Computational Archival Practice: Towards A Theory for Archival Engineering

Kenneth Thibodeau
Independent Researcher
Evergreen, CO, US
email: KThibodeau@Fordham.edu

Abstract—The value of computational archival science is realized only in the delivery of products and services. The ultimate value of archival science is its contribution to the construction of information about the past. Archival engineering offers a systematic basis for delivering value. The paper articulates concepts that can be melded with traditional archival theory to expand the applicable domain and to develop quantified, testable and verifiable archival methods.

Keywords— archival bond, archival edge, archival engineering, archival science; target past.

I. INTRODUCTION

The value of archival science is realized in the benefits it provides to the consumers of archival resources, both goods and services. The potential for realizing that value can be increased by supplementing archival science with archival engineering, a discipline that unfortunately does not yet exist. Archival engineering can be differentiated from archival science using Henry Petroski’s simple assertion, “Science is about knowing, engineering is about doing.” [1] Archival engineering can thus be defined as a discipline which systematically applies archival science to deliver optimal value in the provision of archival goods and services. Archival goods comprise both the materials preserved by archives and information about them. Archival services support the discovery and delivery of these goods and enable evaluation of their qualities, such as relevance, adequacy, authenticity and reliability as related to the interests of consumers.

Archival engineering offers the potential to extend the applicability of archival theory to new domains, beyond the traditional realm of records. The development of archival engineering must start from the basic assertion: what is the consumption of archival resources; what is done with them? In general terms, the use of archival resources increases and improves knowledge of the past. This is not to revert to the view that archival science is an ancillary discipline of history. It encompasses any effort to retrieve, organize, derive or generate information about the past, even the recent past, regardless of the purpose served by such efforts. They can range from simple acts, such as finding contact information for a party involved in an ongoing activity, through ascertaining prior states of activities, such as the negotiation and administration of contracts, or reconstructing prior activities, such as in financial audits and legal investigations, to very complex analyses, such as exploring the evolution of a national culture.

The second question that must be addressed is: what benefits can archival engineering offer for increasing and improving knowledge of the past? Answering this question entails (i) articulating what is involved in knowing the past; (2) exploring how archival science can be applied to contribute to the production and improvement of this knowledge; and (3) identifying how the contribution of archival science can be assessed. Each of these topics is addressed below.

However, given the limitations of this venue, this paper can do little more than suggest possibilities within each of the three topics. Priority is given to introducing and defining concepts that reformulate archival science in a manner appropriate for application in archival engineering.

II. THE PAST

The past never existed. The past is always something that is constructed by thinking, writing or speaking about former times. What is, is intended to be, or has been produced from such construction is called a ‘target past.’ A target past is an object, event, process, action, activity, person, agent or state of affairs, or some combination of these, that existed or occurred at a prior time. The elements included in this definition of target past are themselves defined as follows:

• An event is something that happened.
• A process is a set of related events.
• An action is an event in which at least one agent was actively involved.
• An agent is a person or an object, such as a system or device, that was involved in the performance of an action.
• An activity is a set of related actions; that is, an activity is a process in which all events were actions.
• A person is a human or legal entity.
• A state of affairs is a set whose members are one or more related objects or persons where one or more characteristics, which can include interrelationships, are invariant. The period during which the defining characteristics remain invariant determine the duration of the state of affairs, which can range from instantaneous to long term.
• An object is any of the above or anything else.

III. FRAMEWORKS FOR CONSTRUCTING TARGET PASTS

The range of activities that may be involved in the construction of a target past stretches from finding a piece of information to creating new information. The latter necessarily entails the former, gathering data that is then analyzed or synthesized to produce new insights or understanding. Both what is sought and what is done with what is found are shaped by two, basically independent frameworks. The first, which
may be called purview, derives from the purpose for which a target past is constructed, preconceptions about it and plans for its investigation. The second framework, which may be called historical context, defines an architecture of the target past, including both its structure and the materials that constitute it, based on objects and on relationships among them that are found and deemed relevant within the purview of the target past. The two frameworks are basically independent because (1) objects and relationships that existed in prior times are independent of any effort to discover or analyze them and (2) the purview of a target past is determined after the fact and may involved combining or comparing things that had no prior relationships. For example, disparate and independent historical sources have been used to characterize the ownership of property by women in ancient, patriarchal society. [2] Similarly, topics of target pasts, such as climate change or gender equality, may not have existed or been recognized in the periods being investigated.

While initially independent, the purview and historical context will necessarily merge during the course of construction.

A. Purview

The construction of a target past entails several preliminaries including the identification of the topic and the intent in which the topic is considered. For example, the topic of the historical growth of a forest could be investigated for an environmental analysis or for purposes of wildfire management, or to project timber yields. Each would result in a different construction. Intent refines the perception of the topic by defining the specific characteristics and relationships that are of interest and, thence, the data that are needed to accomplish the intent of the construction. Preparation for construction of a target past also includes identifying and evaluating potential sources of information and methods of data gathering and analysis. In effect, the purview defines a field of vision, which determines whether things are observed or not, as well as a spectrum, which colors how things are seen. Thus, in relation to the target past, purview is constructive, rather than constructed, because it originates with the intent to construct a target past and thus predates and guides the construction. Purview can change in the course of construction, especially in cases that include discovery of unexpected objects or relationships.

The domain of a target past can be defined as the set of objects within the purview of its construction, along with their characteristics and relationships that are of interest given the intent of the construction. The domain constitutes a constructed context. ‘Context’ is often defined in broad terms; for example, as “the interrelated conditions in which something exists or occurs.” [3] For archival engineering, a context can be defined as a set of coherent assertions which clarify a target past. Contextual clarification involves (1) a contextually situated object, abbreviated to contexted object; that is, is an object considered in relation to a set of circumstances that clarify its existence or characteristics, and (2) one or more contextual elements; i.e., things that clarify a contexted object. A constructed context can be either inferred or imposed. An inferred constructed context is one that is derived from the objects in the domain of the target past. A constructed context may be inferred based on information related to the target past available at the start of the activity or discovered in the process. A constructed context is imposed based on the intent of the construction. A constructed context might be imposed when a target past was not an outcome of a single process or a set of interrelated processes. For example, many major developments in science emerge from multiple activities that were independent of one another when they occurred. In such cases, the imposed constructed context is derived from the intention to determine the origins of such developments.

B. Historical Context

Historical context consists of contextual elements that existed at the same time as their contexted objects and were related at that time. Existence at the same time, or contemporaneity, is independent of the date when either the contexted or contextual objects were created. Contemporaneity simply requires some overlap or contiguity in the date span of the related objects. For example, a law may have been enacted centuries before an activity in which it was applied. Nevertheless, its application makes it contemporaneous with the activity. The relationship of a contextual element to a contexted object is independent of any construction of the target past. For example, a record kept in the course of an activity is part of the historical context of that activity regardless of whether it is used in a study of the activity.

Historical context can be represented by a graph where, for any pair of nodes connected by an edge, both nodes and the edge are contemporaneous; that is they either exist at the same time or in immediate succession.

Historical context can be either coherent or coincident. A coherent context is one whose elements are somehow part of the same whole as its contexted object. A special type of coherent context is linguistic context, as defined in systemic functional linguistics. Systemic functional linguistics defines text as written or oral language that is functional, that is, that does some job in some context. Linguistic context is language that is included in the same communication, but does not directly contribute to accomplishing the purpose of the text. Rather, linguistic context associates the text with the situation in which it occurs. [4]

A coincident contextual object is contemporaneous with its contexted object, but was not part of the same whole within the time span of the target past. A common example of coincident context consists of the different records that were kept by two independent agents interacting in an activity; for example, a government agency and a regulated business. While the records of the two agents are related to the same activity and undoubtedly include duplicate copies, they are part of separate archival fonds that are likely to be organized differently and include documents not found in both collections.

C. Total Context

The constructed and historical contexts together constitute the total context of a target past. While the constructed context should be comprehensive and coherent over the entire domain of the target past, historic context can be multifaceted. There can be different contexts for different objects or groups of
objects in the domain and, conversely, a single object may be found in multiple contexts defined by different criteria. Moreover, a single object may be the contexted object in one context and a contextual object in one or more other contexts. Fig. 1 displays a Venn diagram of the possible total context of a target past. The historic context is divided into the coherent and coincident contexts, which have no intersection. Either coherent or coincident historic context could contain subsets that constitute linguistic contexts. An imposed constructed context associates objects within historic context based on criteria established in the purview of the target past, suggested by the outward pointing arrows. Conversely, inferred constructed context is built from relationships found in the historic context, reflected in the inward pointing arrows. Both types of constructed context can intersect either type of historic context. The diagram does not include any intersection of inferred and imposed contexts, but that is a matter for further investigation.

IV. MATERIALS FOR THE CONSTRUCTION OF TARGET PASTS

The raw materials for construction of a target past are not things from the past, but tokens that represent those things, such as a photograph of a person, a video of an action, or a text describing an activity. Even some apparent relics of the past, such as physical artifacts and authentically preserved records, are actually tokens. For example, Fig. 2 shows a page from a book of depositions taken before the Consistory Court of the Diocese of London in the early 1680’s. [5] This digital image is clearly a token for the actual page, but even with the original complications arise from the simple question of what is it. The most obvious answer, that it is a deposition, is false because, in the juridical context, a deposition is the giving of sworn, oral testimony. The document is not the testimony but a transcript — i.e., a representation — of it. The transcript has different features than the original testimony, such as content that identifies the date, place and case in which the testimony was given. Moreover, a study of depositions from English courts in the early modern period has shown that scribes used various visual effects, such as white space, alignment, indentation, and size and embellishment of letters to differentiate elements of content in ways that cannot be presumed to have been features of the oral testimony. [6] Finally, even the original oral testimony was a representation of the matters discussed.

A similar analysis could be constructed for physical artifacts, such as archeological sites, antiques, and even extant buildings. In many cases, even dating artifacts requires external information about them, and understanding how they relate to a target past is likely to require additional outside information. In other words, the artifacts need to be placed in context. Once that happens, the artifact becomes a token of its contextualized self.

A token may be defined by its content, form, expression or a combination of these. Each of the defining facets is a variable, not necessarily a constant. For example, the expression of an internet posting may vary depending on the device used to present it. Analogously, elements of content in digital tokens may vary based on real time input, such as

1. Total Context of a Target Past

2. Deposition, Consistory Court, Diocese of London, 1681-1682.
temperature on a web site that provides weather information. However, in order for some piece of information to be considered as a token, it must have bounded variability. For example, the same content must be displayed in the same way on the same device, or even instance of a class of devices, every time it is retrieved, and a stored view on a database must yield the same data every time it is executed. [7]

The information that can be gleaned from token depends on the perspective in which it is viewed. A painting of a building, for example, could represent the building depicted or the work of the artist who produced it. Generally, the information conveyed by a token depends on the context in which it is considered. Moreover, the defining properties of a token may include its relationship(s) to one or more other tokens, as explained in section V.A, below.

V. ARCHIVAL ENGINEERING AND ARCHIVAL SCIENCE

Archival engineering offers the potential for improving the construction of target pasts and the evaluation of the results by building on concepts in archival science, expanding them to a broader scope, adapting them to encompass unprecedented aspects of digital information, facilitating automated processing, and enabling verification through quantitative testing. Synthesizing the theory of target pasts with concepts of archival theory enables a reformulation of archival science in support of these objectives.

A. Archival Concepts

The core object in archival theory is a ‘record.’ However, there are substantive problems with existing definitions and conceptualization of ‘record.’ The International Standards Organization defines ‘record’ as “Information created, received, and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business.” [8] This definition is overly narrow. The verbs, created and received, do not accurately indicate the many ways an agent may come to have custody or control of records. For example, a computer application may automatically collect data from extensive arrays of instruments that continuously monitor the environment, the planet or other phenomena and aggregate, transform and organize them in longitudinal data sets. The process would more accurately be described as production, rather than creation. ‘Receive’ connotes passivity, but records can be actively obtained; for example, by automated harvesting from the web. The need for generalization has been recognized by the archival community in Italy in its definition of archives, where “created and received” have been replaced with “produced or in some manner acquired” (“prodotti o comunque acquisiti”). [9] But even this expanded terminology is too narrow. Documents that are significant, even critically important in the conduct of business may be neither produced nor acquired in transactions in which they are used, because they are already in the possession of the agent. An extremely common example is that of directives, which influence all transactions in their scope of applicability, but are rarely duplicated and filed together with records of individual transactions.

Another assertion about records limits them to instruments and byproducts of activities, but excludes end products. [10] As in many treatments of the concept of ‘record,’ the rationale for this exclusion confuses a desirable characteristic of records with a defining property. The desirable characteristic is that records provide unbiased information about what was done or intended. More accurately, documents that were instrumental in accomplishing something, or that fell out in the course of accomplishing it have the same bias as the action itself. However, in many cases, instruments used to produce something or documents created as a consequence of its production are incomprehensible absent the end product. The account of what was done is inevitably incomplete without the product of the action.

Probably fruitless arguments about records can be avoided by substituting an alternative construct, that of archival token. An archival token is a type of conceptual object. In contrast to many definitions of ‘record’ that include aspects of materialization, [11] an archival token can have different instantiations. For example, an email could be displayed without significant difference in the message it conveys on a smart phone, a tablet or a large, desktop monitor. The content, form and expression of an archival token are either invariant or predictable based on specified bounded variability. Like records, archival objects are distinguished from other tokens not by inherent properties, such as content or form, but on the basis of their relationships to actions, agents and to other tokens that share the same archival context. Specifically, an archival token is a token that represents one or more objects from a former time in an advantageous manner because of its proximity to its referent both temporally and contextually. Proximity is specified in a known archival context. The components of archival context are an agent or agents who had, or had access to, the token in an activity and the activity in which agent or agents used or intended to use the token. Not all archival tokens are records. The definition of archival token includes the option that an agent may only have had access to the token, and neither produced nor acquired it. For example, an agent may have dependable access to trustworthy information made available on the internet by another party. The agent may see no need to duplicate such information in its records. Nevertheless, if the agent makes significant use of the information in its activity, the information can be as important for understanding that activity as anything kept as a record. The definition of archival token does not include quality criteria because concerns with quality can be addressed by articulating quality parameters, as suggested in section VI below, apart from the definition.

The concept of archival context derives from the idea of the archival bond. As articulated by Georgio Cencetti, the archival bond is a relationship among documents that arises from their use in an activity. [12] This relationship is ontologically prior to, and thus independent of, the keeping of documents as records. Thus, when one agent sends a message to another and the second agent responds, there is a relationship between the two messages regardless of whether either agent keeps them as records. An archival bond grows over time to include all tokens related by their use by an agent in an activity. Optimally, the archival bond is preserved and expressed in the way records are organized in a record keeping system; however, this will not be the case when records are organized
chronologically or on the basis of subject. Even in situations where the organization of records parallels the processes in which they were created, typically in case files, the arrangement of records is likely to express only a fraction of the archival bond. [13] Moreover, the organization of record keeping may be ordained by anticipated need for retrieving sets of tokens together, rather than reflecting prior use of the tokens.

With the concept of archival context, the archival bond is replaced by two constructs: archival link and archival graph. An archival link is the binary relationship of two archival tokens that arises from their use in the same activity. It corresponds exactly to the initial element of the archival bond as articulated by Cencetti, but Cencetti and archivists who have adopted his ideas do not distinguish between the binary relationship between two documents and the total set of relationships among all documents used in an activity, using the single noun, bond (vinculo), to designate both. For purposes of archival engineering, an archival graph is introduced to denote the total set of relationships among archival tokens used in the same activity. An archival graph constitutes a coherent historical context. The archival tokens are the nodes in the graphs and the edges their binary relationships. An archival link is a pair of nodes connected by an edge in an archival graph. In practice, the edges should be labeled to specify the nature of each relationship; e.g., initial message/response, draft/final, et al.

The conceptualization of the archival bond links it to record keeping, [14] but that is an unnecessary constraint not adopted in the case of either archival link or archival graph. [15] Provided they have not been corrupted, ontological relationships between archival tokens exist whether they are preserved in a record keeping system or anywhere else. If copies of two messages are kept as records by both correspondents, they have the same archival link in both record systems; however, the archival graphs in the two systems are likely to be substantially different, as illustrated in Fig. 3, which shows archival tokens used in a hypothetical interaction of two agents. Fig. 3 imagines that an initial agent identified and documented a need, in the needs assessment, that the second agent could satisfy. Consequently, the initial agent composed and sent an initial message to the second agent. On receipt of this message, the second agent invoked its policy for responding to such requests and consulted data it had about relevant available resources. On this basis, the second agent composed its response. In the figure, the large ovals represent the archival fonds of the two agents while the smaller circles represent archival tokens kept as records by each agent. Two tokens connected by a line constitute an archival link and the set of all linked tokens constitutes the archival graph of the interaction. The archival graph can be divided into two subgraphs, each containing only the archival tokens kept as records by one of the agents. Thus, the subgraph of the initial agent includes only three nodes: the needs assessment, initial message, and response, while the subgraph of the second agent contains four nodes: the two messages, the policy document and the resource data. Note that the figure does not indicate that different copies of the messages are kept in the two archival fonds because as a conceptual entity each message is a single archival token, although it is instantiated in two different record keeping systems. Figure 3 illustrates that the concept of archival graph enables the identification of coherent historic contexts that extend beyond the boundaries of an archival fonds. While the figure illustrates a very simple case, the approach can be extended to articulate archival graphs that encompass all archival tokens used in arbitrarily complex activities involving numerous agents. When the target past includes an activity that involved multiple agents, the similarities and differences in the subgraphs of records kept by the various agents can be even more informative than what is revealed by the records of a single agent.

Note further that the subgraphs in Fig. 3 are unlikely to be replicated in structure of the record keeping system of either agent. In the case of the second agent, and in general, it is very unlikely that policy documents and resource data would be filed together with correspondence related to individual transactions conducted in accordance with guidance provided by the relevant policy. In the case of the first agent, the three archival tokens might be filed together if and only if the needs assessment identified a single need or one set of needs that could be addressed by interacting with one and only one other agent. However, if there were needs that required interacting with several other agents independently, it is likely that correspondence with each of the other agents would be filed separately. Nonetheless, the existence and placement of an archival token in a record keeping system provides coherent contextual information, which can supplement or be independent of the archival graph. A record keeping system, then, is a distinct coherent historic context. The act of setting aside a token for possible subsequent use reveals that the agent/record creator deemed the token as having enduring value and its placement within a system indicates what the agent saw as the most important aspect, or aspects in the case of cross referencing or of multiple simultaneous arrangements in digital systems, of its relationship to the agent’s needs or intentions. It may also happen that a token that has not been used by the agent in any way is kept as a record because the agent anticipates possible use, and this anticipation may be imprecise. Nonetheless, the keeping and placement in a record keeping system even in such situations have contextual value.

A record keeping system could be described by a graph. While records are archival tokens and they may be organized in
a way that reflects prior use, in the graph of a record keeping system the edges are defined not my use, but by placement within the order of the system.

While there will often be numerous commonalities between the record keeping system and the archival graph of a given agent, the two constitute distinct graphs of coherent historic context. Beyond the qualitative considerations about filing mentioned above, the two types of graph are likely to differ quantitatively because record keeping systems typically have a tree structure whereas the structure of an archival graph develops ad hoc.

Provided archival tokens are properly preserved, the archival context independently conveys significant contextual information that can be far greater than that derived from the record keeping context. For example, an archival link that originated in one activity could be part of the archival graph of another agent acting in a different sphere of activity. For example, if a party that harvested tokens from a social media platform kept them as records, the archival graph could be drastically different than those of the correspondents who used the social media. As illustrated in Fig. 4, minimally, the harvester would have a document that defined the criteria for harvesting messages and every message collected would have a direct archival edge linking it to that document. Similarly, if the harvester performed an analysis of the collection, each message would have an archival edge incident on the analysis. But what of the message/response links that might be found between pairs of messages in the collection? If the harvester anonymized the messages, those archival links would vanish. But if the messages were retained intact, the links would remain extant, valid and available. However, these edges would not be part of the archival graph of the harvester’s collection unless the criteria for harvesting included collecting message threads. Otherwise, the message/response links would be coincident, not coherent. The message/response link in Fig. 4 is shown with a question mark because of this uncertainty.

This potential for linking relevant tokens can be further extended using the Records in Contexts Conceptual Model proposed by the International Council on Archives. That model provides for defining supergraphs that link archival graphs to elements of coincident historic context, such as biographical or historical information about agents. [16] [17] Another, established method that could be used to supplement archival graphs with coincident contextual information is that of scholarly editions. [18] Obviously, analysis of the coherent context represented by an archival graph could go below the level of archival tokens by using tools of systemic functional linguistics. But concepts from that discipline, such as register, intertextuality, context of situation and contextual configuration, might also be used to identify a different type of edge based on linguistic analysis. In this approach, a text would still be defined a language doing some job, but it would be instantiated across the set of tokens used to complete an activity.

A record keeping system is a distinct coherent historical context that can significantly add to the contextual information available for construction of a target past. Well grounded archival concepts, as traditionally articulated, can be applied in archival engineering by adding provisions to their definitions that facilitate application in automated systems and support quantitative measures of success. The provisions that follow suggest ways to implement and analyze traditional archival concepts within the framework of graph theory.

- **Archival fonds**: the supergraph of the archival bonds of a records creator.
- **Records creator**: an agent or a set of interacting agents who keeps records related to a sphere of activity in which it is active.

The provision includes sets of agents in order to encompass funds that are created in collaborative environments where no single agent is dominant, such as on social media and via internet collaboration tools.

- **Archival bond**: a graph whose vertices are all the records of a records creator related by their use in a single activity and whose edges are the relationships that are expressed in the arrangement of records within a record keeping system.
- **Record**: a persistent token that is used, or intended for use, and kept by a records creator.

Note that persistence here does not entail invariability, no more than a person is invariant over the course of his life. Rather, persistent entails bounded variability as explained above. In interactive computer systems or applications, bounded variability is satisfied if the same inputs produce predictable outputs

- **Record keeping system**: a totally ordered graph of all the records kept by a records creator in accordance with a specific set of organizing principles.

VI. **Evaluation**

Albeit in cursory fashion, we have seen that archival science and the theory of target pasts provide a basis for an engineering discipline using quantified expressions and adaptations of traditional archival concepts. Reformulating archival concepts in graph theoretical terms opens the door to numerous possibilities for quantification. However, it leaves
outstanding the question of how the application of the new and reformulated concepts in archival engineering can be evaluated.

The following criteria are offered as candidates for evaluating how, how much, or how well archival engineering applications improve the construction of target pasts. For each concept, a definition is proposed; a quantitative measure is discussed, and how use of the criterion might improve construction of target pasts is described.

A. Objectivity

Objectivity can be defined as the degree of correspondence between a token and its referent; i.e., the object or objects it informs about.

Objectivity might be quantified on an ordinal scale where zero indicates no correspondence and the highest possible value means complete congruity. Zero would be assigned to a token that is completely unrelated to any aspect of a target past. Intuitively, it might seem impossible to ever determine complete congruity, except for the trivial case of self-referential assertions. Nevertheless, high scores should be assigned to an artifact that is a token of its contextualized self, as described above. The scale for objectivity could be extended below zero to encompass tokens that give erroneous information with the lowest scores assigned to forgeries. Archival concepts such as reliability in the production of a token and the authenticity of its preservation can be applied in assessing objectivity.

Assessing objectivity can benefit the identification and prioritization of potential sources of information in formulation of the purview of a target past and the initiation of its construction. It can also guide inferences made in the conduct of research. Consumers of constructed target pasts could apply the criteria to assess the trustworthiness of a construction.

B. Translucency

Translucency is the possibility of discerning the biases in a token or set of tokens, specifically recognizing how the selection, organization and expression of content in tokens colors what they communicate. Translucency can be seen as the obverse of objectivity because objectivity, in the usual sense of the term, is negatively affected by factors such as personal feelings, interests, interpretations, prejudice or extrinsic intent, that are not coherent with the target past or with one or more of its elements. However, given that archival tokens relate to actions and actions can be presumed to be construed according to the interests or objectives of the agents involved, recognizing the coloration of a token or set of tokens by the agents who produced, acquired or used them may enhance and even be essential to understanding the token within its historical context.

Three possible measures of translucency are (1) the number of different biases identified, (2) the number of occurrences of each type of bias in a token or set of tokens, and (3) the extent to which any bias is counterbalanced. Those measures, however, are crude. Meaningful estimates require even more complex formulations that might, for example, relate instances of bias to independently identified objectives or ulterior motives. An alternative approach might be based on considering translucency as an inferred context and defining indicators of how fully or how precisely the biases of agents are identified.

While the quantification of translucency is difficult, there can be no doubt that elucidating and estimating the biases of agents involved in producing, acquiring, keeping, and organizing relevant tokens would add depth to the construction of target pasts.

C. Richness

While objectivity compares a token to its referent, richness maps a target past to the set of tokens available for or used in its construction. Evaluation of the richness of a set of tokens would build on the objectivity ratings of individual assets and supplement those ratings by identifying the subset of assets relevant to each element or facet of the target past.

Richness might be quantified on an ordinal scale where zero indicates there are no relevant tokens and the highest possible value means the tokens are (at least) completely adequate for the target past. The overall richness rating would be a composite of the ratings for the elements of the target. The top score could be attained in cases where the target past is relatively narrow and simple; for example when the target past is a bilateral currency exchange rate from a prior date, the certified data released by the Board of Governors of the U.S. Federal Reserve System provides all the data needed [19]. Richness is a function that can vary over time. When a target past is initially defined, little or nothing may be known about relevant tokens but the score is likely to increase as the construction proceeds.

Archival engineering can improve construction of target pasts by increasing the richness of appropriate tokens available. Archival engineering can contribute to richness by improving preservation of tokens, increasing dissemination of information about them and facilitating access.

VII. Conclusion

There are rich possibilities for developing archival engineering to enable systematic application of computational archival engineering to increase the realized value of archival science.

REFERENCES


17. Anila Angjeli, Florence Clavaud and Stephanie Roussel. “From Several Datasets to One Graph: designing a prototype for a graph visualization of archival metadata from different institutions.” Society of American Archivists Annual Conference. 2016. https://www.researchgate.net/publication/307173466_From_Several_Datasets_to_One_Graph_designing_a_prototype_for_a_graph_visualization_of_archival_metadata_from_different_institutions
