

# CORPUS NUMMORUM – A Digital Research Infrastructure for Ancient Coins

Ulrike Peter\*<sup>1</sup>, Claus Franke<sup>1</sup>, Jan Köster<sup>1</sup>, Karsten Tolle<sup>2</sup>, Sebastian Gampe<sup>2</sup>, Vladimir F. Stolba<sup>1, 3</sup>

<sup>1</sup> Berlin-Brandenburg Academy of Sciences and Humanities – Berlin, Germany

<sup>2</sup> Big Data Lab, Goethe-University – Frankfurt am Main, Germany

<sup>3</sup> Staatliche Museen zu Berlin Münzkabinett

\*Corresponding author

Correspondence: peter@bbaw.de

## ABSTRACT

CORPUS NUMMORUM indexes ancient Greek coins from various landscapes across collections and develops typologies. The coins and coin types are published on the multilingual website [www.corpus-nummorum.eu](http://www.corpus-nummorum.eu) using numismatic authority data and FAIR principles. As part of the project, the Editor was developed as a multifunctional web app that can fully handle the project's data entry workflow and provides extensive search functions as well as various evaluation options. This open-source tool has a modular structure. As a result, the CN Editor can be quickly extended with new functions or adapted to other object types, which makes it interesting for projects beyond numismatics.

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

## The Corpus Nummorum Project

29 Corpus Nummorum (CN) is a web portal and a powerful research tool for the study of ancient coins from  
30 around 600 BC to AD 300. This brief overview presents the basic concept of the portal [www.corpus-](http://www.corpus-nummorum.eu)  
31 [nummorum.eu](http://www.corpus-nummorum.eu)<sup>1</sup> and its technical strategies, which – as little “minions” – could also be used by projects working  
32 with other categories of mass material, especially those combining image and text.

33 Corpus Nummorum is a joint project of the Berlin-Brandenburg Academy of Sciences and Humanities, the  
34 Münzkabinett Berlin, and the Big Data Lab of the University of Frankfurt.<sup>2</sup> It has been funded by different  
35 grants for over 10 years.<sup>3</sup>

36 CN is involved in international, collaborative efforts to create a corpus of Greek coin types using a Linked  
37 Open Data framework (<https://greekcoinage.org/>). The first version of the overarching ARCH portal for the  
38 coins of the pre-Roman period has just been published (<https://greekcoinage.org/arch/>). CN also participates  
39 in the development of typologies for the Roman Provincial coinage, which is an ongoing international  
40 endeavour (<https://rpc.ashmus.ox.ac.uk/>). Its geographical coverage embraces thus far four historical regions,  
41 including Thrace and Moesia Inferior, as well as Mysia and the Troad on both sides of the Dardanelles.

42 Unlike traditional collection databases, CN operates on different levels, from single coins through coin dies  
43 to coin types and issues. Each coin, die, and coin type receive a unique and stable resource identifier (Uniform  
44 Resource Identifier – URI). The multilingual portal indexes ancient Greek coins from various collections and  
45 develops typologies using numismatic authority data (especially from [nomisma.org](http://nomisma.org)) and FAIR principles.

46 We try to include as many existing coins of those regions as possible from museums and other collections,  
47 but also from auctions and private collectors in order to create as complete as possible complete online coin  
48 type catalogues. But given the fact that ancient coins are available in large quantities and dispersed all over  
49 the world, this is an enormous task. Funded by a series of grants, which had usually a special focus,<sup>4</sup> the portal  
50 offers different degrees of coverage for different regions. For Thrace, as an example, a comprehensive  
51 collection of coins has been compiled since 2012, and nearly complete typologies were developed. Thus, this  
52 region is currently the best represented.

53 As to the adjacent regions, a three-year funding starting in 2018 aimed at digitizing and recording the coins  
54 of Moesia Inferior, Mysia, and the Troad from the Berlin Coin Collection, as well as the plaster casts of coins  
55 from these regions kept in the Berlin-Brandenburg Academy of Sciences and Humanities. Wherever possible,  
56 an attempt has been made to supplement this data with coins from other collections such as Paris, the SHH  
57 collection, etc., but the collection of the material for Moesia, Mysia, and the Troad is far from being  
58 comprehensive for the respective mints. Therefore, achieving a complete typology will take more time.

59 The website <https://data.corpus-nummorum.eu/> gives an idea of the general technical structure of Corpus  
60 Nummorum. Its relationally structured dataset is stored in a MySQL database, accessible via a RESTful JSON  
61 API. It is also mapped to the Resource Description Framework (RDF), the technical standard for the formal  
62 description of resources, and thus additionally searchable via the graph-based query language SPARQL. All  
63 relevant fields are linked to authority data of the repository relevant for numismatics, <http://nomisma.org>,  
64 allowing exchange via Linked Open Data (LOD). The public website <https://www.corpus-nummorum.eu/>  
65 provides the data on coins, dies, coin types, literature, and typologies for general users. The IIIF server  
66 facilitates access to IIIF data of images.

67

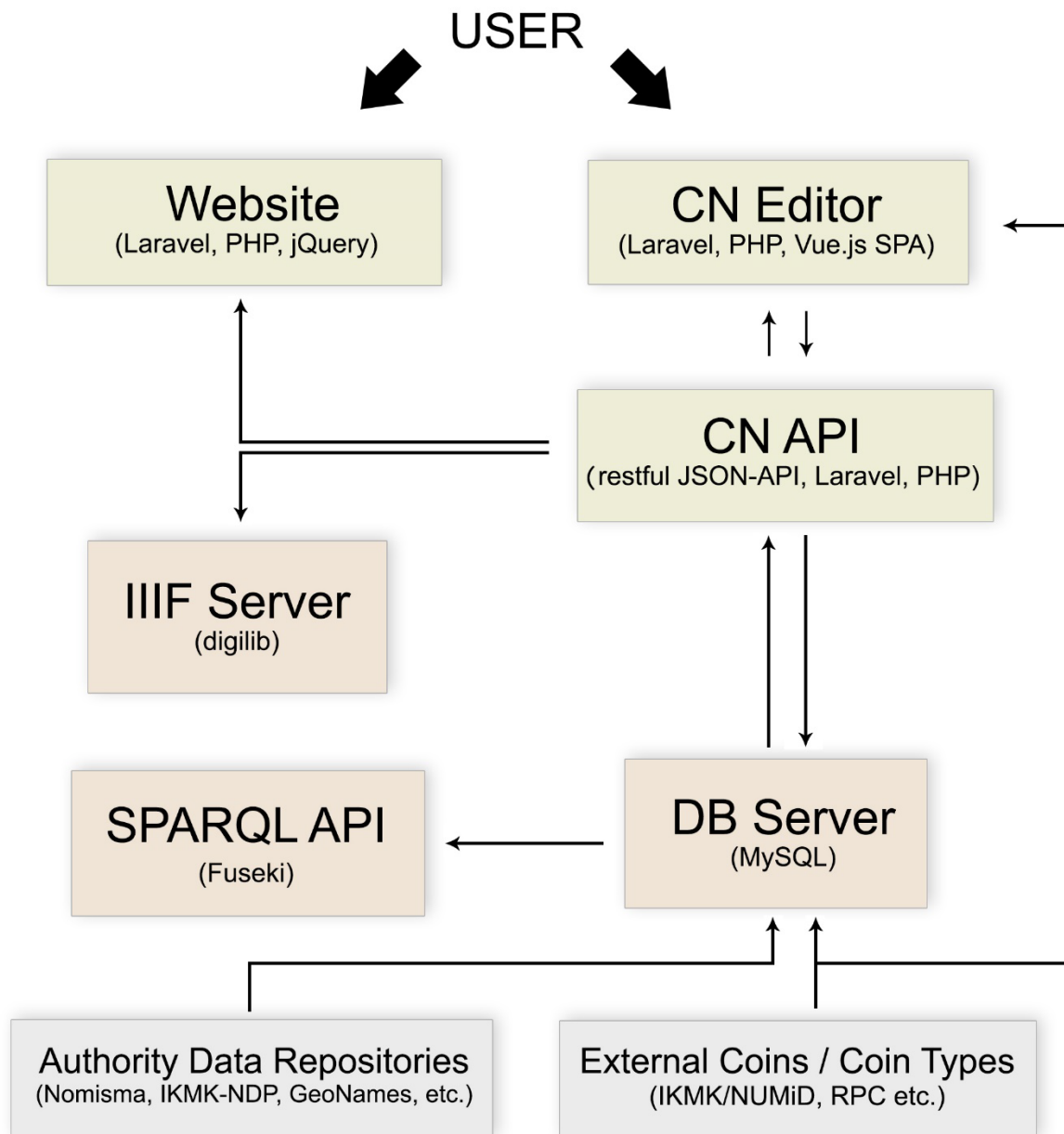
---

<sup>1</sup> All links were checked on May 29, 2023.

<sup>2</sup> 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).

<sup>3</sup> For funding, we especially thank the German Research Foundation, the Federal Ministry of Education and Research, and the Berlin Senate, Department for Science, Health, Care, and Equality (see also 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?cHash=ee3cc81ca50e070a46ac49243f2bdb46>, <https://gepris.dfg.de/gepris/projekt/449148027>).

<sup>4</sup> See n. 3.



68  
69 **Figure 1** – CN infrastructure.

70

71

### CN Editor

72 As part of the project, the CN Editor was developed, which represents a multifunctional web application  
73 for creating and managing the coin and coin type datasets. It is a powerful and useful tool that can serve as a  
74 “minion” for other projects dealing with serial data. Its modular structure makes it easy to adapt and to extend.  
75 The editor is a single-page application (SPA) based on Laravel and Vue.js. It is multilingual, currently supporting  
76 English and German.

77 The CN Editor allows users to modify and edit all data related to coins, coin types, and associated entities  
78 such as mints, material, owner, persons, rulers, denominations, etc. An overview of our complex data model  
79 can be found here: <https://www.corpus-nummorum.eu/resources/database>.<sup>5</sup> The editor allows for entering

<sup>5</sup> Our definition of a coin type is very finely granular, see Peter 2019, 399–403.

80 all information about a coin in a structured way, including e.g. monograms, symbols, and legend directions.  
81 One can group coins according to coin types or create groups or issues. Very convenient for data entry and  
82 structuring are the multi-inherit options, where coins can inherit data from the corresponding coin type.  
83 Furthermore, there is a one-click data import available from many different sources and from various partners,  
84 such as the [IKMK family](#), the [American Numismatic Society](#), or coins in commerce indexed in  
85 <https://www.coinarchives.com/>, <https://www.acsearch.info/>, etc. Existing export interfaces of coin collections  
86 as well as the regex-based full text recognition for auction coin websites are used. In addition, images linked  
87 to the dataset can be imported quickly and easily. The images combining an obverse and a reverse provided  
88 by some portals will be automatically cut into two separate files. All this allows for a very fast addition of new  
89 coins to existing types, for only the metrological and provenance data need to be imported, whereas the type-  
90 specific data will be inherited from the already existing types. Imported data can be adapted to the  
91 requirements of the system.

92 Images are organized in sets that include obverse and reverse pictures of a coin. There can be more than  
93 one image set for a coin, for example, photos of an original coin and a plaster cast. An uploaded image can be  
94 modified by rotating, zooming or moving its contents. However, these modifications will not affect the original  
95 file, which remains unchanged. Instead, these changes are only applied to the thumbnails. In addition, the  
96 modification parameters are stored in a separate JSON file that allows the original to be transformed on the  
97 fly when accessed via our IIF image service (Digilib).

98 According to the regional focus of the portal, CN also records coin hoards with the required parameters  
99 (composition, dating, reference, etc.) and geographical data. Links to hoards published in relevant portals  
100 (<http://coinhoards.org>, <https://chre.ashmus.ox.ac.uk>) are provided.

101

CN Favorites Types Coins Features Individuals Locations Tools Website Hilfe EN

Edit cn hoard 47

Name: IGCH 571

Link: http://coinhoards.org/d/igch0571

Datierung: 130-100 BC

Kommentar

Fundort: Gradeshnitsa

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

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

TYPES (12)

cn type 12884 \*  
cn.thasos\_1\_ed.138  
Linked Coins: 166  
Thasos  
39 mm (68) 16.56 g (166)  
148-90/80 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 20806 \*  
cn.thasos\_1\_ed.120  
Linked Coins: 40  
Thasos  
35 mm (20) 16.59 g (40)  
168/167-148 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 20807 \*  
cn.thasos\_1\_ed.121  
Linked Coins: 13  
Thasos  
34 mm (7) 16.51 g (13)  
168/167-148 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 20812 \*  
cn.thasos\_1\_ed.122  
Linked Coins: 10  
Thasos  
36 mm (3) 16.85 g (10)  
168/167-148 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 20912 \*  
cn.thasos\_1\_ed.139  
Linked Coins: 39  
Thasos  
36 mm (23) 16.49 g (39)  
ca. 148-80 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 20942 \*  
cn.thasos\_1\_ed.127  
Linked Coins: 6  
Thasos  
33 mm (6) 16.23 g (6)  
148-90/80 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 20943 \*  
cn.thasos\_1\_ed.128  
Linked Coins: 3  
Thasos  
34 mm (2) 16.57 g (3)  
148-90/80 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 20944 \*  
cn.thasos\_1\_ed.129  
Linked Coins: 1  
Thasos  
32 mm (1) 16.18 g (1)  
148-90/80 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 20946 \*  
cn.thasos\_1\_ed.131  
Linked Coins: 2  
Thasos  
32 mm (2) 16.57 g (2)  
148-90/80 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 20947 \*  
cn.thasos\_1\_ed.132  
Linked Coins: 2  
Thasos  
31 mm (2) 16.46 g (2)  
148-90/80 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 20949 \*  
cn.thasos\_1\_ed.134  
Linked Coins: 1  
Thasos  
33 mm (1) 16.54 g (1)  
148-90/80 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn type 21027 \*  
cn.thasos\_1\_ed.137  
Linked Coins: 43  
Thasos  
34 mm (21) 16.69 g (42)  
148-90/80 v. Chr.; silver; tetradrachm  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

COINS (1-12 / 76)

cn coin 6154 \*  
Inherited from cn type 20812  
Thasos  
33 mm 16.85 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ ΕΙΘΗΡΟΣ ΘΑΙΩΝ  
Nude Heracles standing facing, head lef...

cn coin 6155 \*  
Linked to cn type 20806  
Thasos  
34 mm 16.9 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

cn coin 6156 \*  
Linked to cn type 20806  
Thasos  
35 mm 16.88 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

cn coin 6157 \*  
Linked to cn type 20806  
Thasos  
33 mm 16.87 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

cn coin 6158 \*  
Linked to cn type 20806  
Thasos  
31 mm 16.49 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

cn coin 6159 \*  
Linked to cn type 20806  
Thasos  
32 mm 16.96 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

cn coin 6163 \*  
Inherited from cn type 20806  
Thasos  
31 mm 16.95 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

cn coin 6164 \*  
Inherited from cn type 20807  
Thasos  
31 mm 16.6 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

cn coin 6165 \*  
Inherited from cn type 20807  
Thasos  
31 mm 16.70 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

cn coin 6166 \*  
Inherited from cn type 20807  
Thasos  
31 mm 16.73 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

cn coin 6167 \*  
Linked to cn type 20806  
Thasos  
31 mm 16.8 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

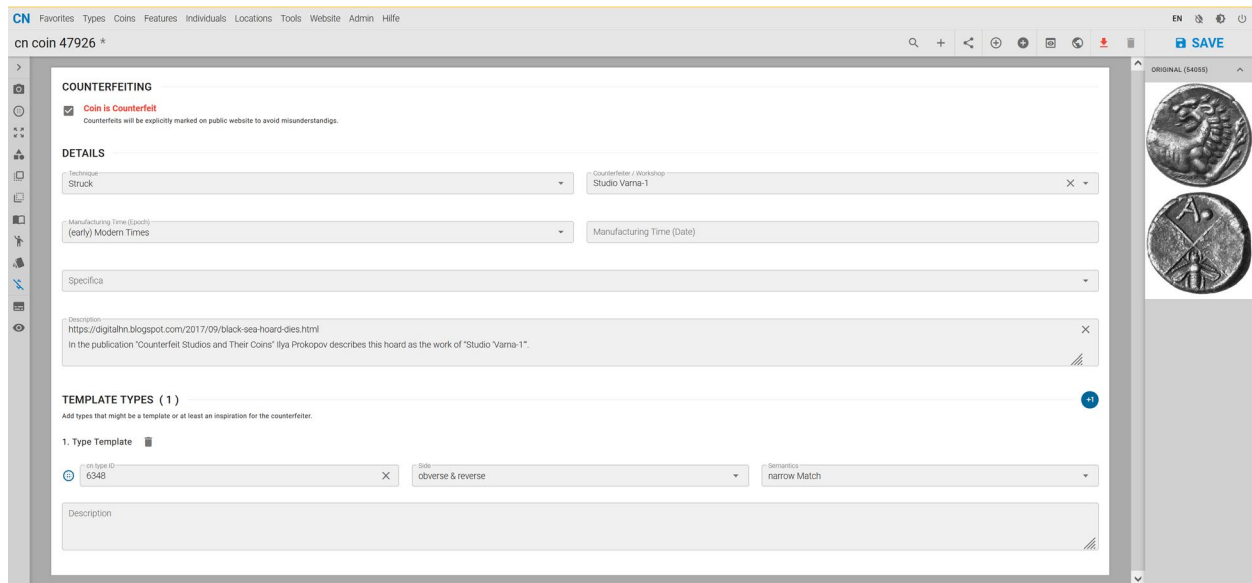
cn coin 6168 \*  
Linked to cn type 20806  
Thasos  
31 mm 16.8 g  
168/167-148 v. Chr.; silver; tetradrachm; Vr...  
Head of (youthful) Dionysus, right, wear...  
R ΗΡΑΚΛΕΟΥΣ - ΕΙΘΗΡΟΣ // ΘΑΙΩΝ // ΕΙ...  
Nude Heracles standing facing, head lef...

102

103

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

104 As of recently, the portal also allows for indexing coin forgeries in order to warn users and to achieve the  
105 exclusion of counterfeits.<sup>6</sup>  
106



107  
108 **Figure 3** – Example of an entry for a counterfeit coin in the CN Editor.

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

112 The CN portal allows users to more easily identify and describe relevant coins. Individual specimens can be  
113 assigned by linking to the CN dies, types, and series, which are hierarchically structured and largely stringently  
114 described. On the other hand, the growth of the CN database makes the identification of new types and the  
115 correction of existing types (e.g. based on better readable monograms or legends on new specimens) possible.  
116 If they want to, collectors and collections can publish their relevant coins directly in the Corpus Nummorum  
117 database. Just one click on the button with their collection will provide an online presentation of their data  
118 and, along with a simple display of records, an output linked to a map and timeline. Coin, die, and coin type  
119 corpora of specific mints can be created with the help of the portal. Numerous search and filter options allow  
120 for addressing different research questions. CN can thus be used as a research database with a variety of  
121 options. However, interested users can also create their own database or independent instance on their  
122 computers or servers using the freely available editor. Thanks to its modular structure it can easily be adapted  
123 to their specific needs and/or expanded with new functions. In the spirit of the Open Access and Open Source  
124 policy,<sup>7</sup> the code and documentation of the CN Editor are available on GitHub for any projects related to the  
125 collection and analysis of ancient coins or other serial data.<sup>8</sup>

126 In the future, we would like to improve and extend the export functions and interfaces so that our  
127 typologies can be used in collection databases to describe the coins there.

## 128 **Versioning System**

129 For the unambiguous citation and further development of the typologies, we have developed a versioning,  
130 or edition, system.<sup>9</sup> The edition number has a specific format. It is composed of cn.[nomisma-mint or nomisma-  
131 ruler\_ID].[version\_number\_ed].[sequence\_number]. That means 'cn' stands for Corpus Nummorum, then

<sup>6</sup> Boteva-Boyanova et al. (forthcoming).

<sup>7</sup> <https://edoc.bbaw.de/frontdoor/index/index/docId/3366>.

<sup>8</sup> <https://github.com/telota/corpus-nummorum-editor> and <https://github.com/telota/corpus-nummorum-editor/blob/public/DOCUMENTATION.md> or see also <https://data.corpus-nummorum.eu/wiki>.

<sup>9</sup> For the importance of versioning the scholarly data, see also Bürgermeister 2022.

132 follows the issuer of the coin in the form of the nomismaID, afterwards the edition number and a unique  
133 number indicating the sort order. As an example, [cn.maroneia.1\\_ed.31](#) means the 31st coin type in the first  
134 edition of the coin typology of Maroneia. The criteria for arrangement are similar to those we would use in a  
135 printed catalogue. These sort orders are posted on our website [https://www.corpus-](https://www.corpus-nummorum.eu/resources/typology-sortorder)  
136 [nummorum.eu/resources/typology-sortorder](https://www.corpus-nummorum.eu/resources/typology-sortorder). All recorded material is available as open source for linking to  
137 other digital data, as well as for visualization and further research inquiries (also via SPARQL).

138 In addition to the edition system, which is based on conscious decisions by the editor, we also use an  
139 automatic versioning system that documents even small changes (such as the correction of a spelling mistake).  
140 If an already published coin or type record is modified, the current state is stored as a separate JSON file. CN  
141 ID and timestamp are used for the file name. All versions can be called directly via the API. It is also possible to  
142 pass any arbitrary timestamp; in this case, the API will determine the version's validity at the given date and  
143 provide the according data. The individual versions are also linked to the editions so that older data states can  
144 always be reconstructed and cited - independent of the constantly changing database.

145

146

### Further Development

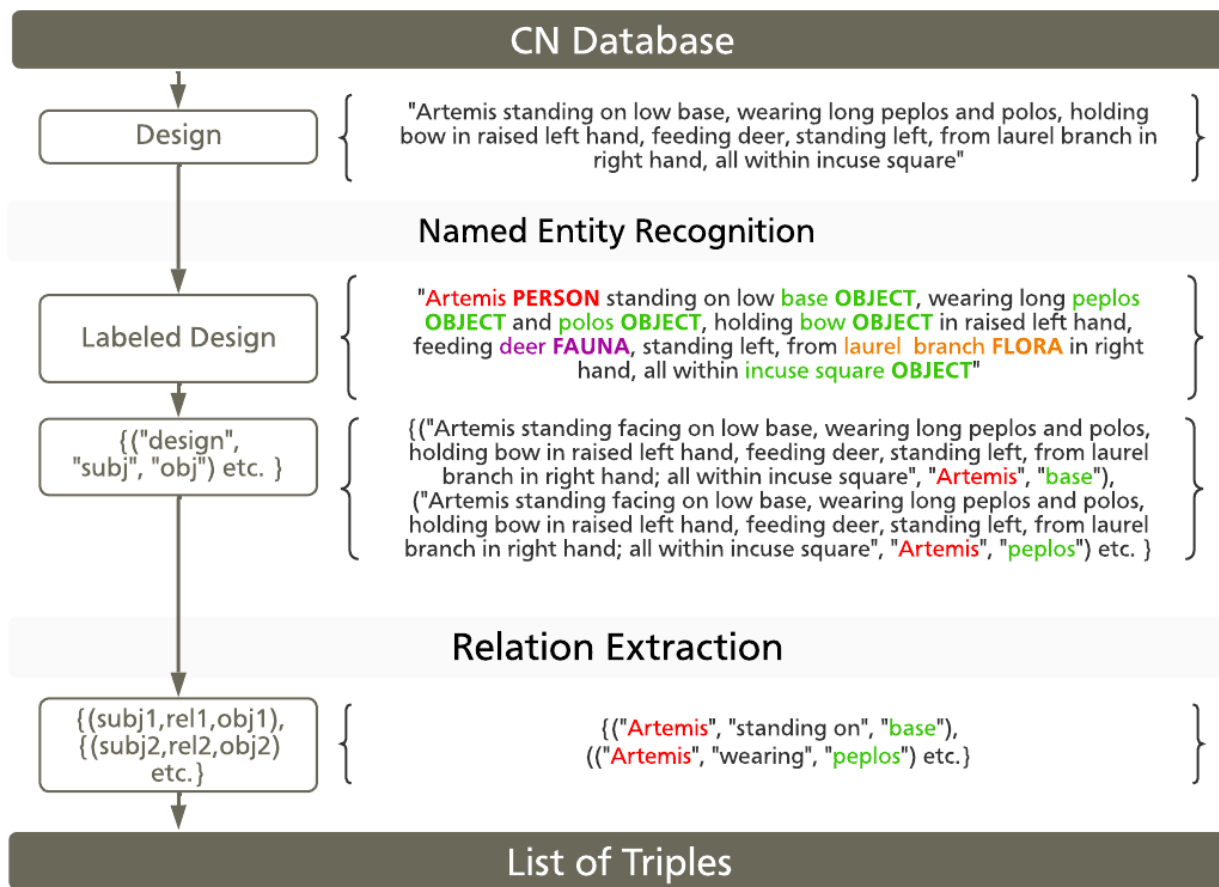
147 At present, CN is being further developed within our projects "Data Quality for Numismatics Based on  
148 Natural Language Processing and Neural Networks (D4N4)" and "Iconography and AI-Methods in  
149 Numismatics". The aim is to establish artificial intelligence research tools for numismatics that can also be used  
150 as services separately or integrated as part of the CN Editor. Our approaches are particularly well applicable to  
151 other object categories that combine images and text and are available in large quantities.

152 These digital tools should help improve the data quality. Furthermore, they should also enable a more  
153 effective recording of the coins already accessible on the web. Text and image recognition can be used to  
154 speed up this process. Since July 2021, we have been testing typological coin assignments using machine-  
155 learning methods. These involve: 1) natural language processing (NLP) methods for multilingual and non-  
156 standardized coin descriptions and linking them to a hierarchical iconographic thesaurus; this combination  
157 shall allow for more sophisticated semantic searches for a given coin and coin type data; and 2) complex image  
158 recognition of entire compositions of designs as well as of their single elements using deep learning.<sup>10</sup>

159

---

<sup>10</sup> For training the computer and its first accuracy results see [Gampe 2021](#). Furthermore, in these proceedings our colleague Sebastian Gampe presents a special paper on this machine learning method, based on our coin data set, which discusses the problems of image recognition models identifying mint classes, which have significantly more images than other mints: [Gampe/Tolle](#) (forthcoming [b]).



160

161  
162

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

163 Convolutional Neural Networks (CNNs) have been trained to add new coins to coin types and mints  
164 available in the database. In addition to coin types, patterns and elements on coins (monograms, persons,  
165 objects, etc.) should also be distinguished.<sup>11</sup> Our high-quality training and testing datasets, which are a  
166 prerequisite for a successful development of machine-learning based image recognition methods such as the  
167 Convolutional Neural Networks (CNN) will be published as well. Currently, the varying degree of preservation  
168 of coins and the similarity of different images offer challenges in the application of AI models, but in principle,  
169 they have immense potential for numismatics in the field of automated image annotation, analysis,  
170 interpretation, and retrieval.

171 On Google Colab, one can get an idea of how these tools already work: for NLP, one can enter or upload a  
172 coin description and see which named entities the program is able to detect (link to GitHub containing the  
173 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) and see the  
174 result for predicting the mint of this coin or the coin type from the CN dataset (link to GitHub containing the  
175 code and the link to Google Colab for testing: <https://github.com/Frankfurt-BigDataLab/IR-on-coin-datasets>).

176 The first promising results have already been published, yet we hope to improve the recognition by  
177 combining these methods.<sup>12</sup> Such a multimodal approach should also allow a monitoring of the data quality.  
178 The project has already successfully deployed quality controls allowing the supervision of the accuracy of the  
179 growing data.<sup>13</sup>

181 Furthermore, our goal is to build an authority data portal for ancient Greek coin iconography: a hierarchical  
182 *Thesaurus Iconographicus Nummorum Graecorum* (ThiNG). We have started with a prototype of a hierarchical

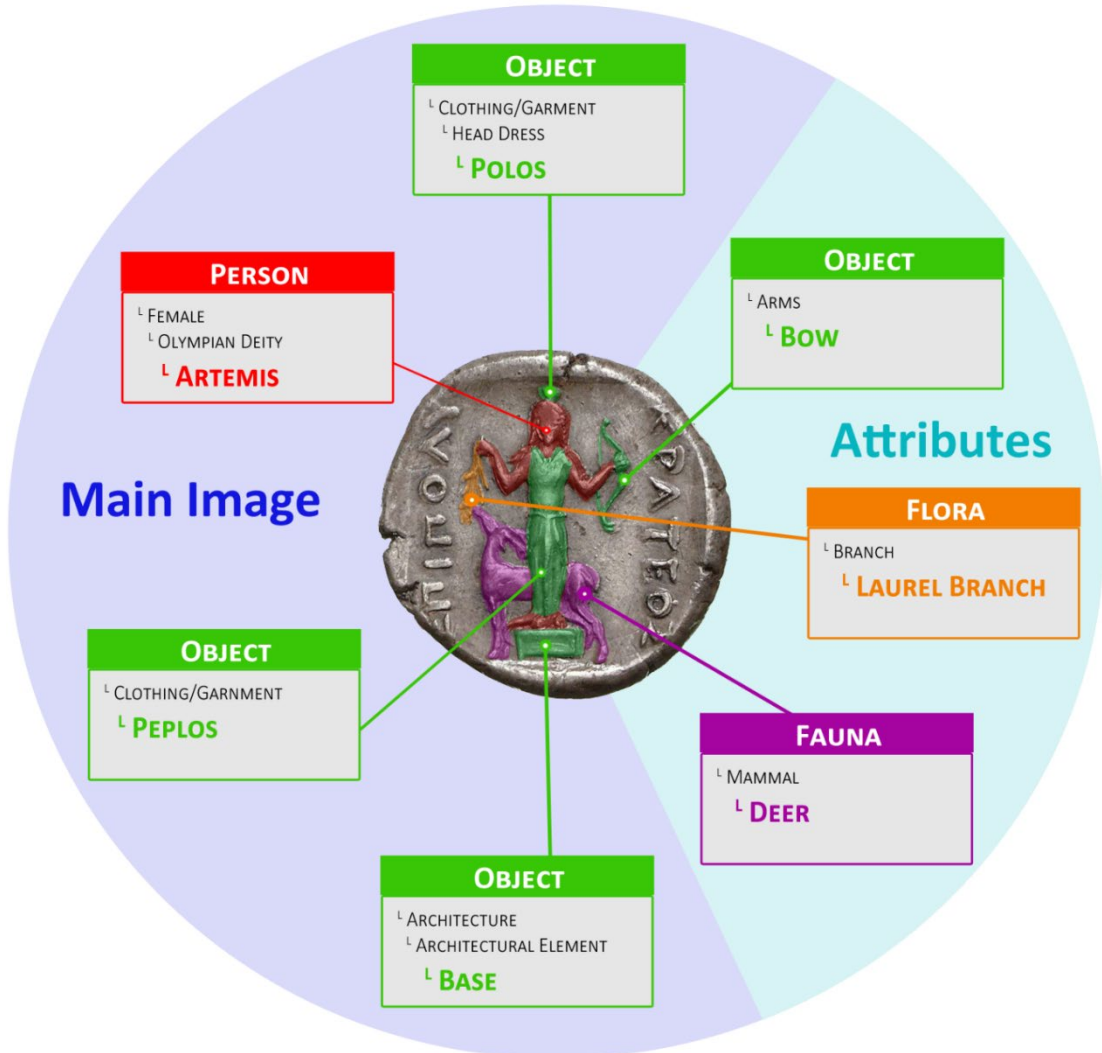
<sup>11</sup> See Luong 2022.

<sup>12</sup> Klinger et al. 2018; Gampe/Tolle (forthcoming [a]).

<sup>13</sup> Based on the nomisma.org/RDF modeling of the CN data, a tool – reusable by others – has already been developed that performs various quality checks (consistency checks) in the form of SPARQL queries; see Peter/Tolle (forthcoming).



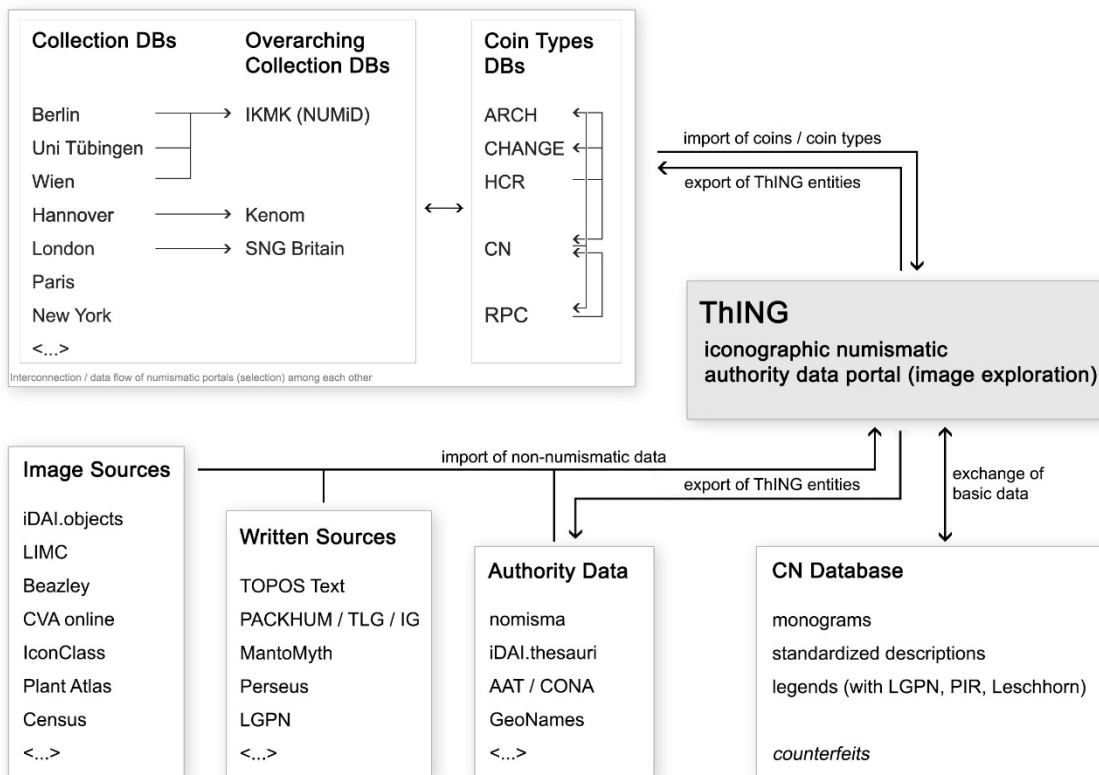
183 thesaurus in English and German. It is primarily based on the CN data with the standardized design  
 184 descriptions, but other coin descriptions and images can be integrated into it via data harvesting in the form  
 185 of LOD. The goal is to index entire image compositions, as well as individual elements. The THING will be a  
 186 constantly evolving portal for standardized iconographic data. Each entry will be illustrated with the  
 187 corresponding coin types and coins. For the purpose of a more detailed identification of different images and  
 188 cross-genre semantic and iconological analyses and research, we foresee a link to similar representations in  
 189 other object classes such as gems, pottery, sculpture, etc. Links to digital image repositories should also include  
 190 resources relevant for the reception of ancient coins.  
 191



192

193  
 194  
 195  
 196

**Figure 5** – Example of a coin image with the labeling which entities should be found through NLP and in future also through object detection. The entities are linked with the hierarchical iconographical thesaurus.



197  
198  
199 **Figure 6** – Schema for the in- and output of the *Thesaurus Iconographic Nummorum Graecorum* (THING).

200  
201 Since the large and constantly growing number of coins available online cannot be marked manually, it is  
202 necessary to train the developed algorithm for semi-automatic marking. The goal is to ensure that different  
203 descriptions of the same image types – especially of more complex compositions, such as the depictions of the  
204 heroic deeds of Heracles – are recognized as such and that the coins can be compared with each other, as well  
205 as with similar scenes on other object categories. New semantic searches will be possible by combining the  
206 NLP and the iconographic thesaurus. In addition, a new graphical user interface should make a semantic search  
207 possible, even for users without SPARQL experience.<sup>14</sup>

208 We encourage you to use our website for your research. At the moment, we already offer more than  
209 28,400 coins and 11,400 coin types online. In the internal database there are an additional 20,000 coins and  
210 more than 6,000 coin types waiting for the last quality control to get published. Check the CN portal regularly,  
211 both for new coins and coin types and for new developments. Our most important button is: participate!

<sup>14</sup> Wirth 2019.

212

213

## Acknowledgements

214 We thank the entire [CN team](#) and the Deutsche Forschungsgemeinschaft and the Berlin  
215 Senatsverwaltung für Wissenschaft, Gesundheit, Pflege und Gleichstellung for funding.

216

## Data, scripts, code, and supplementary information availability

217 Scripts and code are available online: <https://github.com/telota/corpus-nummorum-editor> and  
218 <https://github.com/telota/corpus-nummorum-editor/blob/public/DOCUMENTATION.md>. See also  
219 <https://data.corpus-nummorum.eu/wiki>.

220

## Conflict of interest disclosure

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

223

## Funding

224 The D4N<sup>4</sup> project “Datenqualität für Numismatik basierend auf Natural Language Processing und  
225 Neuronalen Netzen“ is funded by the Deutsche Forschungsgemeinschaft (DFG) in the program “e-  
226 research-Technologien”. The project “Ikonographie und KI-Methoden in der Numismatik (IKKINUM)” is  
227 financed by the Berlin Senatsverwaltung für Wissenschaft, Gesundheit, Pflege und Gleichstellung.

228

## References

- 229 Boteva-Boyanova D, Prokopov I, Köster J, Flueck JE, Grigorova-Gencheva V (forthcoming) RT 8: Ancient  
230 Coins Counterfeits, in: *Proceedings of the XVI International Numismatic Congress, September 2022*.  
231 Warsaw.
- 232 Bürgermeister M (2020) Enabling the Scholarly Discourse of the Future: Versioning RDF Data in the Digital  
233 Humanities, Andrews T, Diehr F, Efer T, Kuczera A, van Zundert J (eds.): *Graph Technologies in the*  
234 *Humanities – Proceedings 2020*, published at <http://ceur-ws.org>, 1– 17 (<https://ceur-ws.org/Vol-3110/paper1.pdf>).
- 235
- 236 Gampe S (2021) *Neuronale Netze zur Bestimmung römischer Kaiser auf Bildern antiker Münzen* (Master  
237 Thesis Universität Frankfurt/M.) ([http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit_Sebastian_Gampe_online.pdf)  
238 [Masterarbeit\\_Sebastian\\_Gampe\\_online.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2022/05/Masterarbeit_Sebastian_Gampe_online.pdf)).
- 239 Gampe S, Tolle K (forthcoming [a]) Combination of machine learning methods of image and natural  
240 language recognition on ancient coin data. *Proceedings Computer Applications and Quantitative*  
241 *Methods in Archaeology (CAA) conference in Kraków 2019*.
- 242 Gampe S, Tolle K (forthcoming [b]) Creating an additional class layer with machine learning to counter  
243 overfitting in an unbalanced ancient coin dataset, in: *Proceedings of CAA conference in Amsterdam*  
244 *2022/23*.
- 245 Klinger P, Gampe S, Tolle K, Peter U (2018) Semantic Search based on Natural Language Processing – a  
246 Numismatic example, *Journal of Ancient History and Archaeology* **5.3**, 68–79. [https://doi.org/](https://doi.org/10.14795/j.v5i3.334)  
247 [10.14795/j.v5i3.334](https://doi.org/10.14795/j.v5i3.334).
- 248 Luong H (2022) Klassifizierung von Motiven auf antiken Münzen mit Mask R-CNN (Master Thesis Universität  
249 Frankfurt/M.) ([http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2023/02/Masterthesis-Huy-](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2023/02/Masterthesis-Huy-Luong.pdf)  
250 [Luong.pdf](http://www.bigdata.uni-frankfurt.de/wp-content/uploads/2023/02/Masterthesis-Huy-Luong.pdf)).
- 251 Peter U (2017) Corpus Nummorum Thracorum – A Research Tool for Thracology and an Example of Digital  
252 Numismatic Collaboration, in: Caccamo Caltabiano M. (ed.) (2017). *XV International Numismatic*  
253 *Congress Taormina 2015. Proceedings*. Rom–Messina, Vol. 2, 1306.

- 254 Peter U (2019) Von Mommsen zum Semantic Web: Perspektiven der vernetzten numismatischen  
255 Forschung – die Münzen der westlichen Schwarzmeerküste online, in: Cojocaru V, Ruscu L, Castelli Th,  
256 Pázsint A-I (eds) (2018), *Advances in Ancient Black Sea Studies: Historiography, Archaeology and*  
257 *Religion. The Proceedings of the International Symposium, Constanța, August 20–24, 2018*. Cluj-Napoca  
258 (Pontica et Mediterranea VIII), 393–418.
- 259 Peter U (2020) Das Münzkabinett und die Berlin-Brandenburgische Akademie der Wissenschaften – Eine  
260 enge Partnerschaft, in: Weisser B (ed.) *Münzkabinett – Menschen, Münzen, Medaillen*, Regenstauf,  
261 285–300.
- 262 Peter U, Tolle K (forthcoming) Corpus Nummorum – Coins, types and data quality control, in: *Proceedings*  
263 *of the 8th Joint Meeting of ECFN and nomisma.org 2019*.
- 264 Peter U, Weisser B (2017) CNT – The web portal for Thracian coins, in: Caccamo Caltabiano M (ed.) (2017)  
265 *XV International Numismatic Congress Taormina 2015. Proceedings*. Rom–Messina, Vol. 1, 247.
- 266 Peter U, Adamik-Köster J, Berthold A, Franke C, Gampe S, Gorys A, Stolba V, Tolle K, Weisser B (forthcoming)  
267 Eine digitale Forschungsinfrastruktur für antike Münzen: neue Entwicklungen im Projekt Corpus  
268 Nummorum, in: Bursche A et al. (eds) *Proceedings of the XVI International Numismatic Congress,*  
269 *September 2022*. Warsaw.
- 270 Wirth A (2019) *SPARQL-Abfrageschnittstelle ikonografischer Daten im Resource Description Framework,*  
271 Bachelorarbeit, Frankfurt/M. ([https://github.com/aliciawrt/IGSApplication/blob/master/SPARQL-  
272 Abfrageschnittstelle%20ikonografischer%20Daten%20im%20Resource%20Description%20Framework.  
273 pdf](https://github.com/aliciawrt/IGSApplication/blob/master/SPARQL-Abfrageschnittstelle%20ikonografischer%20Daten%20im%20Resource%20Description%20Framework.pdf)).