



CORPUS NUMMORUM – A Digital Research Infra- structure for Ancient Coins

Ulrike Peter*¹, Claus Franke¹, Jan Köster¹, Karsten Tolle²,
Sebastian Gampe², Vladimir F. Stolba^{1, 3}

¹ Berlin-Brandenburg Academy of Sciences and Humanities – Berlin, Germany

² Big Data Lab, Goethe University Frankfurt – Frankfurt am Main, Germany

³ Staatliche Museen zu Berlin Münzkabinett

*Corresponding author: Correspondence: peter@bbaw.de

ABSTRACT

CORPUS NUMMORUM indexes ancient Greek coins from various landscapes and develops typologies. The coins and coin types are published on the multilingual website www.corpus-nummorum.eu utilizing numismatic authority data and adhering to FAIR principles. As an integral part of the project, the CN Editor has been developed as a multifunctional web app capable of managing the project's data entry workflow. It offers extensive search functions and various evaluation options. This open-source tool has a modular structure, allowing for swift extensions with new functions or adaptation to other object types, making it appealing for projects beyond the field of numismatics.

Keywords: numismatics, ancient Greek coins, typologies, versioning system

The Corpus Nummorum Project

Corpus Nummorum (CN) is a web portal and a powerful research tool dedicated to the study of ancient coins spanning from around 600 BC to AD 300. This brief overview introduces the fundamental concept of the portal www.corpus-nummorum.eu¹ and outlines its technical strategies, which could prove valuable for projects involving various categories of mass material, particularly those combining image and text.

Corpus Nummorum is a collaborative project involving the Berlin-Brandenburg Academy of Sciences and Humanities, the Münzkabine Berlin, and the Big Data Lab of Goethe University Frankfurt.² The project has received funding through various grants for over a decade.³

CN actively participates in international collaborative efforts aimed at creating a corpus of Greek coin types within a Linked Open Data framework (<https://greekcoinage.org>). The first version of the overarching ARCH portal, focusing on coins from the pre-Roman period, has recently been launched (<https://greekcoinage.org/arch>). Additionally, CN contributes to the ongoing international initiative for the development of typologies for Roman Provincial coinage (<https://rpc.ashmus.ox.ac.uk>). Geographically, CN has thus far covered four historical regions, including Thrace, Moesia Inferior, Mysia, and the Troad on both sides of the Dardanelles.

Differing from traditional collection databases, CN operates on multiple levels—ranging from individual coins and coin dies to coin types and issues. Each coin, die, and coin type is assigned a unique and stable resource identifier (Uniform Resource Identifier – URI). The multilingual portal indexes ancient Greek coins from various collections and develops typologies using numismatic authority data, especially from nomisma.org. With the provision of a Resource Description Framework (RDF) representation, the data is not only accessible via the CN website but also through the mentioned ARCH portal, and as RDF dumps via [Nomisma.org](http://nomisma.org) (see: <http://nomisma.org/datasets>).⁴

It aims at including as many existing coins of the above regions as possible from both public and private collections, as well as from commerce⁵ in order to create comprehensive online catalogues of coin types. The fact that ancient coins are available in large quantities and dispersed all over the world, makes this an enormous task. Supported by a range of grants, each typically tailored to specific themes,⁶ the portal provides varying levels of coverage for different regions. As an example, a comprehensive collection of coins from Thrace has been compiled since 2012, leading to the development of nearly complete typologies. Consequently, this region currently boasts the most comprehensive representation.

Regarding the adjacent regions, a three-year funding initiative, beginning in 2018, focused on digitizing and documenting coins from Moesia Inferior, Mysia, and the Troad in the Berlin Coin Collection. This project also included the digitization of plaster casts of coins from these regions held in the Berlin-Brandenburg Academy of Sciences and Humanities. Wherever possible, efforts have been made to complement this data with coins from other collections, such as the Bibliothèque nationale de France, the [SHH collection](#), etc. However, it is important to note that the collection of the material for Moesia, Mysia, and the Troad is not yet comprehensive for the respective mints. Hence, attaining a comprehensive typology will require additional time.

The website <https://data.corpus-nummorum.eu> provides insight into the general technical structure of Corpus Nummorum. Its relationally structured dataset is stored in a MySQL database, accessible through a RESTful JSON API. Additionally, it is mapped to the RDF, the technical standard for the formal description of

¹ All links were checked on May 29, 2023.

² See also <https://www.corpus-nummorum.eu/team> and [Peter 2017](#), [Peter/Weisser 2017](#). For recent German language presentation of the project, see [Peter et al.](#) (forthcoming).

³ See Section “Funding” below.

⁴ The datasets are machine readable, and the data is published under a CC BY license – see details below. We also promote our portal through other portals, publications, conferences, and social media, aiming not only to link to other sources but also to encourage other parties to include links to our coin types and coins in their datasets. How else CN adheres to the FAIR principles, see further down.

⁵ Studying ancient coinages is inconceivable without material from auctions and private collections, given the significant role these collections have always played in numismatics. The most successful results consistently arise through collaboration among professional numismatists, museums, private collections, and the coin trade. The inclusion of coins from commerce not only significantly expands the datasets but also highlights often rare and unique materials that might otherwise be at risk of being lost. Simultaneously, we strongly condemn illegal excavations and violations of cultural heritage protection. See [Elkins 2022](#) with further literature and publication guidelines, especially p. 60.

⁶ See Section “Funding” below.

resources. This mapping allows for additional searchability via the graph-based query language SPARQL. All relevant fields are linked to authority data of the numismatic repository, <http://nomisma.org>, enabling exchange via Linked Open Data (LOD). The public website <https://www.corpus-nummorum.eu> offers data on coins, dies, coin types, literature, and typologies for general users. The IIIF server facilitates access to IIIF data of images.

To adhere to the FAIR principles,⁷ we employ descriptive metadata for enhanced findability, utilize Persistent Identifiers (URIs for data and DOIs for GitHub repositories) and incorporate standard vocabularies such as nomisma, geonames, etc. Open Access and user-friendly formats facilitate accessibility, with data presented in commonly used, open, and machine-readable formats like RDF, XML, etc., ensuring compatibility with various tools and platforms. Interoperability is maintained through standardized formats and protocols, including the provision of an API enabling seamless interaction with our data for easy access by other systems. Data reuse is facilitated by a CC-BY license. Additionally, we offer comprehensive documentation detailing the structure, content, and usage of our data. Regular updates and curation guarantee accuracy and relevance over time, with any changes or versions clearly indicated (see below).

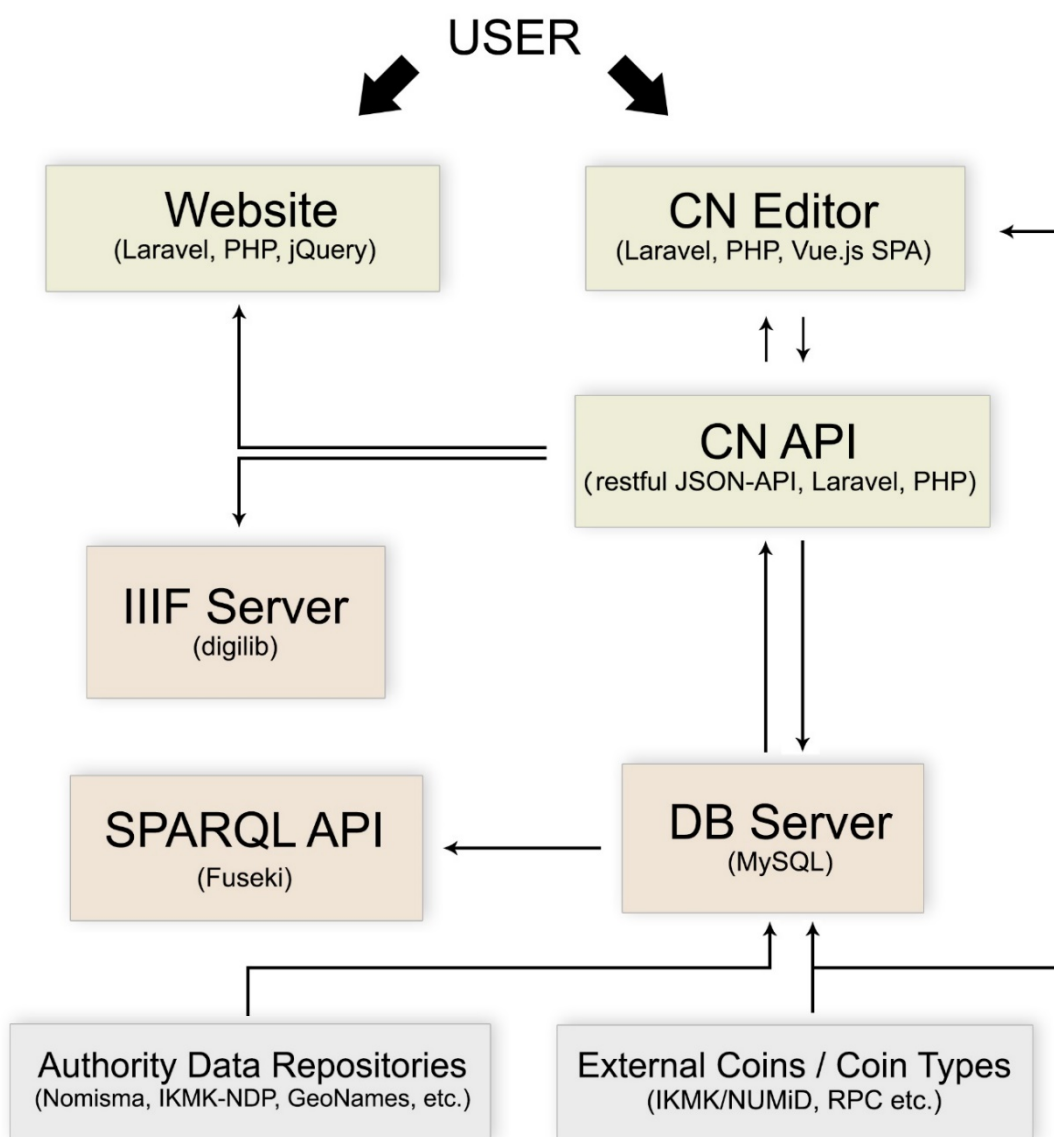


Figure 1 – CN infrastructure.

⁷ For the first publication of the FAIR principles as a guide for enhancing data reusability, see [Wilkinson et al. 2016](#) and [2019](#). We express our gratitude to Ronald Visser for the reference to these publications.

CN Editor

As part of the project, the CN Editor was developed, a multifunctional web application for creating and managing coin and coin type datasets. This powerful and useful tool can also be employed for other projects dealing with serial data. Its modular structure facilitates easy adaptation and extension. The editor operates as a single-page application (SPA), built on Laravel and Vue.js, and currently supports English and German.

The CN Editor enables users to modify and edit all data associated with coins, coin types, and related entities, including mints, material, owners, persons, rulers, denominations, etc. An overview of our complex data model can be found at <https://www.corpus-nummorum.eu/resources/database>.⁸ The editor facilitates the structured entry of comprehensive coin information, encompassing details such as monograms, symbols, and legend directions. Users can organize coins by types or create groups and issues. The multi-inherit options streamline data entry and structuring, allowing coins to inherit information from corresponding coin types. Additionally, there is a one-click data import feature from diverse sources and partners, including the [IKMK family](#), the [American Numismatic Society](#), and coins listed in platforms like <https://www.coinarchives.com> and <https://www.acsearch.info>. Existing export interfaces for coin collections and regex-based full-text recognition for auction coin websites are leveraged. Rapid and straightforward import of images linked to the dataset is supported. Images combining obverse and reverse views from certain portals are automatically separated into two files. This efficient system enables the swift addition of new coins to existing types, as only metrological and provenance data need to be imported, with type-specific data inherited from pre-existing types. Imported data can be customized to meet the system's requirements.

Images are organized into sets that comprise obverse and reverse pictures of a coin. A coin may have more than one image set, such as photos of an original coin and a plaster cast. Uploaded images can be modified by rotating, zooming, or moving their contents. These modifications do not affect the original file, which remains unchanged. Instead, these changes are only applied to the thumbnails. In addition, the modification parameters are stored in a separate JSON file, enabling the original to be transformed on the fly when accessed through our IIIF image service (Digilib).

Aligned with the regional focus of the portal, CN also documents coin hoards, capturing necessary parameters (composition, chronology, reference, etc.) and geographical data. Links to hoards published in relevant portals (<http://coinhoards.org>, <https://chre.ashmus.ox.ac.uk>) are provided.

⁸ Our definition of a coin type is highly detailed; see [Peter 2019](#), 399–403.

CN Favorites Types Coins Features Individuals Locations Tools Website Help EN













Edit on hoard 47

Findspot: Gradeshnitsa

Description: 431 AR. Pot hoard, including 400 coins of Thasos, 1 of Maroneia, 15 of Apollonia, 15 of Dyrhachium

Reference: Gerasimov, BIAB 27 (1964), p. 237

TYPES (12)

 <p>cn type 12884 * cn.thasos_1_ed.138</p> <p>Linked Coins: 166</p> <p>Thasos 39 mm (68) 16.56 g (166) 148-90/80 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn type 20806 * cn.thasos_1_ed.120</p> <p>Linked Coins: 40</p> <p>Thasos 35 mm (20) 16.59 g (40) 168/167-148 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn type 20807 * cn.thasos_1_ed.121</p> <p>Linked Coins: 13</p> <p>Thasos 34 mm (7) 16.51 g (13) 168/167-148 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn type 20812 * cn.thasos_1_ed.122</p> <p>Linked Coins: 10</p> <p>Thasos 36 mm (3) 16.85 g (10) 168/167-148 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn type 20912 * cn.thasos_1_ed.139</p> <p>Linked Coins: 39</p> <p>Thasos 36 mm (23) 16.49 g (39) ca. 148-80 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn type 20942 * cn.thasos_1_ed.127</p> <p>Linked Coins: 6</p> <p>Thasos 33 mm (6) 16.23 g (6) 148-90/80 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>
 <p>cn type 20943 * cn.thasos_1_ed.128</p> <p>Linked Coins: 3</p> <p>Thasos 34 mm (2) 16.57 g (3) 148-90/80 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn type 20944 * cn.thasos_1_ed.129</p> <p>Linked Coins: 1</p> <p>Thasos 32 mm (1) 16.18 g (1) 148-90/80 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn type 20946 * cn.thasos_1_ed.131</p> <p>Linked Coins: 2</p> <p>Thasos 32 mm (2) 16.57 g (2) 148-90/80 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn type 20947 * cn.thasos_1_ed.132</p> <p>Linked Coins: 2</p> <p>Thasos 31 mm (2) 16.48 g (2) 148-90/80 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn type 20949 * cn.thasos_1_ed.134</p> <p>Linked Coins: 1</p> <p>Thasos 33 mm (1) 16.54 g (1) 148-90/80 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn type 21027 * cn.thasos_1_ed.137</p> <p>Linked Coins: 43</p> <p>Thasos 34 mm (21) 16.59 g (42) 148-90/80 v. Chr.; silver; tetradrachm</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>

COINS (1-12 / 76)













 <p>cn coin 6154 *</p> <p>Inherited from cn type 20812</p> <p>Thasos 33 mm 16.85 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ Nude Heracles standing facing, head lef...</p>	 <p>cn coin 6155 *</p> <p>Linked to cn type 20806</p> <p>Thasos 34 mm 16.9 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>	 <p>cn coin 6156 *</p> <p>Linked to cn type 20806</p> <p>Thasos 35 mm 16.88 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>	 <p>cn coin 6157 *</p> <p>Linked to cn type 20806</p> <p>Thasos 33 mm 16.87 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>	 <p>cn coin 6158 *</p> <p>Linked to cn type 20806</p> <p>Thasos 31 mm 16.49 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>	 <p>cn coin 6159 *</p> <p>Linked to cn type 20806</p> <p>Thasos 32 mm 16.56 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>
 <p>cn coin 6163 *</p> <p>Linked to cn type 20806</p> <p>Thasos 31 mm 16.95 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>	 <p>cn coin 6164 *</p> <p>Inherited from cn type 20807</p> <p>Thasos 31 mm 16.6 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>	 <p>cn coin 6165 *</p> <p>Inherited from cn type 20807</p> <p>Thasos 31 mm 16.78 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>	 <p>cn coin 6166 *</p> <p>Inherited from cn type 20807</p> <p>Thasos 31 mm 16.73 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>	 <p>cn coin 6167 *</p> <p>Linked to cn type 20806</p> <p>Thasos 31 mm 16.8 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>	 <p>cn coin 6168 *</p> <p>Linked to cn type 20806</p> <p>Thasos 31 mm 16.8 g 168/167-148 v. Chr.; silver; tetradrachm; Vr...</p> <p>Head of (youthful) Dionysus, right, weari... R ΗΡΑΚΛΕΟΥΣ ΙΔΙΘΡΟΟΣ ΒΑΣΙΔΝ // ΒΑΣΙΔΝ // ΙΔΙ... Nude Heracles standing facing, head lef...</p>

Figure 2 – Example of an entry for a hoard with linked coin types and coins in the CN Editor.

As of recently, the portal also allows for indexing coin forgeries to warn users and prevent the inclusion of counterfeits.⁹ The existence of forgeries has always been a problem in numismatics. Nowadays, distinguishing many forgeries from original coins has become increasingly difficult. Therefore, it is crucial to identify counterfeits to understand which types have been falsified and to improve the overall quality of the data.

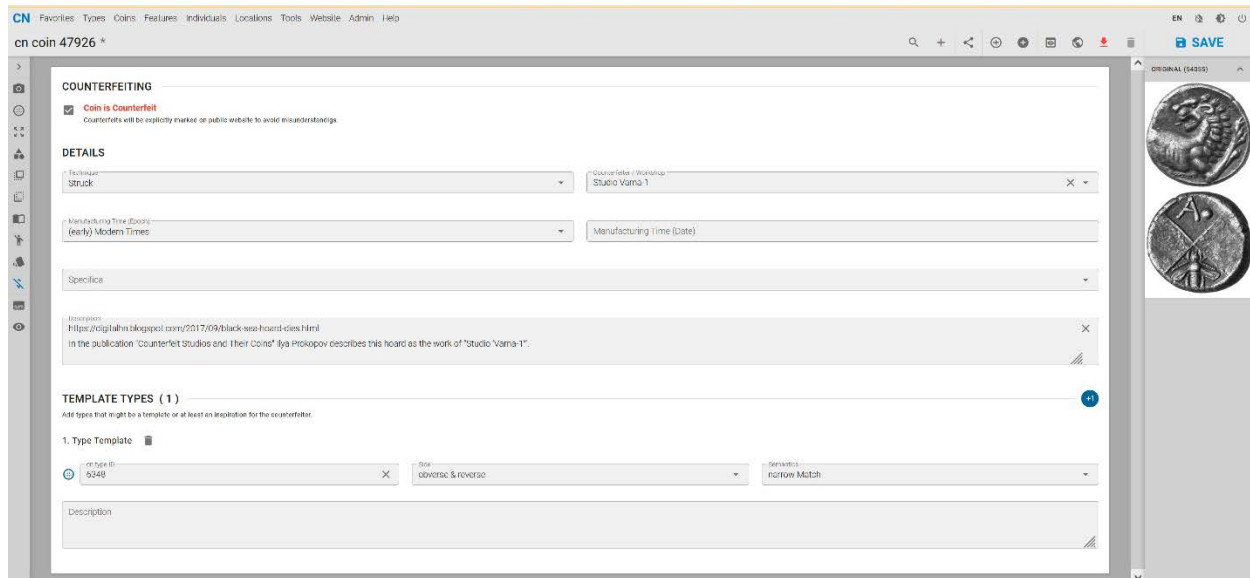


Figure 3 – Example of an entry for a counterfeit coin in the CN Editor.

Adding new coins to our portal is straightforward, and we encourage collectors and other interested parties to join us in sharing their coins. By publishing them on our website, material that might otherwise be lost to scholarly research becomes known.

While access to our web portal is open and unrestricted, a simple registration is necessary to access the Editor (via the 'Participate' button on the website). This allows for the input and modification of data records and information to be assigned to the respective individuals. Additionally, users can be granted varying levels of rights, ranging from basic read-only access (including coins not yet released for the external site) to various write permissions. For example, collectors may be given the right to input and modify only their own coins, providing reassurance that they will not accidentally delete data belonging to others. Other authorization levels encompass modifying all entries associated with the coins. Further rights include creating and modifying authority data and typologies, as well as news and the 'Coin of the Month' feature on the website. The highest level extends to user management through administrator rights.

All published data is accessible via the JSON API, covering coins, coin types, dies, designs, legends, and auxiliary tables such as mints, find spots, materials, etc. The API enables searches across all entities, offering up to 70 different search options and filtering options.

The CN portal facilitates the identification and description of relevant coins. Individual specimens can be assigned by linking to CN dies, types, and series, hierarchically structured and rigorously described. Simultaneously, the growth of the CN database allows for the identification of new types and the correction of existing ones (e.g., based on better readable monograms or legends on new specimens). As mentioned, collectors and collections have the option to publish their relevant coins directly (or via spreadsheet uploads) in the Corpus Nummorum database. A simple click on the button with their collection provides an online presentation of their data, along with a display of records linked to a map and timeline. The portale enables the creation of coin, die, and coin type corpora for specific mints, addressing various research questions through numerous search and filter options. CN functions not only as a research database with a range of options but also allows interested users to create their own database or independent instance on their computers or servers using the freely available editor, which can easily be adapted to specific needs and

⁹ Boteva-Boyanova et al. (forthcoming).

expanded with new functions. In line with the Open Access and Open Source policy,¹⁰ the code and documentation of the CN Editor are available on GitHub/zenodo for any projects related to the collection and analysis of ancient coins or other serial data.¹¹

In the future, we aim to enhance and expand the export functions and interfaces so that our typologies can be utilized in collection databases to describe the coins within them.

Versioning System

For the unambiguous citation and continuous development of the typologies, we have established a versioning, or edition, system.¹² The edition number follows a specific format, composed of “cn” (Corpus Nummorum), followed by the issuer of the coin in the form of the nomismaID (nomisma-mint or nomisma-ruler_ID), the edition number (version_number_ed) and a unique number indicating the sort order (sequence_number). For instance, “cn.maroneia.1_ed.31” refers to the 31st coin type in the first edition of the coin typology of Maroneia. The criteria for arrangement are similar to those used in a printed catalogue, and these sort orders are available on our website at <https://www.corpus-nummorum.eu/resources/typology-sortorder>. All recorded material is open source for linking to other digital data, visualization, and further research inquiries (also via SPARQL).

In addition to the conscious decision-based edition system, we employ an automatic versioning system that documents even small changes, such as the correction of a spelling mistake. If a published coin or type record is modified, the current state is stored as a separate JSON file, using CN ID and timestamp for the file name. All versions can be directly accessed via the API. It is also possible to input any arbitrary timestamp; in this case, the API will determine the version’s validity on the given date and provide the corresponding data. The individual versions are linked to the editions, ensuring that older data states can always be reconstructed and cited independently of the constantly changing database.

Further Development

Currently, CN is undergoing further development as part of our projects, “Data Quality for Numismatics Based on Natural Language Processing and Neural Networks (D4N⁴)” and “Iconography and AI-Methods in Numismatics”. The goal is to establish artificial intelligence research tools for numismatics, which can be used independently or integrated into the CN Editor. Our approaches are particularly suitable for other object categories that combine images and text and are available in large quantities.

These digital tools aim to enhance data quality and facilitate the more efficient recording of coins accessible on the web. Text and image recognition are employed to expedite this process. Since July 2021, we have been testing machine-learning methods for typological coin assignments. This includes: 1) natural language processing (NLP) methods for multilingual and non-standardized coin descriptions, linking them to a hierarchical iconographic thesaurus; this combination enables more sophisticated semantic searches for specific coin and coin type data; 2) complex image recognition of entire design compositions and their individual elements using deep learning.¹³

¹⁰ <https://edoc.bbaw.de/frontdoor/index/index/docId/3366>.

¹¹ <https://doi.org/10.5281/zenodo.10458195> – For technical reasons, the publicly available version of the CN Editor represents an older state and does not include all the functions of our productive instance. A larger release with additional functions and an optimised setup is in preparation. See <https://data.corpus-nummorum.eu/wiki>, for more information about the CN Editor source code and its features.

¹² For the importance of versioning the scholarly data, see also [Bürgermeister 2022](#).

¹³ For details on training the computer and its initial accuracy results, see [Gampe 2021](#). Additionally, in these CAA proceedings, our colleagues Sebastian Gampe and Karsten Tolle present a special paper on this machine learning method, based on our coin dataset. The paper discusses the challenges faced by image recognition models when identifying mint classes, which have significantly more images than other mints: [Gampe/Tolle](#) (forthcoming [b]).

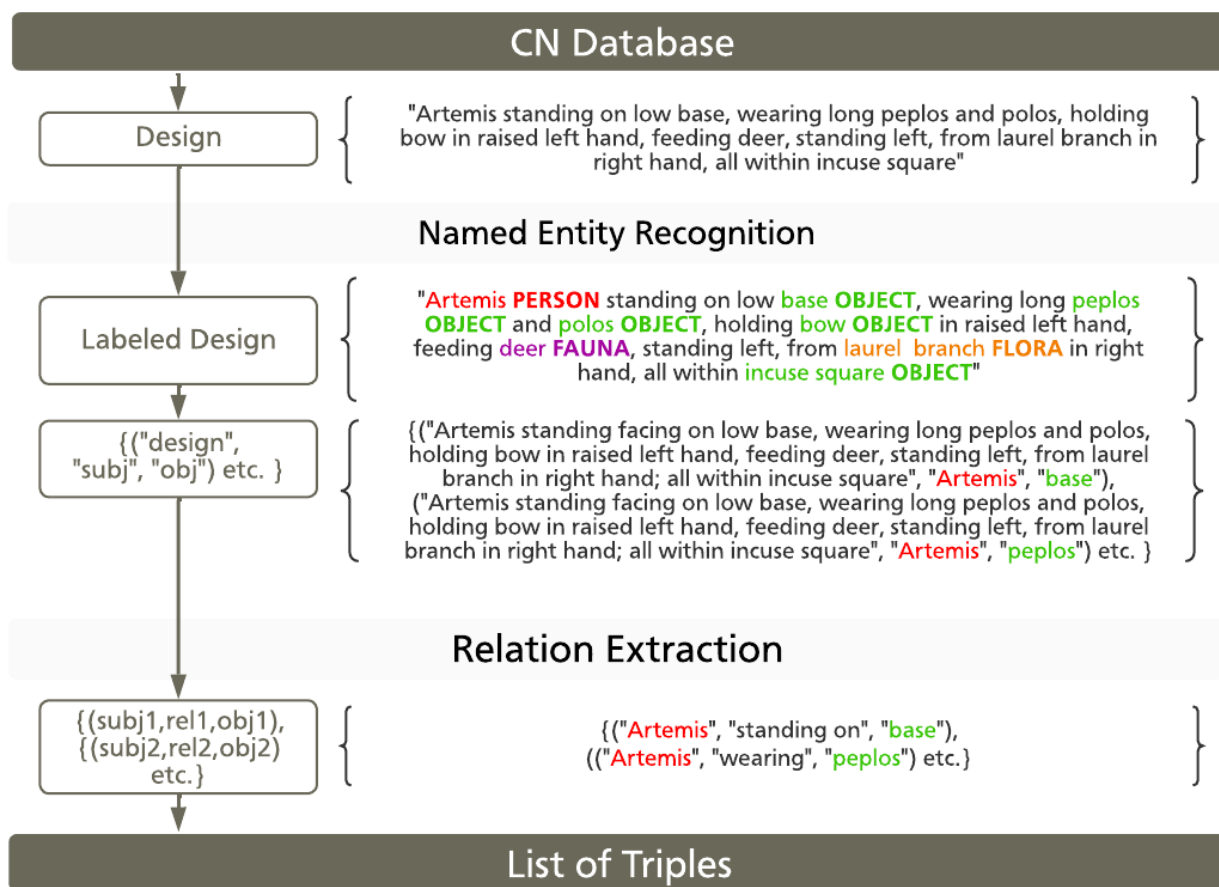


Figure 4 – Example of a description of a coin image and the Named Entity Recognition and Relation extraction of the NLP pipeline.

Convolutional Neural Networks (CNNs) have been trained to add new coins to coin types and mints available in the database. In addition to coin types, the goal is to differentiate patterns and elements on coins, such as monograms, persons, objects, etc.¹⁴ Our high-quality training and testing datasets, essential for the successful development of machine-learning-based image recognition methods like Convolutional Neural Networks (CNN), has been published recently.¹⁵ Currently, the varying degree of preservation of coins and the similarity of different images present challenges in applying AI models. However, these models have immense potential in numismatics for automated image annotation, analysis, interpretation, and retrieval.

To get a preview of how these tools function, one can explore Google Colab. For NLP, one can enter or upload a coin description to see which named entities (NER and RE) the program can detect (link to GitHub containing the code and the link to Google Colab for testing: <https://github.com/Frankfurt-BigDataLab/NLP-on-multilingual-coin-datasets>). Alternatively, one can upload a photo of a coin (obverse and reverse separately) to predict the mint or coin type from the CN dataset (link to GitHub containing the code and the link to Google Colab for testing: <https://github.com/Frankfurt-BigDataLab/IR-on-coin-datasets>).

Promising initial results have already been published, but we aim to enhance recognition by combining these methods.¹⁶ A multimodal approach should not only improve recognition but also enable monitoring of data quality. The project has successfully implemented quality controls to supervise the accuracy of the growing data.¹⁷

¹⁴ See [Luong 2022](#).

¹⁵ [Corpus_Nummorum 2023](#).

¹⁶ [Klinger et al. 2018](#); [Gampe/Tolle](#) (forthcoming [a]).


¹⁷ Utilizing the nomisma.org/RDF modeling of CN data, a tool has been developed that performs various quality checks (consistency checks) in the form of SPARQL queries. This tool is reusable by others. For more detail, see [Peter/Tolle](#) (forthcoming).

Help Projectinfo Language

CORPUS NUMMORUM ONLINE


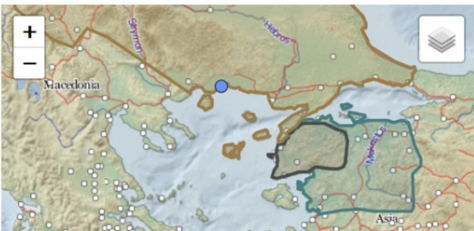
Coins Types Map Collections Literature Resources Participate

Return to Search Results



original - License: CC BY 4.0 - Photographer: Lübke & Wiedemann

Enlarged Download

cn coin 53053

Connected to Type	CN Type 6019
Mint	Abdera Region: Thrace Typology
Date	c. 360-350 BC Classical Period
Obverse	<p>Legend: ΑΒΔΗΠΙ</p> <p>Design: Griffin lying left, right forepaw raised. go to the NLP result of this description</p>
Reverse	<p>Legend: ΕΠΙ ΠΟΛΥΚΡΑΤΕΟΣ</p> <p>Design: Artemis standing facing on a base, wearing kalathos, holding bow in raised left hand, feeding deer, standing left, from branch in right hand; all within incuse square. go to the NLP result of this description</p>
Persons	Magistrate Polykrates (II)
Metrology	<p>Denomination: Double siglos (tetradrachm)</p> <p>Material: silver</p> <p>Weight: 11.39 g</p>
Owner	commerce / im Handel, , Please note that the current ownership is not confirmed. Provenience: Peus 355 (27.04.1998) 697.

Figure 5 – Access to the NLP pipeline is possible directly from the coin description on the web portal.

Additionally, our goal is to establish an authority data portal for ancient Greek coin iconography—a hierarchical *Thesaurus Iconographicus Nummorum Graecorum* (ThING). We have initiated a prototype of a hierarchical thesaurus in both English and German. It primarily relies on CN data with the standardized design descriptions, but integration of other coin descriptions and images is possible through data harvesting in the form of LOD. The objective is to index entire image compositions, as well as individual elements. ThING will evolve as a portal for standardized iconographic data, with each entry illustrated by corresponding coin types and coins. To facilitate more detailed identification of different images and cross-genre semantic and iconological analyses, we plan to establish links to similar representations in other object classes such as gems, pottery, sculpture, etc. Moreover, links to digital image repositories will encompass resources relevant to the reception of ancient coins.

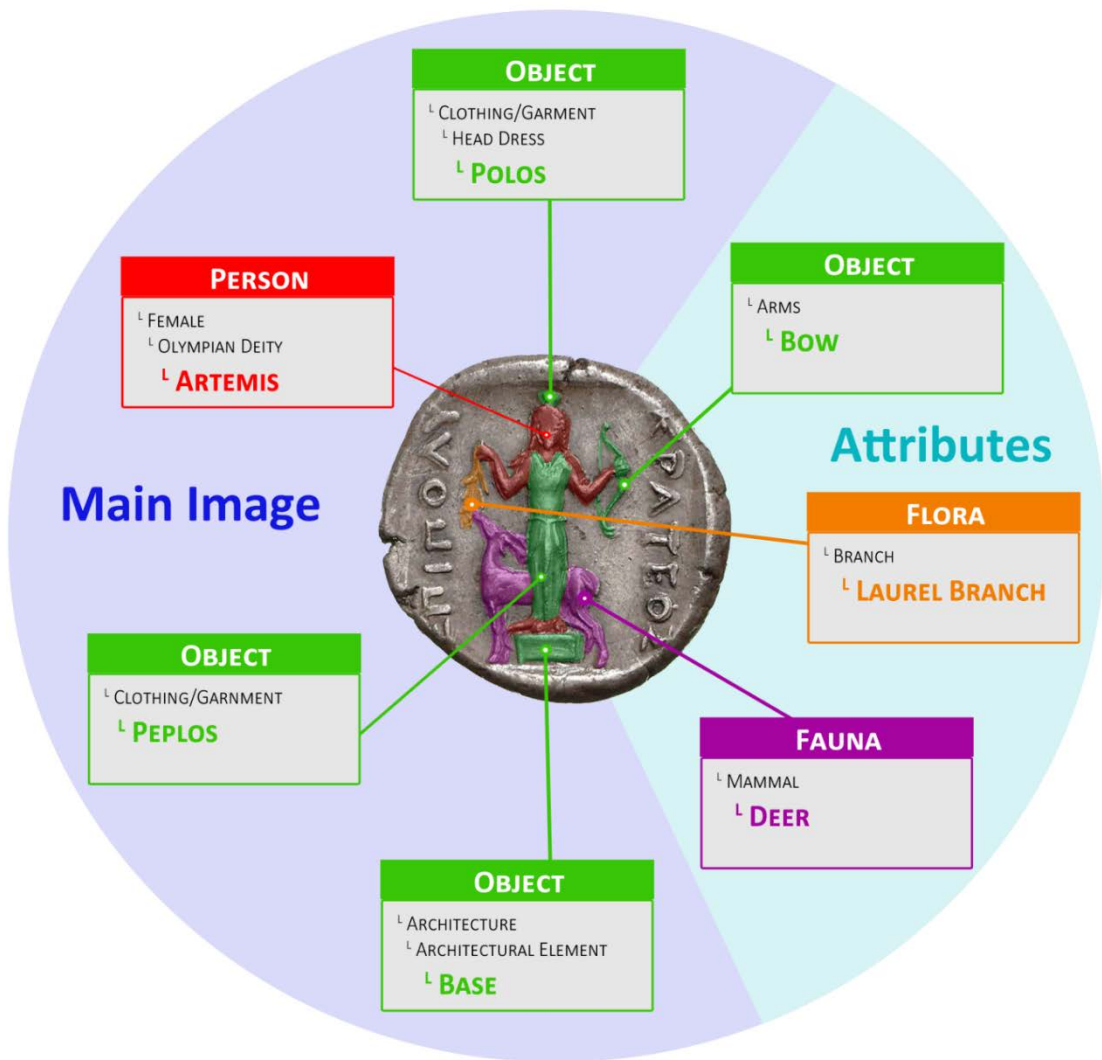


Figure 6 – Example of a coin image with labeled entities, which should be identified through NLP and, in the future, also through object detection. These entities are linked with the hierarchical iconographical thesaurus.

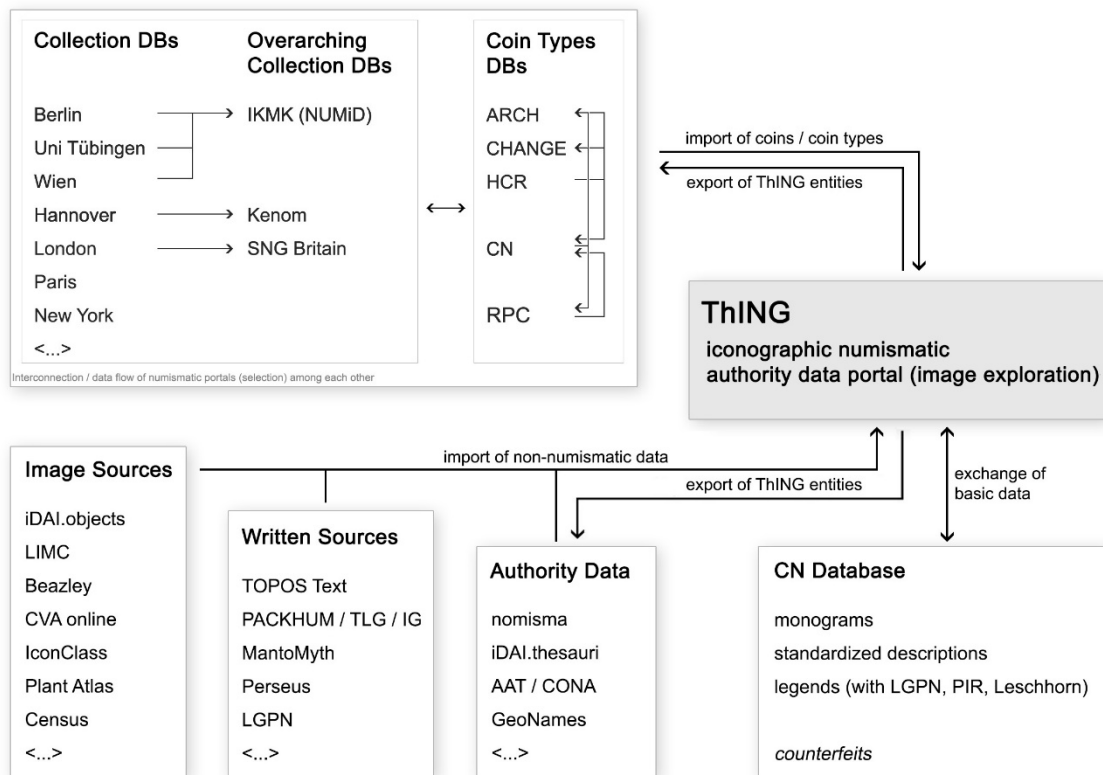


Figure 7 – Schema for the in- and output of the *Thesaurus Iconographic Nummorum Graecorum* (THING).

Given the large and continuously increasing number of coins available online, manual marking is impractical. Therefore, it is necessary to train the developed algorithm for semi-automatic marking. The goal is to ensure that different descriptions of the same image types, especially more complex compositions like depictions of the heroic deeds of Heracles, are recognized. This allows for the comparison of coins with each other and with similar scenes on other object categories. The combination of NLP and the iconographic thesaurus will enable new semantic searches. In addition, a new graphical user interface should make semantic searches possible, even for users without SPARQL experience.¹⁸

We encourage you to utilize our website for your research. Currently, we already offer more than 30,700 coins and 11,800 coin types online. In the internal database, an additional 20,000 coins and over 6,000 coin types are awaiting the final quality control before publication. Check the CN portal regularly for both new coins and coin types, as well as updates on new developments.

¹⁸ Wirth 2019.

Acknowledgements

We thank the entire [CN team](#) for their enthusiasm and support, without which this project could not have been realized. We are also grateful to Ronald Visser, Fleur Kemmers, and an anonymous reviewer who reviewed this paper and suggested a number of amendments.

Data, scripts, code, and supplementary information availability

Scripts and code are available online at <https://github.com/telota/corpus-nummorum-editor> and <https://github.com/telota/corpus-nummorum-editor/blob/public/DOCUMENTATION.md>. See also <https://doi.org/10.5281/zenodo.10458195> and <https://data.corpus-nummorum.eu/wiki>.

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Funding

The D4N⁴ project “Datenqualität für Numismatik basierend auf Natural Language Processing und Neuronalen Netzen / Data Quality for Numismatics Based on Natural Language Processing and Neural Networks” is funded by the Deutsche Forschungsgemeinschaft (DFG) / German Research Council within the “e-research-technologies / Technologien” program. The IKKINUM project “Iconographie und KI-Methoden in der Numismatik / Iconography and AI-Methods in Numismatics” has been supported by the Berlin Senate, Department for Science, Health, Care, and Equality. In its earlier phases, Corpus Nummorum has also received funding from the DFG and the Federal Ministry of Education and Research (see [Peter 2020](#), 296–299 and <https://gepris.dfg.de/gepris/projekt/203766501>; <https://gepris.dfg.de/gepris/projekt/317085821>; <https://www.tib.eu/de/suchen/id/TIBKAT:1809906555/Corpus-Nummorum-Online-die-antiken-griechischen?chHash=ee3cc81ca50e070a46ac49243f2bdb46>; <https://gepris.dfg.de/gepris/projekt/449148027>).

References

- Boteva-Boyanova D, Prokopov I, Köster J, Flueck JE, Grigorova-Gencheva V (forthcoming) RT 8: Ancient Coins Counterfeits, in: *Proceedings of the XVI International Numismatic Congress, September 2022*. Warsaw.
- Bürgermeister M (2020) Enabling the Scholarly Discourse of the Future: Versioning RDF Data in the Digital Humanities, in: Andrews T, Diehr F, Efer T, Kuczera A, van Zundert J (eds.), *Graph Technologies in the Humanities – Proceedings 2020*, published at <http://ceur-ws.org>, 1–17 (<https://ceur-ws.org/Vol-3110/paper1.pdf>).
- Corpus_Nummorum (2023). Corpus Nummorum – Coin Image Dataset (1.0) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.10033993>.
- Elkins N T (2022) Provenance and Legal Issues, in: Alram M, Bodzek J, Bursche A (eds.), *Survey of Numismatic Research, 2014-2020* (Warsaw, Krakow, Winterthur), 55–64.
- Gampe S (2021) *Neuronale Netze zur Bestimmung römischer Kaiser auf Bildern antiker Münzen* (Master Thesis Universität Frankfurt/M.) (http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit_Sebastian_Gampe_online.pdf).

- Gampe S, Tolle K (forthcoming [a]) Combination of machine learning methods of image and natural language recognition on ancient coin data. *Proceedings Computer Applications and Quantitative Methods in Archaeology (CAA) conference in Kraków 2019*.
- Gampe S, Tolle K (forthcoming [b]) Creating an additional class layer with machine learning to counter overfitting in an unbalanced ancient coin dataset, in: *Proceedings of CAA conference in Amsterdam 2022/23*.
- Klinger P, Gampe S, Tolle K, Peter U (2018) Semantic Search based on Natural Language Processing – a Numismatic example, *Journal of Ancient History and Archaeology* 5.3, 68–79. <https://doi.org/10.14795/j.v5i3.334>.
- Köster J, Franke C (2024) Corpus Nummorum Editor. <https://doi.org/10.5281/zenodo.10458195>
- Luong H (2022) Klassifizierung von Motiven auf antiken Münzen mit Mask R-CNN (Master Thesis Universität Frankfurt/M.) (<http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2023/02/Masterthesis-Huy-Luong.pdf>).
- Peter U (2017) Corpus Nummorum Thracorum – A Research Tool for Thracology and an Example of Digital Numismatic Collaboration, in: Caccamo Caltabiano M. (ed.) (2017). *XV International Numismatic Congress Taormina 2015. Proceedings*. Rom–Messina, Vol. 2, 1306.
- Peter U (2019) Von Mommsen zum Semantic Web: Perspektiven der vernetzten numismatischen Forschung – die Münzen der westlichen Schwarzmeerküste online, in: Cojocaru V, Ruscu L, Castelli Th, Pázsint A-I (eds) (2018), *Advances in Ancient Black Sea Studies: Historiography, Archaeology and Religion. The Proceedings of the International Symposium, Constanța, August 20–24, 2018*. Cluj-Napoca (Pontica et Mediterranea VIII), 393–418.
- Peter U (2020) Das Münzkabinett und die Berlin-Brandenburgische Akademie der Wissenschaften – Eine enge Partnerschaft, in: Weisser B (ed.) *Münzkabinett – Menschen, Münzen, Medaillen*, Regenstauf, 285–300.
- Peter U, Tolle K (forthcoming) Corpus Nummorum – Coins, types and data quality control, in: *Proceedings of the 8th Joint Meeting of ECFN and nomisma.org 2019*.
- Peter U, Weisser B (2017) CNT – The web portal for Thracian coins, in: Caccamo Caltabiano M (ed.) (2017) *XV International Numismatic Congress Taormina 2015. Proceedings*. Rom–Messina, Vol. 1, 247.
- Peter U, Adamik-Köster J, Berthold A, Franke C, Gampe S, Gorys A, Stolba V, Tolle K, Weisser B (forthcoming) Eine digitale Forschungsinfrastruktur für antike Münzen: neue Entwicklungen im Projekt Corpus Nummorum, in: Bursche A et al. (eds) *Proceedings of the XVI International Numismatic Congress, September 2022*. Warsaw.
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. (2016) The FAIR Guiding Principles for scientific data management and stewardship, *Scientific Data* 3(1): 160018. <https://doi.org/10.1038/sdata.2016.18>.
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S-A, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. (2019) Addendum: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 6(1): 6. <https://doi.org/10.1038/s41597-019-0009-6>.
- Wirth A (2019) *SPARQL-Abfrageschnittstelle ikonografischer Daten im Resource Description Framework*, Bachelorarbeit, Frankfurt/M. (<https://github.com/aliciawrt/IGSApplication/blob/master/SPARQL-Abfrageschnittstelle%20ikonografischer%20Daten%20im%20Resource%20Description%20Framework.pdf>).