



McEwan, John Alexander, Weronika Grajdura, Christopher D. Hopwood, Karson Million, and Aliénor V. De Smedt. 2024. "Machine Learning and Sigillography: Using Decision Trees to Date British Seal Matrices," in "Digital Approaches to Medieval Sigillography," ed. Martina Filosa, Claes Neufeind, and Claudia Sode, special collection. *Digital Medievalist* 17(1): 1–26. <https://doi.org/10.16995/dm.15125>.



## Machine Learning and Sigillography: Using Decision Trees to Date British Seal Matrices

**John Alexander McEwan**, University of Kansas, United States, [john.mcewan@ku.edu](mailto:john.mcewan@ku.edu)

**Weronika Grajdura**, Saint Louis University, United States, [weronika.grajdura@slu.edu](mailto:weronika.grajdura@slu.edu)

**Christopher D. Hopwood**, Saint Louis University, United States, [chris.hopwood@slu.edu](mailto:chris.hopwood@slu.edu)

**Karson Million**, Saint Louis University, United States, [karson.million@slu.edu](mailto:karson.million@slu.edu)

**Aliénor V. De Smedt**, Saint Louis University, United States, [alienor.desmedt@slu.edu](mailto:alienor.desmedt@slu.edu)

---

The availability of software libraries for machine learning, in combination with the growing number of digital sigillographic information resources, enables scholars to explore how this technology can help date medieval seal matrices. This paper presents and evaluates a pioneering system based on archival seal impression data, courtesy of the Digisig Project. The system constructs a decision tree that can be employed to date seal matrices. To explore the potential practical applications of the method, we compare its outputs with the dates assigned to seal matrices by the cataloguers of the Portable Antiquities Scheme and the Schøyen Collection catalogue. This method does not replace human cataloguers, but it can quickly and cost-effectively identify sections of existing catalogues that might benefit from revision.

---

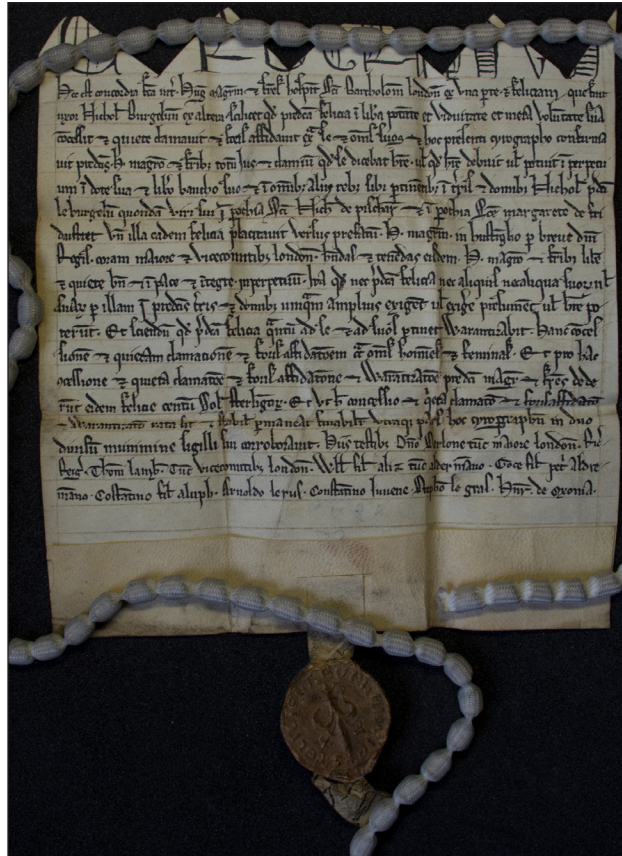
La disponibilité des bibliothèques logicielles pour l'apprentissage automatique, combinée à l'augmentation du nombre de ressources d'information sigillographiques numériques, permet aux chercheurs d'explorer comment cette technologie peut aider à dater les matrices de sceaux médiévaux. Cet article présente et évalue un système pionnier basé sur des données d'impression de sceaux d'archives, grâce au projet Digisig. Le système construit un arbre de décision qui peut être utilisé pour dater les matrices de sceaux. Pour explorer les applications pratiques potentielles de la méthode, nous comparons ses résultats avec les dates attribuées aux matrices de sceaux par les catalogueurs du Portable Antiquities Scheme et du catalogue de la collection Schøyen. Cette méthode ne remplace pas les catalogueurs humains, mais elle peut identifier rapidement et de manière rentable les sections des catalogues existants qui pourraient bénéficier d'une révision.

---



## 1. Introduction

§ 1 During the Middle Ages, people used seals to authenticate, validate and securely close documents (Figure 1).



**Figure 1:** Quitclaim by Felice, formerly wife of Nicholas Burgelun, to Hugh, master of Saint Bartholomew's Hospital, London. Dated 1221–1222. Saint Bartholomew's Hospital Archives, deed 124 (SBHB/HC/1/124). Photograph: John Alexander McEwan.

As seals present words and images, people could also use them to make statements of identity (Bedos-Rezak 2000; Heslop 1987, 114). Consequently, seals offer scholars from many different disciplines—including history, art history, literature, and archaeology—evidence for social, family, and occupational networks, as well as devotional practices, political ideas, gender, visual culture, and other facets of human experience. However, seals must be accurately dated before scholars can make use of their evidence, and dating seals is not always straightforward. Seals survive as both seal impressions, often attached to documents, and as the matrices or stamps used to make seal impressions. Seal impressions attached to documents (Figure 2) can usually be securely dated from information in the document (McEwan 2022, 44–45), but seal matrices are typically

recovered from the ground with little accompanying evidence to show when exactly they were in use. As it is rare for both a medieval seal matrix (**Figure 3**) and a medieval seal impression made from it to survive, seal matrices can be challenging to date in the absence of seal impressions (Linenthal and Noel 2004, xiv).



**Figure 2:** Seal of Felice, formerly wife of Nicholas Burgelun. Stylized lily. Round, 30 mm. Saint Bartholomew's Hospital Archives, deed 124 (SBHB/HC/1/124), dated 1221–1222. Photograph: John Alexander McEwan.



**Figure 3:** Copper alloy seal matrix with suspension loop handle. Round, 20 mm. Lion rampant. Chilcomb, Hampshire. PAS ID: SUR-362D54: <https://finds.org.uk/database/artefacts/record/id/1137604>. Rights holder: Surrey County Council.

Happily, machine learning technology can help.

§ 2 Freely available software libraries for machine learning, in combination with a growing number of digital sigillographic information resources, enable scholars to explore how machine learning can assist in dating seal matrices, such as those recorded by the Portable Antiquities Scheme (PAS). The PAS dataset (PAS 2024) has special importance for students of English seals, partly because of its large size, but also because of the seal matrices it records (Gill 2010). The curators of the PAS dataset regularly add entries, but at the time of preparation of this study, the dataset recorded over 5,500 medieval seal matrices (McEwan 2024a). The PAS dataset includes numerous seal matrices discovered by members of the public, outside the context of formal archaeological excavations. Seal matrices, being made of hard materials such as metal or stone, can turn up in construction debris, on riverbanks, or on agricultural land (Robbins 2014, 11; Anderson 2008). It is likely that many of these seals were discarded or lost by their original owners, and thus the PAS dataset probably includes many seals used by relatively humble people. By contrast, archival collections focus on seals attached to documents, and thus include a disproportionate number of seals used by relatively important and wealthy people, whose affairs were most extensively documented and whose records contemporaries made a special effort to preserve. Therefore, if scholars aim to examine seal usage by people outside the aristocracy (Harvey 1991, 117–118; New 2019, 279–309), PAS data is exceptionally valuable. However, the seal matrices first need to be dated as closely and accurately as possible.

§ 3 Cataloguers generally date medieval seal matrices based on their contents and features. Through research work, cataloguers identify seals that are already securely dated and are similar to the seal matrices in question (Linenthal and Noel 2004, xiv). This paper presents and evaluates a pioneering system, based on machine learning technology, for performing this research and analytical work and proposing dates for seal matrices. Archaeologists have been using computers to sort and organize collections of artifacts for decades (Cowgill 1967; Huggett 2014) and have adopted machine learning tools in that work (Bickler 2021). Thus, this project is innovative not in its use of machine learning to date artifacts, but in its application of machine learning specifically to medieval seal matrices and in its use of archival data in the training process, which thereby bridges the domains of archives and museums. This paper demonstrates this method's potential contributions to scholarship by applying it to the seal matrices in the Schøyen Collection catalogue (Linenthal and Noel 2004) and the PAS dataset. This pioneering machine learning algorithm produces dates that accord with those assigned by cataloguers, but it accomplishes this task in a distinctive way. The specific method outlined in this paper cannot replace human cataloguers due to its limitations, but it can assist cataloguers to date seals and to identify entries in existing catalogues that might deserve revision. Moreover, it is extremely cost-effective.

## 2. Training data

§ 4 The creation of an automated tool for dating seal matrices is a multi-step process, but the first, and perhaps most important, step is the creation of a “training” dataset. The training dataset is the foundation of the project, for it provides the computer with an understanding of the subject in question. In this study, it was necessary that the training dataset include various types of seals from a period of several hundred years in order to introduce the computer to variation within and across time periods. For not only did the seals people favoured change over time, but even within a given period, seals varied in size, shape, and content (Heslop 1987, 116; McEwan 2019). Fortunately, medieval seals survive in large numbers, so there is no shortage of examples. The precise number of seals that survive from northwestern Europe is unknown, but it has been estimated to be in the millions. Hundreds of thousands of seals survive from England alone. The National Archives (previously the Public Record Office), for example, has a vast collection (Jenkinson [1954] 1968, ix); Harvey has estimated that it could hold 250,000 seal impressions, of which perhaps 50,000 are medieval (Harvey 1996, 29). The British Library’s existing catalogue, which describes many thousands of seals, covers only a fraction of its holdings (Birch 1887–1900; McGuinness 1995, 165; Harvey 1991, 117–118). In practice, the easiest seals to study are those that have been described by cataloguers or scholars, either in printed catalogues (New 2010, 129–130; Harvey and McGuinness 1996, 120–121) or in digital formats (Harvey 1996).

§ 5 As this project involves computer processing, the real limit on the size of the training dataset is the number of seal descriptions that can be assembled in a machine-readable format. At the time of writing, there are sigillographic datasets for many places in Europe (for example, see Hablot 2019). These datasets are not currently interoperable, so it is difficult to assemble the data for a large-scale study, but they do enable studies of the seals in various regions. This study uses data from the Digisig project (McEwan 2018; McEwan 2022) to examine seals of the people of Britain. Digisig (McEwan 2024b) aspires not only to enable scholars to discover and search for seals (particularly medieval seals from England), but also to compare collections of seals and to study how seals changed over time. The project’s website and its database have been in development for more than a decade and remain works in progress. At the time of writing, Digisig holds information on more than 45,000 seals in a machine-readable format. This wealth of information offers a solid foundation for a pioneering machine learning project focused on British seals.

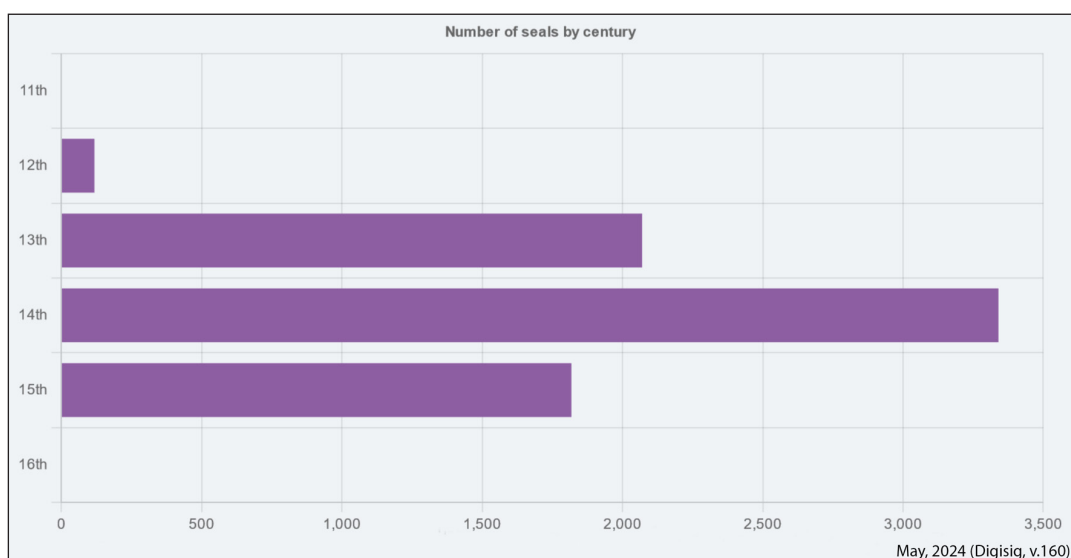
§ 6 Digisig assembles information from published seal catalogues and other information resources that archives and museums provide to researchers. That

information typically includes data on the size, shape (Harvey 1996, 29–36), and visual content (McEwan 2015) of seals, as well as the dates and locations where they were used. As the formats of those resources vary, Digisig converts that information into a single format to facilitate searches. That standardized information is the starting point for the machine learning system. Location data is not included in the training dataset, as this study focuses on one area, but could be used in future studies. The shape, visual content, and date data are used without alteration; however, the date and dimension data require additional explanation, as they are so important to this project. Digisig assigns dates to seals known from seal impressions using an algorithm that gives special importance to the earliest and most securely dated seal impressions. However, cataloguers may assign dates in the form of a specific day, month, or year, or as a span of time, such as “fourteenth century.” As this diversity of formats makes searching and sorting by date more complex, Digisig standardizes seal dates as a single number representing a year. If it is a span date, such as a regnal year that crosses a calendar year, then the year is determined by taking the mid-point of the span (although it preserves information about the breadth of the span). As this project adopts Digisig’s dates, the date format used in this project is the calendar year. Digisig’s values for the dimension of seals also require some adjustment. Cataloguers typically record the height and width of seals separately, and that is how those values are stored in Digisig. However, as area provides a clearer indication of the scale of the seal, in this study those values are replaced with a single number representing the area. The result is a training dataset that includes information on shape, visual content, date, and area.

§ 7 Once the format of the training dataset is established, a sample of seals is selected from Digisig. However, only a fraction of the thousands of seals listed in Digisig are useful for this project. Seal impressions are fragile objects, and many are chipped, fractured, or illegible. When a catalogue description is missing information, then the seal is eliminated from the training set. Furthermore, as the immediate goal of this project is to revise the dating of seal matrices in the PAS dataset, the computer does not need to be trained to date the seals of kings, nobles, archbishops, bishops, abbots, or corporate entities (such as religious houses) that frequently survive in the archives but are exceptional in the PAS dataset. Moreover, the seals of the most privileged members of society have distinctive features, including exceptional size (McEwan 2019, 109–110), that set them apart from the seals of the broader population. As size is a good indicator of date among the seals of those outside the aristocracy, seals of the elite are eliminated not only to focus the project on its intended goal but to avoid the additional

complexities that these seals would introduce. Seals known from impressions with span dates broader than a decade are likewise removed from the training dataset to ensure that the training dataset includes only seal impressions that can be relatively precisely located in time. Seals known from seal matrices (rather than from seal impressions on dated documents) are likewise excluded from the training dataset and set aside for use in the testing or “validation” process (see below). This ensures the quality of the data but reduces the number of cases—especially of those from the twelfth and thirteenth centuries, when it was less common for documents to be dated (Gervers 2000, 14). Once Digisig has been searched for seals known to have circulated in Britain that survive in the form of closely datable archival seal impressions that provide information about the size, shape, and visual content of the seal, and seals of the elite and corporate bodies have been removed, the result is a training dataset of about 7,300 cases.

§ 8 These cases are reasonably well distributed with regard to content, date, and location. Digisig uses a hierarchical classification system for the visual content of seals, which sorts identifiable seals into four top-level categories (McEwan 2015). All four categories are well represented in the training dataset; most seals fall into the “object” (38%) class, with smaller portions in the “animal” (25%), “device” (19%) and “human” (16%) classes. The temporal distribution is also good. The training dataset includes a few twelfth-century cases, but the number of cases climbs dramatically from the beginning of the thirteenth century, when seals used by people outside the aristocracy start to survive in significant numbers (**Figure 4**).



**Figure 4:** Number of seals by century.

Finally, the seals originate from areas throughout Britain (Figure 5).

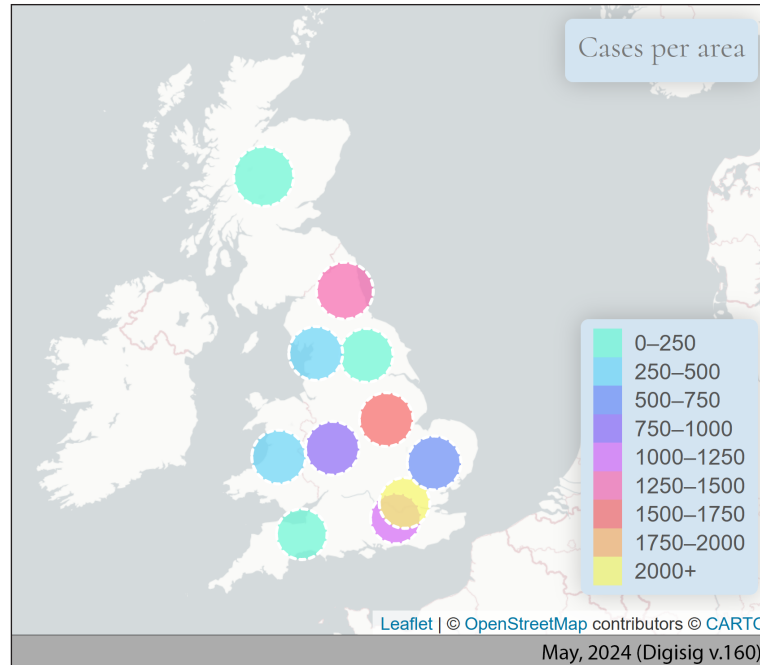


Figure 5: Number of cases by area.

There are significant concentrations of seals from the East Midlands, London, and the Northeast. Only a handful of Scottish seals are included, so this is one area where the dataset could be improved. In sum, the dataset offers good coverage of the period 1200 to 1500, although it is dominated by English seals.

### 3. Training process

§ 9 For this project, we selected the decision tree regressor method, as implemented by Scikit-Learn (Pedregosa et al. 2011). For a pioneering project, it was critical that the results could be easily interpretable and decision trees could be graphed. A decision tree with seven levels was created that divides the seals into 63 groups (leaf nodes) based on a set of decisions (Figure 6).

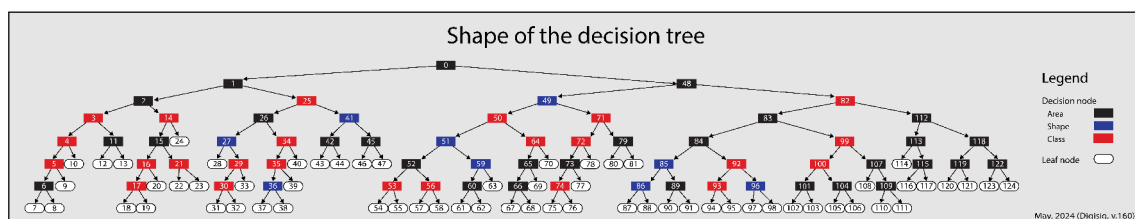


Figure 6: Shape of the decision tree (May 2024).



Each decision evaluates a single feature of the seal, such as the area (black nodes), shape (blue nodes), or class (red nodes). For example, the initial decision in “node 0” considers whether the seal has an area greater than or equal to 191.995 mm squared. By contrast, “node 49” asks whether the seal has the shape “pointed oval” or not. Through a sequence of such decisions, the decision tree provides a mechanism to sort seals into groups, termed “leaf nodes,” of seals with similar features. By examining the sequence of decisions, researchers can better understand how the decision tree sorts seals (Figure 7).

Decision path		
Node	Question	Answer
0	Is size greater than 192.0 mm <sup>2</sup> ?	Yes
48	Is size greater than 492.12 mm <sup>2</sup> ?	No
49	Is it shape/class 'pointedoval'?	Yes
71	Is it shape/class 'radial'?	No
72	Is it shape/class 'standing'?	No
73	Is size greater than 468.41 mm <sup>2</sup> ?	No
74	Is it shape/class 'objects'?	No
75	Prediction: c.1249-1296	

Figure 7: Sample decision path.

§ 10 Some features of seals are more influential than others in determining the groupings. As already discussed, the training dataset includes information about the area, shape, and class of each seal. The decision tree takes into account all these features, but it places special emphasis on the area. The first and second decisions involve the area; thereafter, decisions are based variously on area, shape, or class. Area is prioritized partly because the seals of people outside the aristocracy tended to become smaller in size over the course of the thirteenth and fourteenth centuries, so size is a strong indicator of date. Furthermore, every seal has a size, so the size of the

seal is a feature that can be profitably considered for every seal. By contrast, seals can have varying shapes and a wide range of images, so it is more difficult to sort them based on these features. The result of the training process is a decision tree that gives special consideration to the area of the seal and refines its groupings based on shape and class.

§ 11 Once the decision tree has sorted the seals in the training dataset, the temporal distribution of the seals in each group can be calculated to show when they were in circulation. At the time of writing, Digisig offers scholars a visualization (McEwan 2024c) of the temporal distribution of the seals in the training dataset associated with a group (Figure 8).

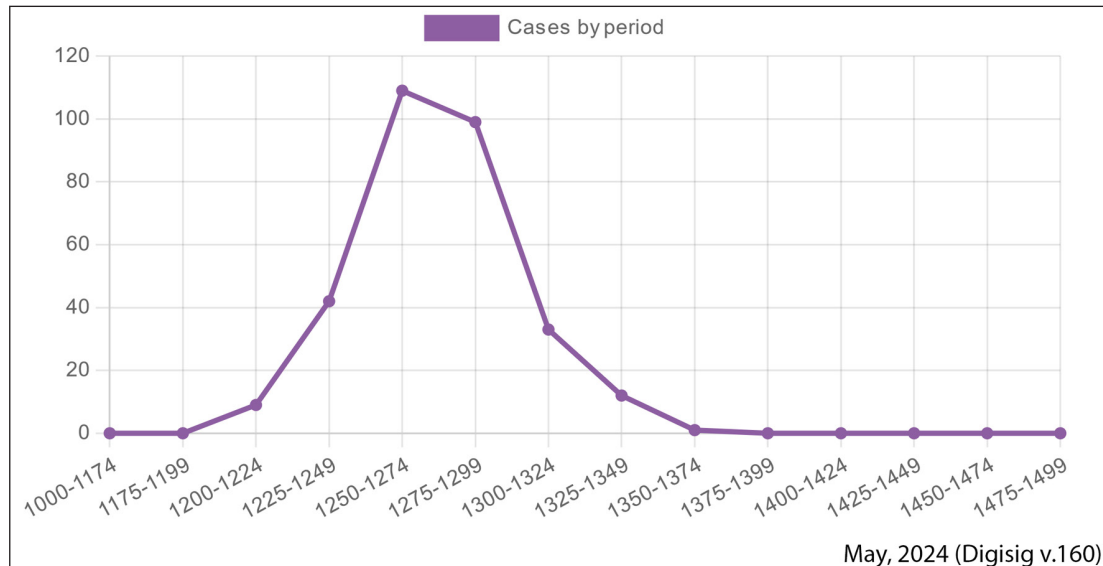


Figure 8: Temporal distribution of leaf node 27.

Typically, the temporal distribution takes the form of a curve, with a few seals earlier and a few later, but most concentrated in a particular period. These groupings could be described by assigning them a single year that approximates the “peak” of the period when the seals in the group were in circulation. This “peak” is useful information, but the span of time encompassing the period when such seals were commonly in circulation is perhaps even more useful to scholars. To help researchers appreciate the period represented by each grouping, Digisig currently creates a span date by dividing the seals from the training dataset in each group into six quantiles and then calculating a span that encompasses the seals in the second quantile through the fifth quantile. The result is a span date that describes the central section of the temporal distribution curve.

§ 12 The groups are based on varying numbers of seals, which were in circulation in a variety of different periods. As already discussed, each group contains a minimum of 20 cases. The median number of cases in a group is 31, the mean is 116, and the maximum is 1,366, so there is a mixture of groupings that represent seals with a rare combination of features and seals with features that are relatively common. Temporally, the groups are well distributed (Figure 9).

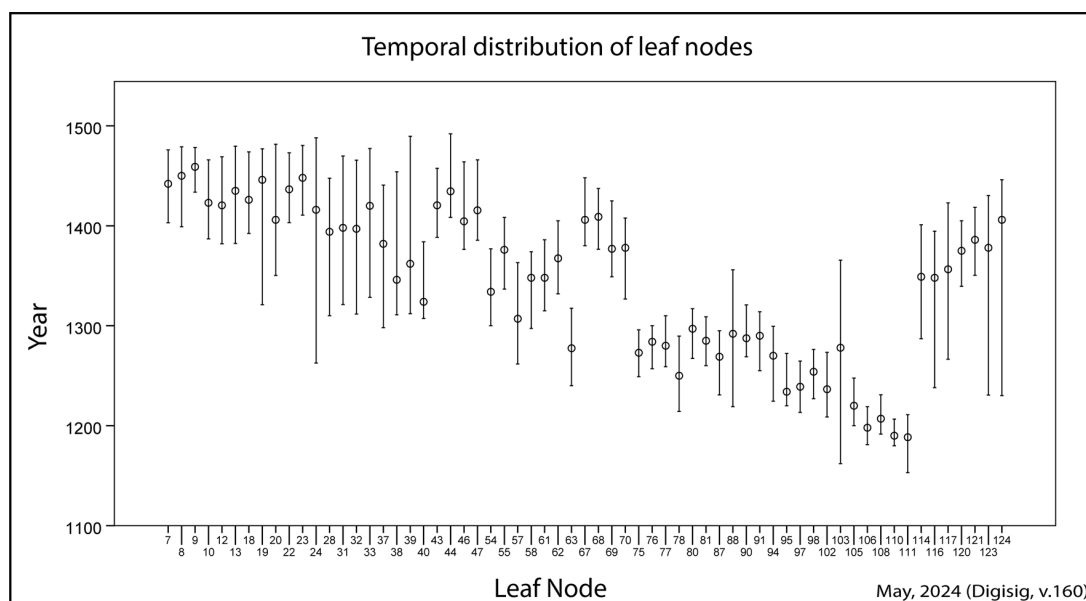


Figure 9: Temporal distribution of leaf nodes.

Of the 63 groups established by the decision tree, the central date, or “peak,” falls in the late twelfth century for three groups, in the thirteenth century for 22, in the fourteenth century for 17, and in the fifteenth century for 17. The span dates of these groups encompass a varying number of years. The minimum span of time is 26 years, the median is 92 years, and the maximum is 226 years, so some groupings represent seals that were in circulation for long periods of time, and others those in circulation for more limited periods. These groupings provide a finite number of ways that any particular seal can be dated, but the groupings are well distributed across the entire period c. 1200–1500.

§ 13 The decision tree thus constructed can now be used to date a seal matrix. Starting at the top of the tree, the decision tree lists a series of questions (called “nodes”) that can be answered yes or no. When the answer is yes, the computer takes the path to the right, but when the answer is no, the computer passes to the left. This process continues until the computer reaches an end to the chain of questions and can thus assign the seal to a group and a date.

#### 4. Testing process

§ 14 The decision tree must be tested to compare its proposed date spans to those determined by cataloguers using more conventional methods. A testing or “validation,” process for a machine learning tool normally involves asking it to make predictions for cases that were not used in the training process, but for which the “correct” answer is already known. For example, medieval manuscript scholars have “trained” a computer to decipher handwriting by presenting it with examples of a particular script, then “tested” it by presenting it with additional examples that it has not seen to determine whether it can identify letters and words correctly (Muehlberger et al. 2019). Thus, a conventional test of the decision tree would present it with seals not previously used in the training process, but whose dates are already established, to measure how closely it could predict those dates. However, the goal of the project is to produce span dates that represent the period when a seal is most likely to have been in circulation, not single dates of production or use. As already discussed, the decision tree was created using seal impressions dated to a particular year. Each date thus represents a single moment in time when a particular seal can be shown to have been in circulation. The decision tree is the result of the computer’s analysis and organization of many individual cases into groups that were in circulation between certain dates; the span date assigned to each group represents the period of time when most of the seals in the group were in circulation. Therefore, the testing process for the decision tree needs to evaluate the span dates, not to assign specific dates to individual seals. To obtain cases for the testing process, we used traditional methods to establish when certain types of seals were in circulation and then compared their span dates with those proposed by the decision tree. We also drew on published catalogues of seal matrices that date seal matrices using span dates.

§ 15 At the time of writing, scholars can use Digisig to gather information on seals that display certain images. In the autumn semester of 2023, a team of undergraduate students (Weronika Grajdura, Christopher D. Hopwood, Karson Million, and Aliénor V. De Smedt) at St. Louis University used Digisig to create histories of a selection of different types of seals. They identified examples of each type of seal, carefully examined the contexts in which they appeared, and studied the development of their features over time (cf. Blair 1943). They then proposed models that scholars could use to date seal matrices with the same features. Their results provide a good point of comparison for the machine learning tool.

§ 16 Seals depicting the hare riding to hunt on the back of a hound (Harvey and McGuinness 1996, 89) are relatively easy for human cataloguers to identify and date because the image is distinctive and because most surviving examples date from a relatively brief period of a few decades. The hare is often depicted blowing a horn and the image is typically accompanied by a legend (text around the outer edge of the seal) that includes a phrase related to hunting, such as the call “Sohou” or “I ride” (Figure 10).



**Figure 10:** Copper alloy seal matrix with a hexagonal sectioned handle. Round, 20.7 mm. Hare riding on a hound. Langtof, East Riding of Yorkshire. PAS ID: YORM-018BAF: <https://finds.org.uk/database/artefacts/record/id/622749>. Rights holder: York Museums Trust.

This image evoked for contemporaries the concept of “the world turned upside down” and was an expression of a cultural movement critical of the social order (New 2016, 110–111). At the time of writing, Digisig registers 38 seal impressions with this image, and these seal impressions represent 24 distinct seals; it is therefore comparatively rare within the training dataset, but the number of cases is still sufficient to sketch its history. The earliest securely dated examples are from the 1290s (**Figure 11**); another case appears in the first decade of the fourteenth century; a surge of cases in the second decade is sustained into the beginning of the 1350s.



**Figure 11:** Hare riding on a hound. Round, 21 mm. Dadford, Buckinghamshire, 1298–1299. Huntington Library, California, STG Evidences, Box 4, item 9. Photograph: John Alexander McEwan.

At that time, the image seems to fall out of regular circulation. If the seals with this image are treated as a group, and the method (division into quantiles) used by Digisig to formulate the span date for the decision tree groupings is employed, the result is the date span c. 1302–1351. Consequently, at the time of writing, the archival evidence indicates that a seal with the hare-riding-on-a-hound image is likely to have been in circulation in the first half of the fourteenth century.



The origins of the motif can be traced to antiquity, but in the Middle Ages, it was Christianized and came to be understood by contemporaries as expressing such concepts as resurrection and sacrifice (Hourihane 2000, 122). Whereas seals with the hare-riding-on-a-hound motif were in circulation mainly in the first half of the fourteenth century, seals depicting the pelican in its piety are currently known to have been used on British seals from the thirteenth century onwards. Digisig registers 163 separate seals known from one or more seal impressions with this image. Several examples survive for each decade from the 1240s through the 1350s. The numbers of cases decline in the 1360s, but the image never goes out of circulation; there are examples for every subsequent decade well into the early modern period. As people used the image of the pelican in its piety over many centuries, the image is of limited value in dating, so cataloguers need to consider other features of these seals, such as size and shape.

§ 19 Based on the seals that survive in archival contexts listed in Digisig at the time of writing, we can establish a provisional history of the development of the pelican-in-its-piety design in British seals. In the second half of the thirteenth century, it was common for seals with this image to be presented with pointed ovals and to have dimensions of approximately 30 × 20 mm (Figure 13).



**Figure 13:** Seal of Thomas de Ware. Pelican in its piety. Pointed oval, 29 × 19 mm. Saint Bartholomew's Hospital Archives, deed 760 (SBHB/HC/1/760), dated 1295. Photograph: John Alexander McEwan.

However, in the early fourteenth century, the dominant shape changes to round (Figure 14).



**Figure 14:** Pelican in its piety. Round, 18 mm. Saint Bartholomew's Hospital Archives, deed 351, seal 2 (SBHB/HC/1/351), dated 7 August 1338. Photograph: John Alexander McEwan.

Then, over the course of the fourteenth century, the typical size of these seals diminishes from 25 mm to 15 mm in diameter (for example, see the National Archives, DL25/2081/1766 [National Archives 2024a]). In the mid-fifteenth century, the size once again decreases to about 13 mm in diameter (for example, see the National Archives, DL 25/3533/3091 [National Archives 2024b]). The shape of these seals also changes, with the appearance of an increasing number of octagonal, square, and even rectangular examples (for example, see Bangor University, Archives and Special Collections PENR/27 [McEwan 2024d]). These changes in the size and shape of the pelican-in-its-piety seals enable the computer to date the seals reasonably well. Just as the decision tree lacks a specific group (leaf node) for the hare-riding-on-a-hound image and instead distributes cases to several different groups, it lacks a group for the pelican-in-its-piety image and similarly places cases in several separate groups (see **Figure 6**). For example, the decision tree situates pointed oval seals 30 × 20 mm with the pelican-in-its-piety image in the group dated c. 1246–1296 (leaf node 75). If the decision tree considers a pelican-in-its-piety seal that is round and 16 mm in diameter, it sets it in group c. 1300–1377 (leaf node 54), which locates it in the same group as most of the hare-riding-on-a-hound seals. However, if the pelican-in-its-piety seal has the dimensions 12 × 10 and a rectangular shape, the decision tree locates it in group c. 1392–1474 (leaf node 18). The decision tree is reasonably successful in sorting most of the pelican-in-its-piety seals chronologically.



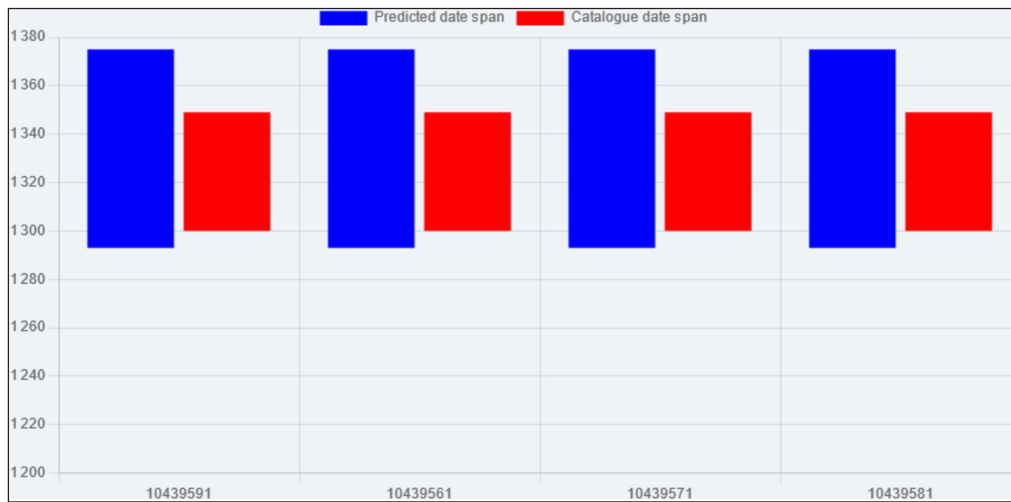
§ 20 The contrasting cases of the hare riding on a hound and the pelican in its piety demonstrate some strengths and limitations of the decision tree method. The decision tree may overlook the significance of images that are comparatively rare but good indicators of date, but it makes very good use of the size data. Indeed, the decision tree rivals a human cataloguer in situations where size provides critical evidence.

§ 21 These small-scale case studies were then supplemented with comparisons of the decision tree's outputs to those in published sigillographic reference works. Assembled in the late twentieth century, Linenthal and Noel's edition of the Schøyen Collection includes 403 seal matrices, most of which are English (Linenthal 2009, 224). Like the seal matrices in the PAS, which will be discussed shortly, the seal matrices in the Schøyen Collection are mostly "non-heraldic personal seal matrices" (Linenthal and Noel 2004, xi) and largely discovered by people using metal detectors (Linenthal and Noel 2004, xvi). To date these seals, Linenthal and Noel state that they relied on the "style" of the seals—"of the device, lettering and type of matrix itself" (Linenthal and Noel 2004, xvi)—so the catalogue reflects the results of traditional scholarly methods of dating such seal matrices. Because the catalogue describes each seal's size, shape, and visual content, we can assess the Schøyen Collection's seal matrices with the machine learning system and compare the decision tree's results to those of Linenthal and Noel. The computer's results contrast with those of Linenthal and Noel in several respects. Linenthal and Noel do not explain how they determined the date of each individual seal. That does not affect the quality of the dates themselves, but it renders the reasoning behind their date selections opaque. By contrast, the decision tree documents how it arrives at a particular date. Moreover, Linenthal and Noel situate each seal matrix in a particular span of time, like the decision tree, but they favour periods that are centuries or fractions of centuries. Indeed, they assign 72% of the seal matrices to a half-century or a full century. By contrast, the decision tree has no attachment to century marks and places the boundaries of spans all over the timeline. However, it mostly uses broader spans of time. Linenthal and Noel assign a seal to a span of time that is half a century or less in 60% of the cases, but the decision tree only does this for 15% of the cases. Thus, Linenthal and Noel tend to place their date spans at conventional but arguably arbitrary points in time, but the date spans they provide are typically narrower.

§ 22 Despite the differences of method, both sets of dates are spans of time and thus can be compared. A convenient but simplistic means of comparison is a Jaccard Index: a measurement of similarity that ranges from 1, when two sets are identical, to 0, when they are entirely different. If the index is calculated for each seal and then the results are averaged across the entire catalogue, the result is 0.4. That figure suggests that there is substantial disagreement between Linenthal and Noel and the decision tree. To some

extent, however, that discrepancy can be explained by the contrasting levels of precision. In 47% of the cases, the decision tree's date spans encompass those of Linenthal and Noel or vice versa, which suggests some consensus on the period of circulation. In two-fifths of the cases, the date spans proposed by the decision tree and by Linenthal and Noel overlap, with disagreement regarding the beginning or end of the span. Only in 12% of the cases do the decision tree and Linenthal and Noel propose date spans that do not overlap. These cases often seem to represent a failure of the decision tree, rather than of Linenthal and Noel. The decision tree struggles with seals that are exceptionally large or small for their eras, but these cases rarely fool cataloguers. Thus, the comparison of the decision tree to Linenthal and Noel demonstrates that while the decision tree can fail, in most instances, it offers date spans that encompass or are in the vicinity of those proposed by expert human cataloguers working under optimal conditions.

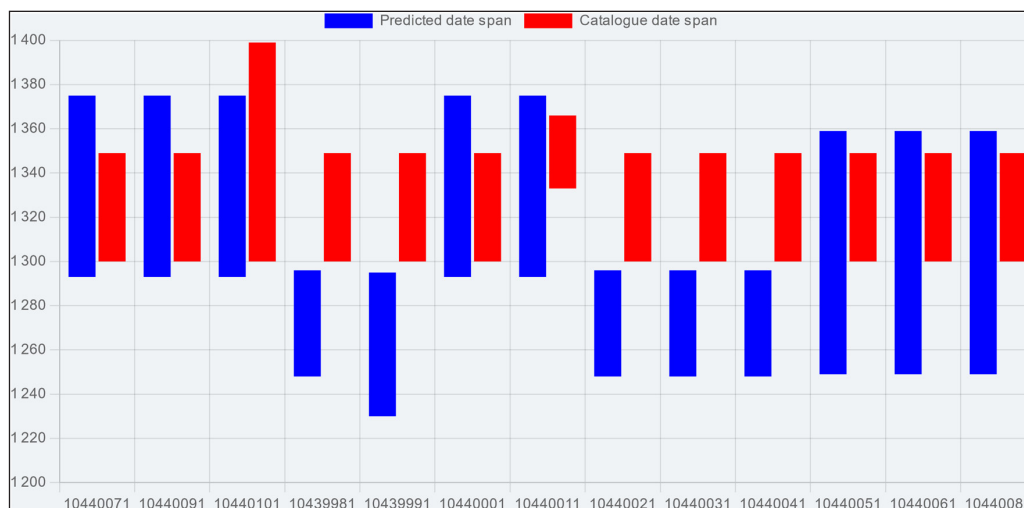
§ 23 With the broad accuracy of the decision tree's dating established, we compared the dates proposed by Linenthal and Noel and by the decision tree for two subsets of seals: those depicting the hare riding on a hound and those depicting the pelican in its piety. The Schøyen Collection includes four hare-riding-on-a-hound seals, all of which are round and which range in size from 16 to 20 mm in diameter. Linenthal and Noel date them to the first half of the fourteenth century, and the decision tree assigns them broader but similar dates (**Figure 15**).



**Figure 15:** Comparison of date spans proposed by the decision tree and Linenthal and Noel, for seals displaying the hare riding on a hound in the Schøyen Collection.

However, the Schøyen Collection also includes thirteen seals displaying the pelican in its piety. Linenthal and Noel date all thirteen to the fourteenth century: one 1300–1400, another to the “mid-fourteenth century,” and the remainder to the first half of the

fourteenth century. Arguably, Linenthal and Noel should have assigned some of the larger pointed oval examples to the late thirteenth century, rather than the fourteenth century, as the decision tree proposes (Figure 16).



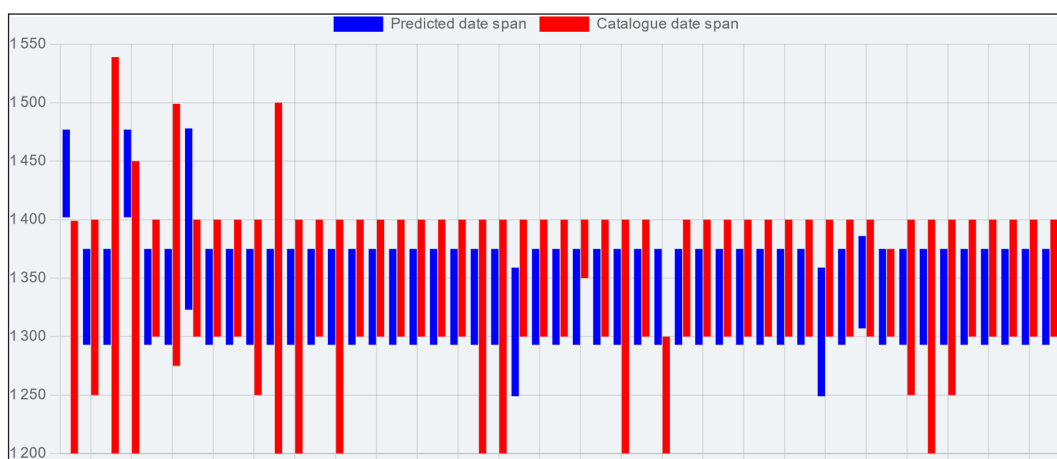
**Figure 16:** Comparison of date spans proposed by the decision tree and Linenthal and Noel, for seals displaying the pelican in its piety in the Schøyen Collection.

The machine learning tool may not be as reliable as Linenthal and Noel, but it can alert scholars to cases that invite reassessment or revision.

§ 24 If the computer, using the decision tree, can offer some assistance to scholars using the Schøyen Collection catalogue, it has the potential to be even more helpful to scholars using the PAS dataset, which at the time of writing contains in excess of 5,500 medieval seal matrices. Established at the end of the twentieth century, PAS records archaeological finds discovered by the public and makes information about those objects available to researchers. From its inception, PAS envisaged making its records publicly available on the internet; the first version of its website came online in 2001 (Pett 2010, 1), and both the website and the database that supports it developed over the following years. Many of the medieval seal matrices documented on PAS end up in private collections, such as the Schøyen Collection, rather than in museums. As most private collections are not catalogued and published, the PAS record for a seal matrix can be the best available source of information for researchers. These records are created by PAS officers, who may take photographs and measurements and then return the artifacts to their owners. PAS officers record all types of artifacts, of which medieval seal matrices are only one small subset, and process tens of thousands of artifacts each year. PAS guides have been created to assist officers in recording various types of artifacts; the guide for recording seal matrices includes advice on the description of a seal matrix's

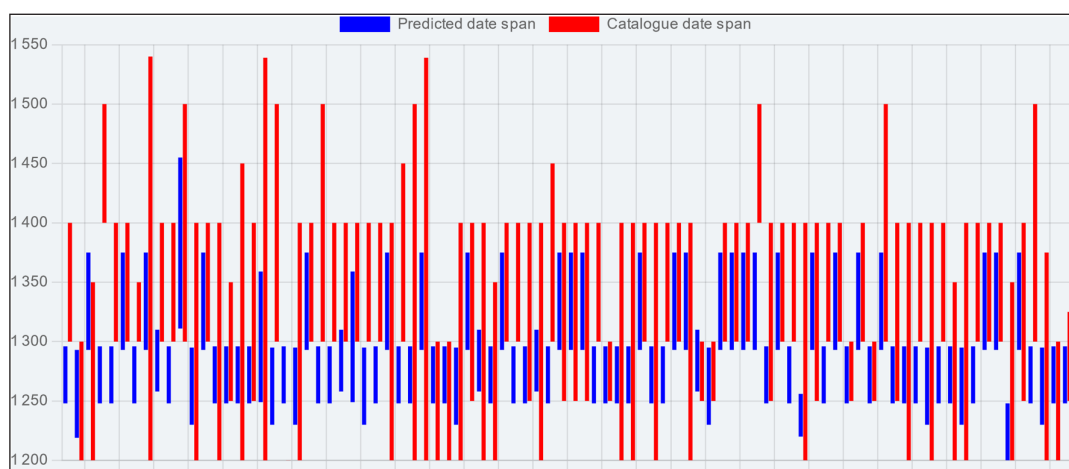
textual and graphical content, its shape and dimensions, and the form of its handle, but offers little information on dating (Geake [2016] 2020). As many different people record each type of artifact, consistency is a potential challenge; however, PAS officers do an excellent job despite the limited time and resources available. The resulting information resource is not as polished as the Schøyen Collection catalogue, but it is not intended to be. It is a dynamic online reference work that is being continuously revised, expanded, and improved.

§ 25 Like Linenthal and Noel, the PAS officers routinely assign seal matrices span dates aligned with centuries. At the time of writing, PAS identifies 48% of the seals with a single century and 21% with a period of two centuries. Since the decision tree, as already discussed, tends to use date spans of less than a century, when the decision tree is fed the information provided on the PAS website about the size, shape, and visual content of each seal matrix, it typically proposes narrower date spans. Nonetheless, the PAS and decision tree dates are generally aligned. Indeed, 50% of the PAS date spans encompass those proposed by the decision tree. This helps to explain why the average Jaccard Index of PAS compared to the decision tree is 0.41; once again the number reflects, in part, the different levels of precision. To better gauge the relative accuracy of the two sets of dates, we can again consider seals depicting the hare riding on a hound and the pelican in its piety. At the time of writing, PAS records 52 separate seal matrices with the hare-riding-on-a-hound motif. Most are dated to the fourteenth century, but one case is dated to 1200–1300, another to 1250–1350, a few to 1250–1400, still others to 1200–1400. By contrast, when the machine learning tool is applied to the same seals, it situates almost all of them in the period c. 1293–1375 (Figure 17).



**Figure 17:** Comparison of date spans proposed by the decision tree and PAS officers, for seals displaying the hare riding on a hound in the PAS dataset.

The decision tree offers more precise dates than the PAS cataloguers and is more consistent, for the PAS cataloguers are liable to assign similar seal matrices different dates. Similar tendencies are evident in the dating of seals depicting the pelican in its piety. The PAS dataset includes some 95 medieval seal matrices with this image, dated to a wide variety of different spans of time: 1200–1300 (31 cases), 1200–1400 (14 cases), 1250–1400 (eleven cases), 1250–1300 (five cases), 1300–1500 (five cases), and eighteen other spans of varying width. The 14 cases dated 1200–1400 are overly broad, since early thirteenth- and late fourteenth-century seals are generally easy to distinguish from each other based simply on size and shape (Figure 18).



**Figure 18:** Comparison of date spans proposed by the decision tree and PAS officers, for seals displaying the pelican in its piety in the PAS dataset.

The comparison of the PAS dates with those of the decision tree suggests that many PAS dates could be refined. The variety of dates that the PAS cataloguers use for similar seals, often with little explanation, also underlines the decision tree’s comparative consistency.

## 5. Conclusion

§ 26 When medieval seal matrices are plucked from the ground, little evidence survives to indicate their dates of circulation. Dating such seals is challenging, not only because medieval people used many types of seals but because their preferences changed over time. The hundreds of thousands of seal impressions preserved in the archives seal impressions provide ample—indeed overwhelming—information on the seals in circulation over many hundreds of years. However, this archival information needs to be organized and analyzed before archaeologists can use it to date seal matrices. Tools such as Digisig can gather various catalogues and information resources and make

them searchable in concert. However, simply providing archaeologists with this service is insufficient, for the analysis of that information remains laborious. Furthermore, the ongoing recording of additional seal impressions provides evidence that prompts us to revise our conclusions—and therefore, existing catalogue entries. Fortunately, the computer can help. Using machine learning, a computer can tease out the subtle and gradual changes in seal sizes, shapes, and images, and their implications for dating seals and revising catalogue entries. A human cataloguer can better discern the “style” of seals and understand the contemporary significance of their images, but human labour is costly and in short supply. A partnership between the computer and the human cataloguer offers a path forward. As this project demonstrates, we can build automatic systems to assist cataloguers in dating medieval seal matrices using training data sourced from archival seal impressions.

---

## Competing interests

The authors have no competing interests to declare.

## Contributions

### **Authorial**

Authorship in the byline is in alphabetical order after corresponding author. Author contributions, described using the NISO (National Information Standards Organization) CrediT taxonomy, are as follows:

Author names and initials:

John Alexander McEwan (JM)  
Weronika Grajdura (WG)  
Christopher D. Hopwood (CH)  
Karson Million (KM)  
Aliénor V. De Smedt (AS)

Authors are listed in descending order by significance of contribution. The corresponding author is JM.

Conceptualization: JM  
Data Curation: JM, WG, CH, KM, AS  
Investigation: JM, WG, CH, KM, AS  
Methodology: JM  
Supervision: JM  
Writing – Original Draft: JM  
Writing – Review & Editing: JM, WG, CH, KM, AS

### **Editorial**

#### **Special Collection Editors**

Martina Filosa, Universität zu Köln, Germany  
Claes Neufeind, Universität zu Köln, Germany  
Claudia Sode, Universität zu Köln, Germany

#### **Recommending Referees**

Hannah Busch, Huygens Instituut, Netherlands  
Georg Voegeler, Universität Graz, Austria

#### **Section Editor**

Morgan Pearce, The Journal Incubator, University of Lethbridge, Canada

#### **Copy and Production Editor**

Christa Avram, The Journal Incubator, University of Lethbridge, Canada

#### **Layout Editor**

A K M Iftekhar Khalid, The Journal Incubator, University of Lethbridge, Canada

---

## References

- Anderson, Michael. 2008. "Medieval Seal Matrices Found at Castle and Castle Mounds in Denmark: What Does Archaeology Tell Us about Their Use?" In *Good Impressions: Image and Authority in Medieval Seals*, edited by Noël Adams, John Cherry, and James Robinson, 71–76. London: British Museum.
- Bedos-Rezak, Brigitte. 2000. "Medieval Identity: A Sign and a Concept." *American Historical Review* 105(5). Accessed June 19, 2024. <https://doi.org/10.1086/ahr/105.5.1489>.
- Bickler, Simon H. 2021. "Machine Learning Arrives in Archaeology." *Advances in Archaeological Practice* 9(2): 186–191. Accessed June 19, 2024. <https://doi.org/10.1017/aap.2021.6>.
- Birch, Walter de Gray. 1887–1900. *Catalogue of Seals in the Department of Manuscripts in the British Museum*. 6 vols. London: Longman and Co.
- Blair, Charles Henry Hunter. 1943. "Armorial upon English Seals from the Twelfth to the Sixteenth Centuries." *Archaeologia* (89): 1–26. Accessed June 19, 2024. <https://doi.org/10.1017/S0261340900015095>.
- Cowgill, George L. 1967. "Computer Applications in Archaeology." *Computers and the Humanities* 2(1):17–23. Accessed June 19, 2024. <https://www.jstor.org/stable/30203945>.
- Geake, Helen. (2016) 2020. "Finds Recording Guides: Seal Matrices." *Portable Antiquities Scheme*. Last modified December 8, 2020. Accessed June 19, 2024. <https://finds.org.uk/counties/findsrecordingguides/seal-matrices/>.
- Gervers, Michael. 2000. "The Deeds Project and the Development of a Computerised Methodology for Dating Undated English Private Charters of the Twelfth and Thirteenth Centuries." In *Dating Undated Medieval Charters*, edited by Michael Gervers, 13–35. Woodbridge: Boydell.
- Gill, David W. J. 2010. "The Portable Antiquities Scheme and the Treasure Act: Protecting the Archaeology of England and Wales?" *Papers from the Institute of Archaeology* 20: 1–11. Accessed June 19, 2024. <https://doi.org/10.5334/pia.333>.
- Hablot, Laurent. 2019. "Le programme SIGILLA, base de données nationale des sceaux des archives françaises." In *Digitizing Medieval Sources. L'édition en ligne de documents d'archives médiévaux*, edited by Christelle Balouzat-Loubet, 129–141. Turnhout: Brepols Publishers.
- Harvey, Paul Dean Adshead. 1991. "Personal Seals in Thirteenth-Century England." In *Church and Chronicle in the Middle Ages: Essays Presented to John Taylor*, edited by Graham Anthony Loud and Ian N. Wood, 117–127. London: The Hambleton Press.
- . 1996. "Computer Catalogue of Seals in the Public Record Office, London." *Janus* (2): 29–36.
- Harvey, Paul Dean Adshead, and Andrew F. McGuinness. 1996. *A Guide to British Medieval Seals*. London: British Library and Public Record Office.
- Heslop, Thomas Alexander. 1987. "English Seals in the Thirteenth and Fourteenth Centuries." In *Age of Chivalry: Art in Plantagenet England, 1200–1400*, edited by Jonathan Alexander and Paul Binski, 114–117. London: Weidenfeld and Nicolson.
- Hourihane, Colum. 2000. "The Virtuous Pelican in Medieval Irish Art." In *Virtue and Vice: The Personifications in the Index of Christian Art*, edited by Colum Hourihane, 120–147. Princeton: Princeton University Press.



Huggett, Jeremy. 2014. "Disciplinary Issues: Challenging the Research and Practice of Computer Applications in Archaeology." In *Archaeology in the Digital Era: Papers from the 40th Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA), Southampton, 26–29 March 2012*, edited by Graeme Earl, Tim Sly, Angeliki Chrysanthi, Patricia Murrieta-Flores, Constantinos Papadopoulos, Iza Romanowska, and David Wheatley, 13–24. Amsterdam: Amsterdam University Press. Accessed June 19, 2024. <https://doi.org/10.2307/j.ctt6wp7kg>.

Jenkinson, Charles Hilary. (1954) 1968. *A Guide to Seals in the Public Record Office*. 2nd ed. London: H.M.S.O.

Linenthal, Richard A. 2009. "Ordinary Lives: Medieval Personal Seal Matrices." In *Recording Medieval Lives: Proceedings of the 2005 Harlaxton Symposium*, edited by Julia Boffey and Virginia Davis, 223–232. Donington: Shaun Tyas.

Linenthal, Richard A., and William Noel. 2004. *Medieval Seal Matrices in the Schøyen Collection*. Oslo: Hermes Publishing.

McEwan, John Alexander. 2015. "The Challenge of the Visual: Making Medieval Seals Accessible in the Digital Age." *Journal of Documentation* 71(5): 999–1028. Accessed June 19, 2024. <https://doi.org/10.1108/JD-12-2013-0163>.

———. 2018. "The Past, Present and Future of Sigillography: Towards a New Structural Standard for Seal Catalogues." *Archives and Records* 39(2): 224–243. Accessed June 19, 2024. <https://doi.org/10.1080/23257962.2017.1353412>.

———. 2019. "Does Size Matter? Seals in England and Wales, ca. 1200–1500." In *A Companion to Seals in the Middle Ages*, edited by Laura Whatley, 103–128. Leiden: Brill.

———. 2022. "New Approaches to Old Questions: Digital Technology, Sigillography, and Digisig." In *Digital Medieval Studies: Practice and Preservation*, edited by Laura K. Morreale and Sean Gilsdorf, 33–48. Leeds: Arc Humanities Press. Accessed June 19, 2024. <https://doi.org/10.1017/9781802700152>.

———. 2024a. "The Portable Antiquities Scheme." *The Digital Sigillography Resource*. Accessed July 2. <https://www.digisig.org/entity/30000047>.

———. 2024b. "Home." *The Digital Sigillography Resource*. Accessed July 2. <https://www.digisig.org>.

———. 2024c. "Seal Dating Tool." *The Digital Sigillography Resource*. Accessed July 2. <https://www.digisig.org/analyze/dates>.

———. 2024d. "Bangor University, Archives and Special Collections PENR/27." *The Digital Sigillography Resource*. Accessed July 2. <https://www.digisig.org/entity/10281872>.

McGuinness, Andrew F. 1995. "Non-Armigerous Seals and Seal-Usage in Thirteenth-Century England." In *Thirteenth Century England. Vol. 5: Proceedings of the Newcastle on Tyne Conference 1993*, edited by Peter R. Coss and Simon D. Lloyd. Woodbridge: Boydell.

Muehlberger, Guenter, Louise Seaward, Melissa Terras, Sofia Ares Oliveira, Vicente Bosch, Maximilian Bryan, Sebastian Colutto, et al. 2019. "Transforming Scholarship in the Archives through Handwritten Text Recognition." *Journal of Documentation* 75(5): 954–976. Accessed June 19, 2024. <https://doi.org/10.1108/JD-07-2018-0114>.

National Archives. 2024a. "Reference DL 25/2081/1766." Accessed July 2. <https://discovery.nationalarchives.gov.uk/details/r/C16101271>.

———. 2024b. "Reference DL 25/3533/3091." Accessed July 2. <https://discovery.nationalarchives.gov.uk/details/r/C16102066>.

New, Elizabeth Anne. 2010. *Seals and Sealing Practices*. London: British Records Association.

———. 2016. "Seals as Expressions of Identity." In *Seals and Society: Medieval Wales, the Welsh Marches and Their English Border Region*, edited by Phillipp R. Schofield, Elizabeth Anne New, Sue M. Johns, and John Alexander McEwan, 105–120. Cardiff: University of Wales Press.

———. 2019. "Reconsidering the Silent Majority: Non-Heraldic Personal Seals in Medieval Britain." In *A Companion to Seals in the Middle Ages*, edited by Laura Whatley, 279–309. Leiden: Brill.

PAS (Portable Antiquities Scheme). 2024. "Home." Accessed July 2. <https://finds.org.uk/>.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. 2011. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12, 2825–2830. Accessed June 19, 2024. <https://www.jmlr.org/papers/v12/pedregosa11a.html>.

Pett, Dan. 2010. "The Portable Antiquities Scheme's Database: Its Development for Research since 1998." In *A Decade of Discovery: Proceedings of the Portable Antiquities Scheme Conference 2007*, edited by Sally Worrell, Geoff Egan, John Naylor, Kevin Leahy, and Michael Lewis, 1–18. Oxford: Archaeopress.

Robbins, Katherine. 2014. *Portable Antiquities Scheme: A Guide for Researchers*. London: The Portable Antiquities Scheme.

