometimes indirect writing style of archaeological reports. However, such limitations
37 do not restrict application of the work in the context of semantic indexing. Future
38 steps include a large scale corpus analysis and evaluation study aimed at expanding
39 glossary coverage and improving the system’s performance with regards to
40 archaeological negation narrative. Longer term aims may involve the system’s
41 generalisation to the broader field of digital humanities and application of semantic
42 modelling solutions capable of addressing negation at an information retrieval level.

43 The paper has revealed a method for adapting the NegEx algorithm to the domain of
44 archaeological grey literature. Our experience has shown that porting of NegEx to a
45 new domain is feasible. The method of modification of the original algorithm was
46 driven by a corpus analysis task, which enabled enhancement and adaptation of the
47 original resources to the new domain. This particular method has given promising
48 results for the domain of archaeological reports though its applicability to other
49 domain remains to be tested. The issue of accommodating negative assertions by the
50 current ontological modelling approaches was also highlighted. Negated findings and
51 facts are important for the research and information retrieval in particular domains,
52 such as the medical and the archaeology domain. Semantic technologies can provide a
53
54
55
56
57
58
59
60

valuable support in modelling and retrieval of such negated findings for enabling new forms of research and information exchange.

Acknowledgments

While we wish to acknowledge support from the ARIADNE project (FP7-INFRASTRUCTURES-2012-1-313193), the views expressed are those of the authors and do not necessarily reflect the views of the European Commission.

References

- Bizer, C., Heath, T. and Berners-Lee, T. (2009), "Linked Data - The Story So Far". *International Journal on Semantic Web and Information Systems* 5(3), 1--22
- Blanco, E. and Moldovan, D., (2011) "Semantic representation of negation using focus detection". In: *49th Annual Meeting of the Association for Computational Linguistics*, Portland
- Chapman, W.W., Bridewell, W., Hanbury, P., Cooper, G.F. and Buchanan B.G. (2001), "A Simple Algorithm for Identifying Negated Findings and Diseases in Discharge Summaries". *Journal of Biomedical Informatics*. 34(5), 301--310
- Crofts, N., Doerr, M., Gill, T., Stead, S. and Stiff, M. (2009), "Definition of the CIDOC Conceptual Reference Model", http://cidoc.ics.forth.gr/docs/cidoc_crm_version_5.0.1_Mar09.pdf
- Doerr, M., Kritsotaki, A., Boutsika, K. (2011), "Factual argumentation-a core model for assertions making". *Journal on Computing and Cultural Heritage*. 3(3), 1—34
- Falkingham, G. (2005), "A whiter shade of grey: a new approach to archaeological grey literature using the XML version of the TEI guidelines". *Internet Archaeology*. 17, (2005), http://intarch.ac.uk/journal/issue17/falkingham_index.html
- Horn L.R. (1989), "A Natural History of Negation". University of Chicago Press, Chicago
- Goryachev, S., Sordo, M., Zeng, Q.T. and Ngo, L. (2006) "Implementation and evaluation of four different methods of negation detection". Technical report. Harvard Medical School
- Grishman, R. and Sundheim, B. (1996), "Message Understanding Conference-6: a brief history". In *16th International Conference on Computational Linguistics*, Copenhagen
- Maynard, D., Peters, W. and Li, Y. (2006), "Metrics for Evaluation of Ontology-based Information Extraction". In *Proceedings WWW Conference 2006, Workshop on "Evaluation of Ontologies for the Web*, Edinburgh.

- 1
2
3
4
5
6
7
8
9
10 Maynard, D., and Funk, A. (2011), "Automatic detection of political opinions in tweets". In:
11 *Making Sense of Microposts Workshop at ESWC 2011*. Heraklion,
- 12 Miller, G.A.(1995), "WordNet: A Lexical Database for English". *Communications of the ACM*,
13 38(11), 39--41
- 14 Nadeau, D. and Sekine, S. (2007), "A survey of named entity recognition and classification".
15 *Linguisticae Investigationes*, 30(1), 3--26
- 16 Navigli R. (2009) 'Word sense disambiguation: A survey', *ACM Computing Surveys*, 41(2),
17 pp.10--11
- 18 O'Hara, K., Berners-Lee, T., Hall, W. and Shadbolt, N. (2010), "Use of the Semantic Web in e-
19 Research. The MIT Press", Cambridge, MA
- 20 Resnik P. (1997) "Selectional preferences and sense disambiguation", In *Proceedings ACL*
21 *SIGLEX Workshop on Tagging Text with Lexical Semantics*, Columbia, Washington, D.C, 4-5
22 April.
- 23 Ruch, P., Baud, R., Geissbuhler, A. and Rassinoux, A.M (2001) "Comparing general and
24 medical texts for information retrieval based on natural language processing: an inquiry into
25 lexical disambiguation". *Med-info* 10, 261--265 (2001)
- 26 Tudhope, D., May, K., Binding, C., Vlachidis, A. (2011) "Connecting archaeological data and
27 grey literature via semantic cross search". *Internet Archaeology*, 30 (2011),
28 http://intarch.ac.uk/journal/issue30/tudhope_index.html.
- 29 Uren, V., Cimiano, P., Iria, J., Handschuh, S., Vargas-Vera, M., Motta, E. and Ciravegna, F.
30 (2006), "Semantic annotation for knowledge management: Requirements and a survey of the
31 state of the art". *Web Semantics: Science, Services and Agents on the WWW*, 4(1), 14--28
- 32 Vlachidis, A. and Tudhope, D. (2012). A pilot investigation of information extraction in the
33 semantic annotation of archaeological reports. *International Journal of Metadata, Semantics*
34 *and Ontologies*. 7(3), 222--235 (2012)
- 35 Vlachidis, A. (2012), "Semantic Indexing Via Knowledge Organisation Resources: Applying
36 the CIDOC-CRM to Archaeological". *PhD Thesis*, University of Glamorgan, UK
- 37 Voorhees E. (1993) "Using WordNet to Disambiguate Word Senses for Text Retrieval", In
38 *Proceedings of the Sixteenth Annual International ACM SIGIR Conference on Research and*
39 *Development in Information Retrieval*, pp.171-180. Pittsburgh, 27 June - 01 July.
- 40 Westbury, C. (2000), "Just say no: The evolutionary and developmental significance of
41 negation in behavior and natural language". In: *3rd Conference The Evolution of Language*,
42 Paris
- 43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60