

Learning through connecting: data management as a basis for meaningful use of legacy data - The African Archaeology Archive Cologne

Apprendre par la partage : la gestion des données comme base d'une utilisation collective des données du patrimoine - The African Archaeology Archive Cologne

Eymard Fäder¹, Friederike Jesse¹, Tilman Lenssen-Erz¹

¹ University of Cologne, African Archaeology – UniK

ABSTRACT. For more than 50 years the African Archaeology at the University of Cologne has conducted field research in various parts of Africa, particularly in the Eastern Sahara and in Namibia. The focus was on environmental history and on rock art, with internationally renowned research schemes like ACACIA (Arid Climate Adaptation and Cultural Innovation in Africa) and 'Rock Paintings of the Upper Brandberg'. Much of it was published in close co-operation with the Heinrich-Barth-Institut, including six voluminous catalogues of rock art and 26 monographs on African archaeology. Most of the materials accumulated in the research projects are not digital in origin and are now digitized and made accessible in the digital online archive AAARc (African Archaeology Archive Cologne). Besides being a repository for tens of thousands of pictures, also whole documentations of excavation are made accessible here. Digitising the repositories of Cologne's African Archaeology within the AAARc project started in 2012, using the DAI (German Archaeological Institute) data-interfaces Arachne, DAI-Gazetteer and DAI-Zenon. Today AAARc opens web access to thousands of fieldwork documentations, pictures and tracings of rock art. Open access interoperability of metadata integrated to picture formats is achieved. Exploiting database management systems of highly varied structure in a digital repository is achieved through collaboration with IANUS (Research Data Centre Archaeology & Ancient Studies), a national long-term preservation digital archive for archaeological data. AAARc offers participation in digital heritage content management, thus perpetuating the long lasting cooperation of Cologne's African Archaeology with a number of African countries. The AAARc online archive aspires to make the full range of archaeological documentations accessible online and not just pictures. Thus AAARc ultimately enables a digital homecoming of complex research data to their countries of origin.

RÉSUMÉ. Depuis plus de 50 ans, l'université de Cologne mène des recherches dans diverses régions d'Afrique, en particulier au Sahara oriental et en Namibie. L'accent a été mis sur l'histoire environnementale et l'art rupestre, avec des projets de recherche de renommée internationale comme ACACIA (Arid Climate Adaptation and Cultural Innovation in Africa) et 'Rock Paintings of the Upper Brandberg'. Les résultats de ces programmes ont été publiés en grande partie en collaboration avec l'Institut Heinrich-Barth, dont six volumineux catalogues d'art rupestre et 26 monographies sur l'archéologie africaine. La plupart des matériaux accumulés dans les projets de recherche ne sont pas d'origine numérique et sont maintenant numérisés et rendus accessibles dans les archives numériques en ligne AAARc (African Archaeology Archive Cologne). En plus d'être un dépôt pour des dizaines de milliers d'images, des documentations complètes sur les fouilles sont également accessibles. La numérisation des dépôts de l'archéologie africaine de Cologne dans le cadre du projet AAARc a débuté en 2012, en utilisant les interfaces de données DAI (Institut archéologique allemand) Arachne, DAI-Gazetteer et DAI-Zenon. Aujourd'hui, AAARc ouvre l'accès en ligne à des milliers de documents de terrain, d'images et de tracés d'art rupestre. L'interopérabilité en libre accès des métadonnées intégrées aux formats d'image est réalisée. L'exploitation de systèmes de gestion de bases de données de structure très variée dans un dépôt numérique est réalisée en collaboration avec IANUS (Research Data Centre Archaeology & Ancient Studies), une archive numérique nationale de préservation à long terme des données archéologiques. L'AAARc offre une participation à la gestion du contenu du patrimoine numérique, perpétuant ainsi la coopération de longue date de l'Archéologie africaine de Cologne avec un certain nombre de pays africains. Les archives en ligne de l'AAARc aspirent à rendre accessible en ligne l'ensemble de la documentation archéologique, et pas seulement des photos. Ainsi, l'AAARc permet en fin de compte un retour numérique des données de recherche complexes vers leur pays d'origine.

KEYWORDS. digital archive ; open access ; African archaeology ; rock art ; cultural heritage management.

MOTS-CLÉS. archives numériques, open access, gravures rupestres, gestion du patrimoine.

Digital environments have changed the way of collecting, processing, presenting and storing archaeological data significantly. Centralised services for long term preservation in an open access setting have been developed in several countries since the 1990s such as the [ADS](#) in the UK, and the similar [DANS](#) in the Netherlands (Schäfer et al. 2015: 128-129). German archaeology started relatively late with comparable initiatives (Siegmund & Scherzler 2015: 14): Online portals such as [ARACHNE](#) provide standardized open access to archaeological objects and research data centers such as [IANUS](#) ensure long-term preservation and long-term digital access to archaeological research.

A common challenge for all these initiatives is based on the exponential growth in available data over different areas of science and scholarship: how can we store data so that they can be re-used in the future? This is not just a question of technical infrastructure; one also has to make sure that data are re-usable at the structural and semantic levels.

Information integration is key to doing research in a computational setting. Funding institutions demand that research data, including results from specific investigations, must be made publicly available in a comprehensible and verifiable form in order to create synergies. Open access and open data are the keywords and the claims for further development should be followed up with practical work. The "[Berlin Declaration](#)", which was formulated in 2003 on the initiative of the Max Planck Society and meanwhile signed by more than 600 academic institutions worldwide, calls for:

“New possibilities of knowledge dissemination not only through the classical form but also and increasingly through the open access paradigm via the Internet have to be supported.”

Therefore, we define open access as a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community. In order to realize the vision of a global and accessible representation of knowledge, the future web has to be sustainable, interactive, and transparent. Content and software tools must be openly accessible and compatible ([Berliner Erklärung](#), [Principles](#)), [Making Data Count](#)).

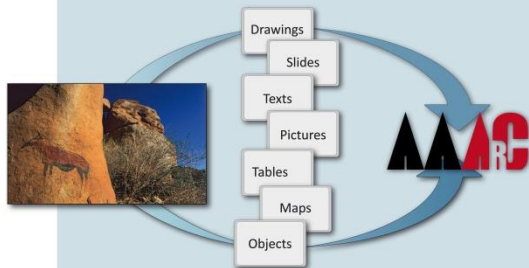
Based on these principles and to fill an existing gap, the African Archaeology Archive ([AAArC](#)) pioneered curating and storing digital objects of most different formats (figure 1). Starting point for the development of this online database were the huge collections of data accumulated during various research projects in African Archaeology conducted by Cologne University's Institute of Prehistoric Archaeology since the 1960s. In order to curate, collect, process, preserve and present the project data, comprising documentations of excavations, photographs, slides, tracings of rock art, drawings, field notes, tabulations as well as physical objects, AAARc started in 2012 with digitizing these work results. By providing long-term preservation and web-based access to the multifarious project data, AAARc helps the scientific community - especially colleagues and students from African countries - to work with this archaeological documentation.

The **African Archaeology Archive Cologne (AAAC)** is a digital archive. It is based on collections resulting from University of Cologne's research in northeastern and southwestern Africa in the past 50 years.

AAAC is a repository complying with the principles of open access. Its long term aim is an open research archive where material from other sources and projects can easily be integrated. Ultimately, **AAAC** enables a digital homecoming of complex research data to their countries of origin.



The Material

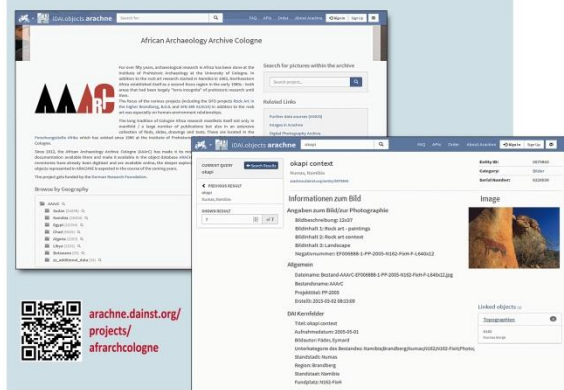


The rich materials require different approaches of curation: there are materials which can be digitised to images (slides, field recordings etc.), genuine digital images (.tiff, .jpg etc.) and non-visual research data (texts, tables, objects etc.).

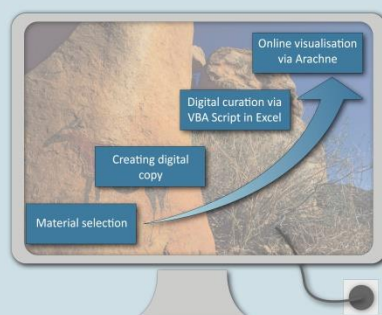
The Result

AAAC currently presents about 77,000 pictures of rock art, excavations and objects in Africa's northeast (Sudan, Egypt, Chad etc.) and southwest (Namibia). Pictures are licensed under Creative Commons CC-by-NC. A database containing information on the sites such as location, chronology, description and bibliography allows in depth exploration of the data available in **AAAC**, presently for more than 1500 sites.

To guarantee sustainability, **AAAC** is hosted by Arachne, the central object database and archive of the German Archaeological Institute and thus is part of the iDAI world. Technical support is provided by Cologne University's Regional Computing Centre (RRZK).



The Process



Every digital object is tagged with metadata in 27 categories concerning geography, origin, content and access permission. **AAAC** metadata is based on established standards (Dublin Core or CIDOC CRM).

The Future

AAAC provides structures for digital data management where new material can easily be integrated. Close cooperation with IANUS (Research Data Centre Archaeology & Ancient Studies) helps to include non-visual archaeological research data and therewith enlarge the information online accessible.

Contact:
Phone: +49 (0)221 55 66 80
Fax: +49 (0)221 55 02 303
Email: aaarc-online@uni-koeln.de

University of Cologne
Institute of Prehistoric Archaeology
African Archaeology
Jennerstrasse 8, D-50823 Köln



Figure 1. This poster was employed to present and introduce AAAC at several congresses. In these exchange forums, AAAC showed that despite the relatively small project size, it has quickly reached eyelevel with well established open digital archives. (Oliver Vogels)

The technological infrastructure of the AAAC project is based on different platforms of the “[iDAI.world](#)” (Deutsches Archäologisches Institut, DAI) which provides efficient services for data ingestion, data curation, metadata enrichment, presentation and long-term preservation with platforms such as [iDAI.gazetteer](#), [iDAI.geoserver](#), [iDAI.objects](#) / ARACHNE and IANUS (figure 2). Regarding technical sustainability, this infrastructure is hosted on a permanent basis by the regional computing centre (RRZK) of the University of Cologne. To enhance the process of ingesting project data into the iDAI.world workflows, the curation tool “Scan-Protocol” was developed by AAAC. “Scan-Protocol”

provides a user-friendly interface for preparing data for the iDAI.world workflows in a single process sequence. A compendium (Wiki) informs about the operational procedures in AAaRC so that the data curation can be accomplished autonomously by everyone interested in contributing to the archive.

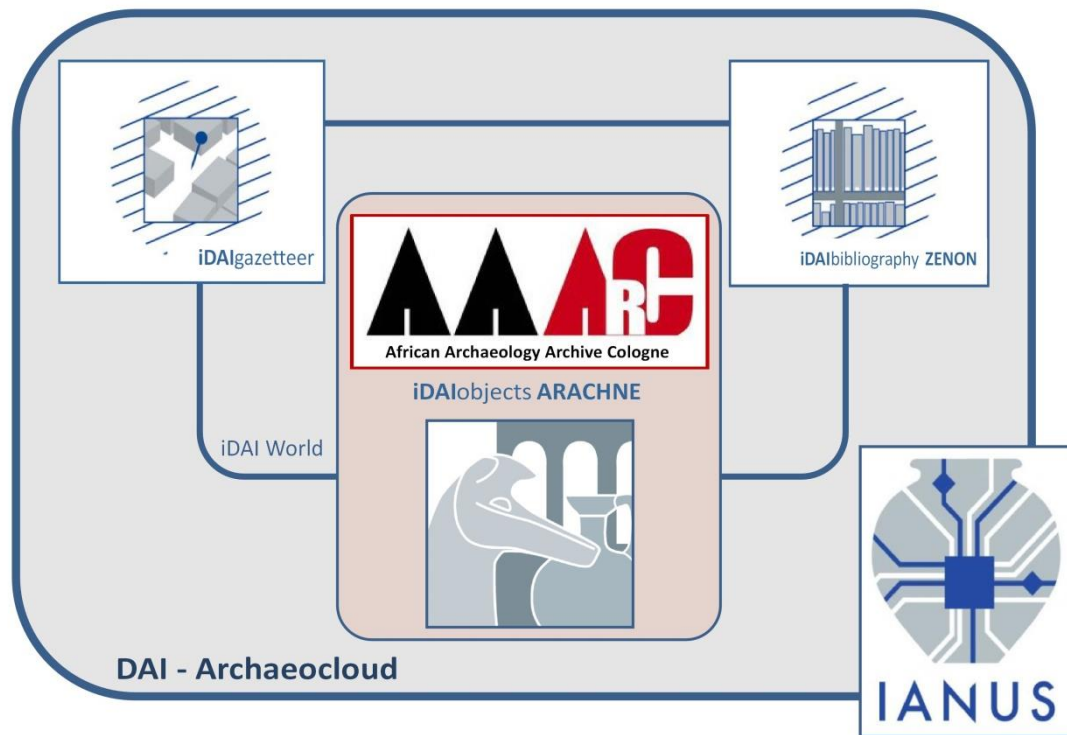


Figure 2. Schematic chart showing the linked open data concept of the platforms of the iDAI world that AAaRC is deploying. (Eymard Fäder)

Thus various standards required definition: from normalized chronology terms, controlled ontology lists and a hierarchical geographical folder structure, to metadata concerning the origin, type and access rights of the digital record (figure 3).

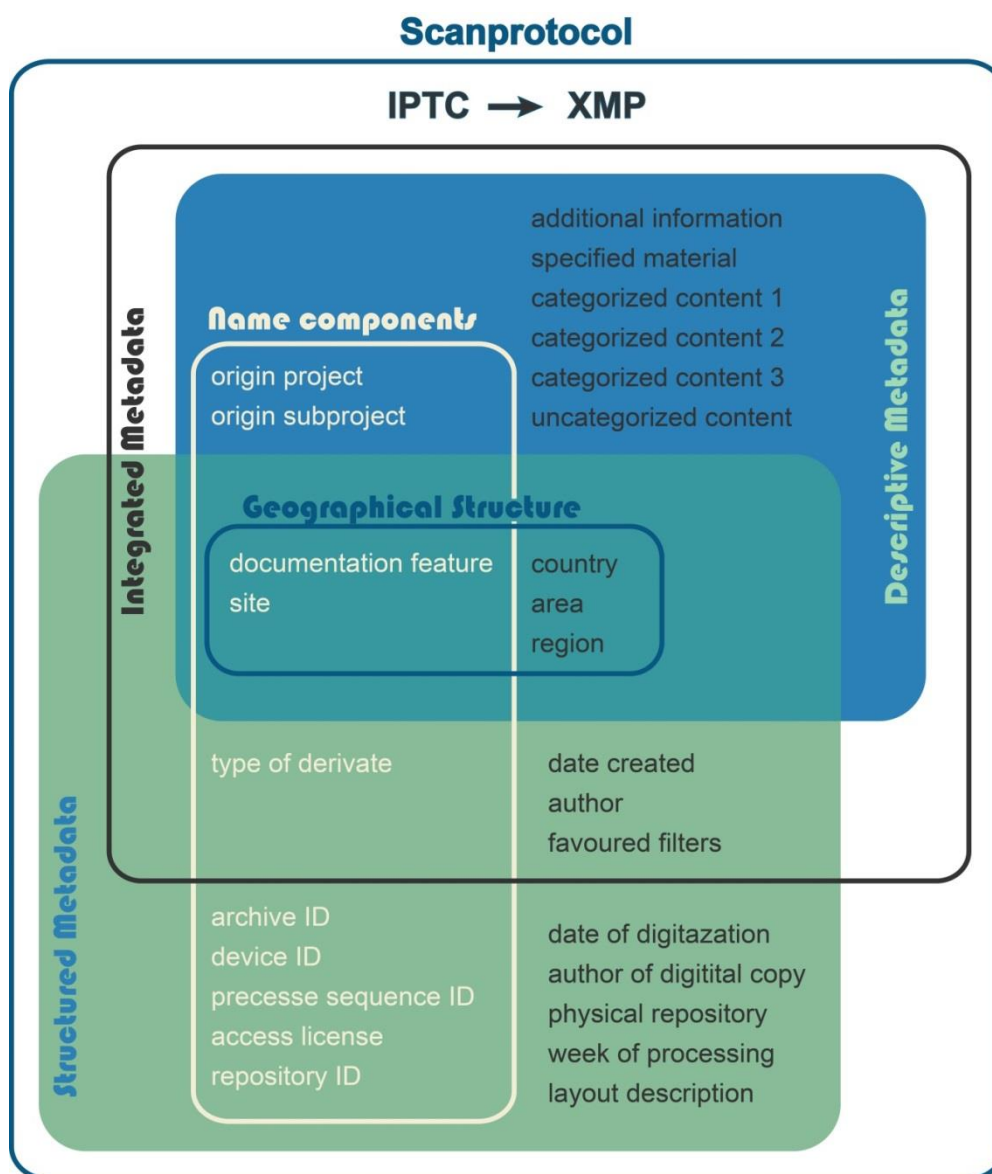


Figure 3. Graphical representation of the AAARc metadata model for pictures that is handled in this complexity by the scan protocol also illustrating how metadata is linked to the filename. This overlap of metadata and filename is a security measure against data loss if one of the two components is damaged, for example, during a transfer or migration. (Eymard Fäder)

These entries are embedded in every digital object by integrated metadata, including geographical information on the location of the find, the research project, the date and originator of the digital object, as well as a detailed description of the content (figure 4). These should by any means be based on established standards such as archaeological cataloguing standards such as [LIDO](#) (Eide et al. 2005, Eide 2008) or [CIDOC CRM](#).

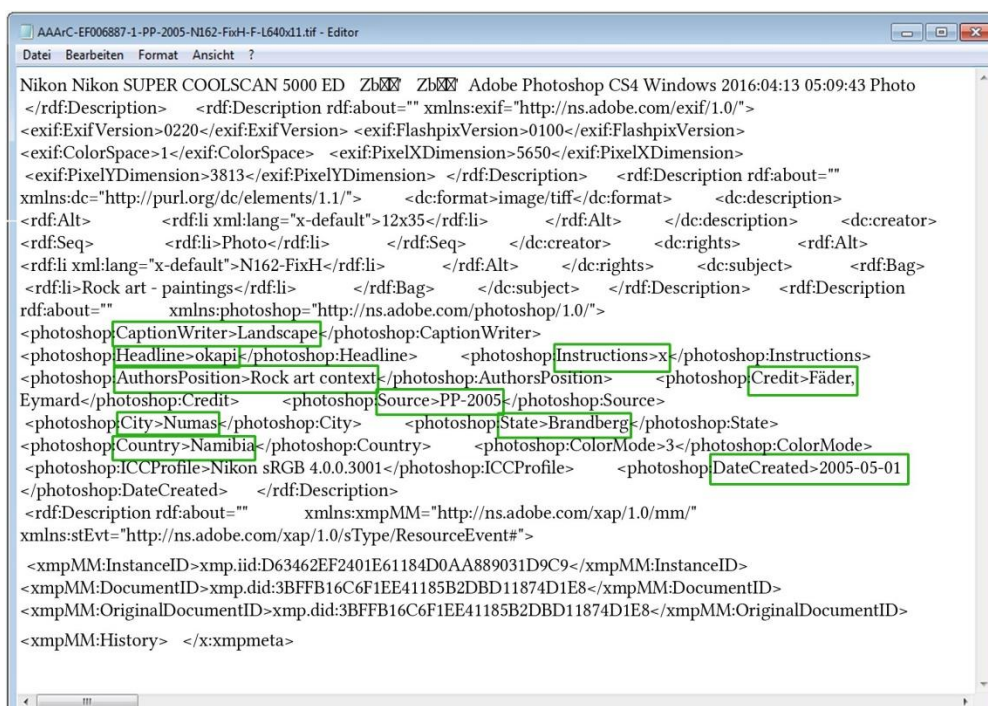


Figure 4. Metadata of a digital picture highlighting the information that, in XMP format, remains inseparably connected to the digital copy in this way. (Eymard Fäder)

Within four years project's time AAARc has made accessible about 70 000 pictures and site information for hundreds of prehistoric sites in Northeastern and Southwestern Africa. The recently funded Priority Programme [SPP 2143 "Entangled Africa"](#) allows to continue and enrich AAARc. AAARc will not only coordinate the data management for the different research projects within the SPP 2143 but also develop further possibilities of use through the project "Learning through connecting". As a node within the SPP 2143, AAARc will coordinate sustainable data ingestion into the repositories of the iDAI-World. For the handling of machine-readable data, particularly complex or non-standardized digital objects, AAARc has established cooperation with [IANUS](#). By monitoring and advising the participating projects, the AAARc team will enable harmonization of metadata linked to the chronological and geographical ontologies used by the projects, but also facilitate a quick growth of the database. Furthermore, the project "Learning through connecting" will integrate different aspects based on archaeology and on cultural heritage informatics: information modelling, developing and running the systems, data import and curation, long term preservation. Given limitations on bandwidth in and out of African countries and in view of occasional political issues connected to control and responsibilities, additionally mirroring of data between servers in Germany and servers in the partner countries should be explored.

Besides providing various coordination infrastructures, the development of an application programming interface (API) to enable researchers to work as flexible as possible with the project data is planned. This will allow public users and researchers a user friendly and efficient way to open their data to colleagues.

References

- AAARc: <https://arachne.dainst.org/project/afrarchcologne> (last accessed 15.10.2018)
- ARACHNE: <http://arachne.uni-koeln.de//drupal/> (last accessed 15.10.2018)
- ADS (Archaeology Data Service): <http://archaeologydataservice.ac.uk/> (last accessed 15.10.2018)
- Berliner Erklärung: <https://openaccess.mpg.de/Berlin-Declaration> (last accessed 15.10.2018)
- [CIDOC CRM](#): <http://www.cidoc-crm.org/> (last accessed 15.10.2018)
- DANS (Data Archiving and Networked Services): <https://dans.knaw.nl/en> (last accessed 15.10.2018)

Eide, Øyvind (2008) "Museum Collections Catalogues and Archives Online: A Reference Resource for Professionals and a Learning Tool for the Public." In Museum and the Internet. Presenting Cultural Heritage Resources On-line Archaeolingua, 2008: 35–51.

Eide, Øyvind, Jon Holmen, and Anne Birgitte Høy-Petersen (2005) "Making the Resources Fit Together. Interconnection of Diverse Archaeological Document Collections." Internet Archaeology 18.

IANUS (Forschungsdatenzentrum Archäologie und Altertumswissenschaften): <https://www.ianus-fdz.de> (last accessed 15.10.2018)

iDAI.gazetteer: <https://gazetteer.dainst.org> (last accessed 15.10.2018)

iDAI.geoserver: <http://geoserver.dainst.org> (last accessed 15.10.2018)

iDAI.objects. <https://arachne.dainst.org> (last accessed 15.10.2018)

Lenssen-Erz, T., Fäder, E., Jesse, F. & J. Wilmeroth, Digital Management of Rock Art: the African Archaeology Archive Cologne (AAArC). Afr Archaeol Rev (2018) 35: 285ff. <https://doi.org/10.1007/s10437-018-9303-5>

Making Data Count:

http://www.dfg.de/dfg_magazin/aus_der_wissenschaft/archiv/knowledge_exchange_workshop_2013/index.jsp (last accessed 15.10.2018)

Principles: DOI: <http://doi.org/10.2312/allianzoa.024> (last accessed 15.10.2018)

Schäfer, F., Heinrich, M., Sieverling, A., Trognitz, M., Foertsch, R., Dally, O., Fless, F. (2015) Forschungsrohdaten für die Altertumswissenschaften – eine kurze Bilanz der aktuellen Situation von Open Data in Deutschland, Archäologische Informationen 38: 125-136.

Siegmund, F. & Scherzler, D. (2015) Open Access und Open Data in der Ur- und Frühgeschichte: Bestandsaufnahme und Ausblick (Open Access and Open Data in European pre- and protohistory: present situation and outlook), Archäologische Informationen 38: 11-19.

SPP 2143, Entangled Africa: <http://www.entangled-africa.org/index.php/en/the-programm/>