The first step towards FAIR-ness in Bulgarian archaeology: The Archaeological Map of Bulgaria in ARIADNE and ARIADNEplus

Nadezhda Kecheva

This article provides an overview of the participation of the National Archaeological Institute with Museum at the Bulgarian Academy of Sciences (NAIM-BAS) as a partner in both the ARIADNE and ARIADNEplus projects and the SEADDA COST Action. The article examines both the workflow and the results obtained when modifying and mapping part of the national site and monuments information system, also known as the 'Archaeological Map of Bulgaria' (AIS AKB). The article's focus is mainly on the technical preparation of data shared through the ARIADNE portal. The main workflow followed throughout the projects included mapping terms to the Getty AAT thesaurus, adding Bulgarian archaeological chronologies to PeriodO, and mapping metadata to the ARIADNE catalogue data model (ACDM) in ARIADNE and to the AO-cat ontology in ARIADNEplus. In 2016, a mapping of AIS AKB to the CIDOC CRM was also undertaken, and the first steps towards creating the geographic information system, 'Archaeological Map of Bulgaria' (GIS AKB), were taken. The experience gathered by the NAIM-BAS team members throughout the whole project was important for moving towards FAIR-ness in Bulgarian archaeology.

1. Introduction

Developments of computer technology and the internet in the new millennium are two of the main reasons for focusing on digital techniques and digitisation in the sciences and humanities, and archaeology is no exception. Digitising existing legacy data is only the first step towards the creation of digital data. Creating easily **f**indable, **a**ccessible, **i**nteroperable, and **r**eusable (FAIR) digital resources requires a lot more work. The National Archaeological Institute with Museum at the Bulgarian Academy of Sciences (NAIM-BAS) was a partner in the ARIADNE (2013-17) and ARIADNEplus (2019-22) projects financed by the European Commission. This participation had two positive outcomes. The first was the experience gained by NAIM-BAS team members in using contemporary European data and becoming familiar with metadata standards following the FAIR principles. The second was the practical aspect. Participation in ARIADNE and ARIADNE plus made information sharing and updating of the archaeological information system 'Archaeological Map of Bulgaria' (AIS AKB) possible. The work process followed the main steps needed for sharing metadata and data in the ARIADNE portal (Bardi *et al.* 2022).

2. Directions of the digital 'Archaeological Map of Bulgaria' in the new millennium

This part of the article is focused on the development of the 'Archaeological Map of Bulgaria' (AIS AKB) in the new millennium. A brief history is necessary since NAIM-BAS participated in both the ARIADNE and ARIADNEplus projects with data exported from the AIS AKB. A detailed history of the digital archiving of archaeological reports in Bulgaria is beyond the scope of this article, but see a brief summary of the current state of archiving in Bulgaria in Kecheva (2021) and overviews of archiving in other countries in Aspöck and Richards (2023) and Jakobsson *et al.* (2021, 2023).

The 'Archaeological Map of Bulgaria' has a long history prior to the growth of computer technology and the development of the internet (Нехризов 2014 , Kecheva 2024). It was created as a catalogue/registry of a 'site and monument' type system containing so-called 'registration cards' for the recorded archaeological sites in Bulgaria. The term 'archaeological site' has many different definitions in archaeological projects and publications worldwide (Kecheva 2024). The accepted definition for AIS AKB is the following: *a particular topographic place in space where, during different chronological periods, human beings have performed various permanent and/or temporary activities*. The AIS AKB information system was created in the 1990s in the DOS operating system, updated in the 2000s in the Windows operating system, and has been accessible through a web application since 2010 (Автоматизирана информационна система 'Археологическа карта на България' 2024). Its accessibility is achieved through an authorised access according to Bulgarian national legislation (*Наредба № Н-2 от 6 април 2011 г. за създаване, поддържане и предоставяне на информацията от автоматизирана информационна система 'Археологическа карта на България' (обн. ДВ. бр. 32 от 19 април 2011 г., доп. ДВ. бр. 7 от 23 януари 2024 г.)* 2024).

The database that is accessible through the web application is visible as a registry and the registration cards are loaded as a list. More than 25,000 archaeological sites have been recorded in the information system as of September 2024. More than a half of them were recorded on paper forms in the 1990s (approximately 15,000 records). Since it was not possible to use GNSS/GPS devices for archaeological site location during that period, these registration cards were not assigned X and Y coordinates. Instead, the archaeological sites were described and approximate locations labelled on topographic maps, with some being more informative than others (Figure 1).

The web application created in 2010 did not require GPS coordinates for listing the registration cards. However, more recently, in the era of widely used geographic information systems (GIS), the inclusion of GPS coordinates (X and Y) became essential. This resulted in the growing need for full digitisation of the paper forms, including the addition of X and Y coordinates to those that were missing this data. Projects, including ARIADNE, provided the opportunity to execute this task over the course of several years, as explained below.
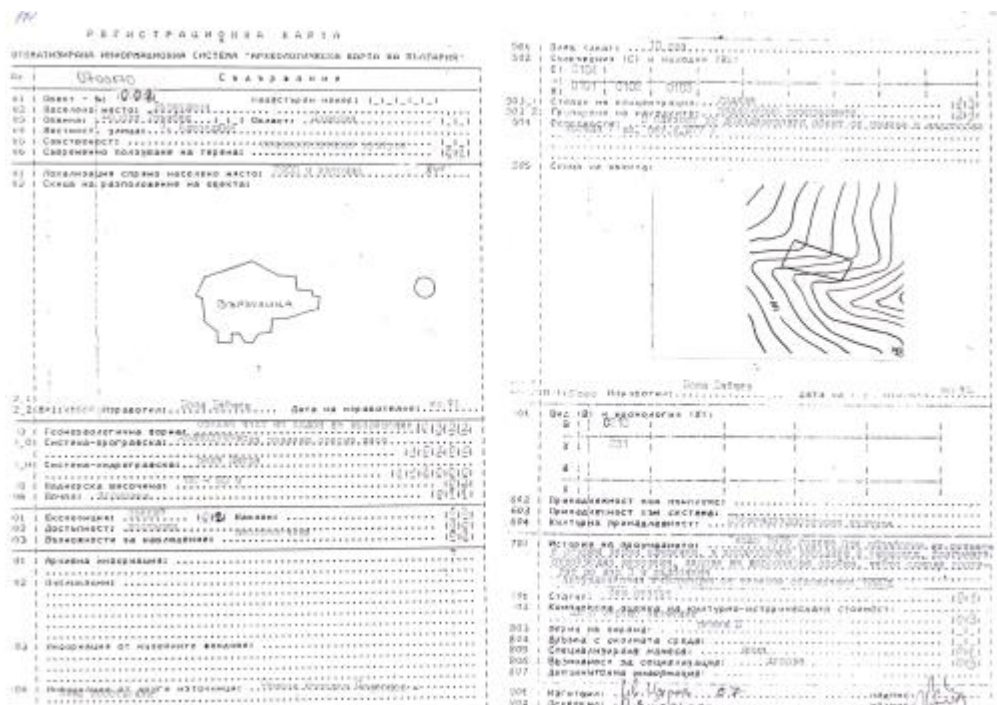
Figure 1: AIS AKB paper registration card No 0700170

3. The 'Archaeological Map of Bulgaria' in ARIADNE and ARIADNEplus

NAIM-BAS was the only Bulgarian institution involved with the ARIADNE and ARIADNEplus projects. A team of different specialists from NAIM-BAS and beyond participated in the work. ARIADNE was the first European project that addressed the 'Archaeological Map of Bulgaria'. Since NAIM-BAS was one of ARIADNE's 'data-provider' partners, the AIS AKB data to be provided had to be modified in the required way.

The modification workflow included activities done in parallel. Some focused on an AIS AKB update and some followed the description of work of the ARIADNE and ARIADNEplus projects.

- Full digitisation of paper registration cards that did not have X and Y coordinates for the location of the archaeological sites;
- Mapping of:
  - Site types, finds, facilities to the Getty AAT thesaurus;
  - Exported modified metadata and added data of the shared archaeological sites to the ARIADNE Catalogue Data Model (ACDM) (ARIADNE support portal 2013);
  - Exported modified metadata and added data of the shared archaeological sites to AO-Cat (Felicetti *et al.* 2023);
  - AIS AKB original metadata to the CRM (CIDOC CRM (2024) Conceptual Reference Model)
- The inclusion of Bulgarian archaeological chronology and periodisation terms in PeriodO (PeriodO client. ARIADNE Consortium. ARIADNE Data Collection PeriodO 2015).

- Bonus activity: the initiation of GIS AKB - the beginning of the 'Archaeological Map of Bulgaria' process of updating as a GIS web application.

3.1. Full digitisation of registration cards with missing X and Y coordinates

The digitisation of the 'Archaeological Map of Bulgaria' legacy data was hard and time-consuming work. The work process included: an inventory of the existing paper registration cards from 1992-2010, division of the registration cards with missing X and Y coordinates between different team members, creation of a standardised spreadsheet for completion (Figure 2), and marking waypoints for the X and Y coordinates of archaeological sites for registration cards that lacked this data. Different sources were used to aid this final process: topographic maps in different scales (from 1:50,000 to 1:5,000), orthophotographs, satellite imagery, cadastral maps, and amateur websites with data for archaeological sites. As a result, in the period 2013-16 all registration cards received X and Y coordinates (Figure 3).

Figure 2: Table demonstrating standardised fields for describing the localisation of X and Y coordinates. Example includes information from AIS AKB paper registration cards Nos 0700170 and 0700171
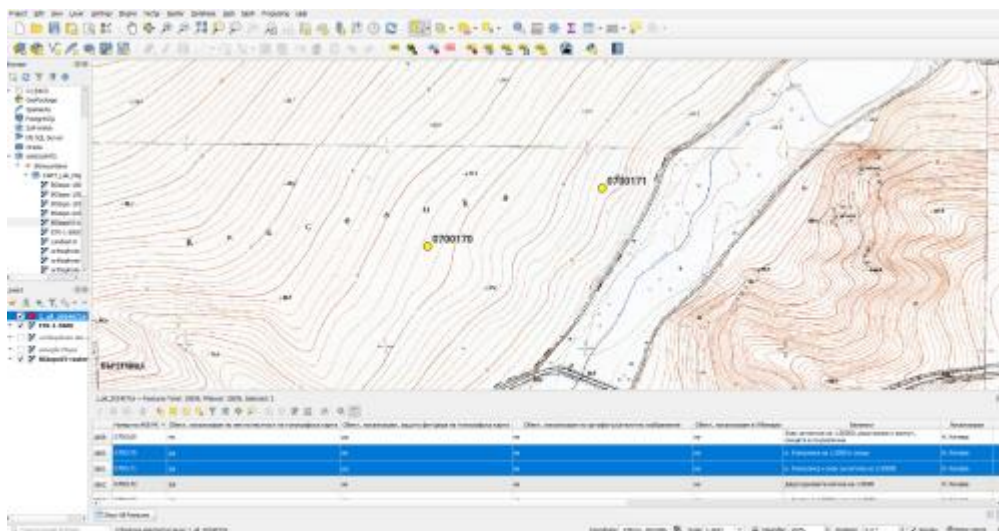

Figure 3: Locations of AIS AKB paper registration cards Nos 0700170 and 0700171 in GIS

3.2. Project requirements

The ARIADNE and ARIADNEplus projects' description of work included sharing metadata and data for 'data providers'. Different work packages covered the preparation of the AIS AKB data. Since AIS AKB still had strict authorised access according to regulatory documents (*Наредба № Н-2 от 6 април 2011 г. за създаване, поддържане и предоставяне на информацията от*

автоматизирана информационна система 'Археологическа карта на България' (обн. ДВ. бр. *32 от 19 април 2011 г., доп. ДВ. бр. 7 от 23 януари 2024 г.)* [2024]), it was decided that only partial data with modified metadata was going to be shared. For the ARIADNE project, data for 'popular' archaeological sites in Bulgaria was shared. This included fortresses, castles, churches, monasteries, tombs, rock-cut facilities, ancient and medieval towns, and megalithic monuments. Additional descriptions in English were specifically provided for each of these entries. For the ARIADNEplus project, data for large burial mounds and tells was added to the portal. The total number of records in ARIADNEplus was around 2000.

In order to fulfil the requirements for aggregating data (Richards [2023]), several work processes were followed. Some of these processes included the mapping of all type, find and facilities terms to the Getty AAT thesaurus as well as the inclusion of all archaeological periods present in Bulgaria in PeriodO. Additional work was done to the two XML exports (defined as one collection (AIS AKB) and many individual resources (archaeological sites)) created from AIS AKB with the shared data. This included modifying and mapping the metadata to the ACDM at first and the AO-Cat ontology later. The CIDOC CRM test mapping in 2016 was also done to the metadata of the original (not modified) AIS AKB registration cards.

3.2.1. Getty AAT mapping

The terms used in AIS AKB were mapped to the Getty AAT thesaurus (Art & Architecture Thesaurus® Online [2024]). The mapping was done using the W3C's SKOS reference (SKOS Simple Knowledge Organization System Reference [2009]) in the first ARIADNE project. The terms were simultaneously translated into English. These included type, finds and facilities sections in the AIS AKB. At this time, the mapping was done in a Microsoft Excel spreadsheet (Figure 4). Later, in ARIADNEplus, it was possible to execute this action online using the Vocabulary Matching Tool ([2024]) in the ARIADNEplus Gateway in D4Science (Binding and Tudhope [2024]).

Figure 4: AIS AKB terms mapped to Getty AAT thesaurus using SKOS in a MS Excel table

3.2.2. PeriodO inclusion

The updated chronologies used in AIS AKB were translated into English, defined by years 'before present' (BP) (2000), and included in another Microsoft Excel spreadsheet. The results are visible on the PeriodO website (PeriodO client. ARIADNE Consortium. ARIADNE Data Collection 2015) with spatial coverage 'Bulgaria'. There are 35 in total, and each has a different permalink with description.

3.2.3. AIS AKB modified metadata mapped to the ARIADNE catalogue data model (ACDM) in 2015-16

The ARIADNE catalogue data model (ACDM) was created during the ARIADNE project (Aloia *et al.* 2017). It was an extension of the Data Catalog Vocabulary (DCAT 2024). The mapping between the modified metadata of AIS AKB and the ACDM was undertaken in a Microsoft Excel spreadsheet using a template (Figure 5). The metadata from AIS AKB included text fields and one-to-one or one-to-many relationships. Thirty-one fields were mapped to the ACDM. Sixteen of them existed originally in AIS AKB and the other 15 were added.

Figure 5: AIS AKB modified and supplementary metadata mapped to ARIADNE ACDM in a MS Excel table

3.2.4. AIS AKB modified metadata mapped to the AO-Cat in 2019-21

The ACDM was updated to the AO-Cat ontology during the ARIADNEplus project (Felicetti *et al.* 2023) and the data aggregation pipeline that was followed is published in Bardi *et al.* 2022. The mapping of the AIS AKB data to the AO-Cat followed a much more user-friendly and clear workflow due to the ARIADNEplus gateway in D4Science, the templates uploaded there such as X3ML Mapping Tool (Mapping Memory Manager (3M) - DYAS 2024, Theodoridou *et al.* 2010), and the experience gathered in the previous ARIADNE project . Three examples of AIS AKB modified metadata to AO-Cat mapping are shown below. They include:

1. XML for the AIS AKB at a collection level (Figure 6),
2. XML of individual resources (burial mounds and tells),
3. XML for an individual resource (e.g., the archaeological site at Ratiara which was originally recorded in registration form 0100003 (Figure 7).

The files from the latter two examples were later combined into one XML file. Twenty-eight fields were exported for the individual resource XMLs (Figure 7). Thirteen fields from AIS AKB were exported and

mapped to twelve of the AO-Cat fields. The data in the other 16 fields from AO-Cat was manually added.

Figure 6: XML for the AO-Cat mapping of the AIS AKB collection as a whole

Figure 7: XML for the AO-Cat mapping of an individual resource (the well-known archaeological site of Ratiaria)

The fields from AIS AKB which were mapped to AO-Cat are as follows:

- name (=ao_cat_has_title);
- specifics (=ao_cat_has_description);
- created (=ao_cat_was_issued);
- modified (=ao_cat_was_modified);
- author (=ao_cat_has_creator);
- card number (=ao_cat_has_original_id);
- type (=ao_cat_is_about);
- imported (=ao_cat_was_created_on);
- latitude (=ao_cat_has_latitude);
- longitude (=ao_cat_has_longitude);
- chronology (only one) (=ao_cat_has_native_periods);
- type (only one), facilities (all), findings (all) (=ao_cat_has_native_subject) (more than one).

'Type' field was mapped twice (to =ao_cat_is_about and to ao_cat_has_native_subject). It is important to emphasise that the 'type' and 'chronology' fields in AIS AKB have a one-to-one relationship. In AO-Cat, only one type and its corresponding chronology (best represented) were used per individual record.

The XML for the collection record (the whole AIS AKB as an information system in one XML entry) was also created (Figure 6). The mappings of the XML (collection level and individual resource level) were done using the X3ML Mapping Tool (Mapping Memory Manager (3M) - DYAS 2024), Theodoridou *et al*. 2010) (Figure 8).
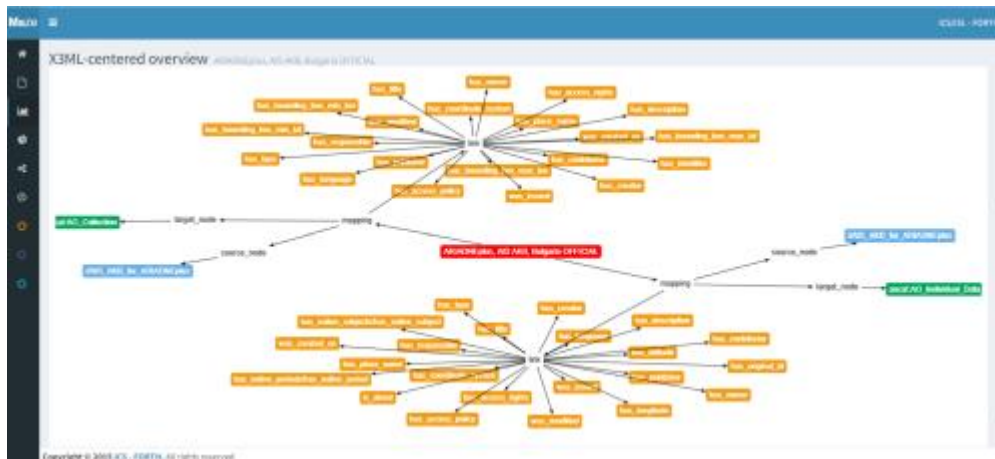
Figure 8: The mappings of AIS AKB collection and individual resources XMLs to AO-Cat visualised in Maze as X3ML-centred overview

3.2.5. AIS AKB registration card fields mapped to the CIDOC CRM - a test in 2016

The original metadata of the AIS AKB registration card was mapped to CIDOC CRM as a test in 2016. Some of the fields (30 in total) were first translated into English, then mapped to CIDOC CRM as part of E53_Place target node (Figure 9).
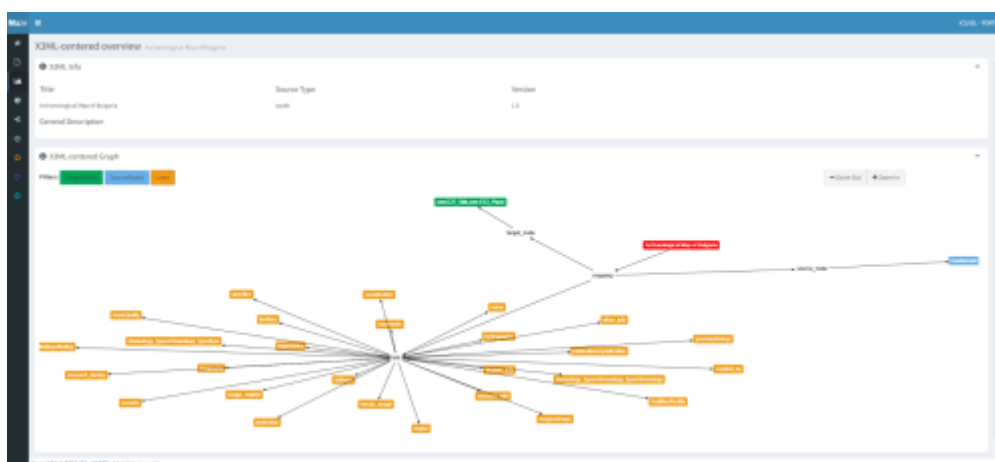

Figure 9: The mappings of the original AIS AKB metadata XML to AO-Cat visualised in Maze as X3ML-centred overview

3.3. Bonus: Geographic information system 'Archaeological Map of Bulgaria' (GIS AKB)

Creation of the geographic information system 'Archaeological Map of Bulgaria' (GIS AKB) started as a side project at the end of 2015 (Kecheva 2019 , 2024 , Географска информационна система „Археологическа карта на България" 2024). Its development included modification of AIS AKB into a web GIS platform. It turned out that a lot more transformations were applied. Geospatial features were created so that different geospatial data (points, lines, and polygons) can be collected. Registration cards continued to exist, but 'AKB types geographic features' were defined (Figure 10). Additional controlled vocabulary and free text fields were added. Mapping of GIS AKB to CIDOC CRM is yet to be fulfilled.

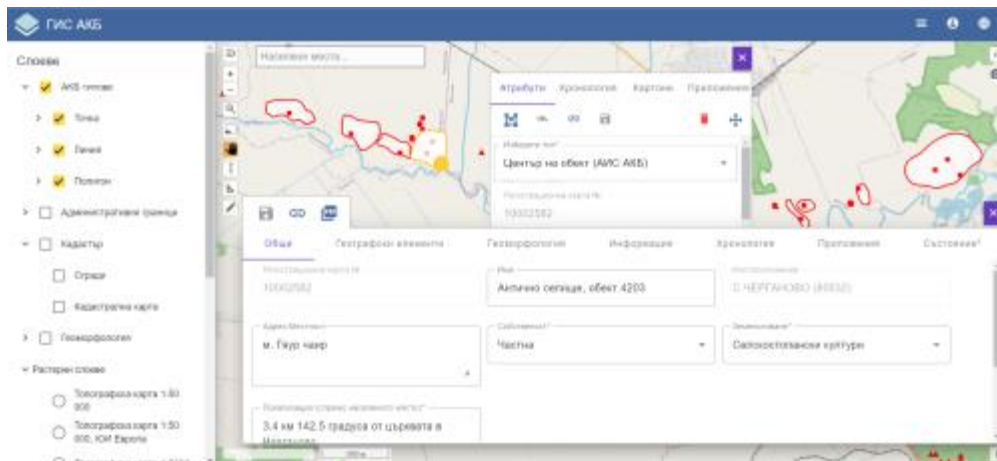Figure 10: Screenshot of GIS AKB demonstrating the registration card and geographic feature windows

4. Bulgaria in the ARIADNE portal

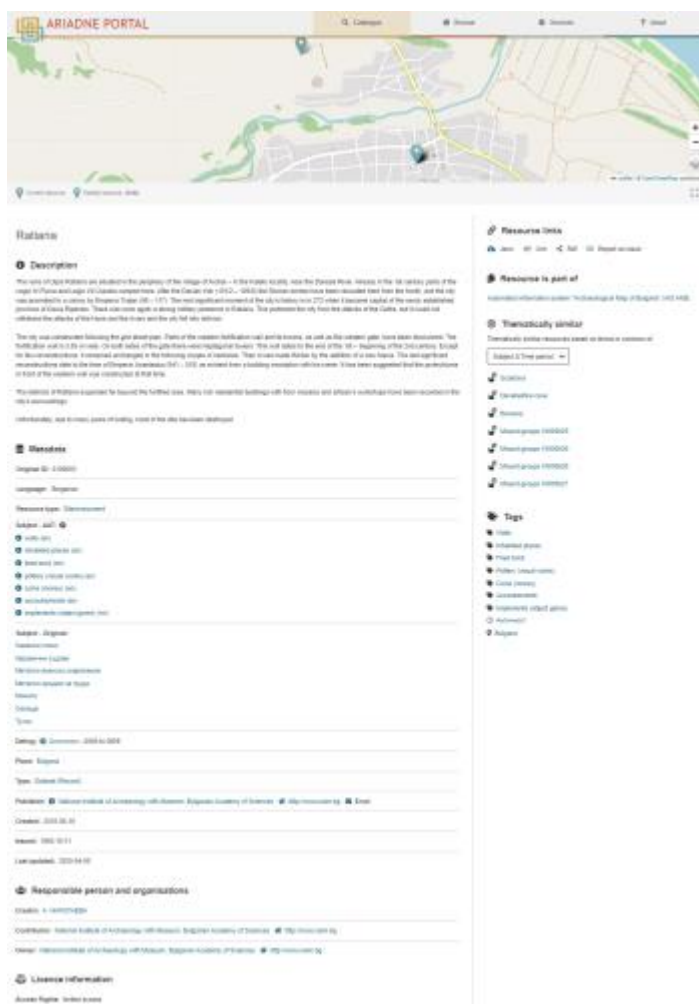The example of an individual resource shown in the XML in Figure 7 is visible in the ARIADNE portal (Figure 11).

Figure 11: The individual resource 'Ratiaria' from Figure 7 shown in the ARIADNE portal

NAIM-BAS is not the only contributor of data for Bulgaria in the ARIADNE portal. When searching by 'Place' for 'Bulgaria', only two publishers were found: NAIM-BAS and Heidelberg Academy of Sciences and Humanities. The latter provided data for 'The Role of Culture in Early Expansions of Humans' (ROCEEH) project with data from three caves (Browse / Where - Ariadne portal 2024). Three other institutions have published data for Bulgaria, but their data did not have any 'Place' assigned. These include The International Association for Classical Archaeology (using FastiOnline records) using 'Site/monument' and 'Fieldwork' resource type, Archaeology Data Service (1 record) with 'Fieldwork archive' resource type, and CENIEH (1 record) with 'Fieldwork report' resource type.

5. Conclusion

The ARIADNE and ARIADNEplus projects offered a lot to the partners: sharing knowledge and experience was the most important benefit. From a practical point of view, the completion of the digitisation of the paper AIS AKB registration cards, mapping of AIS AKB to CIDOC CRM, and beginning to develop the GIS AKB web platform were the most useful and important tasks achieved. The shared metadata and data in the ARIADNE portal allow the FAIR-ness of parts of AIS AKB. Since mapping can be misleading and different researchers can do it individually as they prefer, a lot of thought and work was done to ensure that the shared AIS AKB 'site and monuments' data was as

informative and accurate as possible. This was the reason why the metadata needed modification before mapping. Data's findability is easily checked when searching through the ARIADNE portal, but there is still a long road to go. The author's participation in the ARIADNE and ARIADNEplus projects and SEADDA COST Action was essential for gaining experience in the field of standards, metadata mapping, and the first step towards making part of Bulgarian archaeology more FAIR.

Bibliography

Aloia, N., Debole, F., Felicetti, A., Galluccio, I., and Theodoridou, M. 2017 'Mapping the ARIADNE catalogue data model to CIDOC CRM: Bridging resource discovery and item-level access', *SCIRES-IT* **1**, 1-8. https://doi.org/10.2423/i22394303v7n1p1

ARIADNE support portal [online] 2013 http://ariadne-support.dcu.gr/#model [Last accessed: 1 November 2024].

Art & Architecture Thesaurus® Online 2024 *The Getty Research Institute*. http://www.getty.edu/research/tools/vocabularies/aat/ [Last accessed: 23 August 2024].

Aspöck, E. and Richards, J.D. (eds) 2023 'International data aggregation for archaeological research and heritage management: the ARIADNE experience', *Internet Archaeology* **64**. https://intarch.ac.uk/journal/issue64/

Bardi, A., Binding, C., Felicetti, A., Meghini, C., Richards, J., Theodoridou, M., and Kritsotakis, V. 2022 ;ARIADNEplus Data Aggregation Pipeline: User Guide (2.4)', Zenodo. https://doi.org/10.5281/ZENODO.8060925

Binding, C. and Tudhope, D. 2024 'Vocabulary mapping for archaeological infrastructure', *Journal of Information Science*. https://doi.org/10.1177/01655515241283610

Browse / Where - Ariadne portal [online] 2024 https://portal.ariadne-infrastructure.eu/browse/where?mapq=true&q=bulgaria&fields=location [Last accessed: 11 October 2024].

CIDOC CRM (Conceptual Reference Model) [online] 2024 http://www.cidoc-crm.org/ [Last accessed: 4 August 2024].

Data Catalog Vocabulary (DCAT) [online] 2024 W3C https://www.w3.org/TR/vocab-dcat/ [Last accessed: 4 October 2024].

Felicetti, A., Meghini, C., Richards, J., and Theodoridou, M. 2023 'The AO-Cat Ontology (1.2)', Zenodo. https://doi.org/10.5281/ZENODO.7818375

Jakobsson, U., Novák, D., Richards, J.D., Štular, B., and Wright, H. (eds) 2021 'Digital archiving in archaeology: The state of the art', *Internet Archaeology* **58**. https://intarch.ac.uk/journal/issue58/

Jakobsson, U., Novák, D., Richards, J.D., Štular, B., and Wright, H. (eds) 2023 'Digital archiving in archaeology: Additional state of the art and further analyses', *Internet Archaeology* **63**. https://intarch.ac.uk/journal/issue63/

Kecheva, N. 2019 'Archaeological Map of Bulgaria - transport and pipeline infrastructure projects', *Internet Archaeology* **51**. https://doi.org/10.11141/ia.51.2

Kecheva, N. 2021 'Current state of archiving archaeological reports in Bulgaria: Still more physical than digital', *Internet Archaeology* **58**. https://doi.org/10.11141/ia.58.24

Kecheva, N. 2024 *Archaeological Map of Bulgaria. Non-destructive methods for archaeological fieldwork activities and the application of geographic information systems*, Dissertationes, Vol. 18, Sofia: National Archaeological Institute with Museum - BAS, Prof. Marin Drinov Publishing House of BAS. https://doi.org/10.53250/diss18 (in Bulgarian with English summary).

Mapping Memory Manager (3M) - DYAS 2024 https://en.dyas-net.gr/tools-and-services/mapping-memory-manager-3m/ [Last accessed: 9 November 2024].

PeriodO client. ARIADNE Consortium. ARIADNE Data Collection [online] 2015 http://n2t.net/ark:/99152/p0qhb66 [Last accessed: 28 September 2024].

Richards, J.D. 2023 'Joined up thinking: Aggregating archaeological datasets at an international scale', *Internet Archaeology* **64**. https://doi.org/10.11141/ia.64.3

SKOS Simple Knowledge Organization System Reference [online] 2009 W3C https://www.w3.org/TR/skos-reference/#mapping [Last accessed: 11 October 2024].

Theodoridou, M., Marketakis, Y., Minadakis, N., Samaritakis, G., and Konsolaki, K. 2010 'Mapping Memory Manager (3M). Instance and Label Generator Rule Builder'. https://mapping.d4science.org/3M/Manuals/en/X3ML_Generators_Manual.pdf [Last accessed: 13 November 2024].

Vocabulary Matching Tool [online] 2024 https://vmt.ariadne.d4science.org/vmt/vmt-app.html [Last accessed: 9 November 2024]

Автоматизирана информационна система 'Археологическа карта на България' [online] 2024 http://www.naim-bas.com/akb/ [Last accessed: 19 August 2024].

Географска информационна система „Археологическа карта на България" [online] 2024 https://akb.naim.bg/ [Last accessed: 19 August 2024].

Наредба № Н-2 от 6 април 2011 г. за създаване, поддържане и предоставяне на информацията от автоматизирана информационна система 'Археологическа карта на България' (обн. ДВ. бр. 32 от 19 април 2011 г., доп. ДВ. бр. 7 от 23 януари 2024 г.) [online] 2024 https://www.lex.bg/laws/ldoc/2135726967 [Last accessed: 19 July 2024].

Нехризов, Г. 2014 'Археологическа карта на България - история, достижения и проблеми', Сборник доклади. В дните на българската наука. Кръгла маса на тема: Културно историческо наследство, влияние на науката и образованието върху социално-икономическото развитие на град Сливен и Сливенски регион, 24-25 март 2014 г., Сливен: Строителни конструкции ООД. 15-41.