



Visibility Threshold: Some Considerations on Data Mining Applied to the Study of Eblaite Society

Limiar de Visibilidade: Algumas Considerações sobre Data Mining Aplicado ao Estudo da Sociedade Eblaíta

<https://doi.org/10.21814/h2d.3466>

Erica Scarpa, Ca' Foscari University of Venice, Italy

Como citar

Scarpa, E. (2021). Limiar de Visibilidade: Algumas Considerações sobre Data Mining Aplicado ao Estudo da Sociedade Eblaíta. *H2D|Revista De Humanidades Digitais*, 3(1). <https://doi.org/10.21814/h2d.3466>

ISSN: 2184-562X



Visibility Threshold:

: Some Considerations on Data Mining Applied to the Study of
Eblaite Society

Limiar de Visibilidade: Algumas Considerações sobre Data Mining Aplicado ao Estudo da So- ciedade Eblaíta

<https://doi.org/10.21814/h2d.3466>

Erica ScarpaCa' Foscari University of Venice, Italy

[This article partly draws on materials presented by the author at the workshop 'Thinking Digital in Cuneiform Studies: Methods, Problems, Perspectives,' (Ca' Foscari University of Venice, 27-28 March 2018)]

Abstract

Accounting is a routine activity. Through repetition, the scribes of the Ebla Archives (Syria, 24th cent. BCE) have been able to record thousands of transactions. They organized and stored accounting data referred to more than thirty years of the Palace G activities. The recurring textual patterns characterizing the administrative corpus are a byproduct of this routine-based approach. The ability to see recurring patterns in the textual record is fundamental when dealing with an administrative corpus: however, this ability fails when the patterns are buried in data. In this paper, I argue that theoretical aspects of data mining are not far from theoretical and methodological tenets of the historical approach. Data mining is a useful technique for the identification of document clusters and relevant information which would otherwise remain hidden. Furthermore, textual pattern recognition is critical to address topics such as the study of society: belonging to a category of complex problems, any socio-historical investigation requires dealing with multiple interconnected variables. However, not all research topics require such an approach. I define the line beyond which digital approaches are extremely useful (if not indispensable) as 'visibility threshold'. The position of this interface is relative and subjective.

Keywords

Ebla; Ancient Near Eastern History; Complexity; Systems theory; Data mining

Resumo

A contabilidade é uma atividade de rotina. Através da repetição, os escribas dos Arquivos de Ebla (Síria, século XXIV a.C.) registaram milhares de transações. Eles organizaram e armazenaram dados contábilísticos referentes a mais de trinta anos de atividades do Palácio G. Os padrões textuais recorrentes que caracterizam o corpus administrativo são o resultado desta abordagem baseada na rotina. A capacidade de ver padrões recorrentes nos registos textuais é fundamental quando se lida com um corpus administrativo: contudo, esta capacidade perde-se quando os padrões ficam submersos num mar de dados. Neste artigo, defendo que os aspectos teóricos do data mining se aproximam dos princípios teóricos e metodológicos da abordagem histórica. O data mining afigura-se como uma técnica útil para a identificação de clusters de documentos e informação relevante, que, de outra forma, permaneceriam ocultos. Além disso, o reconhecimento de padrões textuais é fundamental para abordar tópicos como o estudo da sociedade: qualquer investigação sócio-histórica, na medida em que lida com problemas complexos, requer o tratamento de múltiplas variáveis interligadas. No entanto, nem todos os tópicos de investigação exigem tal abordagem. Eu defino como 'limiar de visibilidade' o patamar para além do qual as abordagens digitais são extremamente úteis (se não indispensáveis). O posicionamento desta interface é relativo e subjetivo.

Palavras-chave

Ebla; História do Antigo Próximo Oriente; Complexidade; Teoria dos sistemas; Data mining

1. Introduction

Most 'sentences' found in the administrative documents of the Ebla Archives (mod. Tell Mardikh, Syria, 24th cent. BCE) can be compared to mathematical formulae. They appear hundreds of times with very few variations; these patterns point to specific activities and are in most cases (although not all) characterized by an economic perspective. Behind each sentence are specific events and occasions, each with its peculiarities; these patterns can be defined as 'accounting models' as the scribes used them recursively to refer to different events characterized by the same accounting circumstances.

Accounting is a process characterized by two main moments: recording transactions and reporting information. According to *modern* accounting theory, the input for the accounting procedure is an economic event (Marriott et al., 2002, pp. 1-2); this event generates the production of a record, a document that is generally drafted on the spot since any delay increases mistakes and might compromise the entire process. Accounting in the ancient Near East had very similar objectives to its modern counterpart: collecting data for decision-making.

Through an accounting procedure based on routine, the scribes of the Archives were able to record and summarize an impressive amount of data.

1.1. The Ebla Archives

The Ebla Archives are the most ancient ‘living’ administrative archives found *in situ*¹; no coeval collections of documents of comparable consistency have been found in modern Syria². All tablets have been discovered in the so-called (Royal) Palace G, located on the south-western slope of the City’s acropolis³. Most were discovered in 1975 in loci 2769, 2712 and 2764, although smaller collections of tablets were found in 1974 in locus 2586 (42 tablets)⁴, in 1982 in different loci from the Southern Wing (5 tablets)⁵, and in 2003 close to the so-called Throne Room (locus 8496, 13 tablets)⁶. The tablets were baked in the fire that put an end to the City’s life around 2350 BCE; the documents, although in part damaged by this catastrophic event, have been preserved in their original archaeological context.

At least two of the epigraphic discoveries present the characteristics of organized collections of documents⁷. First, the Great Archive 2769 was a rectangular room measuring 5,10×3,55m housing wooden shelves on both its northern and eastern walls. The western wall might have also been covered by a shelf, although the damage caused by the fire obliterated the area and the central part of the room. Excavation records show that archaeologists recovered more than 10,000 objects in 1975, consisting of complete tablets, fragments, and flakes; most were found in the Great Archive. We can estimate that ca. 3,000-4,000 tablets were originally kept there, although this number is only an approximation⁸.

The scribes of the Great Archive mostly focused on administrative transactions concerning garments, precious items and metals entering and leaving the Palace: the archive contained ca. 500 monthly accounts of textiles, ca. 40 annual accounts of metals, and ca. 80 documents concerning incomes of precious metals. In addition, several lexical documents and cultic texts, as well as at least 60 diplomatic documents, have been discovered. The scribes probably drafted the monthly accounts of garments on the basis of single receipts or small groups of transactions recorded on lenticular tablets during the preceding month(s); the small tablets were subsequently destroyed or recycled. Unexpectedly, the scribes did not adopt any precise dating method, and monthly accounts present only (and not always) reference to the month. However, Archi and Biga (2003) established a relative chronology on the basis of the annual accounts of garments: the numerous cross-references between the monthly and yearly documents constitute the scaffolding on which Ebla’s relative chronology has been established. The age of the Archives is thus a period of roughly 46 years preceding the City’s destruction, which took place around the year 2350 BCE⁹.

The second organized collection of documents, Small Archive 2712, had a shorter life span: its documents have been dated to the last three or four years before the City was destroyed. Ca. 211 tablets have been recovered from this room,

and most concern food allotment for the Palace personnel¹⁰.

A third archival room was probably Trapezoidal Archive 2764, where archaeologists catalogued ca. 600 items; all documents from this room are still unpublished. Furthermore, ca. 100 tablets (reconstructed from 655 fragments) have been discovered in Vestibule 2875, mostly letters sent by Ebla's kings, viziers, and high officials (Catagnoli & Fronzaroli, 2020). Bone styluses and a stone sleeve found there suggest that 2875 was a writing/reading room (Porter, 2012, p. 211; Veenhof, 1986, p. 6).

1.2. The language

Eblaite is a language belonging to the ancient Northern Semitic linguistic continuum¹¹. However, in administrative texts, the scribes adopted a hybrid vocabulary: much of the administrative lexis is Sumerian, although clearly affected by Semitic influence. In fact, local Semitic spellings abound only in diplomatic documents—that is, in a subset of sixty-odd documents mainly consisting of letters, legal documents, and summary reports referring to various subjects. The scribes probably imported the administrative lexis from Mesopotamia through the city of Mari; since there is no self-evident element suggesting that Sumerian native speakers were anywhere to be found in Syria in this historical period, Ebla scribes probably had a limited knowledge of spoken Sumerian. As a result, their Sumerian often differs from the third-millennium Sumerian of Mesopotamia. Because of these differences, lexical compositions (that is, lists of word arranged by topic or sign shape) are fundamental for the understanding of the administrative lexis.

The use of a non-local lexis in the administrative practice has two main consequences: the first regards the interpretation of the documents from a linguistic perspective, and the second its social contextualization. Sumerian and Eblaite are different languages. The former, for example, is ergative and has no markers for feminine and masculine since it only distinguishes between human and non-human. Eblaite, instead, is a Semitic language with a verbal system typical of this language family and with markers for masculine and feminine. Furthermore, the Sumerian lexis imported by Ebla scribes from Mesopotamia had been shaped to describe a very distant society, both geographically and chronologically. How Eblaite scribes domesticated Sumerian words to their use depends on criteria which we are not always able to understand.

1.3. The study of Ebla's society

The administrative record is but one of many lenses through which we can observe Ebla's society. Artistic production, material culture, and architecture are also informative with respect to society, although written sources undoubtedly offer a privileged perspective. However, all documents found at Ebla are 'institutional': the lower City (that is, the area surrounding the acropolis) has yet to be investigated, and no private archive has been discovered so far. The institution

behind the record-keeping system is the Palace: this organization is largely kin-based, with the royal family occupying a prominent position. The second most important family is that of the viziers, who assisted the kings in their duties. During the approximately 46 years documented by the Archives, two kings (Irkab-damu and Iš'ar-damu) and three viziers (ArruLUM, Ibrium, and Ibbi-zikir) managed the City and its surrounding territory.

The events taking place behind the transactions recorded in the Archives are largely linked to these two families and their activities. The accounts can therefore provide a picture of the Palace elite; however, other social groups are consequently underrepresented¹². Nevertheless, these two families constitute the core of the numerous connections linking them to other individuals and kin-groups both in and outside the City. Only by investigating these connections and the social actors involved can an overall picture of Eblaite society be revealed.

As has already been stressed, Ebla's administrative record is characterized by hundreds of 'accounting models'—that is, textual patterns describing a specific transaction concerning the movement of movable properties from the Palace to an individual or group (and, in some cases, vice versa). The central administration was concerned, for example, with someone's death because it had to account for the delivery of precious items for the deceased:

[**TM.75.G.1743 = ARET XX 9**][i] obv.III:10-15: 1 aktum^{túg} 1 íb-II^{túg} gùn / *mí-ga-i*₃ / lú *hu-nu-ba-am*₆ / nagar / *si-in* / É×PAP;
 “1 cloak, 1 embroidered waistband (for) Miga-i, the one of (the household of) Hunabam, the carpenter, for the funerary ceremony.”

In this example, the accounting model includes the expression *si-in* É×PAP, which is generally interpreted as “for the funerary ceremony” and occurs more than two hundred times in the documents of the Great Archive 2769. However, although the funerary meaning related to this context is generally recognized by all scholars, the real meaning of the word É×PAP is still debated¹⁴.

Other accounting models do not refer directly to the specific event or circumstance that prompted the entry. Although the circumstance is clear to the author of the document, it is not self-evident for the modern reader:

[**TM.76.G.530 = ARET I 5**] rev.VII:25-VIII:11: 9à-*da-um*^{túg}-I 9 aktum^{túg} 9 íb-II^{túg} sa₆ gùn / *a-bí* / *en-na-[ba]-al*₆ / *i-in-i* / *má-NI* / *mí-nu*^dKU-*ra* / *a-bù* / *bù-ma-ù* / *en-na-i* / *ru*₁₂-*zi-gú-lu* / *maškim* / *ir-à-ag-da-mu*; “9 cloaks, 9 tunics, 9 embroidered waistbands of good quality, (for) Abi, Enna-[ba], I'in-i(l), Ma-NI, Minu-Kura, Abu, Buma'u, Enna-i(l), Ruzi-gulu, appointees of Ir'ag-damu”.

In this passage, crown prince Ir'ag-damu entrusted nine men (his appointees, *maškim*) with an unknown task for which they all received a complete set of garments including a cloak, a tunic, and a waistband. The Sumerian term *maškim* is one of the most often attested function names in the Ebla Archives, occurring more than 2,300 times; although frequently attested also in Mesopotamian

administrative documents, mainly with a legal connotation, many doubts still persist as to its meaning at Ebla¹⁵. Furthermore, the connection between Ir'ag-damu and the nine men cannot be further specified, as it does not occur in other documents. However, hundreds of other men of different rank and provenance appear in transactions registered with the same accounting model: “[complete set(s) of garments and/or precious items] A(, B, C, D) appointee(s) of E”. Previous studies of ‘maškim’ and its related accounting models did not take into consideration all available data due to the term’s ubiquity in the administrative record.

Over the past five decades, Eblaite language, culture, economy, and society have been widely investigated. However, past studies have not been able to draw a consistent picture of Eblaite society. The reason for this situation lies in the abundance of the written sources. The number of documents published so far is overwhelming and represents a challenge for a holistic approach to the study of many phenomena – especially those relating to society.

A first attempt at describing Ebla’s society was made by Pettinato (1979, pp. 127-170), who divided it into two classes: the “sons of Ebla” and the foreigners, with the former group comprising artisans, merchants, and workers, and the latter other professional figures such as priests, singers, and foreign merchants. This preliminary sketch, offered by Pettinato only a few years after the Archives discovery, is now outdated. A new reconstruction of Eblaite society was offered by Milano in 1990, in a volume devoted to the publication of the first 115 tablets recovered in the Small Archive 2712 concerning food allotment. This outline, however, is concerned exclusively with the social groups mentioned in this particular accounting office and does not take into account the evidence coming from Great Archive 2769.

Scholars have subsequently investigated the roles associated with several professions and social functions¹⁶; however, our understanding of how these roles and functions were interconnected is still in its infancy. Many uncertainties persist about even basic social facts, from how families were structured, to personnel management, and above all about Ebla’s political organization. The way individuals interact with each other in a given society cannot be easily predicted; however, the intricate net of relationships weighs as much as the behavior of the single individual or group. The sets of relationships involving an actor are relevant as much as the single individual. Studies on Eblaite society have largely drawn on a case-study approach focusing on single individuals, events, or the like, which has so far demonstrated its limits for a comprehensive approach to society.

2. Complexity and the study of society

2.1. The meaning of complexity

We often use the word ‘complexity’ to describe the past. Complexity often appears as a feature of ancient societies: in cuneiform studies, we write about complex societies, complex organizations, and complex social structures (Liverani, 2006; Yoffee, 2004, p. 16). Scribes adopted a “complex writing system” and drafted “complex Sumerian literary texts” as part of their “complex curriculum” (Veldhuis, 2014, pp. 2, 14). When it comes to administration, nobody can deny that proto-cuneiform was a “complex accounting mechanism” (Englund, 2004, p. 34), nor that an “expanding and increasingly complex administrative bureaucracy” characterized the Neo-Sumerian period (ca. 2112-2004 BC; Winter, 1991, p. 74). The list could easily be expanded with many more examples. So far, however, there has been little discussion about what we mean when we define something as complex.

Complexity is one of those concepts that has received countless definitions. It characterizes both ‘hard’ and ‘soft’ sciences: in general terms, it is the property of systems made up of many interrelated parts. However, it is not synonymous with complicated: complex problems are not inherently difficult to solve. What makes complex problems difficult to approach is not simply the fact that they have multiple components. The first serious discussions of complex problems emerged during the 1940s with Warren Weaver’s influential essay on the topic (Weaver, 1948). There, Weaver defined new classes of scientific problems: the problems of *simplicity* and those of *disorganized complexity*. Simple problems are characterized by a limited number of variables: a pendulum’s movement depends on its mass, the string’s length, and the angular displacement. Problems of disorganized complexity present multiple variables interconnected in a nonlinear way: predicting the outcome of a coin toss or analyzing the motion of balls on a billiard table are problems of disorganized complexity. In disorganized complexity, variables interact in such a way that the most successful approach to solving problems is to apply probabilistic and statistical methods.

In the middle of the continuum stretching from simplicity to disorganization are the problems of *organized complexity* (e.g., cf. Seising, 2012, p.62, fig.3.5). These problems “involve dealing simultaneously with a sizable number of factors which are interrelated into an organic whole” (Weaver, 1948, p. 539); they are characterized by a large number of dynamic variables, but interconnected in a non-chaotic way. For this reason, statistical and probabilistic approaches are not as effective as in disorganized complexity. Furthermore, organized complexity is self-organizing and adaptive. This means that studying interactions in such systems is like “walking through a maze whose walls rearrange themselves as one walks through” (Urry, 2005, p. 3). Societies are systems belonging to the class organized complexity: “[t]o study society is, by definition, to study complexity” (Castellani & Hafferty, 2009, p. 4).

A society's most basic constituent is the dyad (*i.e.*, a group of two actors), which is often considered its basic unit and the smallest possible human group larger than the single individual. But societies also rely on the continuous change and evolution of these dyadic relationships between the actors; this is because actors move continuously between different social groups (Dudley & Brown, 1981). Time and space are not fixed points for the study of society; social groups are characterized by an irregular diachronic evolution, and actors constantly move from one social group to another. It follows that a society cannot be identified as a given object with space-time coordinates, because at any given moment, only a portion of it is existing (Dudley & Brown, 1981, pp. 315–316). We can only observe several relationships taking place at a given time and in a given place. Although these movements are not chaotic and unpredictable, they can hardly be explained and described only by dyadic pairs. Kinship relationships, for example, can be represented by dyads but also reach levels of complexity that go beyond the interaction between dyadic pairs.

2.2. Dealing with complexity

Breaking something into smaller units is a universal kind of approach: in every discipline, from physics to biology, sociology, and history, when something is too complex, we are used to dividing it into manageable parts. Case-study research has proven to be a powerful research method. It “[has] been around as long as recorded history” (Flyvbjerg, 2011, p. 302) and is one of the most widespread research approaches; however, it also has negative connotations because of its limited scope. Among the common misunderstandings about case studies mentioned by Flyvbjerg, the second is probably the most widespread (p. 304): “one cannot generalize on the basis of an individual case; therefore, the case study cannot contribute to scientific development”. The “casing” activity (Ragin, 1992, p. 217) is a research strategy based on the definition of a unit of study and its boundaries; these, however, can be set in many different ways and according to diverse criteria. A reductionist approach such as “casing” makes analyzing the problem easier. However, it ignores the fact that bringing the pieces back together does not produce knowledge about the system as a whole.

Systems theory is based on the assumption that complex systems depend on nonlinear interactions: one of its most famous mottos states, “the nature of the whole is always different from the mere sum of its parts” (Capra, 1997, p. 29). Systems theory also explains why covering laws—that is, universal laws—do not work when it comes to complex systems: the relationship between *explanandum* (the phenomenon) and *explanans* (the set of statements explaining it) does not take into account nonlinear interactions between the system's variables. Social sciences are now included in the group of disciplines adopting complexity science tools (Crutchfield & Wiesner, 2010, p. 38).

2.3. Staying on top of data

Societies are thus complex systems; it is impossible to understand a society's structure by studying its actors individually. The context of relationships tying individuals to each other would be completely lost. Case-study methodology is an effective approach; however, in the case of complex problems, the "casing" should take into account a data pool exceeding human computing capabilities. In the last two decades, a digitization effort involving Assyriologists from different fields has led to the creation of dozens of databases¹⁷. We now have many data repositories that are high-dimensional (*i.e.*, characterized by a high level of attributes), diverse, and constantly growing. However, beneath the surface of the digitized material lies a submerged iceberg of documents published only on paper, quoted excerpts, and unpublished material. This means that, as the digitization effort proceeds, available digitized data will increase accordingly.

Let us consider, for example, the following research question: are there seasonal patterns in the events mentioned in the administrative documents of the Ebla Archives? Do some deliveries or events take place in specific moments of the year? Some events are certainly not tied to specific moments (*e.g.*, deaths and births), while others take place on specific occasions (rituals). It may also be the case that we are yet unaware that certain events are characterized by a seasonal pattern. In the case of seasonality, there are only two variables: time (*i.e.*, month names) and events (*i.e.*, keywords referring to specific events). But what happens if the variables are more than two? "Casing" should not be dictated by time or effort constraints. However, case-study research inevitably tends to set boundaries around a reduced subset of variables. In the study of society, where the context is even more important than the variables themselves, how can "casing" strategies be defined?¹⁸ How do we establish that a subset of documents or variables are relevant to the study of society if data is buried in thousands of texts?

3. The Data Mining Process

As stated in the introduction, Ebla's administrative documents are based on the re-elaboration of a limited (albeit large) set of textual patterns. These accounting models are arrangements of words (nouns, verbs, prepositions) corresponding to a specific administrative event or circumstance. Each pattern may contain data regarding the event or circumstance it describes, the person(s) involved, the place where the event took place, or the goods that are the object of the transaction. But on a corpus of more than 3,000 published excavation records¹⁹, pattern recognition is not a task that humans can easily carry out²⁰. Data mining is defined as "the process of automatically discovering useful information in large data repositories", and it is used to "scour large databases in order to find novel and useful patterns that might otherwise remain unknown" (Tan et al., 2006, p. 2).

Data mining is not only a tool but a *process*: in this sense, it helps to put into practice what Franco Moretti defines as ‘distant reading’: the automated search for patterns across a corpus too wide to be read and assimilated through the traditional method of ‘close reading’ (Moretti, 2013). However, distant reading through data mining is not a way to avoid reading documents; it is rather a powerful tool to quickly identify the most relevant ones. Close reading must always be part of any philological or historical research: however, distant reading allows the recognition of patterns and relevant dossiers in thousands of documents. In this sense, data mining is a form of ‘intelligence amplification’ – a tool for “increasing the capability of a man to approach a complex problem situation, to gain comprehension to suit his particular need, and to derive solutions to problems” (Engelbart, 1962, p. 1).

Manuals and literature dealing with data mining are mainly concerned with its applications in business, marketing, or science. Its use in the field of the humanities largely depends on the practitioner; there is no clear methodology because this depends largely on the data and specific domains. Detailed tutorials and courses on practical aspects of data mining are plentiful online. In this paper, I would like to discuss a few theoretical aspects of this approach which I consider extremely meaningful for historical research.

In the remainder of this article, I intend to review some of the so-called ‘laws of data mining’ (Khabaza, 2014, pp. 193–194) and offer some considerations about their relevance for a historical approach²¹.

3.1. First Law, “Goals Law” – Objectives are the origin of every data mining solution

Technically, the first law is called “*Business Goals Law*,” and it states that “*business objectives are the origin of every data mining solution.*” This law simply states that data mining “is concerned with solving business problems and achieving business goals” (Khabaza, 2014, p. 194). I omitted the word ‘business’ because it often gives to humanists the false idea that data mining can be successfully performed only in certain contexts, such as management or finance. However, omitting the word ‘business’, data mining laws appear to be much closer on a theoretical level to the historical domain. This first law regards objectives: every data mining project needs a goal because the goal shapes the data mining process. The goal is a research question, which in turn is the origin of the data mining solution; in this regard, a data mining project does not differ much from a traditional research project.

3.2. Second Law, “Knowledge Law” – Knowledge is central to every step of the data mining process

Another fundamental concept in data mining is knowledge: knowing the data and the domain to which it belongs is pivotal to understanding which data is related to the question. During data preparation, knowledge is essential to gathering the

data so that the questions can be both asked and answered. Without knowledge, one cannot ask a question, gather data, prepare it, or find a solution. Data mining can speed up the process of recognizing recurring accounting models; but in all steps of a data mining project, from data gathering strategies to results interpretation, knowledge is a vital component. The data mining process is a cycle continuously informed by knowledge of the domain.

The reason for this depends on a ‘data-reality’ gap: data is only a representation of reality, not the reality itself. There is a gap between what the data *represents* and what reality *is*. Knowledge in the domain is fundamental to bridging this gap. Furthermore, the historical reality we deal with as historians is partial, because the historical record is fragmentary and incomplete – it does not document every single aspect of past reality.

3.3. Third Law, “Data preparation Law” – Data preparation is more than half of every data mining process

Data preparation is the most time-consuming step in a data mining project. It takes from 50 to 80% of the overall time spent on the process. Data preparation usually consists of collecting and cleaning the data for the analysis. In many cases, data mining is performed on a single table, and the table must suit both the tool and the problem. This means it must be understood by the software, and it must be informed by the objective.

In a paper published in 1967, Gelb wrote that approaches to the study of ancient society can be summarized in three steps:

1. Collection of data
2. Primary interpretation of data
3. Secondary interpretation

In Gelb’s view, data is “the original sources as preserved in ancient languages” (1967, p.3). The author also observes that during data gathering, the role of the collector should be adequately assessed: is data collected automatically or cherry-picked by the collector? Should a dataset be comprehensive or specific? Considering the great number of sources available, Gelb suggests that data completeness is utopian and that the completeness of a dataset thus depends on the endurance of the collector. As a consequence, a dataset based on historical data will never be complete; if the incompleteness depends on the sources, there is nothing we can do about it. However, it should not depend on the collector’s selectivity (*i.e.*, subjective judgment), because data relevance and importance in the framework of a specific research question is unpredictable at the moment of collection.

As the data-collection process described by Gelb, data mining is cyclic: the knowledge of the domain informs the whole process, and the data miner can perform each step multiple times until the attainment of a satisfactory result.

As Gelb noted, the approach to the study of ancient societies is not linear: “the process of collecting new data, as they become available, must go on in all stages, while both primary and secondary interpretation may take place at the moment of collecting original data” (p. 3).

3.4. Fifth Law, “Watkins’ Law” – There are always patterns

Patterns are an inevitable by-product of the processes through which data is produced; if there is no pattern, there is no relevant data. Furthermore, looking for patterns means searching for meaningful connections between variables. Since the research question is constantly informed by the iterative process between data gathering and pattern recognition, recurring patterns are always present. As a consequence, recurring patterns depend on the process of data production and on the fact that the research question is informed by an iterative process based on the knowledge of the domain.

From a socio-historical perspective, the accounting models attested in the administrative record of Ebla are the result of specific activities; these, in turn, depend on the interaction between social actors. As previously stated, in organized complex systems, these interactions are always non-chaotic. Social organizations depend on sets of values and norms which permanently regulate the actors’ behavior through habit (Harrington et al., 2006, pp. 280-283). Patterns are an inherent characteristic of human behavior, in particular when it is the result of group effort.

3.5. Sixth Law of Data Mining, “Insight Law” – Data mining amplifies perception in the domain

As already stated, the process of discovering and learning through data mining is interactive: the path from the formulation of the objective to the results is not straightforward. Data mining provides the ability to detect patterns that are not normally visible to human perception; the data miner can then interpret the results integrating them into his or her understanding of the problem. Data mining techniques enable the practitioner to understand the problems better than he or she could have done without them. As Khabaza wrote, “problems are solved by people, not by algorithms” (2014, p.198).

4. Data mining applied to the study of Ebla society

A society is made up of individuals and the relationships linking them together. Hundreds of people appearing in the Ebla Archives took part in the Palace life: some of them even visited Ebla from regions outside the borders of modern Syria. Although what we know about them is strictly related to the role they played within the Palace organization, the scribes of the Archives were able to

record and summarize an impressive amount of data regarding their status and activities. The study of nouns, verbs, and expressions indicating professions, social functions, and events connecting individuals is thus of primary importance for the understanding of society as a whole. Such investigation, however, is hampered by the fact that these words and phrases are scattered over thousands of texts and by our still-incomplete knowledge of Eblaite language.

Since, as already stated, scribes used formulaic expressions to register the transactions, data mining can be used to identify recurring patterns related to the aforementioned words and expression. Accounting models can be identified, for example, using n -grams—that is, repetitive sequences of n elements. Unfortunately, interesting text-mining features such as tokenization, Part of Speech (POS) tagging, and sentence breaking cannot yet be applied to Ebla’s corpus due to the many uncertainties still surrounding its interpretation, from both lexical and morphological points of view²². However, recursive accounting models can be identified by manually instructing the software on at least some of the building blocks characterizing the models themselves, such as the different types of textile, personal names, and geographical names. Such operation highlights, for example, frequent accounting models such as the following (PN = personal name):

garments [and/or] precious items PN_{A-1} PN_{A-2} (PN_{A-n}) maškim
PN_B²³.

The term maškim is generally translated as “agent” and related meanings, such as “delegate” or “representative”: all these translations agree on the idea that the maškim acts on behalf of a superior authority, although any further connotation is still problematic²⁴. This accounting model occurs more than 1,500 times in the published documents. Visualizing the relationship between PN_{A-n} and PN_B in a social network, a previously undetected characteristic of this pattern emerges (see Fig.1)²⁵. All individuals appearing as PN_B (marked by red circles) are high officials, many belonging to important kin-groups involved in the City’s political endeavors; among them are also the viziers and the crown prince. Furthermore, many appointees (= PN_{A-n}) are also personal servants (pa₄-šeš) of influential individuals. In this example, data mining and data visualization aided the recognition of previously undetected characteristics of one particular noun denoting a social function, further clearing its contextual meaning. All in all, this approach shortened the observer’s distance from their visibility threshold.

However, not all accounting models require this kind of approach: in addition to those rarely attested, some are simply not recursive enough to be analyzed through digital means. In addition, for some models, the observer’s knowledge of the domain might suffice, particularly if the administrative lexis is clear and no doubts persist as to its contextual meaning.

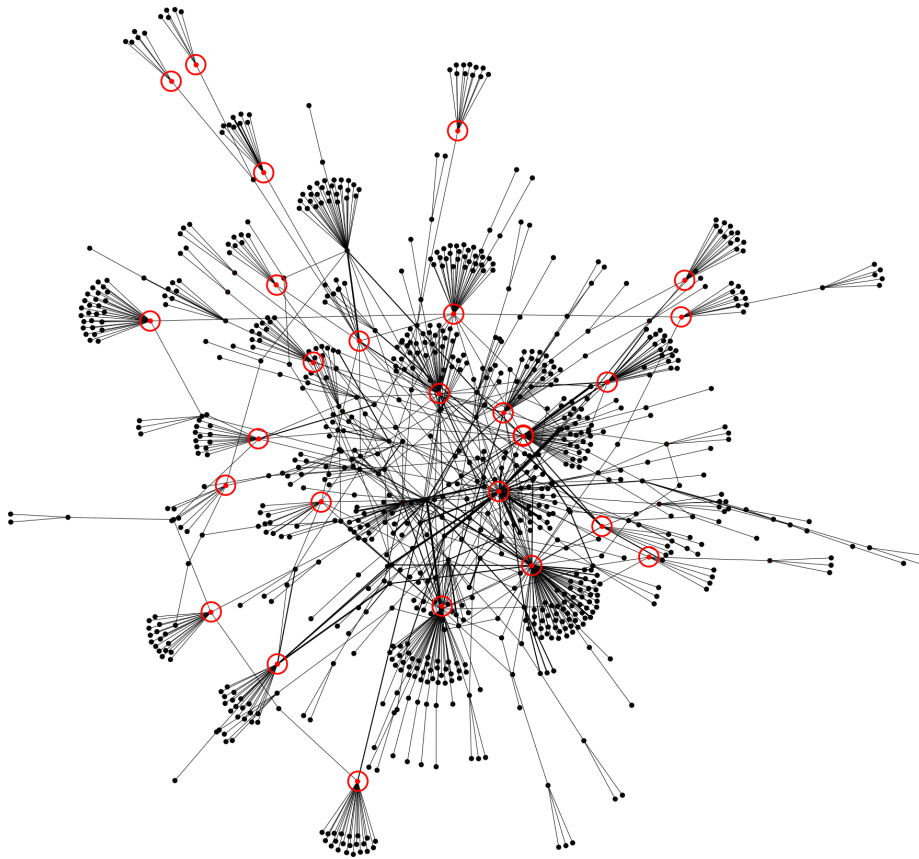


Figure 1: Graph network displaying the accounting model related to the maškim

5. Conclusion

In Weaver’s words, the problems of organized complexity could be solved with “electronic computing devices” and a “‘mixed-teams’ approach” (Weaver, 1948, p. 534). Dealing with complex problems requires new solutions: data mining for pattern recognition is only one of the possible approaches to the study of ancient societies. The ability to recognize patterns and recurring models in textual corpora without digital tools depends on each scholar’s knowledge. However, it also depends on the limits of the human brain—its ability to detect recurring patterns and retain information. On the one side, the visibility threshold is subjective and strictly correlated with knowledge: a higher knowledge of the domain ensures a better ability to detect patterns. Tables, for example, are one of the devices we use to enhance this ability.

In 1801, William Playfair, a Scottish political economist, wrote that

[i]nformation, that is imperfectly acquired, is generally as imperfectly retained; and a man who has carefully investigated a printed table, finds, when done, that he has only a very faint and partial idea of what he has read; and that like a figure imprinted on sand, is soon totally erased and defaced (pp. xiii-xiv).

The difficulty of extracting information from tables is not something recent, but simply something the human brain is not trained to do. Therefore, this threshold is also a fuzzy line beyond which traditional research is no longer effective because we are no longer able to automatically detect patterns (although we might still suspect their presence). This study set out to critically examine how the process of data mining in large part overlaps with the historical approach to the study of ancient societies. Data mining also provides additional tools to identify otherwise undetectable relevant documents or keyword clusters: as previously stated, it helps to augment human intelligence rather than replacing it.

Submitted 2021-06-15 | Published 2021-10-31

Notes

1) I adopt here the distinction suggested by Fales (2003, pp.197-198; drawing on Castel, 1995): the author divides archives found in their original archaeological contexts between ‘living’ (that is, left undisturbed until their discovery) and ‘dead’ (that is, dismissed and removed from their original archaeological context in antiquity).

2) Ca. 150 cuneiform administrative documents have also been discovered in Nabada (Tell Beydar; see Ismail et al., 1996) and Mari (Tell Hariri; see recently Cavigneaux, 2014). However, neither finding is comparable in volume to the Ebla Archives.

3) For a map of Palace G during the age of the Archives, see Fig. 4.1 in Matthiae

(2020, p. 43). For a map of excavated areas at Tell Mardikh (up to 2009), see Fig.1 .10 (p. 14).

4) Matthiae (2008, p. 64).

5) Published in *Archi* (1993).

6) All tablets have been published in *Archi* (2015).

7) See Matthiae (1986) and Matthiae (2008, pp. 63-77), for a detailed description of epigraphic discoveries at Ebla.

8) Most editions of the Ebla tablets are devoted to complete and restored tablets: ca. 600 such documents have been published since the Archives' discovery. However, almost 2,400 fragments have also been published (*Archi & Biga, 1982; Lahlouh & Catagnoti, 2006*).

9) On the relative chronology adopted in Ebla studies, see *Archi and Biga (2003)* and *Archi (2016)*. On the dating of the City's destruction, see *Sallaberger and Schrakamp (2015, pp. 96-100)*.

10) The 115 tablets have been published by *Milano (1990)*. The remaining 117 are yet to be published.

11) On the classification of Eblaite language, see *Catagnoti (2012, pp. 3-4)*, and *Tonietti (2013, pp. 20-21)*.

12) Two social groups are especially neglected in the archival record: male workers and the army. Food allotment for female personnel living in the Palace gives a detailed description of the two main activities involving women: textile production and food preparation. In contrast, no comparable information can be obtained for male personnel (see also *Milano, 1990, pp. 337-338*). Similarly, information regarding the army, its structure and management, is equally scanty. This situation can be in part explained by the fact that women represented the center of many pivotal activities (food preparation, textile production, child-rearing), while men were more involved in activities taking place outside the Palace.

13) Photographs of the documents cited in this paper are available on the *Ebla Digital Archives* website. See, for example, <http://ebda.cnr.it/tablet/view/3098> for [TM.75.G.1743 = *ARET* XX 9] and <http://ebda.cnr.it/tablet/view/6> for [TM.76.G.530 = *ARET* I 5].

14) See *Biga (2007-2008, p. 250)* with bibliography.

15) On the meaning of *maškim* at Ebla, see *Davidovic (1987)*.

16) See, for example, *Biga's* studies on the female members of the royal family, as well as on the roles of several priests; *Catagnoti's* studies on the court's dancers (NE.DI) and acrobats (hUB₂.KI); *Tonietti's* studies on the so-called singers (nar). For detailed bibliographical references on these and other professions, see *Scarpa (2017, esp. p. 42)*.

17) Among these are, for example, the *Cuneiform Digital Library Initiative* (cdli, <https://cdli.ucla.edu>), with more than 345,000 entries, each of which corresponds to a cuneiform document with different geographical provenance and chronology; the *Database of Neo-Sumerian Texts* (BDTNS, <http://bdtns.filol.csic.es>), with its ca. 102,000 Ur III administrative documents editions currently available; the *Open Richly Annotated Cuneiform Corpus* (ORACC, <http://oracc.museum.upenn.edu>), which gathers many projects such as the *Digital Corpus of Cuneiform Lexical Texts* (DCCLT, <http://oracc.museum.upenn.edu/dcclt/>) and its sub-project dedicated to the Ebla lexical material (DCCLT: Ebla, <http://oracc.museum.upenn.edu/dcclt/ebla/>). More than 30,000 digital editions of Old-Babylonian (2000-1600 BC) administrative documents have been published in *ARCHIBAB* (<https://www.archibab.fr/>), while the *Ebla Digital Archives* (EbDA, <http://ebda.cnr.it>) offers digitized editions of the Ebla documents.

18) With regard to kinship relationships, the description of ‘household’ given by Netting is especially interesting. He described them as having “an emergent character that *makes them more than the sum of their parts* [...]; they are a primary arena for the expression of age and sex roles, kinship, socialization, and economic cooperation where the very stuff of culture is mediated and transformed into action” (Netting et al., 1984, p. xxii, emphasis mine).

19) The *Ebla Digital Archives* project (Milano & Maiocchi, n.d.) offers the digitized edition of more than 3,800 excavation numbers, representing more than 90% of the published material. However, this material corresponds to only 30% of the overall epigraphic discoveries from Tell Mardikh.

20) Note, however, that many scholars adopted textual pattern recognition as a research strategy. Tonietti (1989), for example, studied the diachronic distribution of the court ladies mentioned in several lists. She was thus able to establish a chronological sequence based on recurring patterns referring to the ladies’ names. Similarly, Pomponio (2003) studied the meaning of the two related terms TUŠ.LU₂×TIL and šu-du₈ investigating the respective textual patterns. Many historical and philological studies depend on the analysis of recurring textual models.

21) Data mining laws are discussed by Khabaza (2014). Khabaza originally formulated them on his website in 2010 (Khabaza, n.d.).

22) The automatic recognition of sentence boundaries constitutes a pivotal problem: there are no punctuation marks in the cuneiform writing system, and in administrative documents, syntax is in many cases ambiguous. At Ebla, as in other administrative archives, sentences correspond to a single transaction, and their logic boundaries can be observed in the recursive use of similar sentence structures. In Eblaite documents, such boundaries are represented by the list of garments or (precious) items, followed by the recipients, and finally the circumstance of the transaction, as in the following example: [TM.75.G.1288 = *ARET IV 11*] obv.VIII:17-IX:2: 1 túg-NI|NI 1 íb-III gùn^{túg} / *ar-ra-ma-lik* / *kak-mi-um*^{ki} / šu-mu-tag₄ / geštin; “1 NI|NI garment, 1 embroidered waistband,

(for) Arra-malik of Kakmi’um, (who) delivered wine”. This sentence is followed by another starting with a list of garments. This rule, however, has exceptions: small sentences can be chained together and refer to one single occurrence, as in the following example: [TM.75.G.1899 = *ARET XX 16*] obv.X:6-13: 1à-da-um^{túg}-II 1 aktum^{túg} 1 íb-III^{túg} sa₆ gun₃ / ba-da-lum / ha-ra-an^{ki} / 1à-da-um^{túg}-II 1 aktum^{túg} 1 íb-III^{túg} gun₃ / maškim-SU₃ / 1 aktum^{túg} 1 íb-III^{túg} gun₃ / ma-za-lum-SU₃; “1 cloak, 1 tunic, 1 embroidered waistband of good quality, for the deputy (of the king) of harran; 1 cloak, 1 tunic, 1 embroidered waistband, (for) his appointee; 1 cloak, 1 tunic, 1 embroidered waistband for his watchman”; the whole context refers to the visit of the deputy of the king of harran and his retinue to Ebla.

23) It occurs, for example, in the following passage: [TM.75.G.1681 = *ARET IV 19*] obv. I:9-II:4: 5 íb-III^{túg} sa₆ gùn / 5 KIN siki / šu-ma-lik / da-zi-ma-ad / i-ti-a-gú / gú-zi / ha-zi-lu / maškim / I-bí-zi-kir; “5 embroidered waistbands of good quality, 5 KIN-masures of wool, (for) Šu-malik, Dazi-malik, Iti-agu, Guzi, hazilu, Ibbi-zikir’s appointees.”

24) On the meaning of the term maškim in Mesopotamia, see Oh’e (1983, p. 124). On its interpretation in Ebla studies, see Davidović (1987); for a different view, see Sollberger (1986, p. 62 *s.v.* maškim).

25) Social Network Analysis (SNA) has already been successfully applied in cuneiform studies. Wagner et al. (2013) applied SNA to the study the texts of the Murašû archives, in a so-called *two-mode* network in which nodes belong to two categories, individuals and events (in this case, the events coincide with the documents). This approach, pioneered by Waerzeggers (2014), has also been adopted by Maiocchi (2016) on the Ebla corpus with promising results. A different approach is that of Bamman et al. (2013), who applied SNA and probabilistic models to identify the unique individuals involved in the colonial trade network at ancient Kaneš.

Referências bibliográficas

- Archi, A. (1993). Five Tablets from the Southern Wing of Palace G – Ebla. *Udena*.
- Archi, A. (1985). Testi amministrativi: Assegnazioni di tessuti (Archivio L. 2769). *Archivi Reali di Ebla, Testi I*. Università di Roma ‘La Sapienza’.
- Archi, A. (2015). The Tablets of the Throne Room of the Royal Palace G of Ebla. *Archiv für Orientforschung*, 53, 9–18.
- Archi, A. (2016). Ebla and Mari—Years 2381/2380–2369 BC. In J. Patrier, P. Quenet, & P. Butterlin (Eds.), *Mille et une empreintes: Un Alsacien en Orient: Mélanges en l’honneur du 65e anniversaire de Dominique Beyer* (pp. 1–16). Brepols.

- Archi, A. (2018). Administrative Texts: Allotments of Clothing for the Palace Personnel (Archive L. 2769). *Archivi Reali di Ebla, Testi XX*. Harrassowitz.
- Archi, A., & Biga, M. G. (1982). Testi amministrativi di vario contenuto (Archivio L.2769: TM.75.G.3000-4101). *Archivi Reali di Ebla, Testi III*. Università di Roma 'La Sapienza'.
- Archi, A., & Biga, M. G. (2003). A Victory over Mari and the Fall of Ebla. *Journal of Cuneiform Studies*, 55, 1–44. <https://doi.org/10.2307/3515951>
- Bamman, D., Anderson, A., & Smith, N. A. (2013). Inferring Social Rank in an Old Assyrian Trade Network. <https://arxiv.org/abs/1303.2873v1>
- Biga, M. G. (2007-2008). Buried Among the Living at Ebla? Funerary Practices and Rites in a XXIV Cent. BC Syrian Kingdom. *Scienze dell'antichità*, 14, 250–275.
- Biga, M. G., & Milano, L. (1984). Testi amministrativi: Assegnazioni di tessuti (Archivio L. 2769). *Archivi Reali di Ebla, Testi IV*. Università di Roma 'La Sapienza'.
- Capra, F. (1997). *The Web of Life: A New Scientific Understanding of Living Systems*. Anchor Books. <https://doi.org/10.2307/1521798>
- Castel, C. (1995). Contexte archéologique et statut des documents: Les textes retrouvés dans les maisons mésopotamiennes du Ier millénaire av. J.-C. *Revue d'Assyriologie et d'archéologie orientale*, 89, 109–137. <http://www.jstor.org/stable/23281596>
- Castellani, B., & Hafferty, F. W. (2009). *Sociology and Complexity Science: A New Field of Inquiry*. Springer. <https://doi.org/10.1007/978-3-540-88462-0>
- Catagnoti, A. (2012). *La grammatica della lingua di Ebla*. Università di Firenze.
- Catagnoti, A., & Fronzaroli, P. (2020). Testi di cancelleria. Il re e i funzionari: Archivio L. 2875. Parte II. *Archivi Reali di Ebla, Testi XVIII*. Harrassowitz Verlag.
- Cavigneaux, A. (2014). Nouveaux Textes de Mari. Ville II. In P. Butterlin, J.-C. Margueron, B. Muller, & M. Al-Maqdissi (Eds.), *Syria Supplément*, 2, 291–340.
- Crutchfield, J., & Wiesner, K. (2010). Simplicity and Complexity. *Physics World*, 23, 36–38. <https://doi.org/10.1088/2058-7058/23/02/38>
- Davidovic, V. (1987). maskim nei testi di Ebla. In L. Cagni (Ed.), *Ebla 1975-1985. Dieci anni di studi linguistici e filologici. Atti del convegno internazionale (Napoli, 9-11 ottobre 1985)*. Studi Asiatici Series Minor XXVII (pp. 379–93). Istituto Universitario di Napoli 'L'Orientale'.
- Dudley, C. J., & Brown, E. (1981). Social Relativity: The Motion of Groups and Actors. *The Sociological Quarterly*, 22, 313–326. <https://doi.org/10.1111/j.1533-8525.1981.tb00664.x>

- Engelbart, D. C. (1962). *Augmenting Human Intellect: A Conceptual Framework*. Stanford Research Institute.
- Englund, R. K. (2004). Proto-Cuneiform Account-Books and Journals. In M. Hudson & C. Wunsch (Eds.), *Creating Economic Order. Record-keeping, Standardization and the Development of Accounting in the Ancient Near East* (pp. 23–46). CDL Press.
- Fales, F. M. (2003). Reflections on Neo-Assyrian Archives. In M. Brosius (Ed.), *Ancient Archives and Archival Traditions: Concepts of Record-keeping in the Ancient World* (pp. 195–229). Oxford University Press.
- Flyvbjerg, B. (2011). Case Study. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE Handbook of Qualitative Research*. SAGE.
- Gelb, I. J. (1967). Approaches to the Study of Ancient Society. *Journal of the American Oriental Society*, 87, 1–8. <https://doi.org/10.2307/596588>
- Harrington, A., Marshall, B. L., & Müller, H.-P. (Eds.). (2006). *Encyclopedia of Social Theory*. Routledge.
- Ismail, F., Sallaberger, W., Talon, P., & Van Lerberghe, K. (Eds.). (1996). *Administrative documents from Tell Beydar (Seasons 1993-1995)*. Brepols.
- Khabaza, T. (n.d.). Nine Laws of Data Mining. Retrieved June 9, 2021, from <http://www.khabaza.com/9laws>
- Khabaza, T. (2014). From Cognitive Science to Data Mining: The First Intelligence Amplifier. In J. L. Wyatt, D. Petters, & D. Hogg (Eds.), *From Animals to Robots and Back: Reflections on Hard Problems in the Study of Cognition. A Collection in Honour of Aaron Sloman* (pp. 173–179). Springer. https://doi.org/10.1007/978-3-319-06614-1_13
- Lahlouh, M., & Catagnoli, A. (2006). Testi amministrativi di vario contenuto (Archivio L.2769: TM.75.G.4102-6050). *Archivi Reali di Ebla, Testi XII*. Università di Roma 'La Sapienza'. Liverani, M. (2006). *Uruk: The First City*. Equinox.
- Marriott, P., Edwards, J. R., & Mellett, H. (2002). *Introduction to Accounting*. SAGE.
- Maiocchi, M. (2016). Exploratory Analysis of Cuneiform Archives: A Network Approach to Ebla Texts. *Studi Micenei ed Egeo-anatolici Nuova Serie*, 2, 137–158.
- Matthiae, P. (1986). The Archives of the Royal Palace G of Ebla. Distribution and Arrangement of the Tablets According to the Archaeological Evidence. In K. R. Veenhof (Ed.), *Cuneiform Archives and Libraries. Papers Read at the 30e Rencontre Assyriologique Internationale, Leiden 4-8 July 1983* (pp. 53–71).
- Matthiae, P. (2008). Gli archivi reali di Ebla: La scoperta, i testi, il significato. Mondadori.
- Matthiae, P. (2020). *Ebla: Archaeology and History*. Routledge.

- Milano, L. (1990). Testi amministrativi: assegnazioni di prodotti alimentari (Archivio L.2712 - Parte I). Archivi Reali di Ebla, Testi IX. Università di Roma 'La Sapienza'.
- Milano, L., & Maiocchi, M. (n.d.). Ebla Digital Archives. Retrieved January 4, 2021, from <http://ebda.cnr.it/>
- Moretti, F. (2013). *Distant Reading*. Verso.
- Netting, R. McC., Wilk, R. R., & Arnould, E. J. (1984). Introduction: Notes on the History of Household Concept. In R. M. Netting, R. R. Wilk, & E. J. Arnould (Eds.), *Households: Comparative and Historical Studies of the Domestic Group* (pp. xiii–xxxviii). University of California.
- Oh'e, S. (1983). On the Function of the Maskim, I. *Acta Sumerologica Japan*, 5, 113–126.
- Pettinato, G. (1979). Ebla. Un impero inciso nell'argilla. Mondadori.
- Playfair, W. (1801). *The Commercial and Political Atlas: Representing, by Means of Stained Copper-plate Charts, the Progress of the Commerce, Revenues, Expenditure and Debts of England During the Whole of the Eighteenth Century*. Burton.
- Pomponio, F. (2003). La terminologia amministrativa di Ebla: šu-du8 e TUŠ.LÚ×TIL. In P. Marrassini (Ed.), *Semitic and Assyriological Studies Presented to Pelio Fronzaroli by Pupils and Colleagues* (pp. 540–559).
- Porter, A. (2012). *Mobile Pastoralism and the Formation of Near Eastern Civilizations*. Cambridge. <https://doi.org/10.1017/cbo9780511895012>
- Ragin, C. C. (1992). "Casing" and the process of social inquiry. In C. C. Ragin & H. S. Becker (Eds.), *What Is a Case? Exploring the Foundations of Social Inquiry* (pp. 217–226). Cambridge University Press.
- Sallaberger, W., & Schrakamp, I. (2015). Philological Data for a Historical Chronology of Mesopotamia in the 3rd Millennium. In W. Sallaberger, & I. Schrakamp (Eds.), *History and Philology* (pp. 3-136). *Associated Regional Chronologies for the Ancient Near East and the Eastern Mediterranean III*. Brepols.
- Scarpa, E. (2017). The City of Ebla. A Complete Bibliography of Its Archaeological and Textual Remains. Edizioni Ca' Foscari. <http://doi.org/10.14277/978-88-6969-158-4>
- Seising, R. (2012). Warren Weaver's "Science and complexity" Revisited. In R. Seising & V. Sanz González (Eds.), *Soft Computing in Humanities and Social Sciences* (pp. 55–87). Springer. https://doi.org/10.1007/978-3-642-24672-2_3
- Sollberger, E. (1986). Administrative Texts Chiefly Concerning Textiles (L.2752). Archivi Reali di Ebla, Testi VIII. Università di Roma 'La Sapienza'.
- Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to Data Mining*. Pearson Education.
- Tonietti, M. V. (1989). Le liste delle dam en: cronologia

- interna. Criteri ed elementi per una datazione relativa dei testi economici di Ebla. *Miscellanea Eblaitica*, 2, 79–115.
- Tonietti, M. V. (2013). *Aspetti del sistema preposizionale dell'eblaita*. Edizioni Ca' Foscari. <http://doi.org/10.14277/978-88-97735-17-5>
- Urry, J. (2005). The Complexity Turn. *Theory, Culture & Society*, 22, 1–14. <https://doi.org/10.1177/0263276405057188>
- Veenhof, K. R. (1986). Introduction. In K. Veenhof (Ed.), *Cuneiform Archives and Libraries. Papers read at the 30e Rencontre Assyriologique Internationale, Leiden, 4-8 July 1983* (pp.1–36). Nederlands Historisch-Archaeologisch Instituut te Istanbul.
- Veldhuis, N. (2014). *History of the Cuneiform Lexical Tradition*. Ugarit-Verlag.
- Waerzeggers, C. (2014). Social Network Analysis of Cuneiform Archives – A New Approach. In H. D. Baker & M. Jursa (Eds.), *Documentary Sources in Ancient Near Eastern and Greco-Roman Economic History: Methodology and Practice* (pp. 207–233). Oxbow. <https://www.jstor.org/stable/j.ctvh1dn9m>
- Wagner, A., Levavi, Y., Kedar, S., Abraham, K., Cohen, Y., & Zadok, R. (2013). Quantitative Social Network Analysis (SNA) and the Study of Cuneiform Archives: A Test-case based on the Murašû Archive. *Akkadica*, 134, 117–134.
- Weaver, W. (1948). Science and Complexity. *American Scientist*, 36, 536–544.
- Winter, I. (1991). Legitimation of Authority through Image and Legend: Seals Belonging to Officials in the Administrative Bureaucracy of the Ur III State. In M. Gibson & R. D. Biggs (Eds.), *The Organization of Power: Aspects of Bureaucracy in the Ancient Near East* (pp. 59–99). University of Chicago.
- Yoffee, N. (2004). *Myths of the Archaic State*. Cambridge. <https://doi.org/10.1007/s13398-014-0173-7.2>