

Formalization and Reuse of Methodological Knowledge on Archaeology across European Organizations

Cesar Gonzalez, Patricia Martín-Rodilla, Elena Viorica Epure

► **To cite this version:**

Cesar Gonzalez, Patricia Martín-Rodilla, Elena Viorica Epure. Formalization and Reuse of Methodological Knowledge on Archaeology across European Organizations. 44th Computer Applications and Quantitative Methods in Archaeology Conference, Mar 2016, Oslo, Norway. 2016. <hal-01425579>

HAL Id: hal-01425579

<https://hal-paris1.archives-ouvertes.fr/hal-01425579>

Submitted on 3 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Formalization and Reuse of Methodological Knowledge on Archaeology across European Organizations

Cesar Gonzalez-Perez. Incipit CSIC. cesar.gonzalez-perez@incipit.csic.es

Patricia Martín-Rodilla. Incipit CSIC. patricia.martin-rodilla@incipit.csic.es

Elena Viorica Epure. Université Paris 1. Elena.Epure@malix.univ-paris1.fr

Abstract

Archaeological projects vary greatly in size, object of study, timescale and other aspects. Finding the most suitable methodology is often difficult, and an inadequate choice can ruin many months' worth of fieldwork and interpretation. An archaeological methodology should be as adjusted as possible to the project needs, consider knowledge that was successfully applied in the past, and be clearly expressed for better understanding and sharing. These goals are usually pursued informally through the application of tacit knowledge within archaeology organizations, leading to situations where it is difficult to convey what is expected to be done, methodological knowledge is underutilized and rarely reused, and the improvement of methodologies over time is difficult since no explicit knowledge about them exists. Some experiences exist with regard to using situational method engineering (SME) to mitigate these problems. In this paper we present the results of applying SME to the informal methodological knowledge of seven European archaeological organizations. Natural language processing techniques have been used to assist in this process. This work allowed us to obtain variations of established methodologies to cater for different project situations, combine different methodologies for collaborations and other hybrid scenarios, and reason about the methodological choices of different organizations.

Keywords

Methodology. Knowledge reuse. Archaeological practice. Situational method engineering. Natural language processing.

Introduction and Motivation

Our work is motivated by realising three distinct but related facts about organisations working in archaeology. First, it is often difficult to convey what specialists are expected to do in specific situations, especially in relation to new team members or external collaborators. Guiding a large team in a complex project, especially when specialists from other disciplines (such as anthropologists or historians) are involved, is extremely time consuming and prone to errors. Secondly, methodological knowledge is underutilised and rarely reused, especially across organisations. In fact, many archaeological projects reinvent the wheel in terms of methodology, missing the opportunity to take advantage of previous experiences and existing expertise. Third, it is difficult to improve methodologies over time, since little explicit knowledge about

them exists. This means that maladjustments and inadequacies in practices are rarely documented so that the next project can benefit from the acquired knowledge.

In this context, we need to define what we understand by “methodology”. Informally, a methodology describes how people do what they do. More specifically, a methodology is a systematic description of the processes that are followed, the products that are used and created, and the tools that are employed by a group of people to achieve a certain goal (Gonzalez-Perez and Hug, 2013). Long and tedious debates have taken place over the meaning of the term “method” in relation to methodologies. We acknowledge that usage varies from discipline to discipline but, for practical purposes, and in agreement with ISO/IEC 24744 (ISO/IEC, 2014), we will consider these two terms to be synonymous within this paper.

An archaeological methodology, to start with, must address the process to follow. This is often expressed in terms of what tasks are carried out, what particular techniques are employed to achieve them, and how they are grouped into meaningful processes. For example, a task could be “Determine extent of underground features”, possible techniques to accomplish this could be “Ditch surveying” or “Ground-penetrating radar”, and this technique may be grouped with others into the process “Site Analysis”. These are just examples and we do not intend, at this stage, to propose a specific methodology organisation.

In addition, a methodology needs to describe what products are relevant. Process is necessarily performed on something, and for this reason the products that are created, modified or otherwise used by a methodology must be described. This is often done in terms of physical artefacts, documents, models, and other kinds of products. Some of them may be initially available and work as “inputs” to the methodology, such as the physical evidences on which archaeology usually relies. Some others are final or “output” products of the methodology, thus embodying its ultimate purpose, such as the new generated knowledge or a particular report or consolidated artefact. Finally, some products are interim artefacts that do not pertain to the final results or the initial inputs, such as many datasets, photographs and documents. Products and process are intimately related, and a methodology must describe what tasks create, modify, or otherwise use what products, in order to obtain a consistent view of the dependencies that exist between process elements and products.

Finally, a methodology must describe the people, tools, teams and other “producers”, which are in charge of carrying out the process in order to create or use the products. Producers are often expressed in terms of the roles that people may play, in order to decouple the methodology from specific individuals, such as “GIS Specialist” or “Excavation Director”. Other producers are documented in terms of tools, such as “Total Station” or “Database Management Software”, or teams, such as “Fieldwork Team” or “Lab Team”.

Describing an archaeological methodology as outlined above constitutes a very analytical endeavour, and results in a collection of individual but inter-related “atoms” that encapsulate methodological knowledge, such as how to perform a task or what a particular document kind consists of. This componentised information can be stored in a repository and made available to its users in a systematic manner, in order to assist with solving the issues described above. This is, precisely, what Situational Method

Engineering (SME) has been proposing for other disciplines and, more recently, for archaeology too. The next section describes SME in depth.

Situational Method Engineering in Archaeology

The comprehensive study and analysis of methodologies, regardless of their field of application, have been approached for the last couple of decades by the field of situational method engineering (SME). Although SME was born as a discipline inside software engineering (Kumar and Welke, 1992; Rolland and Prakash, 1996), it has since been applied to a variety of fields, such as business process modelling (Gonzalez-Perez and Henderson-Sellers, 2010), archaeology (Gonzalez-Perez and Hug, 2013) and other areas of the humanities (Hug et al., 2011). The fact that it is the word “method” (rather than “methodology”) that appears in “situational method engineering” obeys to historical reasons and, as we said in the previous section, we will consider “method” and “methodology” to be synonyms.

SME acknowledges that each methodology needs to be specifically situated, or adjusted to the project or endeavour to which it is going to be applied. At the same time, it tries to avoid circumstances that involve reinventing the wheel every time, by providing a solid knowledge reuse framework, so that methodologies are never created from scratch. In particular, and from the perspective of SME, a methodology is not a monolithic entity, but an assembly of *method components* that are carefully connected together after being selected from a pre-existing repository. Once a methodology has been created by assembling selected components, it can be enacted on an endeavour (i.e. applied to a project or other activity). During enactment, the performance of each component can be assessed, and the result of this evaluation fed back into the repository in the form of improvements to the components stored there. This way, methodologies that are assembled in the future from the improved components will take advantage from the accumulated enhancements that occur over time, thanks to the ongoing feedback loop. Figure 1 shows an overview of SME.

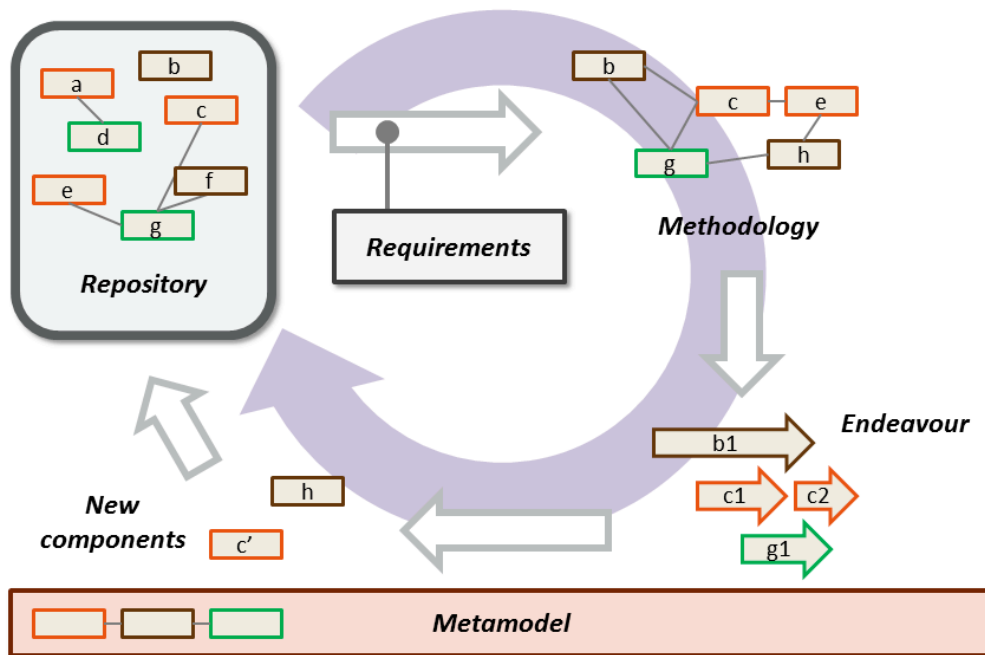


Figure 1. Overview of the dynamics of Situational Method Engineering (SME). The three major processes involved are depicted as boxes. Method components are depicted as small rectangles. The continuous improvement loop is depicted as a light circular arrow in the background.

There are some aspects that must be clarified. Firstly, and as described in (Gonzalez-Perez and Henderson-Sellers, 2008a; Gonzalez-Perez and Hug, 2013), method components are reusable, atomic, self-contained packages of methodological knowledge, i.e. knowledge that is related to how things should be done, what artefacts are involved in doing them, who should do them, or similar methodological aspects. In this sense, method components encapsulate good practices, often distilled from accumulated experience and past errors. Different colours in Figure 1 are meant to depict different kinds of method components. Additional details on this are given below.

Secondly, and even though some of the method components in the repository change little over time, the specific ways in which method components are combined in order to make up methodologies are highly diverse, as are the ways in which said methodologies can be later enacted on specific endeavours. Method construction and method enactment, therefore, rely heavily on methodological requirements, which are often described in terms of what outcomes (such as documents, theoretical models or physical objects) the endeavour is aiming to achieve, the conditions of the environment where the endeavour will take place (such as any time or resource constraints that may exist), and the sociotechnical characteristics of the organisational environment (such as team management style or even personal preferences). Furthermore, by using SME, methodologies do not need to be “frozen” and static, but can be altered and adjusted, even during the course of a project, to cater for

unexpected needs or respond to new information. This is accomplished through the incorporation or elimination of selected method components.

Thirdly, what types of method component are considered, and the ways in which method components can be assembled and enacted are regulated by a formalism called a *metamodel*. The metamodel acts as the grammar that dictates what kinds of combinations of method components are permissible, avoiding meaningless arrangements. The metamodel is not depicted in Figure 1, but it can be thought of as a set of operating rules that permeate everything that one does in SME, along the lines of the grammar rules that underpin a natural language when we use it to talk or write. Thus, having a solid and expressive metamodel is crucial for a successful application of SME. We have adopted the ISO/IEC 24744 (ISO/IEC, 2014) standard metamodel to this purpose. ISO/IEC 24744 implements, among others, the concepts that we describe above in the Introduction in the areas of process (through the WorkUnit class), products (through the WorkProduct class) and people (through the Producer class).

Finally, methodologies composed through SME should, ideally, be assessed for performance at solving the problems it is designed for, and taking into account the relevant methodological requirements. This is a quality-control issue that should be investigated from an empirical point of view and which, unfortunately, lies beyond the scope of this paper.

SME has been tentatively applied to archaeology and related areas in very few occasions. (Hug et al., 2011; Gonzalez-Perez and Hug, 2013) are probably the only published attempts. The *Advanced Research Infrastructure for Archaeological Dataset Networking in Europe* (ARIADNE) project (<http://www.ariadne-infrastructure.eu/>), in addition, has adopted SME as one of the underlying technologies for methodological analysis and study.

Formalizing Methodological Knowledge

ARIADNE established the goal, among others, to understand how different technology-mediated archaeological practices are used across various European organisations. To this purpose, SME and the ISO/IEC 24744 metamodel were adopted. However, ISO/IEC 24744 has been designed for engineering-oriented methodologies, and presented a number of minor shortcomings when applied to archaeology. To solve this, the standard metamodel was extended by using the built-in extension mechanisms. Then, descriptions in natural language were sought from project partners about their usual archaeological practices. The received documents were analysed by hand and through natural language processing (NLP) techniques, thus obtaining the raw input for the SME repository. At the same time, an ISO/IEC 24744-compliant database was constructed to work as repository, and the obtained information stored into it. The following sections describe this work with more detail.

Metamodel Extension

Our work uses an extended variant of ISO/IEC 24744. This extension involves three major innovations that are not part of the original standard. Firstly, the standard provides specific product kinds to describe documents and models, but this was not enough for archaeology. Two new classes were added under WorkProduct:

CognitiveElement, to describe hypotheses, plans and other abstract ideas; and PhysicalObject, to describe physical objects such as a rock, a building or a piece of coal.

Secondly, two Boolean attributes were added to WorkProductKind in the standard, in order to describe the different availability situations of products. The IsExternallyAvailable attribute is used to document whether a product is readily available from external sources during a project (e.g. a public monuments registry), and the IsInternallyAvailable attribute indicates whether a product is internally available in a project (e.g. the team's accumulated experience).

Thirdly, we found out that methodologies are sometimes enacted in archaeology in a non-deterministic manner, so that the steps to take and the products to involve are not pre-established at the beginning of the project, but are defined "on the fly" depending on interim project events. For example, unearthing an unexpected and large feature may mean that the excavation area is redefined and new staff is hired to cope with the extra work. In order to document situations like this, an association Implies was added from WorkProduct to WorkUnit, representing the fact that any particular work product (such as a redesigned work plan or altered hypothesis) that occurs during a project may imply a number of associated work units (such as new tasks or changed techniques).

Methodology Reports

In order to gather information about archaeological practices, ARIADNE project partners were asked to document their usual practices in a concise and informal manner, by using plain English plus tables or figures. They were also asked to focus on the process that they followed (what tasks, activities or techniques they performed), the products that they engage (i.e. what documents, models, artefacts and other relevant things they used or created), and the people in charge of the former (i.e. what teams, roles or tools were employed). Eight partners provided reports describing archaeological practices related to 2D and 3D documentation of features and landscapes, site location analysis, recording during surveying and excavation, stratigraphic analysis, management and treatment of finds, analysis of various kinds of finds (stone, ceramic, wood, charcoal, phytolith, carpological, human anthropological, and archaeozoological), strontium and oxygen isotope analysis, archaeological impact management, and publication of archaeological results.

The reports were analysed, selected and processed by hand and also through NLP techniques.

Natural Language Processing

Given the potentially large number of methodology reports to be analysed, and the large amount of time that this consumes, we decided to apply some natural language processing (NLP) techniques to aid with report analysis and method component extraction. Specifically, we selected unsupervised text mining with semantic analysis and no machine learning. A tool was developed in Python and using the Natural Language Toolkit (NLTK). The tool's final goal was to assist in the identification of methodological information from archaeological sources, so that method components could be easily constructed to be fed into the SME repository database. Previous collaborations with archaeologists (Epure et al., 2015) and online resources (Council

for British Archaeology, 2007) had revealed that the archaeological projects are driven by or operationalized through processes. An example is the excavation process, which consists of established activities, executed chronologically and satisfying various constraints such as two or more activities could take place in parallel, one activity depends on another, etc. All these processes followed during archaeological projects are explained in detail in textual reports, namely under a Methodology section. Consequently, these specific report sections become textual process traces.

Process Mining is the discipline emerging from Business and Information Systems whose goal is to discover processes from database logs (Aalst, 2011). The logs are structured traces of the human interaction with the software application. In our work (Epure et al., 2015), an alternative to the traditional view of process mining is proposed. Specifically, our solution handles unstructured, textual traces and was so far applied to archaeology, a humanities domain where the process participants do not necessarily interact with the software for performing their activities.

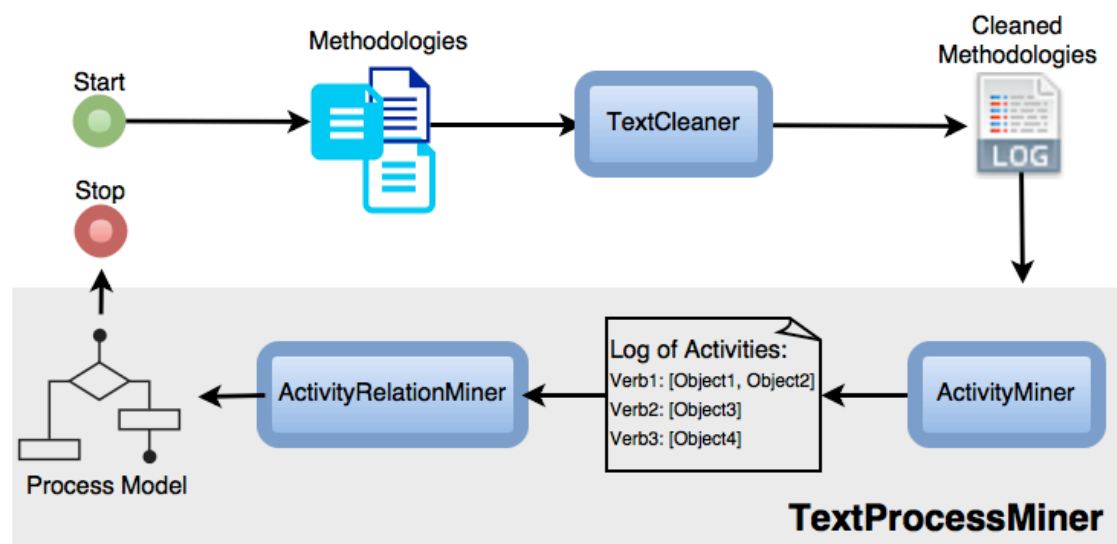


Figure 2. Overview of process mining as carried out for this work.

Our solution, called TextProcessMiner, is presented in Figure 2. It consists of two sub-components: 1) ActivityMiner, which is responsible for the unsupervised discovery of process activities from text; 2) ActivityRelationMiner, which is responsible for the unsupervised discovery of the relations among activities from text, namely sequence, parallelism, mutual exclusion and decisions.

The input to TextProcessMiner is the cleaned methodology text previously produced by TextCleaner. No special requirements exist on this text, apart from it being standard English. During this previous phase, the Methodology section is extracted from the document, its text is split in sentences, and several cleaning actions are performed: a space is introduced before and after punctuation signs, the comments in parentheses are removed if they do not contain verbs, and the sentences with negations are removed. These cleaning actions are necessary in order to standardize the textual format of the archaeological reports from different institutions, and all of them are easily automated.

Further, the first component ActivityMiner discovers the activities from text. An activity is considered a pair formed by a verb plus its objects, such as “use

toothless_bucket”. The property of verbs to take direct objects is called transitivity. Thus, during this phase, the transitive verbs are discovered and their associated objects are identified. This is achieved by using natural processing algorithms for syntactically and semantically tagging the sentences, generating treebanks (Marcus et al., 1993) and word lemmatization. Moreover, a knowledge base containing information about verbs and their transitivity was compiled from two external sources: VerbNet (Snyder and Palmer, 2004) and WordNet (Miller, 1995). Both passive and active verb forms are handled. WordNet and VerbNet have a very broad coverage in relation to standard English. The presence of very domain-specific terms from archaeology, which are not present in WordNet or VerbNet, however, may pose an obstacle, but this issue was not detected during the reported work.

Finally, the second component ActivityRelationMiner takes as input the list of activities produced in the previous phase and the original text, and outputs the process model, that is, the activities and their relations. The default relation is the sequence, because the activities are reported by default in the chronological order of their execution. However, if temporal markers such as “before”, “after” or “last year” are present, they are then considered for sentence re-ordering. The rule-based algorithm was built for identifying decisions, parallelism and mutual exclusion too.

The input to the rule-based algorithm is generated from the original sentence and its corresponding treebank accordingly: the tags of transitive verbs and other key structures as conjunctions, punctuation, and prepositions are kept; all the other content is replaced with “...”. For example, the sentence “The excavations were structured to accommodate the requirements of the developer.” becomes “... 1.VBN to 2.VB”. Further, the algorithm takes each two consecutive verbs within a sentence and identifies their relation by applying a set of predefined rules. Notice that for the relation’s identification, only the activity’s verb is used. Several examples of rules are presented in Table 1. In the case that a verb has multiple objects enumerated with the conjunction “and”, then multiple parallel activities are extracted. Activities become mutually exclusive if the conjunction linking the objects is “or”.

Rule input	Rule output	Explanation of symbols
... 1.VB/VBN/VBD ... , ... 2.VBG ...	1 -> 2 (independent clause)	-> Sequence: activity 1 follows activity 2
... 1.VB/VBN/VBD 2.VBG ...	1 2 (dependent clause)	Parallelism: activity 1 and 2 are executed partially or completely in parallel
... where VB/VBD/VBN ... 2.VB/VBD/VBD ...	1? -> 2 (decision)	? Decision: if activity 1 then 2 in sequence
... 1.VB/VBN/VBD ... or ... in order to 2.VB ... , ... 3.VB/VBN/VBD ...	1 x (2 -> 3) (branches)	x Mutual exclusion: either activity 1 or activity 2 followed by 3. () Precedence

Table 1. Some sample rules being applied to mine activity relationships.

The solution was validated in two steps: firstly, we performed an internal validation in a case study with the lead archaeologist of a project conducted in Villa Magna in Italy (Fentress, 2010). This case study validation was published as (Epure et al., 2015) and allowed us to perform an initial evaluation of the solution performance. The solution was able to create the complete process model performed during the excavations of Villa Magna project with an 88% precision for activity recognition. This means that we were able to discover in an automatic way most of the methodological knowledge presented in the archaeological reports of Villa Magna. The final process model was validated by the lead archaeologist of the project and original author of the analysed report. The TextProcessMiner solution thus established a suitable basis to assist the identification of the reported methodological information in an automatic way.

Secondly, we used the TextProcessMiner solution to analyse multiple reports provided by four ARIADNE project (<http://www.ariadne-infrastructure.eu/>) partners about their archaeological practices. These reports not only included information about excavation processes but also about other archaeological processes carried out within each institution, such as zooarchaeological analysis or 3D modelling methodologies. The results of applying the TextProcessMiner to these reports maintained the success rate in activity identification at over 80%, although additional problems with ActivityRelationMiner were found in relation to the identification of relationships between activities. We also noticed more presence of false positives in activities than in the Villa Magna case. All these aspects might be related to the great diversity of sources, archaeological practices reported, and authors of the reports. However, TextProcessMiner solution still allowed us to obtain initial process models for the four reporting institutions with no false negatives. This means that the solution identifies all the activities reported by the institutions involved without mistake and, in some cases, a few constructs that are not actual activities, which can easily be removed by hand. We can conclude that TextProcessMiner constitutes an effective tool for assisting archaeologists in the extraction of methodological knowledge embedded in reports. This information can be then fed into an SME repository as described in previous sections of this paper.

Nonetheless, improvements were also identified during both validation steps, including the naming of activities with anaphora resolution, the discovery of activities having transitive idioms, and the discovery of complex relations between activities. Also, we should remark that this approach is intended to support rather than replace specialist work, making it faster and easier for them; the intervention of an expert is still needed to validate the output of the tools.

Repository

As a final step, a database was created to work as SME repository by using Microsoft Access 2013 and the Microsoft Jet engine. The structure of this database implements the extended ISO/IEC 24744 metamodel, so method components of any of the necessary kinds could be stored and inter-related. The database was populated with the results from the manual and NLP-assisted processing of the methodology reports, resulting in over 220 individual method components plus multiple associations. Also, each method component is traced back to the source from where it has been “mined”. Figure 3 shows an example of the contents in the repository.

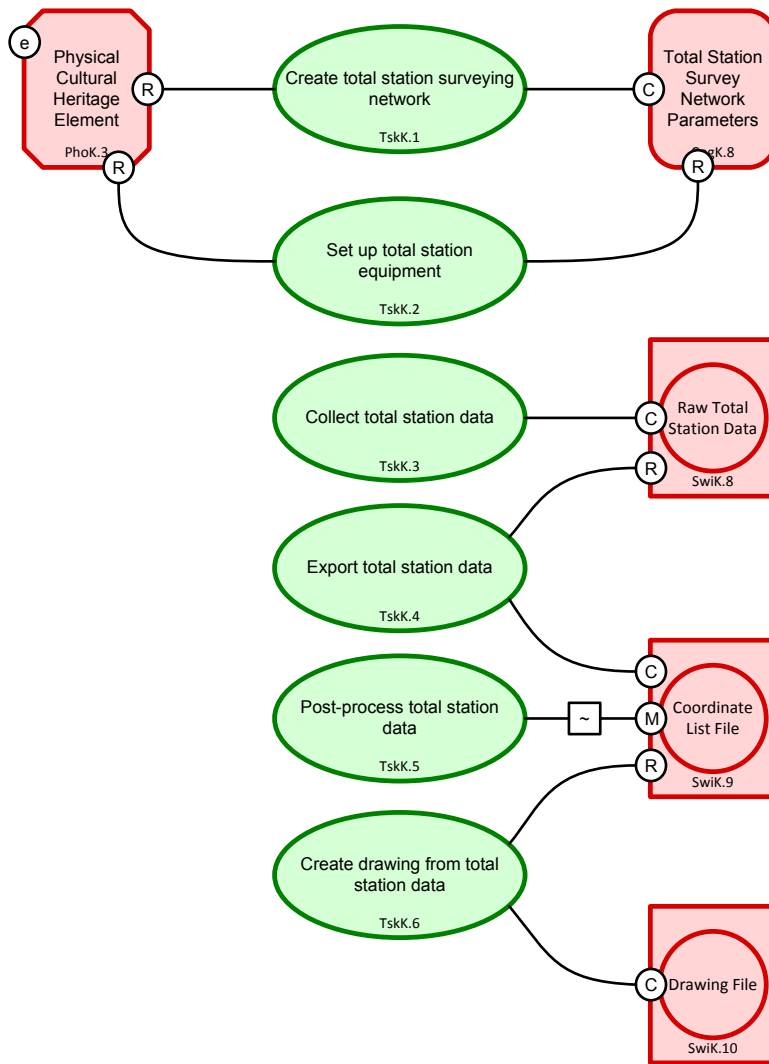


Figure 3. Action diagram for “Total Station and GPS Georeferencing”. The ISO/IEC 24744 graphical notation is used.

Figure 3 is an action diagram for a proposed “Total Station and GPS Georeferencing” process. The ellipses depict tasks that may be carried out as part of the process, whereas the rectangular shapes represent products that are used, created or changed. Each connection between a task and a product corresponds to an *action*, or event of usage. The letters inside the small circles indicate whether a particular task creates (C), reads (R) or modifies (M) a product. Note that no explicit sequencing of tasks is given; in a project, tasks will be carried out as dictated by need and availability of the necessary products, producing in this manner an emergent sequence. For example, the “Set up total station equipment” task needs to read the “Total Station Survey Network Parameters” product, so usually this task may not take place before “Create total station surveying network” has created it. This product-focused and emergent approach has been previously described by (Gonzalez-Perez and Henderson-Sellers, 2008b).

Discussion of Results

Having a repository of consolidated method components presents multiple benefits for archaeological teams concerned about methodology. The first situation under which

benefits were observed was that of obtaining variations of established methodologies to cater for different project situations. For example, a paper-based survey process was easily transformed into a digital-based one by introducing a few new components and replacing the old ones while maintaining most of the process skeleton.

A second scenario under which the repository was very useful relates to the combination of different methodologies for collaboration or other hybrid scenarios. Often, a team working on a project externalises part of their process to collaborating partners. For example, an excavation team may rely on an independent laboratory to carry out radiocarbon dating or other specialist techniques. In situations like these, the methodologies of the “client” and the “provider” can be integrated into a continuous flow while maintaining their relative independence. This allows for either of the parties to alter their internal workings while maintaining the products at the boundary in order to keep the collaboration going.

A third scenario of utility for the repository was related to the dissemination and communication of the methodological choices of different organisations, as well as the reasoning about them. Specifically, having methodological knowledge discretised and arranged into a repository allowed us to describe specific portions of a methodology to any stakeholder in a precise and agile manner, deliver methodological knowledge to all the parties involved, and determine why specific products are being created or used, and detect those that are unnecessary.

Under these circumstances, methodologies become easier to convey to specialists, which is especially important in relation to new team members or collaborators of other disciplines. Also, methodological knowledge, being stored in the repository, is reused systematically every time that a particular component is chosen for a methodology. Finally, since components in the repository are constantly refined and adjusted, new methodologies benefit from the accumulated experience of previous users as they complete the cycle once again.

A few issues still remain. The analysis and processing of text-based methodology reports is extremely arduous, and although NLP tools help significantly, they still have to be augmented to detect work products and producers in addition to processes. Also, specialised tools are necessary for archaeological teams to browse the repository, compose their methodologies, and use them for a project. Without these tools, manual usage of repositories may be too cumbersome for some potential users.

Finally, research in SME tends to blur the distinction between two different but equally interesting goals: mining methodological knowledge out of expert organizations, and systematizing said knowledge into digital platforms for its exploitation. In this paper, we have tackled both at the same time as a proof of concept, but future efforts may focus on one or the other.

Conclusions

In this paper we have proposed situational method engineering (SME) for archaeological methodologies, and described a particular experience involving the manual and NLP-assisted analysis and processing of a number of texts in order to construct a method component repository. This repository has shown to be useful in a

number of scenarios, although the need for software tools makes its exploitation limited at this stage.

Nevertheless, the application of SME to archaeology looks promising, especially when aided by NLP tools that can help us “mine” method components from textual sources. Still, reports may be highly heterogeneous, having a varying range of precision, quality and coverage. A characterization of these aspects should be the focus of future research work.

Acknowledgment

The work described in this paper has been partially funded by the “Advanced Research Infrastructure for Archaeological Dataset Networking in Europe” (ARIADNE) project, FP7-INFRASTRUCTURES.

References

- Aalst, W. M. P. v. d. (2011) *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, Springer.
- Council for British Archaeology (2007). *CBA Research Reports*. Archaeology Data Service. http://archaeologydataservice.ac.uk/archives/view/cba_rr/.
- Epure, E., Martín-Rodilla, P., Hug, C., Deneckère, R. & Salinesi, C. (2015) Automatic Process Model Discovery from Textual Methodologies: An Archaeology Case Study. In Rolland, C., Anagnostopoulos, D., Loucopoulos, P. & Gonzalez-Perez, C. (Eds.) *Research Challenges in Information Science (RCIS), 2015 IEEE 9th International Conference on*. IEEE Computer Society. 19-30.
- Fentress, E. (2010) "Archaeological Fieldwork Reports: Excavations at Villa Magna, 2009". *Papers of the British School at Rome*, 78, 324-325.
- Gonzalez-Perez, C. & Henderson-Sellers, B. (2008a) *Metamodelling for Software Engineering*, Chichester (UK), Wiley.
- (2008b) "A Work Product Pool Approach to Methodology Specification and Enactment". *Journal of Systems and Software*, 81, 1288-1305.
- (2010) Specifying Business Methods with the Work Product Pool Approach. In Loucopoulos, P. & Cavarero, J. L. (Eds.) *Fourth International Conference on Research Challenges in Information Science (RCIS), 2010*. IEEE Computer Society. 383-390.
- Gonzalez-Perez, C. & Hug, C. (2013) Crafting Archaeological Methodologies: Suggesting Method Engineering for the Humanities and Social Sciences. In Earl, G., Sly, T., Chrysanthi, A., Murrieta-Flores, P., Papadopoulos, C., Romanowska, I. & Wheatley, D. (Eds.) *Archaeology in the Digital Era, Volume II; e-Papers from the 40th Conference on Computer Applications and Quantitative Methods in Archaeology*. Amsterdam University Press. 807-820.
- Hug, C., Salinesi, C., Deneckère, R. & Lamassé, S. (2011) Process Modelling for Humanities: Tracing and Analysing Scientific Processes. In Zhou, M., Romanowska, I., Wu, Z., Xu, P. & Verhagen, P. (Eds.) *Revive the Past: Proceeding of the 39th Conference on Computer Applications and Quantitative Methods in Archaeology*. Amsterdam University Press. 245-255.

- ISO/IEC (2014). *Software Engineering - Metamodel for Development Methodologies*. ISO/IEC 24744:2014.
http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=62644
- Kumar, K. & Welke, R. J. (1992) Methodology Engineering: a Proposal for Situation-Specific Methodology Construction. In Cotterman, W. W. & Senn, J. A. (Eds.) *Challenges and Strategies for Research in Systems Development*. Chichester (UK), John Wiley & Sons. 257-269.
- Marcus, M. P., Marcinkiewicz, M. A. & Santorini, B. (1993) "Building a large annotated corpus of English: the penn treebank". *Computational Linguistics*, 19, 313-330.
- Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM*. 38(11): 39-41.
- Rolland, C. & Prakash, N. (1996) A Proposal for Context-Specific Method Engineering. IN Brinkkemper, S., Lyytinen, K. & Welke, R. J. (Eds.) *Procs. IFIP WG8 International Conference on Method Engineering*. Springer.
- Snyder, B. & Palmer, M. (2004) Extending a Verb-Lexicon Using a Semantically Annotated Corpus. *Proceedings of the 4th International Conference on Language Resources and Evaluation*.