Project Description – Project Proposals in the Area of Scientific Library Services and Information Systems

e-Research Technologies

Establishing Semantic Research Data Modelling in Biological Anthropology

Stefan Schlager, Freiburg

This is a shortened version of the original project description, intended for circulation outside the project group. It is not identical with the document submitted to the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG).

Project Description

1 Starting point and preliminary work

Biological anthropology investigates human biology on the background of human evolution, prehistory and history. In contrast to research in medicine and other biosciences, it focuses on the variability of the human condition under varying circumstances in different geographical regions and at different points in time. Major aims are identifying and explaining evolutionary tendencies, understanding human-environment interactions and reconstructing life and population histories.

Human remains are a central research object in biological anthropology, as they are the most direct source of information on past populations. The competence of biological anthropologists in studying human remains leads to anthropological research being ingrained with that of neighbouring disciplines like archaeology, medicine, forensic sciences, demography and ethnology. While skeletal material makes up the bulk of human remains, other forms of preservation include mummified tissue, cultural artefacts containing human body parts and scientific preparations in the form of thin sections and wet or plastinated specimens. Research on these materials draws on a broad spectrum of scientific methods, ranging from morphological and histological assessments through molecular analyses of genes, trace elements and stable isotopes to statistical modelling of virtual representations like CT and surface scans. This heterogeneity poses a substantial challenge for research data management, especially as anthropological results need to be related to very different contextual data, describing taphonomic conditions, soft tissue anatomy, archaeological features and finds, crime scenes and related evidence, diseases and medical treatments or provenance (e.g. with restitution claims targeted at collections of human remains).

Research data management is becoming a pressing issue in biological anthropology, just as in science at large. More and more funding agencies, including the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG)[1], demand sustainable data publication strategies and require grant proposals to contain data management plans (Lämmerhirt 2016). A general outline of the underlying argumentation has been compiled and published by the US National Research Council (NRCCRABS 2003), the emergence of overarching research data infrastructures is moderated by newly created entities like the Council for Information Infrastructures (Rat

[1] http://www.dfg.de/en/research_funding/proposal_review_decision/applicants/submitting_proposal/research_data/; last accessed on 25 July 2017.

für Informationsinfrastrukturen, RfII) in Germany (RfII 2016, 2017). Biological Anthropology has some specific requirements for research data management. Material might only be available for a limited amount of time, e.g. before reburial or restitution and during forensic investigations. The availability of historical and prehistoric human remains is generally limited and its fragility requires to minimize physical investigation. As a consequence, structured documentation is essential. Data standardisation also facilitates pooling information from disparate datasets, a task complicated by the large variety of materials and methods.

The need for standardised data acquisition and documentation has been acknowledged in biological anthropology and led to the formulation of a number of data standards based on commonly used research protocols. The 'Data Collection Procedures' (DCP) manual has had four editions since 1986 (most recently Langley et al. 2016). In reaction to the US 'Native American Graves Protection and Repatriation Act', issued in 1990, standards were compiled to document human remains that qualified for restitution or reburial (Buikstra & Ubelaker 1994). In Great Britain, the Institute for Field Archaeologists issued a manual for standardised analyses of human remains from archaeological excavations (Brickley & McKinley 2004). Following up on a similar project in the US, the Global History of Health Project (GHHP[2]) formed in 2001 and subsequently collected data on about 12,500 skeletons from all over Europe in order to trace the development of pathological conditions from the early Middle Ages to the present (Steckel et al. 2006a; Steckel et al. forthcoming). Uniform data structures were ensured by a common code book (Steckel et al. 2006b). The State Collection for Anthropology and Palaeoanatomy Munich (Staatssammlung für Anthropologie und Paläoanatomie, SAPM) has developed its own data standard (Harbeck 2014) to enforce their policy that researchers have to share the data they acquire from the collection's holdings.

These examples demonstrate the main motivations for data standardisation in biological anthropology: securing scientific evidence, improving verifiability of results, reusing data in multiple studies and pooling data for large-scale investigations. Digital archiving of research data is another objective, which is increasingly recognised. While existing data standards all have their own user bases, none of them is routinely employed by all biological anthropologists, nor has data standardisation become a common routine in anthropological research. Also, standards are not constantly curated, leading them to become quickly outdated in the face of rapid methodological development.

Standardisation of research procedures and data have created a demand for e-research technologies that support anthropologists in their research and provide consistent database input. The Smithsonian Institution (Washington, USA) implemented the standards by Buikstra & Ubelaker (1994) in a database application called Osteoware[3] which they released for general use in 2011. The release was supported by a dedicated grant for the creation of a website and online community[4]. The GHHP created its own data collection software for internal use. The SAPM is currently implementing their standard with the client software AnthroBook[5]. Data collection from hand-held devices is possible with the OsteoSurvey application (Austin 2017).

Osteoware and AnthroBook are static front ends to databases helping researchers to create consistent datasets that comply with the respective data specifications. But while standards formulate reasonable research strategies, most actual study designs deviate from these specifications, depending on concrete research aims. Even if such deviations are only minor and still satisfy the overall standard, this renders the use of single-purpose software in many projects impractical for lack of customisability. OsteoSurvey offers more flexibility but lacks full-fledged database management. It can serve as a versatile data entry interface for any information system but does not constitute such a system itself.

Recent approaches to research data management intend to create interfaces between different data standards. Biological anthropologists working for the US Army are currently developing a

[2]http://global.sbs.ohio-state.edu/; last accessed on 17 June 2017.
[3]http://osteoware.si.edu/; last accessed on 17 June 2017.
[4]https://www.ncptt.nps.gov/blog/2011-13/; last accessed 30 May 2018.
[5]http://xbook.vetmed.uni-muenchen.de/wiki/AnthroBook; last accessed on 17 June 2017.

software called Commingled Remains and Analytics Platform (CoRA)[6] in cooperation with the University of Nebraska (USA). This system provides a database that can be configured to represent different data structures. A major feature is an HTTPS/REST application programming interface (API) through which analytical software applications can connect to the database. Currently, CoRA is serving a specific research project on commingled human remains but the long-term vision for the project is an online platform connecting evolving data standards with a growing range of software tools. Currently, it is not clear if and how the software will be distributed or if CoRA will be offered as a centralised service.

The Institute of Biological Anthropology at Freiburg University is an active contributor to the GHHP (Wittwer-Backofen & Engel 2009; Wittwer-Backofen & Engel forthcoming; Steckel & Engel forthcoming), a fact that has created a well-informed awareness of the great potential analyses of pooled skeletal data hold, but also of the drawbacks it might entail. From 2014 to 2017 the DFG funding scheme 'Standards for Indexing and/or Digitisation of Object Classes in Scientific Collections' provided an opportunity to turn this competence into a digital data standard that was designed to overcome the challenges outlined above. The 'Human Skeletal Collections' (HSC) project (DFG reference: WI 863/9-1) has developed RDFBones, a data standard based on the Resource Description Framework (RDF)[7] and building on several well-established RDF ontologies (Engel & Schlager 2015; Engel, Schlager & Drotziger 2016; Engel & Schlager 2017). RDF is a data modelling language recommended by the World Wide Web Consortium (W3C) that is designed to consolidate disparate sets of information. Its concepts include the representation of resources as instances of classes and properties defining the relations that exist between such instances. Networks of such relations are formulated in specifications referred to as ontologies.

RDFBones provides a core ontology that can be extended by creating subclasses of its concepts. Such extensions allow researchers and institutions to implement their specific research designs with little limitations. Compatibility between datasets collected with different extensions is guaranteed by the overarching structure of the core ontology, the fact that extensions can reuse elements from other extensions and the general potential of RDF for mapping different data structures onto each other. The latter also allows to formulate rules how to consolidate disparate datasets after they were produced according to other diverging standards.

RDFBones is not designed to just annotate research data but to model the entire research process, from study design through inventorying available skeletal material, selection and/or preparation of specimens, data collection, transformation of collected data and drawing conclusions to documentation and publication. This exceeds by far the scopes of previous approaches. The larger documentation effort is justified by increased transparency and reusability of the resulting data. Data from each step of the research process can be used as input for other studies. For example, the scores of age-related skeletal traits established in one study can be reused in another study to apply a different way of modelling age-at-death estimates than originally employed. Similarly, researchers might reuse specimens (e.g. thin sections) that were originally created for another purpose. Even if such appropriations are complex, involving different sources, there is always a coherent documentation from the original skeleton to the final results. A more detailed description of RDFBones is provided in attachment 1 (cf. Section 4).

The RDFBones approach differs from previous ones in some crucial aspects. It does not just address the issue of customisation but also provides a means for consolidating existing databases. Institutions adopting RDFBones just need to map the data models of their existing databases onto the RDFBones core ontology and are spared to recode the entire data. RDFBones also exclusively addresses data annotation to make research data more transparent and reusable. Combined with the fact that RDF data can be stored in simple text files and intuitively intelligible xml formats, this provides a perfect basis for data archival and publishing of primary data. Ontology extensions do not just work as software configuration files but are also highly formalised descriptions of methods that researchers might like to add as appendices to their

[6]See description in section 3 of attachment 4 and the presentation in its appendix E. The CoRA documentation can be found at http://cora-docs.readthedocs.io; last accessed on 16 May 2018.
[7]http://www.w3.org/TR/rdf11-primer/; last accessed on 17 June 2017.

methodological publications. These unique characteristics make RDFBones not just a basis for better data acquisition tools but also for research data managing and the accumulation of sustainable large data stocks.

The advantages of the RDF approach reside in the explicit formulation of data's inner coherences. These specifications dispense with the need for external documentation. On the other hand, more information has to be gathered during data acquisition than it is the case with traditional database applications. But this extra effort can be automated. RDF ontologies can serve as conceptual data models in information systems, specifying how data are represented in the database but also configuring which controls for data display and input need to be offered by the graphical user interface (GUI). This quality is advantageous for implementations of RDFBones as the GUI can adapt to new ontology extensions. In such systems, data input events trigger scripts that automatically create RDF statements based on the input and its context. In a well-designed system, data input requires about the same effort as with traditional database applications but the outcome is a much richer dataset. As RDF is an open standard, it is independent of specific software. Databases and even large parts of GUI configurations can be transferred between different programming frameworks.

During the HSC project it soon became clear that implementation of RDFBones in information systems would be crucial for its adoption by the scientific community and a reference implementation was created. Into this application, research data on the historical Alexander Ecker Collection (Kästner et al. 2011; Wittwer-Backofen et al. 2014) was entered. The reference implementation proved that RDF was a suitable technology for attaining the project aims and that it serves well as a data model for database applications. Transfer of data and ontology to other frameworks was unproblematic. The HSC project also experimented with direct implementation of RDFBones extensions into the statistical framework R[8]. Using the package 'SPARQL'[9], predefined queries can pull data from an RDFBones database and provide them as input for statistical functions. This opens a potential for integrating RDFBones into a number of automated workflows.

The HSC project conducted a detailed survey among institutions holding skeletal collections and held a workshop at Freiburg University in October 2016 (report included as attachment 2, cf. Section 4)[10]. RDFBones was also presented at two international symposiums on data standardisation in biological anthropology in April 2017 (New Orleans, USA; report included as attachment 3) and February 2018 (Seattle, USA; report included as attachment 4, cf. Section 4). These events revealed a high demand for research data management systems among museums, research collections, universities and research projects. There is a general awareness that primary data and related infrastructures will increase in importance in the future. On the other hand biological anthropologists usually lack experience with research data modelling and often make little use of software beyond spreadsheet applications. In Germany, skeletal collections are often under-equipped and struggle to develop research data management strategies for lack of resources. While two American work groups (the Phaleron Bioarchaeological Project and the Forensic Anthropology Center of Texas State University, FACTS) made concrete steps towards managing their research data with RDFBones, several German institutions stated urgent need for data management solutions but did not see themselves in a position to initiate such a project. Open data solutions were largely rejected, while curators of skeletal collections showed interest in putting collection catalogues online for scientists to research available materials before requesting access. Among developers of e-research technologies there is a consent that projects should seek compatibility with other approaches to form a network of scientific infrastructures. It should be kept in mind, however, that most projects are still in development and only Osteoware and OsteoSurvey are currently at productivity level. The HSC project group was approached by the DFG-funded ABCD 3.0 project[11] in February 2017 for collaboration on an ex-

[8]https://github.com/RDFBones/RDFBonesQuery; last accessed on 26 June 2017.
[9]https://cran.r-project.org/web/packages/SPARQL/index.html; last accessed on 18 May 2018.
[10]http://ckan.anthropologie.uni-freiburg.de/dataset/rdfbones-workshop-2016-freiburg; last accessed 26 June 2017.
[11]http://gepris.dfg.de/gepris/projekt/248067007; last accessed on 26 June 2017.

tension to the ABCD (Access to Biological Collection Data) metadata standard, targeted at anthropological collections, to begin in summer 2018.

Section 1 has shown that biological anthropology is lacking common practices meeting the growing demand for research data management, despite a long-standing tradition of research methods and data standardisation. Existing standards are too specialised and inflexible to become broadly adopted, single-purpose software applications like Osteoware and AnthroBook too much focused on specific use cases. While recent developments like OsteoSurvey and CoRA place more importance on customisability, they do not address the important issues of research data consolidation and sustainability. Research data modelling by means of the RDF, on the other hand, provides a sound basis for all requirements of research data management. It allows for reuse and enhancement of existing data, optimised data consolidation and sustainable long-term storage at minimised technical effort. It also provides the key to connect alternative approaches to form a coherent system of software standards and applications. Agencies promoting sustainable data management strategies are institutions curating skeletal collections and databases. But some of these are suffering from weak infrastructures and need empowerment.

1.1 Project-related publications

1.1.1 Articles published by outlets with scientific quality assurance, book publications, and works accepted for publication but not yet published

Schlager S. 2017. Morpho and Rvcg – Shape Analysis in R. In: Zheng G, Li S, Szekely G, editors, Statistical Shape and Deformation Analysis. Academic Press. P 217–256.

Schlager S, Rüdell A. 2017. Sexual Dimorphism and Population Affinity in the Human Zygomatic Structure - Comparing Surface to Outline Data. The Anatomical Record 300(1):226–237.

Sholts SB, Gingerich JAM, Schlager S, Stanford DJ, Wärmländer SKTS. 2017. Tracing social interactions in Pleistocene North America via 3D model analysis of stone tool asymmetry. PLOS ONE 12(7):1–18.

Ketoff S, Girinon F, Schlager S, Friess M, Schouman T, Rouch P, Khonsari RH. 2017. Zygomatic bone shape in intentional cranial deformations: a model for the study of the interactions between skull growth and facial morphology. Journal of Anatomy 230(4):524–531.

Semper-Hogg W, Fuessinger MA, Schwarz S, Ellis E, Cornelius CP, Probst F, Metzger MC, Schlager S. 2017. Virtual reconstruction of midface defects using statistical shape models. Journal of Cranio-Maxillofacial Surgery 45(4):461–466.

Schlager S, Wittwer-Backofen U. 2014. Images in Paleoanthropology: Facing Our Ancestors. In: Henke W, Tattersall I, editors, Handbook of Paleoanthropology. Springer Berlin Heidelberg. P 1007–1034.

Wittwer-Backofen U, Schlager S. 2013. Anthropologische Zugänge zur Provenienz menschlicher Überreste in Sammlungen. In: Stoecker H, Schnalke T, Winkelmann A, editors, Sammeln, Erforschen, Zurückgeben? Menschliche Gebeine aus der Kolonialzeit in akademischen und musealen Sammlungen. Berlin: Ch. Links Verlag. P 224–243.

1.1.2 Other publications

Engel F, Schlager S (2015). A digital framework for managing research data in skeletal collections [abstract]. American J Phys Anthropol Suppl 60:129.

Engel F, Schlager S, Drotziger S, 2016. Integration of contextual information with bioanthropological data from skeletal collections [abstract]. Am J Phys Anthropol Suppl 62:139-140.

Engel F, Schlager S. Make research explicit using RDFBones, an extensible digital standard for research data [abstract]. American J Phys Anthropol Suppl .

2 Objectives and work programme

2.1 Anticipated total duration of the project

This proposal refers to a funding period of two years to accomplish the objectives outlined in Section 2.2. Subsequently, additional funding will be needed to fully accomplish the project's

aims and realise the overarching project strategy (see attachment 0 as listed in Section 4 and notes on project continuation in Section 2.4).

2.2 Objectives

Based on the argumentation outlined in Section 1, we propose to actively promote and establish semantic research data modelling in biological anthropology by means of the RDF and continuous curation of data standards and databases. These changes can only be brought about on the background of an active discussion of research data management in the scientific community. Our aims and a three-stage overarching project strategy to achieve them are formulated in a concept paper included as attachment 0 (cf. Section 4). While immediate steps to improve research data quality in biological anthropology are evident, further perspectives for the creation of data infrastructures are difficult to plan ahead because they depend on science policies that are currently being developed on various levels (summarised in RfII 2016, 2017) and reactions from the scientific community. This proposal suggests funding for immediate measures to spark off the development of research data management structures in biological anthropology and for consultation of experts outside the discipline to concretise further strategic steps. Because the further implementation of the project strategy, including consolidation of the technologies to be developed, is still to be determined, we see our proposal as experimental in nature and situated in the area of application-oriented research and development of e-Research technologies according to section 2.2.1 of the funding guidelines.

The project strategy sets the frame for the immediate objectives of this proposal. They contribute towards the first stage of the strategy (described in section 3.1 of attachment 0), aimed at establishing exemplary use cases successfully employing semantic research data modelling, focussing the development of research data management structures in biological anthropology and to raise the issue with the scientific community. The objectives of this proposal are the following:

- Create a customisable open-source web application implementing RDFBones
- Create a suite of extensions based on the Phaleron Bioarchaeological Project use case
- Deploy the application with the Phaleron Project for software testing
- Release the software in beta status
- Prepare the effective introduction of the application into the research community
- Initiate formal curation of the RDFBones data standard

The main objective of this proposal is to create a server-based application software implementing RDFBones that supports biological anthropologists in performing osteological investigations and automates the semantic enrichment of the research data output. This tool will allow researchers to create semantic datasets without advanced knowledge of the RDF and other related technologies and open the door for semantic research modelling in biological anthropology. In this proposal, the projected software application is referred to by the working title AnthroGraph. AnthroGraph will be designed to be highly customisable with little technical effort and know how by institutions and individual researchers.

In the course of the project strategy elaborated in attachment 0 (cf. Section 4), AnthroGraph might serve the following purposes: (a) render RDFBones usable for biological anthropologists and demonstrate its capabilities, (b) serve institutions and research projects to manage and provide their data outputs on a long-term basis and (c) provide infrastructure for permanent data archival and provision services. As the later stages of this strategy need to be defined in detail, software design has to respect several possible lines of further devleopment. This proposal is exclusively concerned with creating AnthroGraph as a minimal viable product (MVP) with a focus on applicability in typical research scenarios in biological anthropology. At the end of the funding period, the software will be ready for deployment with early adopters (institutions and research projects). Therefore, the scope of this proposal is limited to purposes a) and b) as listed above. Also, the focus will be on software customisability by technologically apt scientists as would be available in realistic use cases with early adopters. Issues that exceed the scope of a

MVP and require advanced programming skills are acknowledged but will only be addressed once the applicability of the software is proven.

The AnthroGraph MVP will satisfy the following criteria:

- **Scope of Application**
  AnthroGraph MVP supports the management of a skeletal collection, both assigned to individual skeletons and as commingled remains. It also provides a generic workflow for standard osteological investigations (e.g. sex and age estimation, body height estimation, palaeopathology). This is comparable to applications like Osteoware and AnthroBook. Their scopes are exceeded by the option to implement new types of investigations and mapping disparate datasets onto each other.
- **Targeted User Groups**
  Both collection management and documentation of investigations with installed ontology extensions can be effected by users of moderate computer literacy. Other tasks like extension writing and system administration require advanced technical skills. It is expected that dedicated postgraduate students will be able to teach themselves how to write extensions and that institutions will pay skilled researchers to write extensions for them in order to implement their research routines. Usability is comparable to applications like Osteoware and AnthroBook but ordinary users should be able to retrieve tabular data by means of predefined queries (as with AnthroBook) and not be forced to code their own queries (as with Osteoware).
- **Scalability**
  AnthroGraph MVP will support managing of mid-range skeletal collections and research data produced from their materials. An example for such a collection would be the SAPM with holdings of about 60,000 skeletons, about 300 other types of specimen and about 30 research projects per year.
- **Output**
  Tabular data output from skeletal investigations will be based on context-sensitive generic SPARQL queries. Additionally, ontology extensions can define their own queries to allow for differently structured output.

More detailed specifications are contained in the scope document provided as attachment 5 (cf. Section 4).

It is essential that the AnthroGraph MVP will be broadly applicable, covering all areas of anthropological research and integrate well with research data management policies established at research-oriented IT facilities. To achieve this objective, specialists from a broad spectrum of backgrounds in anthropology and from research data management will be involved in its development as consultants. In addition to these commitments, AnthroGraph will be deployed as soon as possible in an active research environment for early identification and correction of conceptual and technical flaws, as advised by the RfII (2017, 16/17, 27).

Main test case will be the Phaleron Bioarchaeological Project (PBP)[12] which is dedicated to the conservation and investigation of the approx. 1,500 skeletons from a burial place at the Phaleron delta in Athens (Greece; Lobell 2018). The burials date from the late 8th to the early 5th century BC and follow a broad variety of burial customs, some of which are uncommon or even unique to the site. Anthropological data from the analysis of the human remains will have to be modelled together with the archaeological context. Thematic priorities will be the interpretation of atypical burials, analysis of a large number of exceptionally well-preserved subadult skeletons and molecular analyses of nutrition and geographical origin of the individuals. Much further research on the material, beyond the scope of the PBP itself, is to be expected. It is among the early intentions of the PBP to gather its data in a comprehensive database that will be available to other researchers for future studies. This objective is also professed in the project's data management plan. Work along these lines has already begun and the project employs a software developer to service the collaboration from their side. The project provides a

[12] http://phaleron.digital-ascsa.org; last accessed on 13 July 2017.

good use case as it demands adaptation to a running research project and the creation of a versatile infrastructure open for various lines of possible future research. Initial work on a concept for RDFBones extensions covering the PBP is documented in attachment 6 (cf. Section 4).

The PBP will use AnthroGraph to build an internet platform granting access to a growing database of project outputs. The platform will be permanently sustained for collaborative research projects related to the Phaleron site. It will provide a lasting example of AnthroGraph's implementation in a research project and create the nucleus for a user community.

Apart from the Phaleron Bioarchaeological Project, AnthroGraph will also be deployed at the University of Freiburg. Its IT services (Rechenzentrum – RZ), and especially the eScience Group (Wehrle, Wiebelt & Suchodoletz 2017), are currently developing a research data management system to be integrated into overarching infrastructures like EUDAT[13], that are already in existence. The concept is based on a three-layer hierarchical data storage, accommodating various requirements from temporary file hosting to permanent long-term data archival. Data acquisition and analyses are accompanied by an increased qualification of data, i.e. compliance with digital standards and annotation with metadata. This system follows common practices at other institutions. The RZ will support the project with their expertise, helping to make AnthroGraph compatible with research data management infrastructures and delivering basic input for a strategy to introduce semantic research data modelling in biological anthropology. The deployment of AnthroGraph will serve as a test case for integration of the software into infrastructures for digital archival and evaluate the option to offer the proposed software at Freiburg as a service to research institutions on a more permanent basis. It will have to persist some time after the proposed funding period to serve as demonstration software for perusal by potential adopters and for continued cooperation with the RZ.

After positive experiences with integrating research data management in academic education during the HSC project, the anthropology department of Freiburg University will use the information system installed at the RZ for student project assignments. This will populate the application with research data compiled during the HSC project and recoded for the GHHP. These activities will make the test case more realistic and create further feedback for software improvement.

To facilitate its adoption, AnthroGraph must be freely available and deployable in various research settings. To achieve this objective, the software will be developed and released under a permissive software licence in a public repository. At the end of the funding period, software development will be in advanced beta phase and under active testing by the PBP and at Freiburg University.

Once the MVP will be available, propagation of the software will have to ensue by deployment with a larger number of institutions and projects (cf. sections 3.1 and 3.2 of the project strategy, attachment 0). This phase will be crucial for the success of introducing RDF-based research data management in biological anthropology and requires careful planning and preparation. We suggest uniting leading anthropologists and providers of or specialists for scientific infrastructures in a dedicated workshop to discuss the potentials and propagation of AnthroGraph on the background of similar approaches elsewhere. The meeting will evaluate the applicability of the software on the basis of an advanced alpha version and search for viable strategies how to deploy it with institutions that are capable of providing its services to large numbers of researchers. Systems of research data provision and long-term storage will also be addressed with substantial input from the RZ.

Another objective of the workshop will be the integration of existing and upcoming infrastructures and the resolution of a roadmap towards better research data management in biological anthropology. Among the issues that need to be addressed here are the coordination of current projects producing and maintaining standards and e-research technologies, compatibility between technologies and infrastructures, curation of standards and structures supporting regular exchange. This output will be placed as an article in a major scientific journal in order to carry the discussion into the anthropological research community. A more detailed workshop concept is provided as attachment 7 (cf. Section 4).

[13]https://www.eudat.eu/; last accessed on 18 July 2017.

The creation of AnthroGraph will establish the need for formal curation of the RDFBones data standard. The funding scheme under which the HSC project was conducted did not require initiating mechanisms for continuous development and improvement, nor were such considerations immediately necessary in the absence of its practical application. This situation will change with the development of AnthroGraph, requiring routines for bug reports, feature requests and releases of new versions. These tasks are outside the scope of this proposal and have to be undertaken on a voluntary basis by a consortium of scientists, comparable to similar bodies guarding other standards[14]. Among the objectives of this proposal, however, is the initiation of such a consortium.

2.3 Work programme and proposed research methods

In addition to the applicant ('**principal investigator**', **PI**), the working group will be composed of two researchers and two research assistants. One of the researcher positions requires a background in biological anthropology (referred to as '**anthropological researcher**', **AR**), the other one experience with web-based technologies (referred to as '**researcher charged with software development**', **SDR**[15]). The main **research assistant** is referred to as '**RA**', an additional assistant will temporarily help with organisation and execution of the workshop ('**WRA**').

In the course of the proposed work programme (Figure 2.1), the two researchers will jointly develop a software design for AnthroGraph, based on the scope document (attachment 5, cf. Section 4) and an analysis of use cases provided and supervised by the project consultants (cf. Section Fehler: Verweis nicht gefunden). Afterwards, the SDR will focus on implementation of the design and software testing while the AR will create a suite of ontology extensions based on the concrete use case of the PBP and develop a strategy for the introduction of AnthroGraph and semantic research data modelling into the anthropological research community. The PI will manage the project, supervise its progress, represent it to the outside on national and international meetings and work on securing the project's continuation. Contributions of the RA will comprise recurring organisational tasks, maintenance of project resources and helping with documentation. The WRA will help with organising a workshop to discuss and finalize the propagation concept and editing the resulting publication. The individual phases of the work programme (not to be confused with the stages of the project strategy related in attachment 0) are described in more detail below.

AnthroGraph will be developed as an adaptation of the ResearchSpace software[16] which the British Museum (London, Great Britain) has conceptualised since 2010 and developed over the past three years[17]. The software was released in November 2017 in a public repository[18] under the GNU Lesser General Public License (LGPL v2.1). ResearchSpace is a software for building web-based information systems presenting the holdings of research collections, especially in the domain of art and culture, and offering various ways to logged-in users to interact with this information and create own content. It is not particularly designed for research data management. The main point for the development of AnthroGraph is that it also represents the open-source community version of the metaphactory framework[19] for semantic web-based applications. It provides a hierarchical structure of page templates representing semantic contexts of RDF classes that can contain features for input and representation of data. Experimentation with the metaphactory framework during the HSC project showed that page template configuration is a task that can be performed by apt researchers.

Realisation of the work programme will require advanced knowledge of RDF modelling and SPARQL query design of both researchers, a skill that is also highly valuable in many scientific

[14]E.g. the Ontology for Biomedical Investigations (OBI) with a regular schedule of weekly conference calls.
[15]Note that this is not necessarily a trained software developer.
[16]https://researchspace.org/; last accessed on 15 June 2028; not to be confused with the eLab software Rspace developed at Edinburgh (Great Britain).
[17]Personal communication with Dominic Oldman, head of the ResearchSpace development team at the British Museum, on 2 May 2018.
[18]https://github.com/researchspace/researchspace; last accessed on 15 June 2018.
[19]http://www.metaphacts.com/application-areas/cultural-heritage; last accessed on 15 June 2018.

research scenarios. The SDR ideally has fundamental programming skills in Java and TypeScript in order to effect adjustments to the framework code.

ResearchSpace operates by its own ontology based on the CIDOC CRM, an ontology for representing cultural heritage management. In AnthroGraph, RDFBones will replace this ontology as the main data model and provide the basis for all application workflows. Some features specific to ResearchSpace (e.g. the clipboard, assertions and arguments or image annotation) rely on the original ResearchSpace ontology. To what extent these features can be profitably used with AnthroGraph will be determined during software design (phase 2 of the work programme). As RDFBones also relies on the CIDOC CRM for managing research collections and the objects they contain, integration of the two ontologies would be straightforward. Creation of the MVP will be largely restricted to the customisation of page templates in order to create an application that proves useful in scientific investigations. More complex use cases, e.g. data repositories managing multiple independent collections and projects, will require changes to the ResearchSpace framework itself, calling for professional software programmers. These can only be effected in later project stages and in cooperation with institutions that can sustain such efforts. A more detailed assessment of ResearchSpace is provided as attachment 8 (cf. Section 4).

The proof-of-concept implementation of RDFBones created during the HSC project (cf. Section 1) was realised with VITRO[20], the underlying framework of the VIVO software for the creation of websites representing academic institutions[21]. The software was chosen as one of the few open-source RDF frameworks that allowed for the creation of data entry forms. Also, the idea to configure the graphical user interface with RDF statements was appealing. In practice, it turned out that repurposing VITRO into a usable platform for research data managing required too many alterations to the framework itself. Institutions adopting the software would have to invest into customisation of the software that would neither contribute to, nor be supported by the general development of the VITRO framework. With the open-source release of ResearchSpace, there is now a much more practicable solution available that also offers good perspectives in the long run. The British Museum will continue active development of the software. Currently, funding is secured for another three years[22]. We assume that the high customisability and performance of ResearchSpace and its applicability in all fields of research will lead to adoption by numerous research projects and institutions in the near future. While much of its configuration can be carried out by non-professional programmers (see above), professional support is available on a commercial basis from the producer of the metaphactory framework where this is required. The British Museum also recognises the difficulty to integrate commercial enterprises in publicly funded projects and is considering to lend itself as a collaborator on such projects in the future[23]. These points of consideration suggest that ResearchSpace will be a sustainable resource for scientific applications in the long run and a better option for realising AnthroGraph than VITRO.

The following describes the individual phases of the work programme outlined above in detail:

2.3.1 Phase 1: Infrastructures (4 months)

The aim of phase 1 is to develop, create and document a series of work routines and technical infrastructures that enable project members to collaborate on software production and documentation, process user feedback and communicate project updates to various target groups. Tasks include the development of concepts for information flow, resource and task management and the establishment of a project website, a code repository, a bug tracking mechanism and a documentation system. Infrastructures need to be designed for long-term operation by a changing developer and user community to achieve the aims and objectives defined in the project strategy (attachment 0, cf. Section 4) and Section 2.2. Persistent infrastructures need also to be created for the ongoing curation of the RDFBones data standard, consolidating structures developed during the HSC project. This work will include the initiation of a consortium of volunteer developers and the development of routines for processing bug reports, feature requests and

[20]http://vitro.mannlib.cornell.edu/; last accessed on 17 June 2017.
[21]http://vivoweb.org/; last accessed on 17 July 2018.
[22]Personal communication with Dominic Oldman on 2 May 2018.
[23]Personal communication with Dominic Oldman on 2 May 2018.

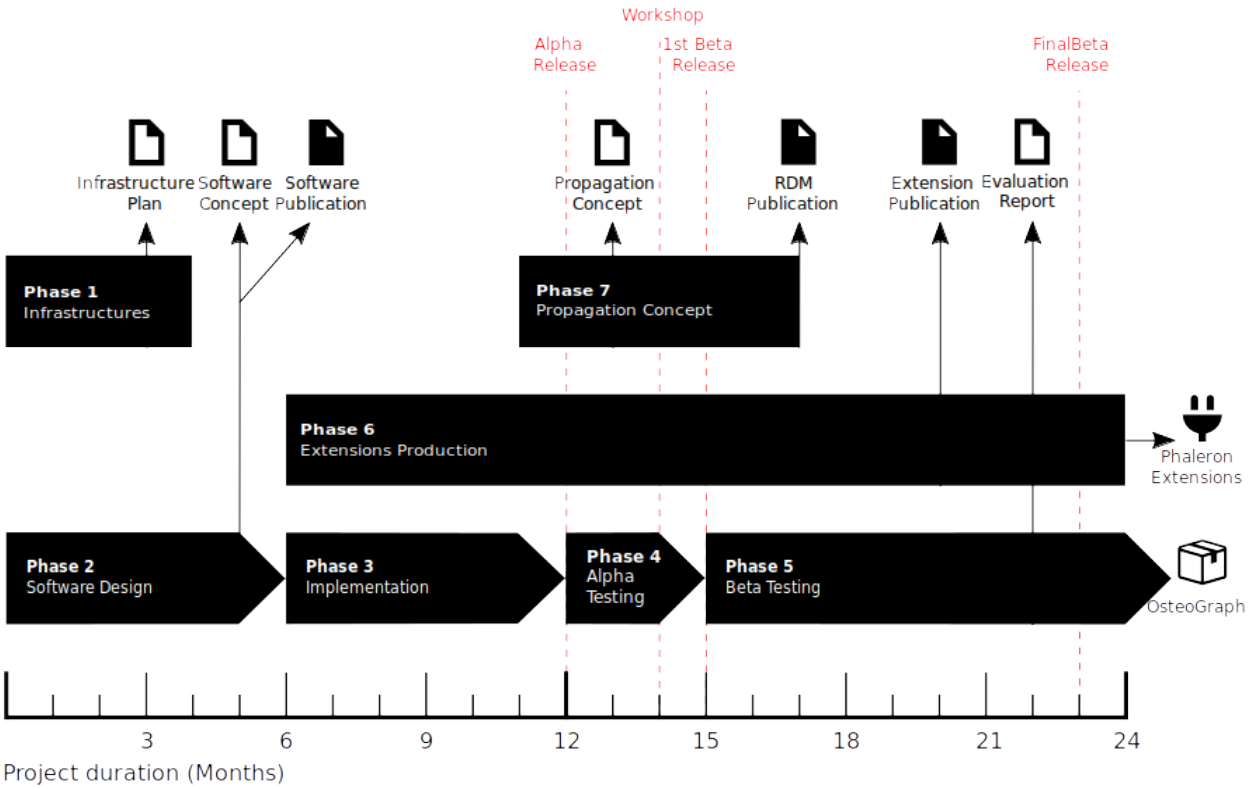

Figure 2.1: Work programme with phases, major outputs and events (red). White document symbols represent informal project resources, black symbols publications in scientific journals.

version releases. An additional task in phase 1 are early preparations for the workshop in phase 7, including appointing date and participants and securing the venue.

Output: Phase 1 will establish a *system of interlinking infrastructures* for the development and propagation of project results. The duration of the phase was determined to give project routines time to consolidate before publication of an *infrastructure plan* that will be mandatory for all project transactions and collaborations.

Assignments: The AR and the PI will collaborate on conceptualisation, the production of the infrastructure plan and workshop preparation. They will receive support from the RA who will also be charged with setting up large parts of the infrastructures.

Evaluation: Project routines will have to be continuously optimised throughout the funding period and during subsequent project stages. In the interest of sustainable project coordination, these changes will be documented in a series of updates to the infrastructure plan to be formally released with notes on changes.

2.3.2 Phase 2: Software Design

The aim of phase 2 is to develop detailed and concrete specifications for the AnthroGraph software to be developed in phases 3 and 4. These will be documented in a software concept listing static pages and page templates for ontology classes that need to be created, presenting page mock-ups, outlining trajectories along which different types of users will navigate through the application and specifying the types of operation sequences that will be supported by the software. These contents will be structured as clearly defined tasks to be processed during implementation. The software concept will also contain technical solutions to these tasks in the form of predefined code blocks or proof-of-concept implementations along with relevant documentation. In addition to the core application, the software concept will also contain specifications for RDFBones extensions, possible assets specific to AnthroGraph extensions (e.g. custom page templates) and on their integration into AnthroGraph.

Input to software design are the AnthroGraph scope document (attachment 5) and the project-specific assessment of the ResearchSpace framework (attachment 8, cf. Section 4). The design

of page templates and supported workflows will be based on an analysis of use cases from all areas of osteological research in biological anthropology which are provided by the PBP and by the project consultants (cf. Section Fehler: Verweis nicht gefunden). A production plan, detailing requirements and procedures in the development of the AnthroGraph MVP, is included as attachment 9 (cf. Section 4).

The approach to implementing RDFBones in a software application developed in phase 2 will be published in a journal article. This paper will not promote AnthroGraph but rather discuss the possibilities of semantic research data modelling in biological anthropology and general options to support it with software applications. A central advantage of this approach is it's portability to other scientific disciplines and the perspective to model osteological data together with contextual data from other knowledge domains. An open-access publication is necessary to reach not just anthropologists but also researchers in related scientific communities. In addition to this paper, the software concept will also be published as a poster presentation at a meeting of the Society for Anthropology (Gesellschaft für Anthropologie, GfA).

Output: Phase 2 will produce the *software concept* as a document to act as a binding guideline for software implementation in phases 3 and 4. It will also bring forth a manuscript for a *journal article* to be handed in at an open-access publishing outlet (e.g. PLOS ONE) and a *poster presentation*.

Assignments: The AR and the SDR will collaborate on the creation of the software concept. The AR will draw up and analyse use cases and define scientific requirements while the SDR will focus on technical implementation. For the production of the manuscript they will be joined by the PI. The RA will provide support with literature search, obtaining and processing input from consultants, formatting and proof reading.

Evaluation: The software concept will be sent to the project consultants for review and feedback. It will be tested and improved in software implementation during phases 3 and 4.

2.3.3 Phase 3: Implementation

The aim of phase 3 is to implement the specifications concerning the core application from the software concept developed in phase 2.

The phase is primarily concerned with producing the core AnthroGraph MVP, without any extensions loaded. If software specifications turn out to be impracticable, improved solutions are recorded in new versions of the software concept in order to maintain the document's authority.

Phase 3 will terminate with an alpha release of AnthroGraph through the public repository. It will be installed at the RZ for testing purposes and this installation will be continuously updated until the end of the project.

Output: Phase 3 will produce an *initial version* of the AnthroGraph MVP, ready for ontology extensions to be loaded and alpha testing to be performed. There will also be an updated version of the *software concept*.

Assignments: Software implementation will be carried out by the SDR, supervised by the PI.

Evaluation: The product of phase 3 will be evaluated by alpha testing in phase 4 and the experts workshop in phase 7.

2.3.4 Phase 4: Alpha Testing

The aim of phase 4 is to ensure that the software created in phase 3 is generally deployable.

During alpha testing, AnthroGraph will be subjected to test routines and inspection, mainly from within the project group.

1. By the beginning of phase 4, two to three extensions will have been produced as outputs of phase 6. These will be loaded into the core application to determine if this can be performed free from error.

2. Sample data related to the loaded extensions is imported to test if it is displayed and processed correctly by the application. All implemented queries are executed to check for correct output.
3. The sample datasets are also typed in manually to check if data entry alongside osteological investigations is well supported. These entries will be accompanied by uploads of images and documents related to the entered data.
4. The installation of the alpha version will be made available to the project consultants and other anthropologists with a vested interest in data standardisation and selected researchers with the PBP in order to receive qualified feedback in the form of bug reports, feature requests and suggestions for product enhancement.

Before the end of phase 4, feature requests issued by then will be prioritised and it will be decided which of them can be implemented with the MVP.

Phase 4 will terminate with the first beta release through the public repository. At this stage the software will be feature complete and all further work restricted to testing and improvement. In addition to the installation at the RZ, the software will also be deployed for the PBP. This instance will be continuously updated until the end of the project.

Output: Phase 4 will produce a feature complete *beta version* of AnthroGraph ready for testing by end users.

Assignements: Test routines will be carried out by the PI and the AR. The RA will support them with manual data entry. The SDR will perform bug fixes and effect final feature implementations.

Evaluation: The output of phase 4 will be evaluated by beta testing in phase 5.

2.3.5 Phase 5: Beta Testing

The aim of phase 5 is the elimination of software faults in AnthroGraph and to bring the application to a mature beta stage.

Researchers at the PBP will enter skeletal data via the available extensions throughout the phase. At the beginning of phase 5, they will be coached in the usage of AnthroGraph and the performance of feedback at the Wiener Laboratory at Athens. As more and more extensions will be produced in phase 6, the scope of test entries will gradually increase. Feedback will be provided in the form of bug reports, feature requests and suggestions for improvement via the infrastructures set up in phase 1. Additionally, there will be a structured survey of user experiences among all participants in beta testing during the period of month 18 to 21 of the project's duration. All feedback will be summarised in an evaluation report that will also make recommendations for future development.

Based on the feedback from beta testing, software faults will be continuously eliminated and improvements made as long as this will be compatible with the production schedule. This will result in a series of improved minor versions of AnthroGraph. The final version during the funded period will leave the software in a well-documented state allowing further development by new developers and present AnthroGraph with a comprehensive software design. It will be highly stable and a good basis for quick development of a fully productive application.

Output: Phase 5 will produce a *mature beta version* of the AnthroGraph software. A secondary product will be the *evaluation report*.

Assignments: The SDR will process feedback from beta testing by eliminating software faults and optimising the code base. The structured software evaluation will be carried out by the SDR and the AR together who will also collaborate on the evaluation report, supported by the RA.

2.3.6 Phase 6: Extensions Production

The aim of phase 6 is to produce six to nine RDFBones ontology extensions to be loaded into the AnthroGraph software. These are necessary to render the application operational and as exemplary cases in order to demonstrate to researchers how to create their own extensions.

The extensions created in phase 6 will support the establishment of an information system providing osteological data from the PBP. This is necessary in order to use that project as a test case in phase 5. The PBP operates data acquisition procedures that are modifications of the standards set forth by Buikstra & Ubelaker (1994), a popular guideline in biological anthropology. Extensions, therefore, will first be developed as implementations of these standards before additional versions specifically adapted to the PBP will be derived from them. This procedure has the benefit of providing more generic extensions to the research community and of demonstrating how existing extensions can be adapted to specific research projects. Phaleron-specific extensions may include page templates for input masks of a different design than the one provided by the core application. This will demonstrate another adaptive feature of AnthroGraph.

In compliance with the RDFBones principles, extensions are seen as independent micro projects with their own production cycles. Because phases 5 and 6 run in parallel, individual extensions might still be implemented while the AnthroGraph application has already reached beta status.

All AnthroGraph extensions contain ontology extensions of RDFBones. These can be published as highly formalised specifications of methods (cf. Section 1). To give an example of this practice, at least part of the extensions created for the AnthroGraph project will be published in an anthropological journal. Which extensions will be most suitable for such a paper will have to be determined during extensions development. Possible options are a more generic publication of an implementation of Buikstra & Ubelaker (1994; e.g. in the International Journal of Osteoarchaeology) or a contribution of more limited scope to a specialised journal (e.g. the International Journal of Paleopathology).[24] Additionally, extensions will be presented on a poster at an annual AAPA meeting.

Output: Phase 6 will produce a suite of *AnthroGraph extensions* rendering the application immediately operational and customising it for use in the PBP. A secondary output will be the manuscript for a *publication of extensions* to be submitted to a dedicated journal and a *poster presentation*.

Assignments: The AR will conceptualise and produce the general versions of extensions. Customisation of the Phaleron-specific versions will be undertaken in collaboration with the software developer at the PBP. The RA will help with documentation and communication with the PBP.

2.3.7 Phase 7: Propagation Concept

The aim of phase 7 is to develop an effective strategy for the introduction of semantic research data modelling in biological anthropology with the help of AnthroGraph and to coordinate it with other approaches to research data standardisation and management.

The central means to achieve this objective is an experts' workshop with leading anthropologists involved with research data management and representatives of non-profit information infrastructure facilities and organisations. The workshop will deal with the following topics:

- Future of research data management and publication of primary data in academic and non-academic research institutions.
- Synthesis of approaches to research data standardisation and management in biological anthropology. Interfaces between the employed technologies and coordination of efforts. Perspectives for the creation of research data repositories.
- Applicability of RDFBones and AnthroGraph in anthropological research. Potentials for interdisciplinary research.

A more detailed description of the workshop concept is provided as attachment 7 (cf. Section 4).

The workshop will produce a concise summary of current developments within biological anthropology and science at large and a vision for research data management within the discipline in

[24] An alternative publication outlet is the recently founded Journal of Computer Applications in Archaeology.

the future. This output will be condensed into a joint publication by workshop participants to be placed in a biological anthropology journal (e.g. American Journal of Physical Anthropology).

In order to achieve the desired output, the workshop will need careful preparation. This will involve obtaining conceptual input from all participants before the workshop to create a basis for discussion. In respect of the AnthroGraph project this will be a concept paper on the propagation of semantic research data modelling in biological anthropology and the role of AnthroGraph in this process. After the workshop, this document will be enriched with relevant workshop results and form the basis for additional funding proposals to advance the introduction of semantic research data modelling. Workshop outcomes will also be presented on a poster at an AAPA annual meeting.

Output: Phase 7 will bring about the first *publication of a concise perspective for research data management* in biological anthropology and a *poster presentation*. It will also form a strategy for the propagation of semantic research data modelling by planting instances of AnthroGraph (*propagation concept*).

Assignments: The PI and the AR will write the propagation concept and organise the workshop in collaboration with the Research Data Management Group at Freiburg University. They will also edit the joint publication emanating from the event. The WRA will organise the technicalities of the workshop and support the publication's editing process.

2.4 Measures to meet funding requirements and handle project results

The HSC project has shown that RDFBones is capable of amending current problems with research data standardisation and management and that the RDF approach is regarded as practicable by peer researchers. This is evidenced by the analyses carried out as part of the project and by the project workshop (cf. Section 1), but even more so by its reception within the scientific community. The joint symposiums at New Orleans and Seattle have initiated fruitful cooperation with other project groups, representatives of which are among the consultants brought forth by this proposal. Even while RDFBones was still at a conceptional stage, two institutions, the PBP and the FACTS, took concrete steps towards implementation of RDFBones. Skeletal collections in Germany declared urgent need for collection and research data management systems but also reported lack of funding and infrastructure for in-house developments. With the creation of AnthroGraph we intend to provide a resource that renders setting up modern information systems a realistic option for an increasing number of institutions and projects.

The key innovations that RDFBones brings to osteological research are annotated data structures, independence of specific application frameworks and the option to map existing data structures onto each other. While other projects in standardised data acquisition are focused on the provision of research tools, RDFBones offers a basis for data exchange and long-term storage. Applications like AnthroGraph can work as a hub between other tools, mainly by translating their data models into each other. The workshop held in phase 7 of the work programme (Section 2.3.7) will develop a plan how to build a unified research infrastructure from existing projects. AnthroGraph will be able to play many roles in such a system but they will most likely involve data preparation for long-term storage and use in specific projects. Integration with systems for research data archival is a central topic within the proposed project group which comprises representatives of the Research Data Management Group at Freiburg University.

The information system created at the PBP will be continuously maintained by that project, providing a lasting example of usage for AnthroGraph. Also, the Biological Anthropology Department at Freiburg University will continue to use the system installed there for managing their own research data and as a technology demonstration for potential adopters of AnthroGraph. How the main project output, AnthroGraph, will be maintained beyond the funding period cannot be said precisely at this stage. A sustainable strategy will be developed during the workshop held in phase 7 of the work programme. In this respect we consider the proposed project to have an experimental orientation. As a consequence, this proposal deliberately leaves the software in beta phase, expecting that additional funding will be necessary for its final establishment. Currently it is envisaged that a quickly increasing number of projects and institutions will employ AnthroGraph because of its many advantages and the project's exposure through the

PBP and the publication of workshop results. Eventually, one or several adopting institutions should take the lead of a user community supporting the continuous development of the software. This process is expected to gain support from a growing user base of the ResearchSpace framework. Another business model would be to have large research foundations operate instances of AnthroGraph and provide its usage as a paid service to smaller institutions. Such solutions would require substantial modifications of the ResearchSpace framework that can only be accomplished by institutions with their own IT infrastructures. While we expect further funding of the software's propagation to be necessary, its maintenance should be secured through research funding and/or institutional resources in the medium run.

The development of RDFBones was made possible by three years of funding from the DFG. After its remarkably good reception in the scientific community we propose to create AnthroGraph in order to deliver the benefits that RDFBones holds and to prevent its abandonment at an early stage. We expect an increasing interest in the application, resulting in additional requests for its implementation in research projects and institutions. We deem the project to be successful if evaluation by the PBP will be positive and further opportunities for similar collaboration projects will arise. An additional criterion is the development of a workable strategy to further pursue the aims set forth by the project strategy, resulting in a proposal for continuative funding.

All resources developed by the proposed project will be made openly available. Publications will be open-access or appear in major anthropological journals. Details on the provision of resources are given in the work programme.

This proposal is the resubmission of a previous application to the ″e-Research Technologies″ programme submitted on 4 August 2017 and rejected on 5 April 2018 (attachment 10, cf. Section 4). One reviewer with a background in anthropology deemed the project's objectives worthy of support while all three reviewers (two with a technical background) found too many uncertainties concerning their implementation for the proposal to be accepted. We have addressed most of the criticism by limiting the scope and duration of the project proposed here, focusing on the production of the MVP. Software production is limited to technologies that apt researchers can handle and serves applicability in research rather than deployment with large organisations. As requested, we have included additional information about the production plan and the employed technologies (see also attachments). A major point of criticism was the lack of a concept to ensure the development of an active user and developer community. We have taken this up by making the issue a major research objective, tackled in phase 7 of the work programme. With the workshop featuring a distinguished panel and a related journal article we intend to create exposure for research data management in general and AnthroGraph in particular. We also intend to establish research data management as a genuine concern of anthropological research as opposed to a service rendered to individual anthropologists. Introduction and propagation of AnthroGraph in the anthropological research community is consciously delayed for later project stages when the workshop will have brought about a concise strategy. In their note of rejection, the DFG encouraged us to resubmit a modified version of our proposal. We believe to have addressed the major suggestions brought forth by the reviewers and that the current proposal is a considerable improvement to the former.

2.5 Information on scientific and financial involvement of international cooperation partners

The PBP has an independent budget for developing and maintaining a data platform for the project and hires a software developer for its realisation. They provide an active use case and testing environment for the development of AnthroGraph. In return, we intend to provide the AnthroGraph MVP as a basis for the data platform and help with conceptualising extensions that also support the PBP. The cooperation is exclusively based on mutual benefit and does not involve financial transactions of any kind.

3 Bibliography

Austin AE. 2017. OsteoSurvey: An Open-source Data Collection Tool for Studying Commingled Human Remains. *American Journal of Physical Anthropology Supplement 64*:105.

Brickley M, McKinley JI, ed. 2004. *Guidelines to the Standards for Recording Human Remains*. Institute for Field Archaeologists.

Buikstra J, Ubelaker DH, ed. 1994. *Standards for data collection from human skeletal remains*. Fayetteville: Arkansas Archaeological Survey.

Engel F, Kramis S .2010. Documenting Skeletal Trauma: A Proposal for a Standardised Routine. In: C. A. Buhl, F. Engel, L. Hartung, M. Kästner, A. Rüdell, C. Weißhaar, eds., *Proceedings of the 4th Meeting of Junior Scientists in Anthropology*. Freiburg: University Library. pages 35-43.

Harbeck M. 2014. *Anleitung zur standardisierten Skelettdokumentation in der Staatssammlung für Anthropologie und Paläoanthropologie München*. URL = http://www.sapm.mwn.de/attachments/article/249/AnleitungSkelettdokumentation2014.pdf.

Kästner M, Ortolf S, Rüdell A, Möller D, Wittwer-Backofen U. 2011. The Alexander Ecker Collection in Freiburg. *Documenta Archaeobiologiae 8*:275-284.

Lämmerhirt D. 2016. *Briefing Paper: Disciplinary differences in opening research data*. URL = http://www.pasteur4oa.eu/resources/209#.WXdSTidLdpg.

Langley NR, Jantz LM, Ousley SD, Jantz RL, Milner G. 2016. Data Collection Procedures for Forensic Skeletal Material 2.0. Knoxville: Department of Anthropology, University of Tennessee. URL = https://fac.utk.edu/wp-content/uploads/2016/03/DCP20_webversion.pdf.

Lobell JA. 2018. Ancient Athens' Other Cemetery. *Archaeology* 71(4):48-53.

NRCCRABS. 2003. *Sharing Publication-related Data and Materials: Responsibilities of Authorship in the Life Sciences*. Washington D.C.: National Academies Press (US).

RfII – German Council for Scientific Information Structures 2016. *Enhancing Research Data Management: Performance Through Diversity. Recommendations regarding structures, processes, and financing for research data management in Germany*. Göttingen: RfII. URL = http://www.rfii.de/?wpdmdl=2075.

RfII – German Council for Scientific Information Structures 2017. *An International Comparison of the Development of Research Data Infrastructures: Reports and Suggestions*. Göttingen: RfII. URL = http://www.rfii.de/?wpdmdl=2792.

Steckel R, Larsen CS, Sciulli P, Walker P. 2006a. The Scientific Value of Human Remains in Studying the Global History of Health. In: Lohman J, Goodnow K, ed. *Human Remains and Museum Practice*. New York: Berghahn Books. p60-70.

Steckel RH, Larsen CS, Sciulli PW, Walker PL. 2006b. *Data Collection Codebook*. URL = http://global.sbs.ohio-state.edu/new_docs/Codebook-01-24-11-em.pdf.

Steckel RH, Larsen C, Robers C, Baten J, eds. Forthcoming. *The Backbone of Europe: Health, Diet, Work and Violence over Two Millennia*. Cambridge: Cambridge University Press.

Steckel RH, Engel F. Forthcoming. Climate and health: Europe from the pre-Middle Ages to the nineteenth century. In: Steckel RH, Larsen C, Robers C, Baten J, eds., *The Backbone of Europe: Health, Diet, Work and Violence over Two Millennia*. Cambridge: Cambridge University Press. Chapter 12.

Wittwer-Backofen U, Engel F. 2009. The History of European Oral Health: Evidence from Dental Caries, Dental Abscesses, Antemortem Tooth Loss. *American Journal of Physical Anthropology Supplement 48*:275.

Wittwer-Backofen U, Kästner M, Möller D, Vohberger M, Lutz-Bonengel S, Speck D. 2014. Ambiguous provenance? Experience with provenance analysis of human remains from Namibia in the Alexander Ecker Collection. *Anthropologischer Anzeiger 71*:65-86.

Wittwer-Backofen U, Engel F. Forthcoming. The history of European oral health: evidence from dental caries and antemortem tooth loss. In: Steckel RH, Larsen C, Robers C, Baten J, eds., *The Backbone of Europe: Health, Diet, Work and Violence over Two Millennia*. Cambridge: Cambridge University Press. Chapter 4.

4 Additional information

This document is submitted with the following attachments:

0. Attachment_00-ProjectStrategy.pdf
   ″Project Strategy. Aims and Opportunities for Action to Introduce Semantic Research Data Modelling in Biological Anthropology″
1. Attachment_01-RDFBonesArticleManuscript.pdf
   ″RDFBones: An Infrastructure for Digital Standardisation in Physical Anthropology″ Manuscript submitted to the journal ″Anthropologischer Anzeiger – Journal of Biological and Clinical Anthropology″ in March 2018
2. Attachment_02-RDFBonesWorkshopFreiburg2016-Report.pdf
   ″Digital Standards for Research Data from Human Skeletal Collections: Report on the RDFBones workshop, Freiburg 7/8 October 2016″
3. Attachment_03-AAPA2017-SymposiumReport.pdf
   ″Skeletal Standards: Documentation Software, Databases, and Online Digitization Resources Available to Researchers: AAPA 2017 Symposium Report″
4. Attachment_04-AAFS2018-SymposiumReport.pdf
   ″Digital Data Standards, Analysis and Archiving in Forensic Anthropology: Workshop at the 2018 AAFS Meeting in Seattle (Washington, USA)″
5. Attachment_05-AnthroGraphMVPScopeDocument.pdf
   ″Scope of AnthroGraph: Requirements for the Development of a Minimum Viable Product″
6. Attachment_06-CollaborationRDFBonesPhaleronBioarchaeologicalProject.pdf
   ″Concept for RDFBones Extensions Covering the Phaleron Bioarchaeological Project: Skeletal Inventory and Material Preservation/Conservation″
7. Attachment_07-AnthroGraphWorkshopConcept.pdf
   ″Workshop Concept: The Future of Research Data Management in Biological Anthropology (working title)″
8. Attachment_08-ResearchSpaceAssessment.pdf
   ″Assessment of the ResearchSpace Framework: Functionality and Necessary Adaptations for the Production of AnthroGraph″
9. Attachment_09-AnthroGraphProductionPlan.pdf
10. Attachment_10-InitialSubmission_Reviews.pdf
    Initial submission ″Establishing RDF-based research data standards in biological anthropology″ and letter of rejection with reviews
11. Attachment_11-VM_Pricing_RZ_Uni.pdf
    ″ESX Virtualisierung für die Universität″