IT- and machine learning-based methods of classification: The cooperative project ClaReNet – Classification and Representation for Networks

Chrisowalandis Deligio[2], Caroline von Nicolai[1], Markus Möller[1], Katja Rösler[1], Julia Tietz[1] Robin Krause[2], Kerstin P. Hofmann[1], Karsten Tolle[2], David Wigg-Wolf[1].

1 Römisch-Germanische Kommission des Deutschen Archäologischen Instituts.
2 Big Data Lab, Goethe University Frankfurt.

Abstract

The classification of archaeological finds and their representation are shaped by various object epistemological approaches and changes of medium. With ever increasing digitisation, there are now new possibilities of classification, for example using methods of automatic image recognition, as well as the representation of finds on the web with linked open data.

ClaReNet, a cooperative project of the Römisch-Germanische Kommission (German Archaeological Institute) and the Big Data Lab (Goethe University Frankfurt), funded by the **Bundesministerium für Bildung und Forschung (BMBF; Federal Ministry of Education and Research)**, tests the possibilities and limits of new digital methods of classification and representation. To this end, traditional approaches of typification and the recording of attributes in numismatics and archaeology are compared with IT-based methods of classification, including deep learning, using the examples of three Celtic coin series that were each chosen to address particular research questions and problems. This work is accompanied by a science and technology study (STS) PANDA, which focuses on **P**ath dependencies, **A**ctor **N**etworks and **D**igital **A**gency.

This paper briefly introduces the approach of object epistemologies before considering Celtic coins as scientific objects and the history of research on them with regard to classifications and typologies. Using the example of a series of coins from *Armorica* (Brittany), we will present how deep learning classifies the coinage, how this may differ from classifications by numismatists, and the lessons that are to be learned from this exercise. From an STS perspective, we analyse the actor network that emerges during image data processing. The paper concludes with a reflection on the transformation of numismatic practices resulting from the IT-based methods used in the project, as well as an outlook on further possibilities for research into the classification and representation of Celtic coins.

keywords: Celtic coinage, machine learning, object epistemologies, knowledge practices, science and technology study, die studies, classification.

Introduction

Humans gain knowledge about the world through objects, both in everyday life and in science. Our ways of knowing and our approaches change in the process, not only because of different conceptions of the world and related questions, but also because of research practices and technologies, as well as the empirical examples and observations chosen, all of which interact. This in turn will always affect the classification and representation of objects of knowledge, in our case Celtic coins. For example, there are cultural evolutionist approaches that establish typological series on the basis of imitation and degeneration, whereas cultural history approaches have described the spatio-temporal occurrence of certain coin types in typographies and tried to assign them to historically transmitted tribal designations. Our approach, in contrast, is based on production-technological observations – such as die studies – that create fine chronological sequences of coins.

These and other approaches to producing knowledge about and with objects are explored within the framework of an emerging transdisciplinary field of research known as object epistemologies. Epistemology, from the ancient Greek ἐπιστήμη *(epistḗmē,* 'knowledge'), is the theory of knowledge, which addresses questions about the preconditions for knowledge, the production of knowledge and other forms of beliefs. Object epistemologies investigate how and why past and present discourses about things developed, for what reasons knowledge about things circulated, and which conceptualizations were inherent in such knowledge. They also study the relationship between such knowledge and epistemic or scientific practices (Hilgert et al. 2018).

The project ClaReNet, short for "Classifications and Representations for Networks. From types and characteristics to linked open data for celtic coinages", which officially started on 1 February 2021, is a cooperative project of the Römisch-Germanische Kommission of the Deutsches Archäologisches Institut (RGK) and the Big Data Lab of the Goethe University Frankfurt. It is funded by the German Federal Ministry of Education and Research in the framework of research on the theoretical, methodological and technical development in the digital humanities[1].

For the project, three examples of Celtic coin series with very different characteristics were each chosen to address particular research questions and problems (fig. 1)[2]:

The 'bushel' *quinarii* from southern Germany and northern Switzerland are characterised by the successively more abstract image of the head on their obverses and the combinations of various symbols around the horse on the reverse[3]. Several classifications have already been established for them using traditional methods of numismatics. The different classifications will be re-assessed with methods of artificial intelligence (AI) and systematically compared

1 For further information on the project see the ClaReNet blog https://clarenet.hypotheses.org; on the funding initiative https://www.geistes-und-sozialwissenschaften-bmbf.de/de/Digital-Humanities-1710.html.

2 For more detailed information on the three coin series and their relevance for the project see the relevant pages of the project blog: https://clarenet.hypotheses.org/numismatics.

3 Numismatic terms are explained in the *Glossar* of the project blog: https://clarenet.hypotheses.org/glossar_num.

with respect to performance on the data. The aim is to observe whether there are similarities and, if so, to see how the classifications can be harmonised.

Fig. 1: Examples of the coin series examined in the project. The differences in shape and form also become clear within the individual coin series (photos: Münzkabinett der Staatlichen Museen zu Berlin, Bernhard Weisser; graphics: M. Möller).

For the second series, the *monnaies à la croix*, the representations of the head on the obverse and the symbols in the fields of the cross on the reverse of the are analysed by computer-assisted classification methods. The focus lies on understanding the combination of features in the designs of the coins. The starting point is Eneko Hiriart's recent classification (Hiriart 2017). In addition to the images, the descriptions of the coins provided by E. Hiriart will also be integrated into the analysis by applying methods from Natural Language Processing[4].

The third series, the billon stater coinage which is traditionally attributed by numismatists to the *civitas* of the *Coriosolitae* in modern Britanny (*Armorica*), offers a very different challenge arising from the availability of a massive stock of digital images of some 70,000 coins from the recently discovered coin hoard of Le Câtillon II (Jersey, UK). The aim is to assist the numismatists working on classifying the coins and to observe how far AI methods can work with the dataset independently. The dataset[5] is sorted by the AI using unsupervised

4 Natural Language Processing is a subfield of AI, focusing on analysing and interacting with human language (Natural Language Processing with PyTorch, D. Rao and B. MacMahan, 2019, Chapter 2).
5 On our definition of dataset see https://clarenet.hypotheses.org/it-teil4.

methods[6], and the central question here is to what extent the computer and human classifications will correspond.

Using these examples, ClaReNet tests the possibilities and limits of new digital methods of classification and representation. The project consists of three strands:

1. Traditional approaches of typification and the recording of attributes in numismatics and archaeology will be compared with IT-based methods of classification, including deep learning and automatic image recognition[7].
2. Representation is addressed by setting up a virtual union catalogue, called Online Celtic Coinage (occ.dainst.org), for the coin series under investigation. The catalogue will be linked to publicly accessible online collections and research databases, following the paradigm of portals based on the vocabulary and ontology of nomisma.org[8], such as Online Coins of the Roman Empire (OCRE)[9]. This will require the development of an ontology to cover the specificalities of Celtic coinage.
3. The third strand consists of a science and technology study (STS) to document the production and circulation of knowledge, as well as to instigate reflection on changes in cognitive processes resulting from the use of digital tools and algorithms. A particular focus of the study is on **P**ath dependencies, **A**ctor **N**etworks and **D**igital **A**gency (PANDA)[10].

In this paper, we focus on the chances and limitations of machine learning methods for the automated classification of large numbers of coins, using the example of the coinage of the *Coriosolitae*. For this purpose, we first look at how Celtic coins have been studied and classified in the past, before turning to the question of automated classification. In the STS, we analyse the path dependencies, actor networks and digital agencies involved in the image data processing with the JupyterLab. In an outlook, upcoming research is outlined and initial overarching findings are presented.

Celtic coins as scientific objects in the past

Coins are generally defined as flat metal objects stamped with an impression on both sides. They are mobile, multi-medial objects that combine form, material, image and (in the case of Celtic coins only sometimes) text. In regard to typologies and classifications, it is important to understand how the coins we study were manufactured: the blanks and dies were made by hand, and they were also struck by hand. In this way, they differ considerably from today's standardised, machine-made Euro coins, which are all identical in shape and form. Celtic

6 For further information on supervised and unsupervised methods see https://clarenet.hypotheses.org/it-teil1.
7 For definitions of the IT-terminology used here see "What's the Difference Between Artificial Intelligence, Machine Learning and Deep Learning?": https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/.
8 https://nomisma.org/.
9 https://numismatics.org/ocre/.
10 https://clarenet.hypotheses.org/sts-teil3.

hand-made coins, on the contrary, even those belonging to the same type and struck with the same dies, can differ significantly from one another (fig. 2).

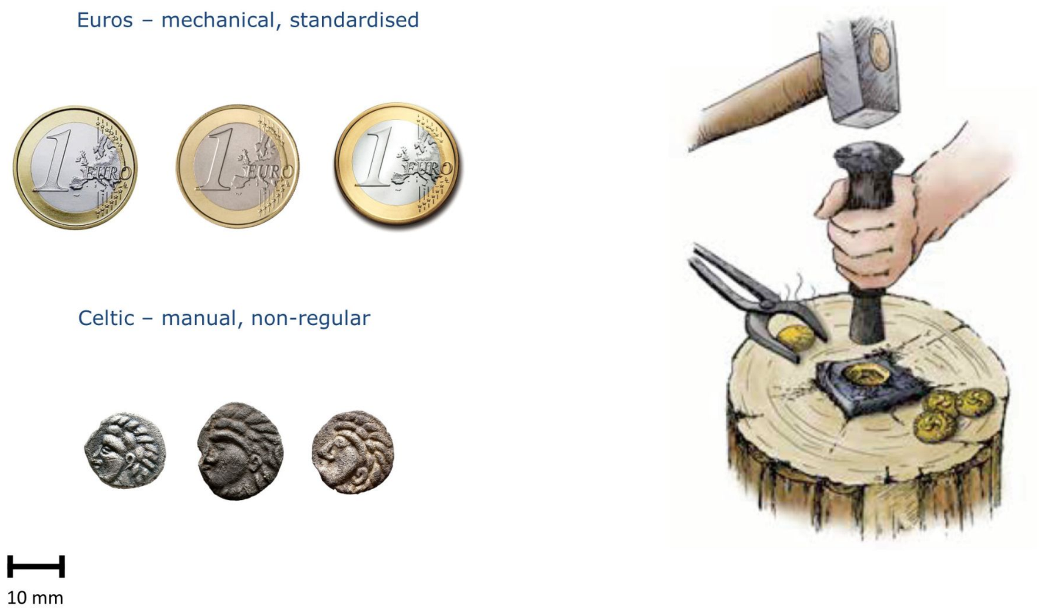

Fig. 2. Left: modern, standardised Euro coins vs. Celtic, hand-struck bushel *quinarii*. Right: manufacture of Celtic coins with a hand-held die (graphics left: M. Möller, photos: Archäologische Staatssammlung München; drawing right: Sabine Schade-Lindig, hessenARCHÄOLOGIE).

The traditional approach: classifications and representations in Celtic numismatics

This form of manufacture presents particular challenges for classifications and typologies[11], and it should be noted that in numismatics there is no universal or consistent concept of what exactly constitutes a type or a class. Both terms are used irregularly and are often confused. Furthermore, it was only in the second half of the 19th century that Celtic coins became the subject of serious scientific interest. Previously, they were regarded as vile and degraded derivatives of Greek and Roman originals (Forrer 1908, 1). In a mixture of cultural evolutionist and cultural historical approaches, they were classified according to material, design and the geographical distribution of finds. Relative chronology was established on stylistic criteria, absolute chronology on the basis of datable Mediterranean prototypes they were presumed – often incorrectly – to copy (Forrer 1908, 16-21). Moreover, until well into the 20th century, the opinion prevailed that hardly any two Celtic coins were identical, i.e. the number of variants was almost as great as the number of individuals (e.g. Blanchet 1905, 71). Consequently, a “disparité indéfinie des espèces gauloises” (an unlimited variety of Gaulish coins) (Colbert de Beaulieu 1973, 14) was assumed. It was considered almost impossible to identify two coins with identical impressions.

11 https://clarenet.hypotheses.org/clareterms_typologie.

From past to present

The earliest attempt at producing a comprehensive reference work illustrating types of Celtic coins was Henri de La Tour's ‘Atlas de monnaies gauloises’ produced in 1892 (de La Tour 1892). The Atlas was based on the catalogue of Celtic coins in the collection of the Bibliothèque nationale de France in Paris published three years earlier by Ernest Muret and Anatole Chabouillet (Muret / Chabouillet 1889). The coins were attributed to tribes known from written sources (fig. 3), mainly on the basis of the geographical distribution of the coin finds. But such attributions do not necessarily reflect the actual authorities behind the production of coins in the pre-Roman Iron Age. We now know that not only tribes but also confederations, *pagi* and powerful individuals minted coins (Ziegaus 2010, 17-18. On the *pagi* see Roymans 1990, 23-26). Today a tendency can be observed to replace tribal affiliations with references that are considered to be more neutral, such as names for types based on a find spot (e.g. ‘Martbeger Typ’) or the iconography (e.g. ‘potins au sanglier’), or else, above all by British numismatists, on geographical areas (e.g. Haselgrove 1987 and 1999, Leins 2012). But even here there is a certain reluctance to completely abandon tribal attributions.

Fig. 3: Billon coins attributed to the *Curiosolitae* or *Coriosolitae*, represented in the “Atlas” of de La Tour (de La Tour 1892, pl. XXII).

Representation of coins in Celtic numismatics

If de La Tour's Atlas was originally conceived as an illustrated and selected representation of the collection of the Bibliothèque nationale de France in Paris, it has since served as a reference work with the drawings of individual coins being used as representatives of individual coin types or classes. This practice of using a single image as a representative (called “ideal type”, “modèle type” or “dessin idéalisé” in numismatics) still persists today, for example in the most recent typology for the coinage of Iron Age Britain, Ancient British Coinage (Cottam et al. 2010) (fig. 4) or E. Hiriart’s corpus of the *monnaies à la croix* (Hiriart 2017, 20); this in spite of the fact that it does not do justice to the individuality of Celtic coins and the wide variation between examples of the same type. It suggests that “types” were fixed and standardised, a concept that was adopted from the much older disciplines of Greek and Roman numismatics and is reflected in reference works such as Roman Imperial Coinage (e.g. Mattingly 1923). For many coinages produced in the Celtic world, what we in fact find is a development of the imagery on the coins over time, one that does not always proceed in a straight line as a typological series in the sense of Oscar Montelius. Sometimes it is gradual, sometimes in more obvious steps, or it can be in networks (cf. Hofmann 2014).

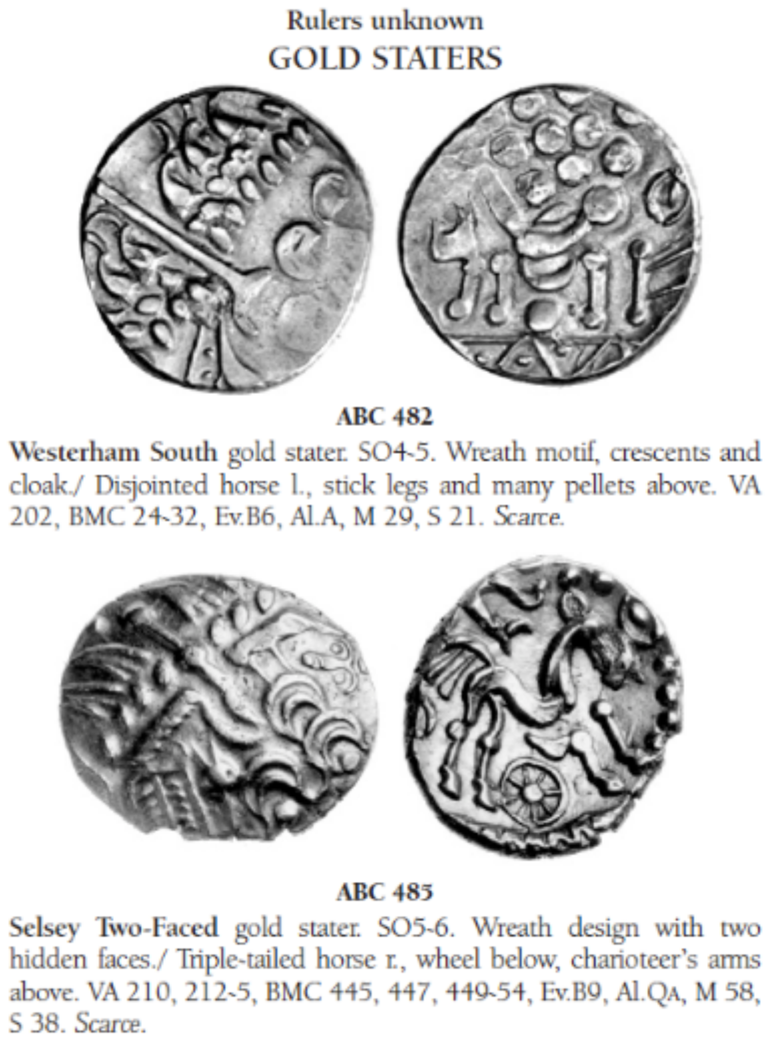

Fig. 4: Representation of British gold staters in the reference work "Ancient British Coins" (Cottam et al. 2010, 47), here with regional affiliation (Code SO for Southern region). Note the desire to still name at the top a minting authority if possible.

It's all about context

A change in mindset that departs from the purely antiquarian approach to Celtic coins has come about since the 1980s, much influenced by Iron Age archaeology. Above all, archaeological contexts are now included in the interpretation of coin finds, which has led to significant revisions in the chronology of Celtic coinage. As Colin Haselgrove states, „Detailed analysis of association, stratigraphy and site context are fundamental if coinage is to be systematically related to other components of material culture and society. Moreover, both archaeological context and cultural context are fundamental to the interpretation of every coin find." (Haselgrove 1987, 11-12; see also Polenz 1982, 48-49). Much work is now focussed on producing ever finer chronologies for Celtic coins, not just for the sequence of types, but also the sequence of production within types (fig. 5).

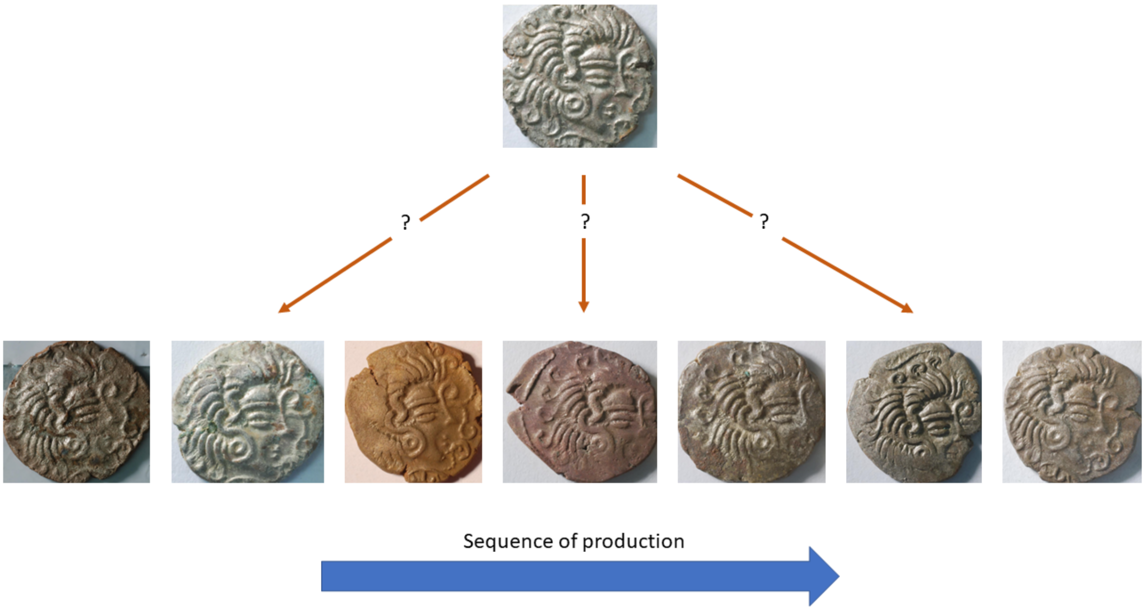

Fig. 5: Visualisation of a sequence of production for staters of the *Coriosolitae* (photos: Jersey Heritage; graphics: C. Deligio).

Die studies: ‘charactéroscopie’

So-called die studies can be used to determine the chronological order in which dies were used and coins were produced. Die studies were developed as early as the 19th century in Greek numismatics (de Callatay 2007), but were only used in Celtic numismatics for the first time in 1948 by Jean-Baptiste Colbert de Beaulieu (Colbert de Beaulieu 1948), before finally being firmly established in the field by him in the 1970s. Previously Celtic coins had been considered an “impenetrable chaos” unsuited for such studies (Colbert de Beaulieu 1973, 17. On the history of the use of die studies see de Jersey 2016). J.-B. Colbert de Beaulieu's work was strongly influenced by French structuralism, the predominant intellectual movement in France at the time, the aim of which was to uncover the fundamental elements and regularities within the respective fields of research. Consequently, J.-B. Colbert de Beaulieu also referred to his set of methods as “numismatique structurale”. The most important one was the analysis of die links, which J.-B. Colbert de Beaulieu called “charactéroscopie” (from Ancient Greek χαρακτήρ, coinage, and σκοπιά, observation). This method essentially consists of identifying the specific characteristics of each individual coin die, and assigning to them the individual coins produced with it in order to create groups (Colbert de Beaulieu 1973, 40-41).

While the mints of imperial Rome were tightly structured institutions that could produce dies of the same coin type that look very similar, it is assumed that the process of die production and minting in the Celtic world was less centralised, often resulting in dies with the same motif that look very different and are thus more easily distinguished. Behind the use of die studies to determine the sequence of production of a coinage is the assumption that, if coins are indeed struck with the same die they should share a common origin in time and/or place. Furthermore, since the dies for the obverse (in English the “heads” side) and reverse (“tales”) used will normally not both have been replaced at the same time, but when one or the other became too worn or broken, it is possible to establish the sequence in which dies were used and a group of coins produced (Goebl 1972).

Praxis: How does a die study work?

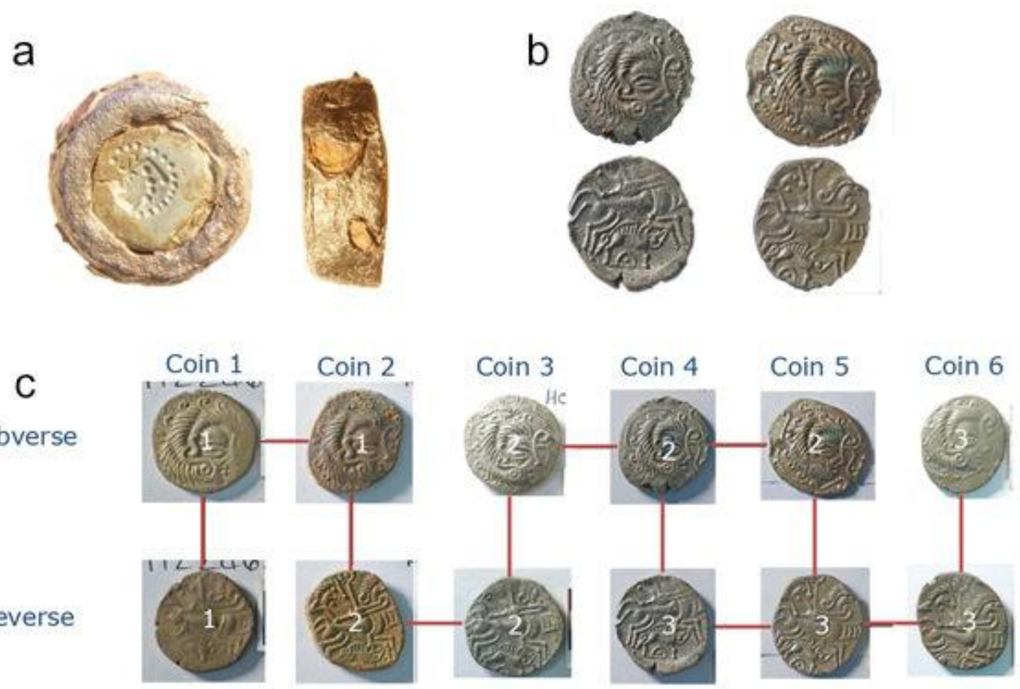

Fig. 6: a. Immobile lower die for a quinarius of the Treveri from the Donnersberg, Germany; b. Class VI stater of the *Coriosolitae*; c. Visualisation of a die study of staters of the *Coriosolitae* (photos, a: Römisch-Germanisches Zentralmuseum Mainz; b,c: Jersey Heritage; graphics: D. Wigg-Wolf).

Fig. 6c illustrates how a die study works. It shows a coin (coin 1) which was produced with dies that we shall call obverse 1 and reverse 1. From a technical point of view, the reverse is struck by the upper, mobile die, the obverse by the lower, fixed die (see fig. 2; 6a). In our example, there is a second coin (coin 2) produced with obverse 1 but with a different reverse die, reverse 2. A third coin was produced with reverse 2, but with a new obverse die 2. Sometimes we find groups of coins struck with the same dies, and in our example coins 4 and 5 were both produced with obverse 2, but with a new reverse 3. Coin 6 is produced with the same reverse 3, but now with a new obverse 3, and so on… Such sequences can tell us a great deal about the way in which the production of coinage was organised.

More than a lifetime's work

Die studies carried out by eye are extremely time consuming and tedious. This is why we are looking at how machine learning and automatic image recognition can help with this task, and what new ways of sorting coins are emerging as a result. For the coinage of the *Coriosolitae*, one of the three coin series studied in the project, we have access to images of nearly 70,000 coins from the hoard of Le Câtillon II (de Jersey 2016; de Jersey 2018; de Jersey 2020). Some coins from the hoard were found by two metal detectorists in 2012 in a field on the island of Jersey, some 15 metres from the location of another coin hoard, Le Câtillon I, which was discovered in 1957 (Gruel 1990). The discovery of the coins resulted in further excavations that led to the discovery of the main body of the second Le Câtillon hoard. It is the largest Celtic coin hoard ever found, and was lifted as a block weighing approximately 750 kg, before being painstakingly excavated in the Jersey Heritage laboratory (fig. 7).

Fig. 7: The hoard of Le Câtillon II on the island of Jersey contained nearly 70,000 coins. Its analysis poses a great challenge for numismatics (photo: P. de Jersey, States of Guernsey, Culture and Heritage).

The coinage of the *Coriosolitae*

Apart from the coins, the find also contained eleven gold torques, silver and gold jewellery, silver and copper ingots, glass beads, a spearhead of a type dated to the Bronze Age, as well as a worked stone. The coins are mainly staters that are attributed to the *Coriosolitae*, but also some of other polities from the region of northern France and Britain. The coins of the *Coriosolitae* are made of billon, an alloy of copper and silver in which the amount of silver is less than 50 %. A typical motif on the obverse of the coins is a stylised head surrounded by strings of beads. On the reverse, there is often a human-headed horse with a charioteer, surrounded by symbols such as a boar, a winged figure, a multi-spoked wheel and a lyre, as well as others that are not always easy to identify with our modern perspectives. The staters weigh between 6 and 7 g. The coin hoard has been dated to the 40s or perhaps the 30s of the 1st century BC on the basis of the accompanying finds. That means that the find dates to the period after the Roman conquest of Gaul (de Jersey 2016; de Jersey 2018). The doyen of French numismatics, J.-B. Colbert de Beaulieu presented the first typology of the staters of the *Coriosolitae* in 1957. He divided them into six classes (I-VI) on the basis of the shape of the nose and the eyes of the head on the obverse (Colbert de Beaulieu 1957). In the late 1970s and early 1980s, Katherine Gruel, then a doctoral student of J.-B. Colbert de Beaulieu, further developed his typology and added other variants (VIa, VIb, Va, Vb, IVa, IVb) to this classification, based mainly on the symbols on the reverse (Gruel 1981). Her typology, as well as the relative chronology of the coin series based on it, were the result of one of the first such studies to employ IT-based statistical methods, and are still considered to be valid today. The discovery of the hoard of Le Câtillon II now offers the possibility of testing the validity of the typology, based on a very large number of coins. For this purpose, we are analysing the coin hoard in cooperation with the domain expert Philip de Jersey, who is currently working on it.

Classification with Artificial Intelligence: machine learning

Obtaining the necessary information on nearly 70,000 coins and classifying them requires a great deal of working time. In particular, P. de Jersey estimates that the ongoing die study would take several more decades. Our goal was therefore to find methods to assist him and to evaluate the extent to which machine learning can help in the process of classification. Furthermore, the performance of AI on the dataset will be observed and AI used to question the underlying classification.

There are two approaches to machine learning that we have tried so far:

1. Unsupervised approach: this approach works with data that is not yet labelled, e.g. only the coin photos are used. If labels have already been assigned, the analysis can also be done independently of them.
2. Supervised approach: in this approach the data is labelled with additional information from a domain expert, who defines, for example, to which group, or in our case to which die a coin belongs (this labelled data is known as the ground truth).

In the following, the basics of these methods are explained before presenting the preliminary results of our studies.

Convolutional Neural Networks

In the field of image classification, so-called Convolutional Neural Networks (CNNs) are today state-of-the-art[12]. Other numismatic projects have already benefited from the use of CNNs and promising results have been achieved[13].

CNNs are machine learning algorithms intended for image data and usually used in supervised learning tasks. A CNN consists of several layers, typically convolutional layers followed by a pooling layer and dense layers (explained below). A CNN can be assembled as desired, but there is also the possibility of using structures that are already successful[14]. In our case, we use a well-known structure, VGG16 (Simonyan / Zisserman, 2015). The network consists of 16 layers, 13 of which are convolutional and the remaining three dense. Pooling layers are not counted individually and follow after a number of convolutional layers (for a visual representation of VGG16 see: https://towardsdatascience.com/illustrated-10-cnn-architectures-95d78ace614d#c5a6).

The **convolution layer** (based on the VGG16 approach), which is the name-giving layer for the algorithm, works by sliding an area (size 3x3) over the input, applying an element-wise dot product. This sliding area is also called *kernel* or *filter* and holds different weights. These weights are trained and adapted by the CNN in the training phase. During training based on ground truth data, the result of the overall network is compared with the given annotation within the ground truth. In a back propagation phase the relevant weights are amplified if the training was successful in selecting the correct outcome, or weakened if not. Fig. 8 shows what the output of a single convolutional layer could look like.

Typically, **pooling layers** are set after convolutional layers. A pooling layer eliminates pixels in an image while preserving the semantics of the image content (Moroney 2020, 35-36). There are different types of pooling such as 1) average pooling, where the average in the area being considered is calculated, 2) min(imum) pooling, where the smallest value, or 3) max(imum) pooling, where the highest value is chosen. The VGG16 CNN uses 2x2 max pooling layers. By doing this, the amount of information is reduced while the features are kept (fig. 9).

A **dense layer** is a fully meshed layer, where every neuron is connected to all previous neurons (Di et al. 2018). For each of these connections, weights are again used and trained during the training phase. Therefore, dense layers are usually set at the end of a CNN, due to the extensive computational resources needed. For example, when used as the last layer the size corresponds to the number of classes. This is then also called the classification layer (fig. 10).

12 https://paperswithcode.com/sota/image-classification-on-imagenet?p=centroid-transformers-learning-to-abstract.
13 E.g. http://www.bigdata.uni-frankfurt.de/cnt-nomisma-org/.
14 https://pytorch.org/vision/stable/models.html.

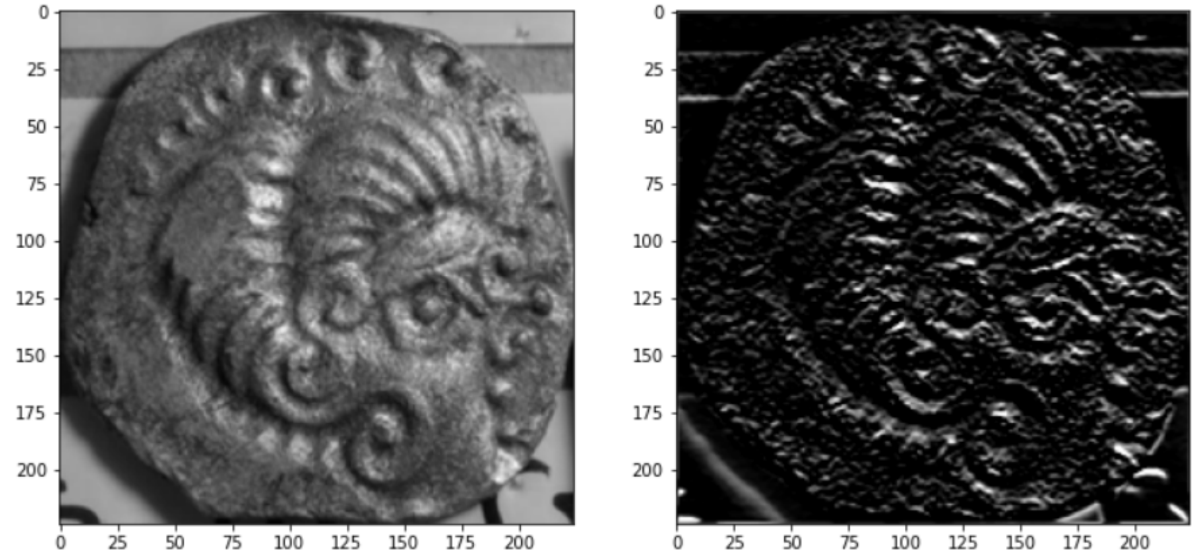

Fig. 8: Output of a convolution layer trained to detect horizontal edges. (photo: Jersey Heritage; graphics: C. Deligio, Big Data Lab).

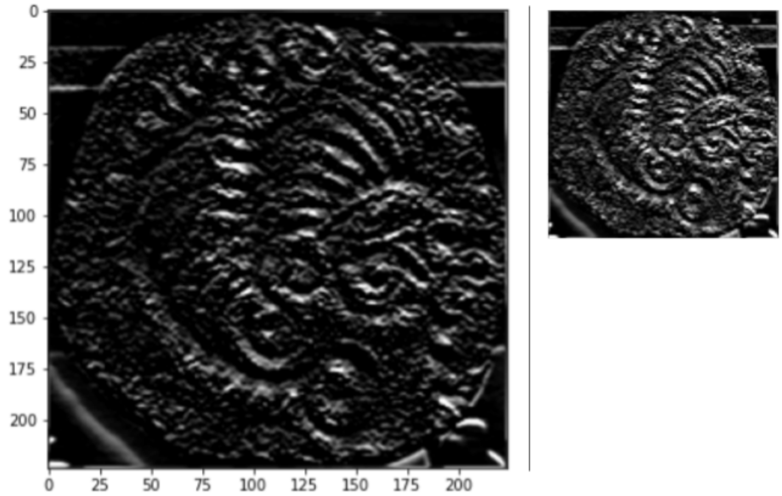

Fig. 9: Example of the input (left) and output (right) of a 2x2 max pooling layer (C. Deligio).

Taking all weights into account, the VGG16 architecture consists of about 138M parameters and the trained model needs a corresponding amount of space (the VGG16-model of ImageNet has a size of 528 MB[15]).

15 https://keras.io/api/applications/.

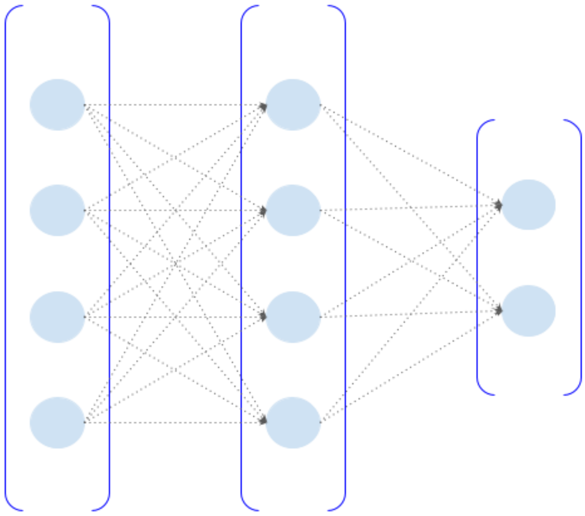

Fig. 10: Example of 3 connected dense layers (C. Deligio).

Clustering: k-Means

The cluster algorithm we use in combination with CNN is k-Means (Patel 2019), which belongs to the group of unsupervised algorithms. Instead of learning based on human-provided data (ground truth), as in supervised learning, the goal is to independently recognize common patterns or rules in the dataset. For k-Means the input data of objects are transformed into a vector representing a point in an n-dimensional space. This allows the algorithm to define distances between the points and to group them in so-called clusters. This means that objects of one cluster are more similar (based on the distance function and the way the data are transformed into their vector) to each other compared to those in other clusters. For k-Means the number of expected clusters (k) must be specified, although finding an optimal value for k presents significant difficulties if there is insufficient information on the distribution in the dataset.

While this seems ideal for setting up a classification, unfortunately the results can also be very unexpected, especially for images, where each pixel will be handled independently and no structures such as edges or larger structures are identified that can contextualise the pixel, as CNNs do. Therefore, a combination of CNN and cluster algorithm is recommended. In this way, the CNN extracts the features while the cluster algorithm groups them together.

Clustering the coinage of the *Coriosolitae*

For the coins of the *Coriosolitae*, we decided to initially use unsupervised machine learning methods, for which we need only images without additional information, in order to create a pre-sorted dataset. We opted for an algorithm known as DeepCluster, where deep neural networks are combined with clustering algorithms, as developed by Mathilde Caron and her colleagues (Caron et al. 2018). The method combines a CNN and a clustering algorithm, such as k-Means, to build classes. The most important criteria for choosing this algorithm was that it does not require any further information from a domain expert and the approach outperformed competing algorithms based on existing challenges like ImageNet (Caron et al. 2018, table 3).

In the DeepCluster approach[16], the images are distributed by the cluster algorithm and the resulting clusters are passed on to the CNN as pseudo-classes. By repeating the iteration process multiple times, the classification should improve. The CNN searches for similar features in the clusters, and the extracted features are then used as input for the cluster algorithm, which in turn forms a new distribution (fig. 11).

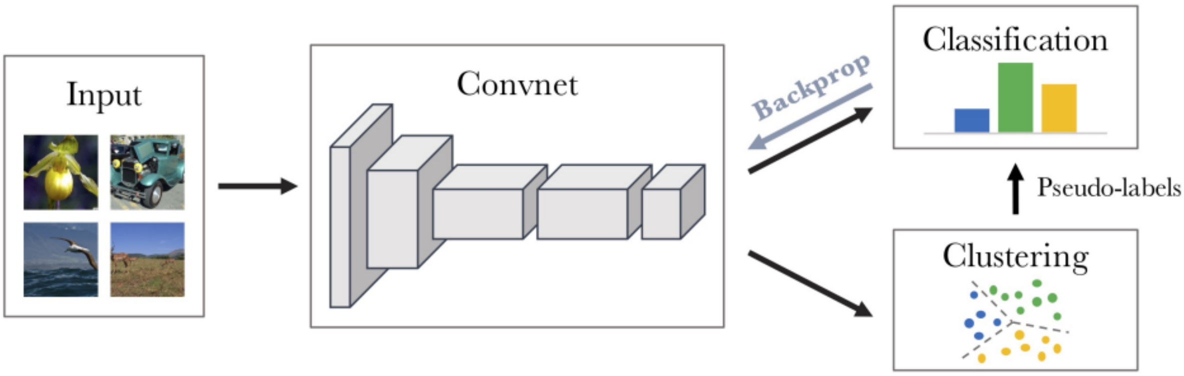

Fig. 11: Illustration of the DeepCluster method approach (after Caron et al. 2018, fig. 1).

Image data preprocessing

Before feeding the CNN with image data, preprocessing steps are needed to prepare the data. In our case, each image includes the coin, a scale and the assigned ID. In order to avoid the algorithm being biased by the scale or other elements of the image not representing the coin, we first cut out the coin (fig. 12). To do so, we used Object Detection to automatically detect the coins on the images[17]. The detected object is marked by a so-called bounding box, and the image is then cropped to the area of the bounding box.

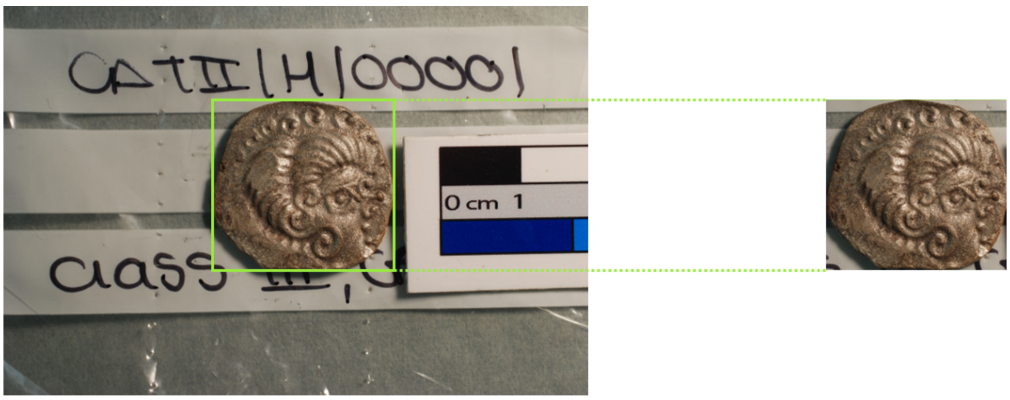

Fig. 12: Object detection methods are used to automatically crop the coins on the photos provided by our project partners (photos: Jersey Heritage; graphics: C. Deligio).

Before the cropped images are made available to the CNN, scaling follows. It is possible to choose any size, but the larger the images are, the more processing power is typically required. Also, if a pre-trained network is used, it is recommended to choose the size that it has been trained to. For our process, the images are scaled to 224x224 pixels (height *x* width),

16 https://github.com/facebookresearch/deepcluster.
17 https://github.com/tensorflow/models/tree/master/research/object_detection.

which corresponds to the trained size of the VGG-16 network. When scaling, care should be taken not to lose the proportions and level of detail.

DeepCluster experiment – unsupervised

In order to use the algorithm, the number of clusters, k, has to be defined. Caron et al. 2018 recommend choosing a significantly larger k than there are classes, and since we have an idea of the expected number of classes for the *Coriosolitae* coins, namely six according to the numismatic classification (see above), a value of k=50 was chosen for the results presented here.

For our first experiments, due to the limited processing power initially available, we used only some 25,000 coins, each with an obverse and reverse image. The aim was that with the selected algorithm the images should now be grouped according to similarity. However, we also intended to explore the possibilities of employing an unsupervised method, as this step is important from the point of view of challenging the classification of the coins. If the clusters formed are similar to the underlying classification, this will confirm the classification to a certain extent.

Of the 25,000 coins, the domain expert P. de Jersey determined that 1072 belong to one class of coins, namely class VI. Using this information, we could track the coins identified by him in the clusters created and evaluate how the algorithm grouped or distributed them.

Coin	Class
	II
	II
	I

Fig. 13: Table with class assignments used to evaluate the cluster (photos: Jersey Heritage; graphics: C. Deligio).

In our experiment, one of the 50 clusters consisted of 1162 images containing 860 of those 1072 coins identified by the domain expert to belong to class VI. This means that at least 74% of the verified coins in this class were successfully grouped together by the algorithm.

Fig. 14: left, a cluster with a high similarity; right, a mixed cluster (photos: Jersey Heritage; graphics: C. Deligio).

After this first experiment, the expert provided us with an additional table with information about the entire dataset that had been created by him and his team of volunteers who have worked on photographing, identifying and classifying the coins. The new table allowed us to evaluate the result more accurately, as it was now possible to check all clusters. We identified 17 clusters with obverse images and 12 clusters with reverse images that contained 70 % or more coins that were assigned to one of the six classes. It should also be noted that the method was able to cluster coins that were in a poor state of preservation, i.e. fragmented, corroded, etc. It also successfully separated obverses and reverses into different groups! However, it should be noted that some other clusters generated by the algorithm were less convincing.

Overall, we are very pleasantly surprised by the results. We are thus optimistic that we can help the domain expert in classifying the 70,000 coins in the hoard and contributing to the revision of existing classifications. The clusters that have a high correspondence with the expert's identifications are relevant for further research. Since they have been confirmed by human and machine, they can provide a good starting point for an approach based on supervised methods. In addition, the similarity of the distribution by the machine also provides some confirmation for the validity of the existing classification of the coins by numismatists.

Furthemore, the clusters containing coins with a poor state of preservation could also be helpful. Some of these coins cannot be classified at all, or only by experts well acquainted with the material. Thus, the identification by the machine of such problematic coins could help to avoid the time-wasting step of giving such coins to team members who do not have the appropriate knowledge. The same is true for supervised approaches we plan to set up.

Die study experiment - supervised

In another approach we used a supervised method that also combines a CNN with the k-means algorithm, but in a different way. The aim in this case was to recognize the products of individual dies by referring to the die study developed by our domain expert P. de Jersey. In this approach a CNN was trained using the existing die study based on physical observation and measurement as a ground truth. The features (vectors) forming the results of the last convolutional layer were extracted by the trained model and then used as input for the k-means algorithm.

In a first test conducted by Robin Krause as part of a bachelor thesis (Krause 2020), 1200 coins of the *Coriosolitae* belonging to class VI were used as input. When he carried out this initial work, only 50 % of the coins had then been assigned to individual dies by the domain expert. Subsequently the attributions of all of the coins were provided, facilitating an evaluation of the initial test.

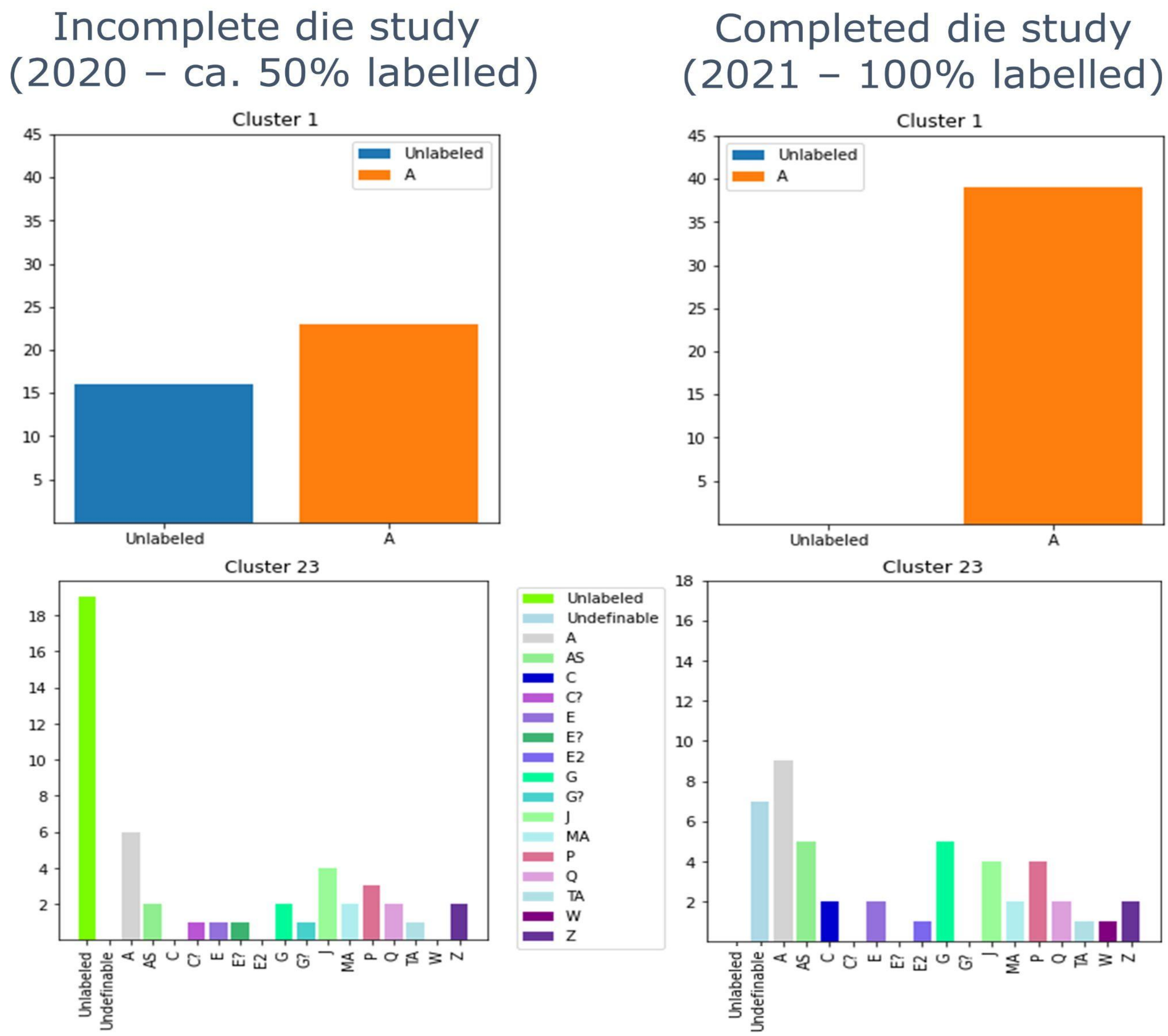

Fig. 15: Supervised clustering approach combining a convolutional neural network with the k-means algorithm. Upper row: cluster containing only one die class. Lower row: cluster containing multiple die classes (graphics: R. Krause).

Fig. 15 shows the result of this approach. On the left the results for two clusters are evaluated against the ground truth at the time R. Krause did the initial work, i.e. when only 50 % of the coins had been attributed to a die. On the right the same clusters are evaluated after P. de Jersey's die study was completed. In the upper two diagrams it is clear that in cluster 1 the algorithm successfully grouped 38 coins produced with the same obverse die together. On the other hand, dies for which only a few examples existed in the dataset were poorly grouped, as in cluster 23. It is therefore indeed possible to distinguish the individual dies with the help of AI, but a reasonable number of previously verified images is required for the method to be successful. Just how many depends, however, on the given problem. We intend to apply this approach to the full set of images of the staters of the *Coriosolitae*.

PANDA watching Jupyter: An actor-network analysis of image data preprocessing

The production and circulation of knowledge in the project ClaReNet is analysed by a study on "**P**ath Dependencies, **A**ctor-**N**etwork and **D**igital **A**gency" (**PANDA**) that is part of the project. The notion of path dependency, which denotes past decisions that create a long-term commitment to a certain way of doing things (cf. Mahoney 2000; Beyer 2005), can be used to describe archaeological practices that transform traces of the human past into data and thus manifest the path of research on and with them. It is often non-explicit, perhaps consisting of quickly made determinations which, when asked, are often vaguely described as "that's just how it's done" (see also Hofmann 2018, 202). With the help of Actor-Network-Theory (ANT) (cf. Belliger / Krieger 2006) we describe what parts the archaeological objects, the techniques at hand and the human actors have in the research process, with particular attention to the *digital agency* of software, technical devices, table entries and the like (Rammert / Schulz-Schaeffer 2002; Hofmann et al. 2019).

On the following pages we present some of the analytic possibilities of this study by means of the short example of image data preprocessing (see above, p. 15 here), in order to illustrate the kind of path dependencies we can identify, what the actor-network looks like and what in this case *digital agency* means.

JupyterLab: A central digital platform

As can be seen in Fig. 12, the photographs made by P. de Jersey's team not only show the coins, but also "elements of the image not representing the coin", including a shadow caused by the method of photographing the coins using a digital camera and LED light[18].

In order to avoid the AI falling back on non-relevant information in the image, such as the dark blue or black bar at the bottom of the scale, image data preprocessing is needed, i.e. the

18 The light source is placed "at about 300 degrees on the horizontal axis from the coin and 30-40 degrees vertically above the coin" (Email correspondence ClaReNet with Jersey Heritage, Response N. M. 18.07.2022.) The project leader of ClaReNet, David Wigg-Wolf, commented on the angle as being "unbelievably oblique", transcript of Jour fixe 02.11.2021, Line 1646. According to the ability to take photos of coins by the team members, the images differ in quality. Therefore, the team did some editing in Adobe Photoshop to adjust the light, and to straighten and crop the images where necessary for evaluation by human viewers, but not as much as was done in preparation for the DeepCluster. Finally, the team members stored the finished images in files. P. de Jersey sent the images to ClaReNet per cloud access.

background information of the 70,000 images of the *Coriosolitae* has to be cropped, a task that can only be fulfilled within a manageable time frame with the help of digital methods. Thus, the project's computer scientist, Chrisowalandis Deligio, started working on developing a solution within his field of knowledge on AI. As will be seen below, path dependencies play a powerful role in this working process and several intra- and interactions (Barad 2003), as well as interdependencies between him and technical actors can be observed.

C. Deligio employs JupyterLab (https://jupyter.org/) as a central digital platform, a coding environment in which he can use a range of digital tools with his own coding skills[19]. In order to be able to work in JupyterLab, C. Deligio first assembles tools, i.e. devices with functionalities known to him, of which some are already installed on his computer and some he has to download from the WWW. He draws on the open-source library Tensorflow (https://www.tensorflow.org/), which is already installed and contains defined mathematical terms and functions for machine learning models, and selects Object Detection API[20], a basic framework of functions which provide quick training on object detection models. With its help, he aims to train a CNN to detect the coins on the images. He imports its script into JupyterLab (Fig. 16).

C. Deligio then selects 125 suitable images[21], that is, images on which he identifies coins with different states of preservation and condition, and separates them into a folder that he can access via JupyterLab. The number of images results from his assessment that he is able to edit and evaluate this number of images himself within a manageable period of time.

He leaves JupyterLab and uses another tool publicly available on the WWW, the heartexlabs/labelImg software[22], to annotate the 125 coin representations by positioning a bounding box exactly at the rim of the coin (Fig. 12). The software creates a file for every image by using the coordinates of the pixels which form the corners of the bounding box and setting the associated label "This is a coin".

Now C. Deligio begins a process of examination, trial and testing carried out in JupyterLab, using 100 coin images for training and 25 for testing. This process C. Deligio calls "creating and training of a model", whereby "training" refers to recurring checking and correction of the output and the term "model" denotes a CNN that has to learn certain skills, in this case finding coins in photos.

19 The environment works with the language Python, a standard for coding that is unfortunately not yet certified. This "unofficial standard" (C. Deligio, Transcript of Meeting IT and PANDA 09.08.2021, Line 408) could be an initial point for a potential future path dependency, as digital sustainability is not guaranteed.
20 https://github.com/tensorflow/models/tree/master/research/object_detection.
21 Since the dataset is consistent and the coin is usually in the centre with the scale to the right, experience indicated that 100 coin images would adequately represent the full dataset for the given problem, and 25 to test the performance of the model. This proved sufficient to handle the task.
22 https://github.com/heartexlabs/labelImg.

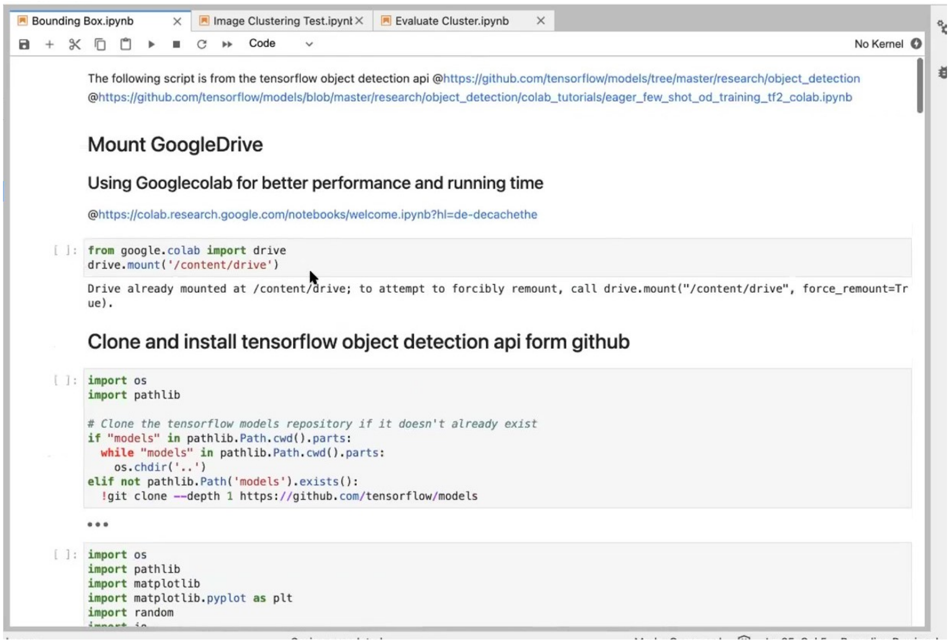

Fig. 16: The coding environment JupyterLab with the Tensorflow Object Detection API. (screenshot: J. Tietz)

Training involves coding commands that C. Deligio has to give. For example, he causes the model to save the result in a so-called “dictionary”, which he can view in JupyterLab. This dictionary, a data type with functions for data management, has been modified by C. Deligio in advance into a file repository with certain representational formats. It presents to him lists containing the file ID, the pixel coordinates (corner values) of the bounding boxes in which the model has identified the coin ({“Cat-0001”: {coin:[0,0,300,300], coin:[...}}}, i.e. Coin = (x1,x2,y1,y2)), as well as a probability indication on the accuracy of coin identification.

At the end of training, C. Deligio writes a code command that accesses the dictionary, uses the coordinates of the model plus the input image in the folder and crops the images, using an existing function to reduce the resolution of the images to 224 x 224 pixels. This self-made tool completes the set of autonomously working digital tools for the detection of the coins in the 70,000 images and the cropping of superfluous information.

From the perspective of PANDA

By observing and reconstructing this working process, the PANDA study is able to make path dependencies and the digital agencies of the actor network in this preprocessing of the images transparent. Path dependencies quite obviously occur in the method of object photography conducted by P. de Jersey’s team. The coin images, depicting the coin properly aligned and placed on its storage bag close to the inventory number and class assignment noted on the bag, as well as a scale for size estimation, contain much of the information archaeologists and numismatists need at this stage (with the exception of information on material and weight). This form of representation was chosen by P. de Jersey and his team as a consequence of two path dependencies: 1) the coin alignment follows the standard form of representation in numismatic catalogues, 2) displaying additional object information follows standard practices for recording archaeological finds. But in the context of AI, the result of 2) is to produce *bricolage* (Lévi-Strauss 1966, 29; Rheinberger 2001, 245), with the effect that the photo has to be reduced to the coin image. The term *bricolage* can be used to describe the kind of knowledge production in which certain knowledge and tools are used in trial-and-error procedures and which are specific to place, actor and time. If the experiences gained are then standardised through publications such as these, on Github or as best practice guides, new path dependencies can emerge. On the other hand, the standardised alignment of the coins in accordance with the natural orientation of the motif on them is an advantage in the context of AI, as its uniformity facilitates the training process.

In our case, the *bricolage* (Lévi-Strauss 1966; cf. Schreiber 2015, 203-206) involves the following different actors working together in the setting[23] of the digital platform JupyterLab. C. Deligio, the specialist for AI, his computer, digital images of coins, the designs on the coins, off-the-shelf programming functions, as well as the length of time available, all interact in a momentary net held together by the given task. While C. Deligio conducts his tools with the help of coding commands that comprise his agency, it is important to remember that the whole process is a result of the current level of development of DeepCluster algorithms, which ideally need unambiguousness in order to be trained with the computer available, as well as the amount of time and images at hand.

The introduction of digital technologies is changing the meaning of representations. CNNs are new image recipients with specific needs that will be taken into account in the future. How algorithms and digitisation will affect the distribution of agencies in the classification of coins can be addressed by actor network theory and path dependencies. But not only CNNs, also the preprocessing of data must be taken into account, as we have been able to show here.

Outlook

The next step of the ClaReNet project will be to look at two other coin series, the bushel *quinarii* from southern Germany and Switzerland and the *monnaies à la croix* from southern France (fig. 1). The main challenge posed by the bushel *quinarii* results from the fact that the depiction of the head on the obverse grows more and more abstract over time, gradually becoming a tuft or bushel. Due to this gradual nature of process, it is not as easy to divide the bushel *quinarii* into classes as it is for the coinage of the *Coriosolitae*. This development poses great challenges not only for numismatics, but also for machine learning techniques. We thus hope to combine different approaches in order to harmonise the existing classifications, of which there are several, and make fluid transitions in the development of the designs on the coinage apparent, while at the same time taking manufacturing processes into account. Another challenge will be classifying coins that are difficult to assign, for example because the symbols on the reverse that characterise individual classes are often not visible on the coins. The question will be whether it is possible to classify such coins with AI-based methods and whether the numismatic community will then accept the decision.

23 The term is used in actor-network theory to denote a network of human and non-human constellations recognised and explored by an observer (Latour 1992; Belliger / Krieger 2006, 44; 399).

For the *monnaies à la croix* we are faced with the challenge that there are a large number of types that are defined by the combination of the symbols in the fields of the cross on the reverse, but generally relatively few, or sometimes even no, images available to us for each one. To address this challenge, we intend to extract information from the type descriptions and to use the information gained in this way as a starting point for image recognition methods. One option we have in mind is to use the extracted information in order to connect similar type descriptions, thus effectively providing more images per type group.

Furthermore, the introduction of machine learning methods not only has great potential for work on classifications. It also provides a particularly suitable moment for fundamental discussions about research practices, changes in object epistemologies and of media, and the interplay of theories and methods in the context of concrete empirical case studies. This requires – especially in interdisciplinary projects – conceptual work. Therefore, we are currently trying to present in a transparent manner the different terminologies and concepts used by experts, for example classification and typology, in a wiki called ClaReTerms, which is publicly accessible on the project blog (https://clarenet.hypotheses.org/clareterms). It is only through the combination of the history of science, STS and current method evaluation of the different subjects involved that it is possible to reflect on research paths in all their complexity and on their impact.

The digital revolution in particular promises new insights into the classification of mass products such as coins. Our initial findings suggest that these no longer even need to be highly standardised. However, close cooperation and intensive exchange between AI, various scholars and domain experts will always be a basic prerequisite in order to obtain results that are of value for the experts, in our case archaeologists, numismatists and computer scientists. AI can and will not replace human experts, but is a new important player in the field, and one that will change the game.

Literature

Barad, Karen. 2003. 'Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter'. *Signs* 28 (3): 801–31. doi:10.1086/345321.

Belliger, Andréa, and David J. Krieger, eds. 2006. *ANThology: ein einführendes Handbuch zur Akteur-Netzwerk-Theorie*. ScienceStudies. Bielefeld: Transcript-Verl.

Beyer, Jürgen. 2005. 'Pfadabhängigkeit ist nicht gleich Pfadabhängigkeit! Wider den impliziten Konservatismus eines gängigen Konzepts / Not all path dependence Is alike – A critique of the "Implicit Conservatism" of a common concept'. *Zeitschrift für Soziologie* 34 (1): 5–21. doi:10.1515/zfsoz-2005-0101.

Blanchet, Adrien. 1905. *Traité des Monnaies Gauloises*. Paris: E. Leroux.

Caron, Mathilde, Piotr Bojanowski, Armand Joulin, and Matthijs Douze. 2019. 'Deep clustering for unsupervised learning of visual features'. *ArXiv:1807.05520 [Cs]*, March. http://arxiv.org/abs/1807.05520.

Colbert de Beaulieu, Jean-Baptiste. 1948. ‘La trouvaille de monnaies celtiques de Saint-Jacques-de-La-Lande’. *Revue Belge de Numismatique* 94: 15–76.

———. 1957. ‘Le Trésor de Jersey-11 et la numismatique celtique des deux Bretagnes’. *Revue Belge de Numismatique* 103: 47–88.

———. 1973. *Traité de Numismatique Celtique 1. Méthodologie des ensembles*. Vol. 135. Ann. Litt. Univ. Besançon. Besançon.

Cottam, Elizabeth, Philip de Jersey, Chris Rudd, and John Sills. 2010. *Ancient British Coins: ABC for Short*. Aylsham, Norfolk: Chris Rudd.

De Callatay, François. 2007. ‘L’historique de l’étude des liaisons de coins (XVIIIe-XXe s.)’. *Bulletin de la Société Française de Numismatique* 62 (4): 86–92.

De La Tour, Henri. 1892. *Atlas de Monnaies Gauloises préparé par la Commission de Topographie des Gaules et publié sous les auspices du Ministère de l’Instruction Publique*. Paris: Plon & Nourrit.

Di, Wei, Anurag Bhardwaj, and Jianing Wei. 2018. *Deep Learning Essentials*. Packt.

Forrer, Robert. 1908. *Keltische Numismatik der Rhein- Und Donaulande*. Strasbourg.

Gruel, Katherine. 1990. ‘Le Trésor du Câtillon (Jersey 11). Réexamen à la lumière des fouilles et études plus récentes'. In *La Bretagne et l’Europe préhistorique: Mémoire en hommage à Pierre-Roland Giot*, edited by Jean L’Helgouach and Pierre-Roland Giot, 293–98. Revue Archéologique de l’Ouest. Supplement 2. Rennes: Langouet.

Haselgrove, Colin. 1987. *Iron Age Coinage in South-East England: The Archaeological Context*. BAR British Series 174. Oxford: B.A.R.

Haselgrove, Colin. 1999. ‘The development of Iron Age coinage in Belgic Gaul’. *Numismatic Chronicle* 159: 111–68.

Hilgert, Markus, Kerstin P. Hofmann, and Henrike Simon, eds. 2018. *Objektepistemologien. Zur Vermessung eines transdisziplinären Forschungsraums*. Berlin Studies of the Ancient World 59. Berlin: Edition Topoi.

Hiriart, Eneko. 2017. *Catalogue des monnaies celtiques 2. Les monnaies à la croix*. Paris: Bibliothèque nationale de France, Musée d’Archéologie nationale. http://editions.bnf.fr/catalogue-des-monnaies-celtiques-2-les-monnaies-%C3%A0-la-croix.

Hofmann, Kerstin, Susanne Grunwald, Franziska Lang, Ulrike Peter, Katja Rösler, Louise Rokohl, Stefan Schreiber, Karsten Tolle, and David Wigg-Wolf. 2019. ‘Ding-Editionen. Vom Archäologischen (Be-)Fund Übers Corpus ins Netz’. In *e-Forschungsberichte des DAI, Fasc.2 (2019)*, Fasc.2, 1-12. Berlin: Deutsches Archäologisches Institut.

Hofmann, Kerstin P. 2014. ‘Auf der Suche nach der Jastorf-Fibel: Die ältereisenzeitlichen Plattenfibeln Norddeutschlands: Eine Leitform?’ In *Das Jastorf-Konzept und die vorrömische Eisenzeit im Nördlichen Mitteleuropa: Beiträge der Internationalen Tagung zum einhundertjährigen Jubiläum der Veröffentlichung ‘Die ältesten Urnenfriedhöfe bei Uelzen und Lüneburg’ durch Gustav Schwantes 18.-22.05.2011 in Bad Bevensen*, edited by

Brandt and Björn Rauchfuß, 129–42. Veröffentlichungen des Helms-Museums, Hamburger Museum für Archäologie und die Geschichte Harburgs. Hamburg: Archäologisches Museum. https://publications.dainst.org/journals/efb/article/view/2236/6674.

Jersey, Philippe de. 2016. ‘Colbert de Beaulieu, the Coriosalitae and the Jersey Hoards’. *Revue Belge de Numismatique et de Sigillographie* 162: 159–78.

Krause, Robin. 2020. ‘Clusterbildung keltischer Münzen basierend auf Convolutional Neural Networks’. Johann Wolfgang Goethe-Universität Frankfurt am Main.

Lévi-Strauss, Claude. 1966. *The Savage Mind. The Nature of Human Society Series*. Chicago.

Mahoney, James. 2000. ‘Path dependence in Historical Sociology’. *Theory and Society* 29 (4): 507–48. doi: 10.1023/A:1007113830879.

Mattingly, Harold, and Edward Allen Sydenham. 1923. *The Roman Imperial Coinage. Vol. I: Augustus to Vitellius*. London: Spink & Son.

Moroney, Laurence. 2020. *AI and Machine Learning for Coders: A Programmer's Guide to Artificial Intelligence*. 1st ed. O'Reilly Media.

Muret, Ernest, and Anatole Chabouillet. 1889. *Catalogue des Monnaies Gauloises de la Bibliothèque nationale*. Paris: Plon Nourrit.

Polenz, H. 1982. ‘Münzen in Latènezeitlichen Gräbern Mitteleuropas aus der Zeit zwischen 300 und 50 vor Christi Geburt’. *Bayerische Vorgeschichtsblätter* 47: 27–222.

Rammert, Werner, and Ingo Schulz-Schaeffer, eds. 2002. *Können Maschinen handeln? Soziologische Beiträge zum Verhältnis von Mensch und Technik*. Frankfurt/Main ; New York: Campus.

Rao, Delip, and Brian McMahan. 2019. *Natural Language Processing with PyTorch: Build Intelligent Language Applications Using Deep Learning*. First edition. Beijing ; Boston ; Farnham: O'Reilly. http://digitale-objekte.hbz-nrw.de/storage2/2019/03/03/file_11/8270927.pdf.

Rheinberger, H. J. 2001. *Experimentalsysteme und epistemische Dinge. Eine Geschichte der Proteinsynthese im Reagenzglas*. Göttingen.

Roymans, Nico. 1990. *Tribal Societies in Northern Gaul. An Anthropological Perspective*. Cingula 12. Amsterdam: Najade Press.

Simonyan, Karen, and Andrew Zisserman. 2015. ‘Very Deep Convolutional Networks for Large-Scale Image Recognition’. arXiv. doi:10.48550/arXiv.1409.1556.

Ziegaus, B. 1994. *Das Geld der Kelten und ihrer Nachbarn. Sammlung Josef Schörghuber. Ausstellung Vom 28. Oktober 1994 bis 8. Januar 1995 in der Prähistorischen Staatssammlung München*. Ausstellungskataloge Der Prähistorischen Staatssammlung 26. München: Prähistorische Staatssammlung München.

Acknowledgements

We thank the Bundesministerium für Bildung und Forschung for funding the ClaReNet project and the following institutions as cooperation partners:

Bibliothèque nationale de France, Département des Monnaies, médailles et antiques
Archäologische Staatssammlung, München
Inventar der Fundmünzen der Schweiz, Bern
Université Bordeaux Montaigne, IRAMAT CRP2A
States of Guernsey, Culture and Heritage