IDEAH • Vol. 2, Iss. 1 (DHSI 2019 & 2020)

Sumerian Networks: Classifying Text Groups in the Drehem Archives

Anya Kulikov¹, Adam Anderson¹, Niek Veldhuis¹

¹UC Berkeley

Published on: Jul 28, 2021

DOI: 10.21428/f1f23564.6da394df

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

Project Goals and Research Question

The goal of the *Sumerian Network* project has been to build reproducible socio-economic networks from the Ur III archives, and to further refine these models to more accurately reflect the actors and entities active in these unprovenanced archives over the 80-year period in the 21st century BCE. Beginning with ca. 15,000 transliterated texts from the site of Drehem (known in antiquity as Puzriš-Dagān), which are curated online in three databases (the *Open Richly Annotated Cuneiform Corpus* (ORACC), the *Database of Neo-Sumerian Texts* (BDTNS), and the *Cuneiform Digital Library Initiative* (CDLI)), we applied various classification methods in order to delineate sub-archival data sets, known as text groups in the scholarly literature. In order to make our study reproducible, we used iPython Jupyter Notebooks (Pérez and Granger 2007) to describe the tools and methods we use in connection with the code and dataset, as well as a tool for generating a series of empirical network models.

The primary question which we pursue in this article is how one can use reproducible and replicable workflows for discovering the optimal classifications of the text groups from an unprovenanced archival context. We describe how we leverage existing scholarship to help validate our findings, both in terms of published work as well as through workshops with hands-on tutorials.

Our results show that the key factor for success lies in building reproducible and replicable workflows. This allows for the combination of classification methods with scholarly input. For the most recent results of these reproducible network models, see our Jupyter Book, [Sumerian Networks](#).

Background

In April of 2017, as our Data Science Discovery project was just beginning, we organized a series of multidisciplinary conferences at UC Berkeley, with the most recent having taken place in [October of 2019](#). These workshops brought together an international group of scholars, including: historians, archaeologists, Assyriologists, computer scientists, and data scientists, all with the purpose of discovering new computational methods for quantifying the immense cuneiform archives of the Ur III (also known as the Neo-Sumerian) Period, ca. 2100-2000 BCE. These workshops and the Sumerian Network research project have been successful thanks to the combined support of the Data Science Discovery Program, the D-Lab, and Digital Humanities at UC Berkeley, the latter in the

form of a collaborative research project grant made possible by the Andrew W. Mellon Foundation.

The initial workshop provided examples of how to use quantitative methods to generate network graphs representative of ancient textual archives, which in turn helped to articulate certain workflows for constructing network graphs (see [Sumerian Networks Discovery](#); [Escobar](#); [Pearce](#)). By bringing together the researchers who worked on the primary databases of digitized texts from the Neo Sumerian Period, we discovered new pathways for harmonizing these data sets and leveraging the added metadata, which each uniquely possessed to build a more robust model.

(ManuelMolina, [BDTNS](#); Niek Veldhuis, [ORACC](#)). Subsequent workshops brought the scholars back together after further stages of the project were completed, including the first iterations of a socio-economic network representative of a subset of the Neo-Sumerian texts, with ca. 15,000 texts from the archives from Drehem (ancient Puzriš-Dagān).

Figure 1a: Two Views of the Text Groups in the Drehem Archives

Figure 1b: Two Views of the Text Groups in the Drehem Archives

The Figure 1a is a network in which the nodes represent a text with links corresponding only to tablets which were given a text group designation (nodes = 14,603; edges = 14,594). Figure 1b is a network of named entities (i.e., names, places, institutions) which have links to those who are co-mentioned in a given text (nodes = 43,025; edges = 61,098). Colors indicate text groups: domesticated animals (light blue: 67.8%), dead animals (brown: 15.4%), unknown (violet 6.7%), queen' s archive (green 6.2%), wild animals (red:

2.2%), precious objects (olive: 1.2%), leather objects (royal blue 0.3%), wool (dark green: 0.3%). Figure 1b shows the groups of actors with the highest betweenness centrality as bridges between the eight text groups identified in our current classification.

The Ur III Period: History and Texts

In ancient Mesopotamian history the term Ur III refers to the 21st century BCE, when a dynasty based at the city of Ur (in the deep south of today's Iraq) managed to bring the traditional Babylonian city states under its rule, unite the entire area, and dominate large areas to the north and east from which tribute and booty were extracted. The period is famous among cuneiformists because of the very large number of texts produced in a relatively brief period of time. Current estimates hold that there are at least 100,000 documents that deal predominantly with the income and expenses of the state and its provinces (the old city states). For this reason, the Ur III state has sometimes been described as bureaucratic and obsessed with administrative detail (it is not uncommon to find a document that registers the death of a single lamb - on a total of tens of thousands of animals). There are good reasons, however, to reject the label “bureaucratic” in the Weberian sense. The state was run as a family business, with the king personally owning the property of the state and family members put in charge of key positions, which fits better in a “patrimonial” model (Michalowski 1987, 55). Be that as it may, the administrative offices of the time produced more documents than in any other period of Mesopotamian history.

The study of these documents is hampered not only by their sheer number, but also by their modern history. Countless texts were looted in the late nineteenth and early twentieth century and sold in large or small batches in Europe, the United States, and other countries (see Molina 2008, 20-42; Walther Sallaberger and Aage Westenholz 200-203). After the Gulf War in 1991 and the American invasion of Iraq in 2003 new waves of looting hit the archaeological sites of the area, resulting in a true tsunami of artifacts, including large numbers of tablets from the Ur III period. Without archaeological context it is impossible to know how documents were archived and how they were grouped. We have to rely on internal criteria - textual and paratextual - to assign documents to royal, provincial, or private archives and classify these texts in meaningful groups. The random distribution of the texts over museums and collections all over the world only complicates this problem.

Starting in the late 1990s, several projects started to collect transliterations, metadata, and images (line drawings and photographs) of Ur III documents, enabling scholars to consult and compare documents on their screen, even if they were many thousands of

miles apart physically. Today, there are three such projects that essentially cover the same set of Ur III data, each with its own strengths and weaknesses.

	<i>BDTNS</i>	<i>CDLI</i>	<i>ORACC</i>
coverage	Ur III	All of cuneiform	project-based
lemmatization	no	no	yes
images	yes	yes	no

Table 1: Projects which cover Ur III data.

ORACC is an umbrella project, allowing project directors to define a corpus of cuneiform texts. Its Ur III corpus of some 72,000 documents is part of the electronic Pennsylvania Sumerian Dictionary project; the data are imported from CDLI and have been cleaned and lemmatized by Niek Veldhuis (UC Berkeley) and Steve Tinney (University of Pennsylvania). CDLI started in the late nineteen nineties with editions and images of third millennium administrative texts (among which the Ur III texts are by far the largest group). Over time, it broadened its scope as a clearing house for metadata and images of cuneiform texts from all periods. BDTNS concentrates on Ur III administrative texts only. The smaller scope of the BDTNS project allows its director, Manuel Molina, to pay much more attention to cleaning the metadata (provenance, date, secondary literature).

Based on these considerations, we decided to use the strengths of ORACC (lemmatization) as well as the strengths of BDTNS (up-to-date metadata) and mash the two data sets accordingly.

Figure 2a: <http://oracc.museum.upenn.edu/epsd2/admin/ur3/P135928>

- HMA 9-02611¹**
- obverse
1. 1 udu
n: udu[sheep]N
 2. ki šu-i₁-li₂-ta
ki[place]N; Šuili[0]PN
 3. ur-mes [i₁-dab_s]
Urmes[0]PN; dab[seize]V/t
 4. giri₁ x-[...]
niri[via]N; u
- Reverse
1. u₄ 3-kam
ud[day]N; n
 2. Iti ezem-{d}nin-a-zu
itud[month]N; Ezemninazu[1]MN
 3. mu us₂-sa ma₂-dara_s
mu[year]N; us[follow]V/t;
- Madara[0]ON
- “1 sheep, from Šuili, Urmes took. Via [...].
- Day 3, month Ezemninazu, year after the
- year when the Madara (boat was caulked).”

Figure 2b:
<http://oracc.museum.upenn.edu/epsd2/admin/ur3/P135928>

Ur III document from Puzriš-Dagān (modern Drehem) in line drawing with transliteration and interlinear lemmatization. The document is dated to ‘day’ , ‘month’ , and

‘year’ . Year names can be mapped to regnal years, in the present case the third year of Šu-Suen, the fourth king of the Ur III dynasty (HMA 9-02611; Hearst Museum of Anthropology, University of California at Berkeley).

Metadata: Provenance and Date

The great mass of Ur III documents come from three large archives: the provincial archives of Umma and Girsu, and the royal archive of Puzriš-Dagān. Since the great majority of documents do not derive from regular excavations, the provenance is most commonly based on internal data such as month names, proper nouns, commodities, and spelling habits (for the criteria, see Sallaberger 1999, 207). For most documents this yields a secure attribution, but there is a category of texts for which the classification is unclear or in doubt. More difficult is the categorization of documents in smaller files or text groups. The documents from Puzriš-Dagān deal overwhelmingly with domestic animals (sheep, goats, and bovids), but smaller groups of texts may be attributed to the dead animals file, the queen’s file, the shoe-and-treasure file, or the wool file. A separate text group deals with offerings of a single lamb or sheep to the main gods of the state by dignitaries or military officers from peripheral areas of the state. For example, the group deals with Mari, where a number of the queens of Ur were known to have come from (see T.M. Sharlach 2017, 63, 92). Recognizing and classifying such text groups in a reliable way is a prerequisite for analyzing the data and for understanding the administrative, political, financial, and diplomatic realities behind them. Each of these text groups is characterized by a specific set of functionaries who appear frequently. Dead animals, for instance, are delivered to a man called Šulgi-irimu in more than a thousand Puzriš-Dagān documents. Identifying such groups makes it possible to analyze the social networks of these people in more detail. On the other hand, the various bureaus or “offices” at Puzriš-Dagan did not work in isolation, nor were they disconnected from the other royal and provincial centres of the period. A grass-roots approach is to start reconstructing smaller networks and build up to larger and more complex systems. The smaller text groups also help in disambiguating namesakes.

These texts are unique in that almost every tablet is dated to a particular day, month, and year in which the transaction took place. This means that by reconstructing these series of trades in a statistical model, our analysis can reflect a dynamic network, with directed links and time series data. Further, when the named entities are aggregated, the model is much more than just a list of people recorded in total, because along with this list come their hierarchical relationships, family ties, roles, and responsibilities. The resulting model is able to accommodate an ever-increasing amount of additional data, as the

databases continue to be populated with new texts digitized on a monthly basis. By implementing a reproducible workflow, we can pose an infinite number of questions for this complex society by analyzing a granular empirical model.

Reproducible & Replicable Workflow

If quantitative tools and text analysis prove to be effective in the scientific discovery of new archaeological sites, it will be because the resulting models were reproducible and replicable, and that a community of specialists in Assyriology and archaeology engaged collaboratively with the process of standardizing the series of tools and methods.

Individual projects attempting to quantify the entities of the Ur III period have demonstrated novel implementation of ML algorithms, but too often generate data with high error rates (e.g. false positives, see Yudong Liu et al. 2015, 1450) and exhibit a general lack of consultation of the scholarly literature written in German, which provides the systematic markers for the named entities, including their professions and titles, along with the typology of transactions used in the Sumerian language (e.g. Sallaberger 1999, 214; 2003, 52; Paola Paoletti 2012, 37-41; Liu 2017, 13). By beginning with the guidance of Ur III specialists, both in print and in person, these discussions helped to establish the proper workflow for building reproducible network models. The above-mentioned markers written in the Sumerian texts (e.g., roles and activities performed by named entities) could then be used to identify the relationships between people and place names mentioned in the texts, many of which are still unknown to scholars in the field.

After close inspection of the cutting-edge work in named-entity recognition (NER) (see Luo et al. 2015; Liu et al. 2016), we determined that supervised methods informed with specialist input provide a superior result to unsupervised approaches, or semi-supervised NER, and that a hand-made name authority catalog would be necessary to properly deduplicate the named entities in the ancient archives of the Ur III Period. This objective was the first prerequisite for our project, and by September 2017 the name authority was completed for the ca. 15,000 Drehem texts by John Carnahan. An initial script harvested all proper nouns from the corpus and tried to produce normalized forms according to rules of Sumerian grammar and orthography. This left a good number of name forms that needed hand editing. The resulting name authority contains more than 6,000 name forms, corresponding to over 3,300 normalized names and representing approximately 60,000 name instances. The same name may be represented by multiple name forms because of spelling variations, nominal morphology marked on the name (such as ablative or genitive), or both. Importantly, the names do not directly map to individuals

because at this stage namesakes have not been disambiguated in any way. Using text analysis and machine learning tools, we then generate a network graph, in order to statistically model the socio-economic system of trade as documented in the Drehem archives.

The purpose of such a network would be to have a quantifiable and multi-dimensional model with which we could pose a series of questions and reproduce verifiable results. Due to the fact that the Drehem archives were found in a looted context, it is impossible to reconstruct the archive as it would have been found in situ. However, due to the presence of certain text types, we hypothesize, along with most scholars in the field, that the archives were sorted into multiple sub-groupings in order to organize these volumes of administrative texts. This hypothesis implies the contours of an extensive administrative structure behind the clearly documented accounting practices. If this hypothesis proves to be correct, then the question remains, how was the Ur III state capable of keeping in sync the records of trade across the many city-states? An even more basic question central to our study has to do with the scale of this system, that is, how many people, places, and institutions were documented in these records? Because the Drehem texts are representative of the royal administrative archives, we begin here in our quantitative analysis to better understand the scale and reach of the Ur III state, and to learn the main actors of the network of officials and royal family members at the center of the administration during this period.

While specialists in the field of Ur III studies have been working toward answering such questions (see Sallaberger 1999, 207; Steven Garfinkle 2015, 149; Sharlach 2017, 195), without a clearly defined reproducible method, such scholarship will remain isolated, and based on a system of trust, rather than testing. The fact is, that there will never be a ‘gold standard’ or ground truth to these questions, especially as the complete body of texts still remains to be discovered. For this reason, we state that the reproducibility of our methods is our highest priority, and we describe a detailed method for applying the same tools and methods, so that any future scholar can engage with the same workflow and replicate new advancements on the computational foundations which we build.

Archival Classification Methods

Although the Ur III tablets have been studied for almost a century now, this corpus remains difficult to work with, due to the wide distribution of texts. In order to make a systematic study, it is necessary to work with the primary databases of Ur III texts, which have collected the texts for scholarship, whether they are stored in public museums or private collections hidden all over the world (Molina 2008, 21-42).

Of course, it is typical for ancient archives to include texts with dubious provenience; however, the Neo Sumerian archives are especially difficult to organize into archives. This is due to two types of looting which has impacted the Ur III textual record. First, the more general looting in the region before the site could be officially excavated in 2007 is evidenced from at least the late 19th century (see Molina 2008, 19). This long-standing looting of the Sumerian sites was compounded by a second looting of Iraqi sites, which occurred after the 2003 Gulf war, and included a more systematic looting of the Iraq Museum (see Neil Brodie 43). One of the worst cases of looting at an archaeological site was the ancient city of Umma (modern Umm al-Aqarib) in southern Iraq. Images of Umma show a site pockmarked with looters holes. Although the majority of the Drehem archives are without proper archaeological context, an excavation from 2007, conducted by the State Board of Antiquities and Heritage of Iraq, determined that the provenance of the ca. 15,000 tablets is accurate, and that the present archive was the result of an administrative center of the Ur III dynasty (Al-Mutawalli and Sallaberger 2017, 158ff.).

Ultimately, our methods are used to address an accounting issue, in that we have so many texts that we cannot really say where they come from, with at least 10-15% of the corpus dubious at best. Therefore, archival designations of the tablets from Puzriš-Dagān in the Ur III Period, to the degree which they have been reconstituted by specialists in the field of Assyriology, are the result of scholarly deduction. We have chosen to leverage the content of each tablet in order to determine the contours of the archive by clustering, using a series of classification methods.

Binary Classification

- Basic description: As a binary, this classifier can select one of two options as an archival category. While this is a simple process, it fits with the historical reality in which each text belonged to a distinct archive. The classifier decides based on whether a text has a single designation or multiple, continuing in a loop until all texts are given one class.
- How we use it: We apply this classification method in order to verify the archival designations applied to the Drehem texts in the databases. While the final counts have differed only slightly with those from the database, the few texts which were kicked out seemed to reflect the textual characteristics of archives outside of the Drehem corpus.

The first categorization algorithm works as a binary, which means that it cannot consider multiple archival categorizations. This approach works well for these historical archives, since a given text could only have existed in one archive at a time. Thus, we treat each text as belonging to a distinct archive, and let the classifier decide if a text has a single designation or multiple designation. One example was found in text [P125630](#) (PDT 1,

214b), which was labeled as belonging to the queen's archive in the scholarly literature (Picchioni 1975, 157). Because our classifier used a simple dictionary method, we could see that the text only concerns small domestic animals (sheep, goats, etc.), and should therefore belong to the domesticated animals text group, rather than the queen's archive.

In this process we found subsets of the 15,000 texts that belong to different archives. For example, the text [P135153](#) (TRU 389) is a very atypical text, and has been assigned to Drehem in the database, although it may not have been from Drehem at all. For this reason, we created an 'unknown' text group, so as not to filter out texts which we are currently uncertain if they belong. The results of this work helped to show which texts were given the Drehem label in the database, but which may not have come from Drehem at all.

Lastly, this classification method also helped us to identify certain texts labeled as 'unknown' in the database. After discovering four entries, which initially had an 'unknown' label, we could determine that these were in fact seals with inscriptions. For example, [P430140](#) is a composite seal impression, but not a cuneiform tablet. We then extended the training to all the texts with the designation 'seal', in order to discover and classify all similar seals, which came to 309 texts.

After the general parameters of the Drehem archive were determined, the next task was to further delineate the sub-archives, dossiers, and text groups. The methods which we pursued began with bayesian models and evolved into non-probabilistic supervised learning models with cross validation. In simpler terms, we used words and other features found in the texts to help calculate probabilities of texts belonging to a certain archive. In addition to that, we used some texts that we knew were labeled correctly and used supervised models with them. We then used cross validation, a technique used to make sure data is not overfitted.

Text Groups (Textgruppe) within Archives

It is important to note that the majority of the scholarly designations for text groups of the tablets, and to some degree the larger archival designation of the tablets in the databases, were determined using a few basic heuristic observations, in a similar manner to the one that Markus Hilgert (2003) describes:

- Sorting the texts by their dates (i.e., month, and year names) with reference to a reconstructed state calendar, used in the Neo Sumerian Period (Hilgert 13);
- Noting the combination of the materials (i.e., precious metals & stones) and persons (i.e., precious objects made for the king and elites) (Hilgert 14);

- Finding out what we know of the time and place when the tablets were purchased from the markets (Hilgert 2).

In a recent publication about the administrative practices of the Ur III Period, Liu gives a long list of bureaus with a chronological frame, along with animals and commodities for each (2017, 407-408, 412). While he goes to great lengths to identify many text groups, he unfortunately confuses the difference between archives and bureaus and other formal designations. In order to ensure that we make the proper distinction, we focused our initial assessment on the so-called queen' s archive (or Šulgi-simtī archive), a text group which has received some scholarly attention, and therefore has some texts identified, but which is by no means fully accounted for (see Sallaberger 1999, 377-390; Sharlach 2017, 189-210; Junna Wang and Yuhong Wu, 51-61; Weiershäuser 2008, 94-96.).

While we know that the queen' s name is Šulgi-simtī, her archive is difficult to find, because there is not a very clear list of keywords found across her documents. She primarily deals with animals and precious objects, which are also found in other text groups (e.g., domesticated animals and treasure texts). Therefore, in order to establish a baseline for this text group we began with the scholarly literature (e.g., Sallaberger 1999, 377-390, Wang and Wu 2011, 51-61) to classify the texts belonging to the queen' s archive (ca. 250 tablets), which we could then use for training. As the graph in Figure 3 indicates, the classification methods we describe below were able to allocate an additional ca. 750 texts to the queen' s archive.

Classification Methods

Figure 3: Distribution of text groups across the Drehem archives.

Figure 3 provides the total counts for the text groups resulting from the series of classification methods. While these results are put forth without a gold standard form of validation, they have been scrutinized to some degree by the leading scholars in the field and are provided in our open source Jupyter Book, [Sumerian Networks](#).

As we do not have a gold standard to use to determine whether a lemmatized text falls under a particular text group, we must make our best guess and classify archives using unsupervised learning methods. Unlike supervised learning, there is no labeled data to compare our results with, i.e., no gold standard to determine accuracy. Therefore, we extract the archival information by using key terms from lemmatized texts, such as “sheep” or “goat” or “bull” for “domesticated animal” (see Table 2).

Label	List of Words
Domesticated animal	"[ox]", "[cow]", "[sheep]", "[goat]", "[lamb]", "[~sheep]", "[equid]" # account for plural.

Wild animal	"[bear]", "[gazelle]", "[mountain]" # account for "mountain animal" and plural.
Dead animal	"[die]" # find "die" before finding domesticated or wild.
Leather object	"[boots]", "[sandals]"
Precious object	"[copper]", "[bronze]", "[silver]", "[gold]"
Wool	"[wool]", "[~wool]"
Queen' s archive	"[]" # this is what we are training/testing to find out.

Table 2: The labels we initially trained for classification included (1) domesticated animals, (2) wild animals (3) dead animals, (4) leather objects, (5) precious objects, (6) wool, (7) queen' s archive.

If the words of the text match, we label it with the archival attribute and repeat this process for every text. We have created six possible labels for archives, which include “domesticated animal” , “wild animal” , “leather object” , “precious object” , “wool” , and “queen' s archive” .

Naive Bayes Classification

- Basic description: The Naive Bayes Classifier uses a conditional probability model and assumes that all the features are independent from one another, given the class variable. The model predicts the probability of a lemmatized text belonging in an archive class based on its features (Rish).
- How we use it: With variable feature length alternating between unigram, bigram, and trigram models. The results provide three measures to determine the fit of the model and for optimization: 1) **Accuracy**: (true positives + true negatives) / total; 2) **Recall**: true positives / (true positives + false positives), where high recall means that an algorithm returned most of the relevant results; and 3) **Precision**: true positives / (true positives + false negatives), where high precision means that an algorithm returned substantially more relevant results than irrelevant ones.

In order to test how well the unsupervised ML classifiers were able to detect the different dossiers within the Drehem archives, we took a subset of the queen' s archive, which belonged to queen Šulgi-simtī. While it is still uncertain exactly how many texts belong to this sub-archive, Veldhuis' s scholarly expertise suggested there were probably

upwards of 800-1000 texts in the queen' s archive in total. After going through the list of texts in the scholarly literature (Sallaberger 1999, 377-390), we had an initial training set of ca. 250 text. While we knew there were more, it was very labor intensive to find those through close readings. Because we only had a small number of texts for training, the problem arose that the Bayesian classifier ran through the entire corpus of 15,000, but it only identified a few queen' s texts with each run (ca. 1-2%). This meant that the Bayesian classifier was incentivized to have a low number of hits, making the accuracy extremely high. We decided to take the whole corpus into consideration, and check the results by hand, with 525 texts for training and 299 texts for testing. We then labeled 75% of the transactions, just by hand-coding.

In order to build the Bayesian model, we first included 220 texts which were known to belong to the queen' s archive, and we then included 200 other transactions. The test set was 300 transactions, to predict whether these texts would fall into the queen' s archive or not. Our initial pass was rather low, with 0.76 accuracy, 0.75 recall, and 0.82 precision. Following this result, we decided that looking at the data with n-grams would provide further optimization. Veldhuis suggested trying unigrams, bi-grams and tri-grams, but no more than that. The Naive Bayes classifier found many more texts assigned to the queen' s archive (75%), and improved accuracy to 92%. After trying out the bi-gram, tri-gram, and unigram+bi-gram models, the precision rose to 95%. While these scores were impressive, our main concern was that 256 (queen' s) + 200 (non-queen) texts was still not enough data, especially considering that each text consisted of ca. 50 words on average. In other words, the sample was as conservative as our model turned out to be, and ideally, we would have had ca. 1,000 queen' s texts for training.

Support Vector Machine (SVM)

- Basic description: Support Vector Machine works by classifying a text without using any priors, which was the limitation of Naive Bayes. SVM counts the best line between classes (Nobles 2006).
- How we use it: we wanted to compare SVM with the Naive Bayes classification, and we found it to have higher accuracy in general, and to be especially useful in reducing the number of texts assigned to an unknown class.

When we initially ran the SVM model, it suggested that the queen' s archive compromised about 7% of the total texts. This helped us find additional smaller archives within the Drehem texts as well. Compared to old stats, accuracy rose from 90% to 93.3%, recall up to 93.8%, precision up to 90.2%. This means that the size of the predicted queen' s archive dropped from 7.5% to 6.5%.

K-Fold Cross-Validation

- Basic description: K-Fold Cross-Validation is used to examine the efficiency of a model by seeing how it performs on new data. The data sample itself is limited. K is the number of ways in which the data is split. K must be chosen appropriately so that we don't get too high of variance or bias (see Davide Anguita et al. 441).
- How we use it: for parameterization, we used 10-fold classification with all the data as training, which resulted in 95% accuracy. We then filtered everything out and only added based on the label (e.g., wool), with domesticated animals being the largest category.

The keywords used for the five main text groups were as follows (see Figure 3):

- “Precious objects” uses terms like ‘gold’ , and ‘silver’ , etc.
- “Leather objects” has words like ‘shoe’ , ‘sandal’ , and ‘hide’ , etc.
- “Domesticated animals” includes the words ‘bull’ , ‘sheep’ , and ‘goats.’
- “Dead animals” may have the same words but must include the word ‘dead’ to qualify.
- “Wild animals” has more exotic animals, such as ‘bear’ , ‘gazelle’ , and ‘camel’ etc.

While it is uncertain whether there were any livestock at Drehem (Christina Tsouparopoulou iii), it is clear that the royal archives at Drehem made an exact accounting of all the animal exchanges (live or dead) all over the Ur III state (see Molina 2016). For taxation, all the records were concentrated in Umma, but the records for the precious objects for the Crown were kept in Drehem. Livestock is counted in at least three different text groups: domesticated animals, wild animals, and dead animals. The former group is by far the largest and could perhaps be broken down even further by the size of the animals (i.e., large bovine, and small ovine, as in Sallaberger 1999, 197), however as we will describe below, our classification methods did not detect this distinction from the lists of domesticated animals on the tablets.

Classifier	Accuracy	Precision	Recall	Classes
Multinomial Naive Bayes (Unigram)	0.899	0.861	0.923	6
Naive Bayes (Unigram + Bigram)	0.916	0.885	0.935	9

Multinomial Naive Bayes (Trigram)	0.892	0.854	0.915	10
SVM	0.933	0.902	0.938	8

Table 3: Summary of results based on method.

Running the Naive Bayes classifiers by counting single words as tokens, generally improved the accuracy (except for the trigram model). However, the best result of this type came from a combination of the Unigram + Bigram model, which increased the accuracy from 90% accuracy to 91.6%; recall went up from 92% to 93.5%, and precision rose from 86.5% to 88%. Our final clustering runs with modifications of the Naive Bayes models suggested 9 clusters instead of the previous 10. Estimates of the percentage of queen' s texts dropped from 5.8% to 5.5% for Naive Bayes.

While the combination of Unigrams and Bigrams seemed to optimize the Naive Bayes model, ultimately, we reached our highest precision scores using the non-probabilistic binary linear classifier, SVM. This method was especially effective in combination with k-fold cross-validation (where $k = 10$), which allowed us to map the inputs into high-dimensional feature spaces. The result was actually a smaller number of classes, but this also seemed to fit nicely with the scholarly literature (see Sallaberger 1999, 240).

Phase I Conclusions

Although the Sumerian Networks Project is still ongoing, after four years of research it has demonstrated how a careful combination of historical sources with quantitative analysis can produce reproducible multi-dimensional socio-economic models for a century of data from the Neo Sumerian archives (2100-2004 BCE). Working in harmony with the online databases of Ur III texts ([BDTNS](#), [ORACC](#) & [CDLI](#)), the resulting network graphs are capable of juxtaposing the intersubjective individualism of each named entity within their social, institutional, geographical, and temporal dynamics, by visualizing these entities in a relational framework which reveals both the structural positionality of each individual along with their cultural elements as identified textually (Mische 90). Such networks of empirically attested relations may then be measured statistically, leveraging mathematical formalization along with temporal and geospatial quantification, to the individuals and locations of otherwise unknown names in the historical archives of the Ur III texts. The combination of these elements, with sufficient textual documentation, has given new insights into the events and latent organizational structures resulting from the sum of the individual actors, institutions, and geographic names, as they are counted in

the network model, to interpret “the social within the cultural” (John W. Mohr and Harrison C. White 486).

After applying a series of classification methods, we were able to identify the most salient text groups within the Drehem archives. However, without having a gold standard to know when the optimal number classes are reached, these results will require ongoing fine-tuning and, more importantly, scholarly supervision through close readings. Fortunately, the tools and methods we’ve applied to this task are open source and available for replication and iteration, as the same methods may be applied to the Ur III corpus as a whole.

Therefore, we propose that the open-source tools and methods we have built and described here can be used by the scholarly community to test the current designations of all archives in the Neo Sumerian Period within their respective databases. Further, based on our findings we suggest that the text groups which our model has identified may, in fact, reflect the ontological structure within the Ur III bureaus or institutions responsible for making accurate accounts of the commodities and livestock exchanged on a daily basis. These detailed accounts form an ancient database which the ancient academics used for economic analysis and forecasting (see Piotr Steinkeller 80). This means that in visualizing the network of entities designated by professional titles and comparing these titles with the archival designations we currently know for the documents at Puzriš-Dagān, the emerging result is the extant socio-hierarchical structure inherent in the administration of the Ur III bureau system.

Phase II Next Steps

While the archival designation of the Ur III archives is an important first step, one of the underlying questions which we have pursued from the beginning has to do with the Ur III prosopography. As our network models demonstrate, we have been successful in parsing all the named entities in the Drehem texts, using a hand-coded name authority. However, without a detailed method for name disambiguation, the number of homonyms makes the prosopography somewhat dubious, and underemphasizes the importance of certain individuals whose names are attested most frequently. In order to facilitate the process of disambiguation, we have identified a number of textual features, which we extract systematically for the parameterization of our models.

The classification algorithms we have used so far indicate that a quantitative approach to the boundaries of specific text groups can be determined. However, without scholarly supervision the accuracy of these assessments for the cutoff value is imprecise (i.e.,

metric to choose to make something a cluster). The default settings described in this study are a start, but further work is required to make a more comprehensive classification of the Ur III texts as a whole. Therefore, the subsequent phases of the project include the identification of persistent features which we can derive from the textual records for each named entity (see Figure 4). This process is also performed using Jupyter Notebooks, in order to ensure the accuracy and reproducibility of the methods we use for disambiguation.

Figure 4: Drehem Textual Features

The above graphical model delineates the current feature space for the Sumerian social network derived from the Ur III Period, Sumerian Administrative archives from Drehem (c. 15,000 Puzriš-Dagān, 2100-2004 BCE). Features include the following attributes:

A: Archive (see archival classification methods above)

C: Commodity (i.e., the goods which are traded)

D: Date (i.e., a date formula written on each tablet, which date most transactions to the year, month, and day)

F: Family (i.e., patronymics are seldom listed, but this may be a useful feature for disambiguation)

GN: Geographic Name (i.e., when a GN is mentioned in relation to a person or transaction)

N-N: n-Neighbors (i.e., the text written before and after each PN, which we intend to use for collocating PNs and GNs)

NA: Name Authority (i.e., a catalog of names to texts, which will allow us to normalize spellings)

PC: Processed Commodity (i.e., the numerical count and measurement of each commodity listed along with any attributes, e.g., 6 goats for milking, in which we count the 6 goats and include the milk as attributes of these commodities)

PD: Processed Date (i.e., a conversion of the written formula (string) to an integer value)

PG: Processed GN (i.e., geospatial coordinates [Lat, Lon] for each known GN)

PN: Proper Name

Prof: Profession

Pt: Patronym (~30 years boundary for PD)

R: Role

Works Cited

Al-Mutawalli, Nawala and Sallaberger, Walther. “The Cuneiform Documents from the Iraqi Excavation at Drehem.” *Zeitschrift für Assyriologie* 107 (2017), 151-217.

Anguita, Davide, et al. “The ‘K’ in K-fold Cross Validation.” Proceedings, *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 2012, Pp. 441-446. Bruges, Belgium.

www.esann.org/sites/default/files/proceedings/legacy/es2012-62.pdf.

Brodie, Neil. “The Market Background to the April 2003 Plunder of the Iraq National Museum.” In *the Destruction of Cultural Heritage in Iraq*. Edited by Joanne Farchakh Bajjalay and Peter G. Stone. Boydell Press. 2008, ISBN 9781843833840.

Garfinkle, Steven J. “Ur III Administrative Texts: Building Blocks of State Community.” Chapter 6 in *Texts and Contexts: The Circulation and Transmission of Cuneiform Texts in Social Space*, edited by Paul Delnero and Jacob Lauinger. De Gruyter, 2015, pp. 143–65. doi.org/10.1515/9781614515371-006.

Gomi, Tohru. “Shulgi-simti and her Libation Place (ki-a-nag).” *Orient*, vol. 12, 1976, pp. 1–14. doi.org/10.5356/orient1960.12.1.

Hilgert, Markus. *Drehem Administrative Documents from the Reign of Amar-Suena*. Oriental Institute Publications (OIP). Cuneiform Texts from the Ur III Period in the Oriental Institute, Volume 2:121, 2003. oi.uchicago.edu/research/publications/oip/cuneiform-texts-ur-iii-period-oriental-institute-volume-2-drehem.

Kang, Shin T. “The Role of Women at Drehem” , in *Neo-Sumerian Account Texts from Drehem*. Ed. C.E. Keiser. BIN 3, 2-11. Yale, 1971, New Haven. babylonian-collection.yale.edu/sites/default/files/files/BIN%203-Keiser%2C%20Clarence%20E_%20-1971-%20Neo-Sumerian%20Account%20Texts%20from%20Drehem.pdf

Liu, Changyu. *Organization, Administrative Practices and Written Documentation in Mesopotamia during the Ur III Period (ca. 2112-2004 BC): A Case Study of Puzriš-Dagan in the Reign of Amar-Suen*. Ugarit Verlag, 2017, Münster. www.academia.edu/32943644/Organization_Administrative_Practices_and_Written_Documentation_in_Mesopotamia_during_the_Ur_III_Period_c_2112_2004_BC

Liu, Yudong, Hearne, Hames, and Conrad, Bryan. “Recognizing Proper Names in Ur III Texts through Supervised Learning” . Proceedings, *The Twenty-Ninth International*

Florida Artificial Intelligence Research Society Conference. Association for the Advancement of Artificial Intelligence (www.aaai.org) , 2006. www.aaai.org/ocs/index.php/FLAIRS/FLAIRS16/paper/viewFile/12976/12626.

Luo, Liang, Liu, Yudong, Hearne, James, and Burkhart, Clinton. “Unsupervised Sumerian Personal Name Recognition” . *Florida Artificial Intelligence Research Society Conference*, 2015, Florida. www.aaai.org/ocs/index.php/FLAIRS/FLAIRS15/paper/view/10406.

Michalowski, Piotr. "Charisma and Control: On Continuity and Change in Early Mesopotamian Bureaucratic Systems." Pp. 55-68 in R. D. Biggs and M. Gibson, eds., *The Organization of Power: Aspects of Bureaucracy in the Ancient Near East*. Chicago: University of Chicago Press, 1987.

Mische, Ann. “Relational Sociology, Culture, and Agency” . *The SAGE Handbook of Social Network Analysis*. Edited by John Scott and Peter J. Carrington. Sage Publications: London, 2011. methods.sagepub.com/book/the-sage-handbook-of-social-network-analysis/n7.xml.

Mohr, John W. and White, Harrison C. “How to Model an Institution” . *Theory and Society* 37, 485-512, 2008. [dx.doi.org/10.1007/s11186-008-9066-0](https://doi.org/10.1007/s11186-008-9066-0).

Molina, Manuel. “Archives and Bookkeeping in Southern Mesopotamia during the Ur III period” . *Comptabilités*8, 1-19, 2016. journals.openedition.org/comptabilites/1980.

———. “The Corpus of Neo-Sumerian Tablets: An Overview” . *The Growth of an Early State in Mesopotamia: Studies in Ur III Administration*. Edited by Steven J. Garfinkle and J. Cale Johnson. Biblioteca del Próximo Oriente Antiguo 5, 19-46. Consejo Superior de Investigaciones Científicas, 2008, Madrid, www.academia.edu/10126771/The_Corpus_of_Neo_Sumerian_Tablets_an_Overview_en_S_Garfinkle_J_C_Johnson_eds_The_Growth_of_an_Early_State_in_Mesopotamia_Studies_in_Ur_III_Administration_Biblioteca_del_Pr%C3%B3ximo_Oriente_Antiguo_5_Madrid_2008_pp_19_53.

———. *Database of Neo-Sumerian Texts* (BDTNS). 2020, bdtns.filol.csic.es/.

Noble, William S. "What is a support vector machine?" *Nature Biotechnology* 24.12, 2006, pp. 1565-1567.

Paoletti, Paola. *Der König und sein Kreis: das staatliche Schatzarchiv der III. Dynastie von Ur*, 2012, Madrid: Consejo Superior de Investigaciones Científicas. ISBN:

9788400094911.

———. *Elusive Silver? Evidence for the Circulation of Silver in the Ur III State*. Kaskal 5.

Rivista di storia, ambienti e culture del Vicino Oriente *Antico*, vol. 5, 2008.

www.academia.edu/347477/2008_Elusive_Silver_Evidence_for_the_Circulation_of_Silver_in_the_Ur_III_State.

Fernando Pérez, Brian E. Granger, *IPython: A System for Interactive Scientific Computing*, Computing in Science and Engineering, vol. 9, no. 3, pp. 21-29, May/June 2007,

doi:10.1109/MCSE.2007.53. URL: ipython.org

Picchioni, S.A., "Miscellanea Neo-Sumerica, II: Collazione a M. Çiğ - H. Kızılyay - A.

Salonen, Die Puzrish-Dagan-Texte der Istanbul Archäologischen Museen, Teil I: Nrr. 1-725", *Oriens Antiquus* 14 (1975) 153-168

Rish, Irina. "An empirical study of the naive Bayes classifier." *IJCAI 2001 workshop on empirical methods in artificial intelligence*. vol. 3, no. 22, 2001.

Sallaberger, Walther. "Schlachtvieh aus Puzriš-Dagān" . *Jaarbericht Ex Oriente Lux*, No.

38, 2004. Leiden. [www.assyriologie.uni-](http://www.assyriologie.uni-muenchen.de/personen/professoren/sallaberger/publ_sallaberger/wasa_2004_puzris-dagan.pdf)

[muenchen.de/personen/professoren/sallaberger/publ_sallaberger/wasa_2004_puzris-](http://www.assyriologie.uni-muenchen.de/personen/professoren/sallaberger/publ_sallaberger/wasa_2004_puzris-dagan.pdf)
[dagan.pdf](http://www.assyriologie.uni-muenchen.de/personen/professoren/sallaberger/publ_sallaberger/wasa_2004_puzris-dagan.pdf).

———. *Der kultische Kalender der Ur III-Zeit*. Untersuchungen zur Assyriologie und

Vorderasiatischen Archaeologie, Band 7/1. Berlin: Walter de Gruyter, 1993. [epub.ub.uni-](http://epub.ub.uni-muenchen.de/6382/)
[muenchen.de/6382/](http://epub.ub.uni-muenchen.de/6382/).

Sallaberger, Walther and Aage Westenholz. *Mesopotamian: Akkade-Zeit und Ur III-*

Zeit. Orbis Biblicus et Orientalis 160/3. Vandenhoeck & Ruprecht Göttingen, 1999.

www.zora.uzh.ch/id/eprint/151632/1/Sallaberger_Westenholz_1999_Mesopotamien.pdf.

Sharlach, T. M. *An Ox of One's Own: Royal Wives and Religion at the Court of the Third*

Dynasty of Ur. Studies in Ancient Near Eastern Records 18, 211-238. Edited by Gonzalo

Rubio. De Gruyter, 2017, Berlin. doi.org/10.1515/9781501505263.

———. "Shulgi-simti and the Representation of Women in Historical Sources." *Ancient*

Near Eastern Art in Context: Studies in Honor of Irene Winter. CHANE 26, 363-368. Ed. M.

Feldman and J. Cheng. Brill,

2007, Leiden. [www.academia.edu/36052634/](http://www.academia.edu/36052634/Shulgi_simti_and_the_Representation_of_Women_in_Historical_Sources) [Shulgi_simti_and_the_Representation_of_Women_in_Historical_Sources](http://www.academia.edu/36052634/Shulgi_simti_and_the_Representation_of_Women_in_Historical_Sources).

Steinkeller, Piotr. “The Function of Written Documentation in the Administrative Praxis of Early Babylonia.” In *Creating Economic Order: Record-keeping, Standardization, and the Development of Accounting in the Ancient Near East*, vol. 4, Edited by Michael Hudson and Cornelia Wunsch, CDL, 2004. Bethesda, Maryland.

Tinney, Steve, et al. *ORACC: The Open Richly Annotated Cuneiform Corpus*. 2014. oracc.museum.upenn.edu/.

Tsouparopoulou, Christina. *The Material Face of Bureaucracy: Writing, Sealing and Archiving Tablets for the Ur III State at Drehem*. University of Cambridge, 2009. ethos.bl.uk, ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.611396.

Wang, Junna and Yuhong, Wu. “A Research on the Incoming (MU-TÚM) Archive of Queen Šulgi-simti’ s Animal Institution.” *Journal of Ancient Civilizations*. vol. 26.1, pp. 41-61, 2011. chinaqikanwang.com/thesis/detail/1097967.

Weiershäuser, Frauke. *Die königlichen Frauen der III. Dynastie von Ur*. Göttingen University Press, 2008. doi.org/10.17875/gup2008-504.