The Form of the Content: The Digital Archive Nahuatl/Nawat in Central America

Laura Matthew
Michael Bannister

Follow this and additional works at: https://epublications.marquette.edu/hist_fac

Part of the History Commons
DHQ: Digital Humanities Quarterly

2020
Volume 14 Number 4

The Form of the Content: The Digital Archive Nahuatl/Nawat in Central America

Laura Matthew  <laura_dot_matthew_at_marquette_dot_edu>, Marquette University
Michael Bannister  <bannisms_at_gmail_dot_com>, Independent Programmer Analyst

Abstract

The digital archive “Nahuatl/Nawat in Central America” (NECA) assembles and makes publicly available a growing corpus of Nahuan-language documents produced in Spanish Central America. Many are fragments within larger Spanish-language documents and difficult to locate in the archive. NECA has succeeded in bringing attention to this understudied corpus but has so far failed to attract users to its transcription and translation tool. We consider the reasons for this creative failure based on user data, and suggest that the specialized skills and distinct academic communities needed to move this project forward require other workspaces, including the non-digital, in advance of online collaboration.[1]

Some thirty years ago, in The Content of the Form, Hayden White reminded his fellow historians of the extent to which history's content is dictated by the form of its presentation. Annals, chronicles, biographies, narrative, and discursive analyses all entail “ontological and epistemic choices with distinct ideological and even specifically political implications” [White 1987, ix]. Here, we adapt White’s title to make a similar point about the digital archive El Náhuatl/Náhuat en Centroamérica or in English, Nahuatl/Nawat in Central America (NECA; http://nahuatl-nawat.org).[1]

Whereas White focused on how content may be influenced by its form independent of the creator's intent, here we examine the curatorial decisions we made regarding NECA's form in order to intentionally impact the reception, use, and utility of its content.

NECA assembles a corpus of handwritten, colonial-era texts produced in Central America in variations of the related Mesoamerican languages Nahuatl and Nawat, from eight repositories in Guatemala, Mexico, Spain, and the United States. It emphasizes the fact that these oft-ignored documents exist, and encourages their collaborative study across national, scholarly, community, and disciplinary lines. Neither goal is neutral or apolitical, although the significance of studying these texts may vary depending on whether the user is an Indigenous rights activist from Mexico City or Los Angeles, a linguist of Mayan languages from Guatemala, a native speaker from Guerrero, a primary school teacher from El Salvador, or a doctoral candidate from Europe, etc.

In this essay we explain our rationale for creating a digital archive of Nahuatl texts from Central America in the first place, arguing that NECA’s content should be studied not only by individuals analyzing particular texts for the purposes of geographically or disciplinarily bounded research and revitalization projects, but also collaboratively and more experimentally as a standalone corpus. We then review the ontological and epistemic as well as technical choices we made in the project's design to encourage this outcome. NECA's form attempts to prod users towards a variety of actions both within and outside the digital archive. The success or failure of the affordances we created to increase the usefulness and usability of the site, and thus to direct the user toward specific activities, can be measured in the site's analytics. These indicate not just where the digital environment we created is working well or can be improved, but also where it may not be the best workspace available — or at least, not yet.

The Content: Why Nahuatl in Central America?

Nahuatl, best known as the language of the Aztec empire, was spoken by tens of millions of people in the early sixteenth century. It is not a single language but a range of mutually intelligible "Nahuan" variants ranging from northern Mexico to Nicaragua since at least the second half of the first millennium A.D. (see Figures 1 and 2). Many Nahuan languages have died out, especially in the last 150 years. Others persist but are threatened by continued and increasing contact with and preference for European languages such as Spanish and English. Today, there are approximately 1.5 million native speakers of Nahuatl in Mexico and the United States diaspora, and around 200 native speakers of the related language Nawat in the Izalcos and Santo Domingo de Guzmán areas of Sonsonate and in Tacuba, Ahuachapán, both in western El Salvador (http://www.unesco.org/languages-atlas/index.php).[2] Nahuan languages in Guatemala, Belize, Honduras, and Nicaragua have largely ceased to exist.
When the Spanish arrived in 1519, central Mexico was the most urbanized, politically powerful, and densely populated part of Mesoamerica. The Spanish made the defeated Aztec capital of Tenochtitlan the bureaucratic heart of their own nascent empire, and engaged Indigenous intellectuals in a remarkable, sometimes violent merging of Mesoamerican and European writing systems [McDonough 2014] [Townsend 2016]. This produced a significant amount and variety of Nahuatl written in Roman script that has been studied extensively, for centuries in Mexico and more recently in the United States and Europe.

This large corpus of Nahuatl documentation from central Mexico has spawned a number of digital projects with a variety of aims, such as increasing access to lesser-known texts and making databases of glyphic and linguistic information searchable online for comparative study. For instance, the Compendio Enciclopédico Náhuatl (http://cen.iib.unam.mx/) links linguistic data from approximately twenty historical and modern
Nahuatl dictionaries with separate databases of information from pictorial and alphabetic texts. The *Nahuatl Dictionary of the Wired Humanities Project* at the University of Oregon (https://nahuatl.uoregon.edu) allows users to search for attestations, headwords, and themes associated with any string of letters in English, Spanish, or Nahuatl, in order to compare usages in early modern Nahuatl from central Mexico as well as contemporary Nahuatl from the Huasteca, Veracruz. Also from Oregon, the Early Nahuatl Library and the Mapas Project (https://enl.uoregon.edu) make available images, transcriptions, and English translations of around 100 Nahuatl texts with annotations from a variety of archival and published sources. Axolotl (https://axolotl-corpus.mx) similarly depends on the published and unpublished work of established scholars to cross-reference approximately 30 colonial-era books in Spanish-Nahuatl translation.

Significant colonial-era Nahuatl language documentation also exists from outlying regions of the former Aztec empire. Like the Aztecs, the Spanish used central Mexican Nahuatl as an imperial *lingua franca* [Dakin and Lutz 1996] [Herrera 2003] [Gasco 2017] [Herranz 2001]. *Nahuatl* — native and non-native speakers of the language whose speakers were — constituted a crucial link in the chain of translation from other Mesoamerican languages to Nahuatl to Spanish or Latin and vice versa, making them key actors in diplomacy, Catholic evangelization, and the application of Spanish law. Aztec practices assimilated by Nahuatl speakers at the edges of uncontacted territory may have been the groundwork for Nahua-Spanish invasion and colonization of independent regions such as Michoacán, Oaxaca, the Yucatán, and Central America [Carrasco 1999] [Navarrete 1996] [Voorhies and Gasco 2004]. In the United States in the 1990s, a historical methodology called the New Philology began to analyze records of Spanish bureaucracy written in Nahuatl not only in central Mexico, but also in regions where it acted as a second language of translation [Restall 1997] [Restall 2003] [Terraciano 2001] [Christensen 2013].

In Central America, Nahuatl's usefulness as a tool of empire was augmented by its mutual intelligibility with Nawat and other Eastern Peripheral Nahuan languages natively spoken in today's Chiapas (Mexico), southwestern Guatemala, and El Salvador (Arúz 1960) [Rivas 1969] [Campbell 1985] [Fowler 1989] [Reyes García 1961] [Navarrete 1975] [Knab 1980] [Gasco 2016]. Comparatively little attention, however, has been paid to Central American documents in colonial-era Nahuatl languages. This is partially due to an apparent lack of material. Such appearances, however, are deceiving. The largest repositories of colonial-era documents from Central America outside of Spain are located in Chiapas and Guatemala, which have both of which have significant Mayan populations. Mayan language documents from these regions are therefore highly valued, highlighted in archival catalogs, and may even be removed from their original context to become standalone documents. [By contrast, documents in Nahuatl languages are fragmentary, rarely noted as such, and often remain hidden inside bundles of Spanish-language legal papers. Historians of Spanish Guatemala typically rely on scribal Spanish translations of Nahuatl language text, while Maya linguists and language revitalization activists tend to view historical writing in Nahuatl languages as a colonial-era imposition that has little to offer their project of fortifying Mayan languages for future generations and recovering Mayan historical and sacred texts.

In neighboring El Salvador, by contrast, Nawat — the only surviving natively spoken Nahuatl language in Central America — is simultaneously valued as a part of the national patrimony and discriminated against in everyday life. In 1932, Salvadoran state forces massacred tens of thousands of peasants, most of whom Nawat speakers are, in response to an uprising against coffee plantations. Fearful of further repression, survivors avoided speaking Nawat in public or teaching it to their children [Lindo-Fuentes et al. 2007] [Gould and Laura-Santiago 2008]. This generational trauma, combined with deep-seated social and economic prejudices against indigeneity and a heavy emphasis on Spanish in the education system, has brought Nawat in El Salvador to a critical point of endangerment in the twenty-first century. Research on historical Nawat has therefore taken a back seat to the urgent task of recording and teaching modern Nawat [Lemus 2004] [Lara Martínez 2015]. In Nicaragua and Honduras, where Nahuatl languages are no longer spoken, Nahua heritage is also nationally valorized but historically hazier and thus far, not well documented [Bonta 2009] [Lara Martínez and Maccallister 2014] [Brinton 1883] [McCafferty 2015].

For all these diverse and contradictory reasons, few Central Americans have studied historical documents in Nahuatl languages from their own region (although this is beginning to change; see Romero 2017, Cossich 2012). Indeed, it has long been assumed that hardly any such documentation existed. The most basic goal of NECA is to correct this false impression. Our central claim, however, is not merely that these documents exist, but that they are worth studying.

Linguistically, Central American documents in Nahuatl languages bring an entirely new data set to debates about the historical evolution of Nahuatl languages, especially in areas beyond the imperial center. Linguists generally agree on the basic dialectal features of the two main branches of Nahuat, Eastern and Western, and of the urban, imperial Nahuatl developed in fifteenth- and sixteenth-century Mexico-Tenochtitlan [Dakin and Canger 1985] [Canger 1988] [Canger 2011] [Hansen 2014]. How Nahuatl languages from Central America fit into these typologies is less settled. Lyle Campbell (1985) viewed the central Mexican characteristics of Nahuatl-language colonial-era documents from Guatemala as the product of contact with the central Mexican allies of the Spanish. Karen Dakin's broader analysis of 20 letters in Nahuatl from sixteenth-century Santiago de Guatemala [Dakin and Lutz 1996] and 14 other documents mostly from Chiapas [Dakin 2009] [Dakin 2010a] led her to posit an *archaic* Nahuatl language that predated and continued to be used in Central America alongside the Aztec/Spanish *koiné*. Dakin considers this a unique southern Postclassic *lingua franca* quite distinct from the Aztec *koiné*, linking it to pan-Mesoamerican Zuyuan ideology [López Austin and López Luján 2000] and possibly earlier Nahua-Maya interactions [Dakin 2010b]. Sergio Romero (2014) sees the same texts as evidence of local, precolumbian Nahuan vernaculars. NECA makes possible significant advances in these linguistic debates, by more than doubling the number of identified sources and making high quality images of them accessible online.

NECA is also notable for its range of dates and genres: catechisms, wills, letters to Spanish officials, town council memos, bills of sale, community annals, tributary rolls, judicial testimony and denunciations, land titles, musical manuscripts, and confreterry books from the mid-sixteenth to the early eighteenth century. Religious texts in Indigenous languages are a foundational genre in Mesoamerican studies, and have been analyzed for the cadences of Mesoamerican ceremonial speech as well as the intense and sometimes antagonistic back-and-forth between European and Indigenous intellectuals [Burkhart 2011] [Sell et al. 2008] [Sparks et al. 2017] [Doebeg and Swanton 2008]. Grammars, vocabularies, catechisms, and other Mesoamerican language texts produced by Catholic friars also provide valuable linguistic information, sometimes unwittingly. The clerical author of the late seventeenth-century Guatemalan sermon *Teodamachilitz in yuilliz auh in ymiquiliz Tutemaguzicatizim Lusu Christo* now held at the John Carter Library at Brown University in the United States, for instance, noted the existence of a vehicular or "vulgar" Nahuatl
used alongside Nawat and the central Mexican koine in Guatemala. The cleric aspired to write his sermon in the "vulgar" dialect but frequently slipped back into the central Mexican variety with which he was more familiar [Madajczak and Hansen 2016] [Romero 2014].

Bureaucratic documentation generated mostly by Indigenous nahuatlotos, conversely, tends to imitate the prestigious central Mexican koine and to adopt Spanish legal formulae, but also employs less standardized orthography that reflects local speech patterns and the decreasing influence over time of the Catholic church on translation norms [Menz 2009] [Lockhart 1991] [Pizzigoni 2007] [Olko and Sullivan 2013]. Historians have used such bureaucratic and legal documentation to track political, sociocultural and linguistic changes in Mesoamerica as a result of European colonialism, and to uncover regional and subregional variations of the language. They have done so by systematically assembling, transcribing, translating, cataloguing the characteristics of, and comparatively analyzing various corpora of Nahuatl documents. This methodology holds great promise for NECA. With transcriptions and translations — to date, an aspirational goal — we would be able to create a database of dialectal and other linguistic features, locations, genres, scribes' names, year of creation, etc., which would surely yield new insights into the history of Nahuatl's diffusion, scribal and ecclesiastical networks, relationship to geography, and other avenues of future research.

Beyond philology, translations and transcriptions of the documents assembled by NECA would enrich the social history of the region. The vast majority of lives revealed are of non-native speakers of Nahuan languages: African urbanites, Oaxacan plantation workers, Maya choir masters and cofradia officials, French merchants, and innumerable Indigenous political leaders: Mam, 'Kiche', Tzeltal, Tojolabal, Jakalteko, Kaqchikel, etc. Contact points between friars, Spanish administrators, and local authorities are also plentiful in these documents. Family relations simmer underneath accusations of adultery, bigamy, and incest. Inventories and wills track the material culture of everyday life and the globalization of Mesoamerican commerce. Witchcraft, land and inheritance disputes, and the forced labor of women all make an appearance. The input of scholars and community members who may not have Nahuan language skills but who bring deep expertise in Mayan and Central American history, anthropology, archaeology, geography, and art history is crucial for contextualizing such information and incorporating it into larger narratives.[3]

To our assertion of NECA's potential for advancing Nahuan linguistics and Central American history, we add the possibility of supporting Nawat revitalization efforts in El Salvador. Diverse and overlapping intercultural and intergenerational campaigns have been underway in that country since the early 2000s, including a "language nest" primary school immersion program [Lemus 2018], university classes in Nawat as a second language (http://www.uca.edu.sv/escuela-de-idiomas/cursos-nahuat), regional initiatives such as Tushik (http://tushik.org/) and the Colectivo Tzunhejekat (https://www.facebook.com/Tzunhejekat), and social media hubs (https://www.facebook.com/groups/33974937500/). Increasing native speakers' access to historical documents written in Nahuan languages has proven valuable in other revitalization and decolonization efforts, from the workshop-and-publication model in the Polish-Mexican project Revitalizing Indigenous Languages [Olko and Sullivan 2014] to the Ticha digital archive of historical Zapotec documents from Oaxaca discussed by Broadwell et al. in this special issue (https://ticha.haverford.edu/en).

Preliminary discussions with Salvadorans involved in Nawat revitalization indicate that while there may be a place for NECA in the future, for now the urgency of recording and promoting modern Nawat overshadows interest in historical documents. How NECA might contribute to Nawat revitalization is uncertain, in part, because the linguistic identification of so many of our documents remains unclear and the majority are from Guatemala, where Nawat was historically spoken but is no longer. Again, further study via transcriptions and translations is needed in order to clarify how the NECA corpus may speak to the case of Salvadoran Nawat. In the meantime, we hope that NECA's expression of international scholarly interest in Central American Nahuan languages, free access to downloadable, high-quality images of colonial-era documents for anyone with an internet connection, and public witness to the long history of Nawat in El Salvador stands as a one more "symbol of cultural identity and pride ... [the] first step in any language revitalization process" [Lemus 2008, 8].

The Form: Going Digital

NECA began with a list of over 40 documents compiled by Sergio Romero (University of Texas at Austin) and Laura Matthew (Marquette University), in collaboration with a dozen other colleagues, for an encyclopedia project that never materialized. As Romero and Matthew sought alternate ways to publish the list, new items continued to surface. It became clear that given the number of Nahuan language documents that go unrecorded in archive catalogs and the extent to which scholars tend to run across them unexpectedly, the list could easily grow longer and a traditional print publication would quickly become outdated. Simply posting the list online might stimulate interest, but the need to travel to physical archives represented a significant barrier to serious engagement since those with the most capacity to read early modern manuscripts in Nahuan languages tend not to live or work in Guatemala and Chiapas, where the main repositories of NECA's documents are located. Working with programmer Michael Bannister, and with permission from the original repositories, Matthew decided in 2015 to create a digital archive of high-quality images using Omeka, the popular open-source content management system for digital collections from the Roy Rosenzweig Center for History and New Media (CHNM) at George Mason University. For the remainder of this essay, "we" refers to Matthew and Bannister as the sole creators and curators of NECA.

Our first curatorial decision was conceptual: to restrict the archive's geographical range to Central America as defined by colonial-era administrative boundaries. This meant including documents from Chiapas but not from neighboring and similarly multilingual places like Oaxaca, where Nahuatl also functioned as a vehicular lingua franca [Terraciano 2001] [Swanton 2008]. Segregating Oaxaca from Central America seemed in some ways artificial and over-determined by the same national, academic, and disciplinary boundaries NECA aspires to overcome. But such boundaries are both real and significant. At a practical level, we accumulated items from Chiapas but not Oaxaca by default, because Chiapas's colonial records were sent to the judicial Audiencia of Guatemala while Oaxaca's were sent to Mexico. Linguistically speaking, a Central American focus also directed attention to the contact points between Mayan and Nahuan languages. We did not want the Central American material to be prematurely absorbed into the considerably more developed academic literature on Nahuatl in Oaxaca and elsewhere in Mexico, without a proper understanding of the local contexts that produced it.

We also took seriously Justyna Olko's and John Sullivan's assertion that "more research on this topic [of local and regional differences and their relation to standardization] is greatly needed; especially useful would be a systematic comparison between regions as well as between higher and
lower-ranking scribes/authors within a given locality” [Olko and Sullivan 2013, 192]. A distinctly Central American corpus creates the possibility of comparative study with data sets from other multilingual, borderland, and outlying regions where Nahua was and is spoken, such as Oaxaca, Jalisco, Veracruz, and Guerrero [Canger 2017] [Olko and Sullivan 2014] [Guion et al. 2010] [Yañez Rosales 2017]. Finally, by drawing a line around Central America we hoped to direct attention towards and raise awareness of the ongoing, severely underfunded, but multi-pronged efforts to revitalize Nawat in El Salvador.

As we began to build the site, created and solicited feedback from an advisory board, and presented at conferences in the United States, Guatemala, and El Salvador, overlapping and mismatched interests in the NECA corpus became increasingly apparent. Historians, anthropologists, and archaeologists working in Central America were enthusiastic about sharing their archival references and interested in the information the documents contained, which they often could not read. Linguists and philologists working primarily in Mexico were interested in the dialectal features of the documents but were unfamiliar with their Central American context and history. Scholars and activists working on Nahuan languages in Central America expressed interest but lacked the financial and human resources to engage NECA without diverting valuable attention from existing projects, especially those supporting revitalization of Nawat in El Salvador.

We began to think about how NECA’s structure could more actively facilitate communication across these disciplinary, regional, and national borders. Unlocking the information inside the documents would be the essential first step for any kind of macro-analysis of the entire corpus, computational or otherwise, and for connecting scholars with similar interests and complementary skills. Could we help scholars find not just the documents, but each other? Could we create an online workspace that encouraged scholars to share their expertise and begin to generate data for comparative and collaborative analysis?

Taking inspiration from crowdsourcing projects such as Colored Conventions (which has since retired this feature) ([https://web.archive.org/web/20150322130256/http://coloredconventions.org] and DIY History ([https://diyhistory.lib.uiowa.edu]), we added the transcription plugin Scripto, and created an “Add a Document” feature using a Simple Contact Form plugin to encourage contributions of new documents. A separate, linked Wordpress site ([https://nahuatl-nawat.wordpress.com]) became the project blog and discussion space.

The backbone of Omeka is the items list, supported by Dublin Core-based metadata. Most metadata elements are obvious: date, title, source, etc. Nevertheless, each element reflects a curatorial decision made by us with certain goals in mind. We added new metadata elements for the number of “folios” to emphasize the variety and extent of the corpus, and for at-a-glance decisions by users about whether or not to transcribe; “sample text” to spark the potential transcriber’s and/or translator’s interest; “location” with the modern countries, states, and/or departments in addition to the colonial-era information to allow for sub-regional searches and future experiments in mapping; “date of creation” of the item itself to keep a record of the corpus’s growth; and the “contributor” of the document in order to acknowledge her or his research and participation.

Metadata omissions also reveal the synergy between form, content, and curation. A primary goal of NECA is to encourage the linguistic study of a larger corpus of Nahuan documents from Central America than usual, and eventually to gather the results in a database of linguistic features for comparative analysis. Some of our documents conform to a single, clear Nahuan variant. Most, however, present a mix of attributes, as one might expect of writing produced by non-native speakers in a context of ongoing (or decreasing) standardization, colonial power dynamics, and the adoption by Indigenous people of foreign writing technologies. This linguistic heterogeneity makes the NECA corpus an exceedingly valuable resource for exploring the history of Nahuan languages at linguistic borderslands [Madajczak and Hansen 2016, 239]. We chose not to create a metadata element that prematurely assigned the documents a reductive linguistic label until we have more data through transcription and analysis. We also wanted to avoid a situation in which non-linguists might interpret such labels as more definitive than they really are.

Decisions about the items themselves predetermine what researchers can and cannot do with them. Most of NECA’s items are fragments within larger documents — sometimes, much larger. On a mostly non-existent budget, we faced issues of server space, labor, and funding: photographers require payment, repositories may charge publication fees. Additionally, in this first iteration of the project we were focused on access and translation. We therefore chose to publish only the Nahua1l portions of any given document, for both practical reasons and in order to attract Nahua1l translators. This decision has consequences. For better or worse, it denies the user access to any Spanish translation that might have appeared in the original document. It also separates the fragment from its larger documentary context, digitally replicating the same de-contextualization that has been suffered by many Mayan-language documents. A fuller understanding of the document’s creation and information can only be achieved by consulting the original document in relation to its archival context. Data sets of people, places, and other kinds of information contained in the digital archive — for instance, paying attention to geographical location or scribal networks — will also remain incomplete without access to the full original. Researchers will have to return to the physical archives in order to get the whole picture, and we run the danger that they will not [Putnam 2016].

Finally, anticipating the user experience led to some programming alterations. Omeka’s automatically generated citations omitted the original archive; we changed the code to cite the document’s physical repository and archival signature first, followed by NECA and the date of access. To guide users towards specific activities, we turned Omeka’s “featured items” into a “sample transcription” and “featured collections” into “document teams.” Omeka’s built-in internationalization combined with the plugin Locale Switcher made the site bilingual, allowing users to choose in real time whether to view the site in Spanish or English. Because we had significantly altered the standard Omeka framework with new navigation headings, metadata categories, etc., Spanish versions had to be added to the internationalization code, as did all Spanish translations of all the text within the transcription tool Scripto. However, these changes affected only the user interface, not the items’ metadata. Assuming that most of our users would be competent in Spanish but not necessarily in English, we decided to make Spanish the primary language of the site (and in doing so, officially baptized it as NECA: in Spanish, El Nahuatl/Nawat en Centroamérica). All metadata is in Spanish regardless of the interface language, and simple pages unaffected by the plugin privilege Spanish at the top with anchors to an English translation below.[7]
At every structural opportunity we emphasized the collaborative, open nature of the project and minimized our own gatekeeping. Conversations during beta testing between anthropologist Janine Gasco and historians Julia Madajczak and Agnieszka Brylak inspired us to create mechanisms for interdisciplinary document teams to work on single items. Contributors of new citations are individually added to the “About Us” page as well as to their items’ metadata. Transcribers and translators are encouraged to register for Scripto with their full name so they can be properly identified in the versioning of transcriptions and translations and credited in future publications, as we require under our Creative Commons Attribution-Non Commercial 3.0 U.S. license. NECA is not a crowdsourcing project, but it does invite researchers to share their documents, modern nahuatlato share their translations, and academics and community members to share their ideas transnationally and interdisciplinarily. Through its design, the website attempts to make the case that this is worth everyone’s while.

The Form of the Content: If We Build It, Will They Come?

Archival research and the transcription and translation of idiosyncratic documents written in difficult handwriting, often in foreign languages, requires patience, time, resources, and above all, advanced skills that accrue over the years. Doctoral degrees, job offers, tenure, and future funding depend on demonstrating the fruits of this individual labor. There is nothing wrong with claiming the privacy to work, and what we have labeled “document teams” in NECA can also form via email, conferences, special journal issues, and edited volumes. If NECA’s first iteration – the digital archive – produces a flurry of new publications and dissertations created outside our platform, this will be a positive result.

NECA nevertheless encourages scholars to go beyond individual documents and to work beyond their comfort zones. It identifies common research interests across disciplines and national and academic communities, and presents the opportunity to share citations, translations, and knowledge in a public forum; to compare notes online; and eventually, given transcriptions and translations, to create databases, analyze the corpus as a whole, and experiment with different digital and computational tools. The NECA corpus is large and geographically varied enough to reveal not only the dialectal features of Nahuan languages in Central America, but also the documents’ production related to colonial settlement, ecclesiastical influence, social and political networks, the economy, and geography. We see great future value, especially, in thinking through NECA’s data using spatial analysis and mapping tools. Bringing linguists and translators of Nahuatl together with non-nahuatlato scholars of Central America has the potential to advance all this research further, faster. We built NECA to nudge people in this collaborative direction. The question is, will they come?

So far, the answer is yes and no. NECA’s analytics from Reclaim Hosting show that since the digital archive went online in July 2016, it has received the most intensive and consistent use (measured by bandwidth used, the ratio of pages to hits, and annual location data) from Mexico, El Salvador, and Guatemala, as well as Spain, Germany, Poland, and France in Europe — the last three being major centers of Mesoamerican and Nahuatl studies — and the United States, Brazil, and Canada in the Americas. Presentations at the University of Texas at Austin in March 2017, the Congreso de Estudios Mayas in Guatemala City in July 2017, the Asociación Centroamericana de Lingüística annual meeting in San Salvador in August 2017, the American Historical Association annual meeting in January 2018, and the Sociedad Mexicana de Historiografía Lingüística in Mexico City in October 2018, each produced temporary bumps in the number of unique visitors and/or intensity of use, which then tapered off.
Austin presentation acted as an official launch of the project with the power of social media behind it, resulting in an eighteen-fold increase in unique visitors immediately afterwards (March-April 2017). Subsequent presentations in Guatemala and El Salvador produced the most remarkable user data in the site's history thus far. In the two months following (July-August 2017) — and with no official social media push — the number of unique visitors to the site quadrupled. More importantly, the bandwidth and pages-to-hits ratio indicated significantly more searching through the site's most complex pages, such as those containing document images, than after the Austin presentation. The Central Americans' more intensive use is visible in the contrast between their relatively low number of unique visitors (yellow) relative to pages, hits, and bandwidth (blue and green):

<table>
<thead>
<tr>
<th>Month</th>
<th>Unique Visitors</th>
<th>Number of Visits</th>
<th>Pages</th>
<th>Hits</th>
<th>Bandwidth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan 2017</td>
<td>79</td>
<td>342</td>
<td>8,710</td>
<td>10,206</td>
<td>36.55 MB</td>
</tr>
<tr>
<td>Feb 2017</td>
<td>69</td>
<td>331</td>
<td>5,283</td>
<td>7,869</td>
<td>157.66 MB</td>
</tr>
<tr>
<td>Mar 2017</td>
<td>1,241</td>
<td>1,678</td>
<td>27,117</td>
<td>34,264</td>
<td>446.55 MB</td>
</tr>
<tr>
<td>Apr 2017</td>
<td>654</td>
<td>1,599</td>
<td>32,651</td>
<td>36,107</td>
<td>240.19 MB</td>
</tr>
<tr>
<td>May 2017</td>
<td>252</td>
<td>926</td>
<td>28,501</td>
<td>29,996</td>
<td>91.43 MB</td>
</tr>
<tr>
<td>Jun 2017</td>
<td>121</td>
<td>429</td>
<td>30,350</td>
<td>31,495</td>
<td>75.27 MB</td>
</tr>
<tr>
<td>Jul 2017</td>
<td>135</td>
<td>504</td>
<td>26,167</td>
<td>28,402</td>
<td>160.66 MB</td>
</tr>
<tr>
<td>Aug 2017</td>
<td>742</td>
<td>1,361</td>
<td>31,595</td>
<td>49,117</td>
<td>791.17 MB</td>
</tr>
<tr>
<td>Sep 2017</td>
<td>329</td>
<td>1,029</td>
<td>29,679</td>
<td>32,849</td>
<td>333.77 MB</td>
</tr>
<tr>
<td>Oct 2017</td>
<td>281</td>
<td>1,054</td>
<td>16,251</td>
<td>19,029</td>
<td>142.62 MB</td>
</tr>
<tr>
<td>Nov 2017</td>
<td>269</td>
<td>927</td>
<td>9,389</td>
<td>11,301</td>
<td>144.45 MB</td>
</tr>
<tr>
<td>Dec 2017</td>
<td>225</td>
<td>747</td>
<td>10,947</td>
<td>13,060</td>
<td>94.75 MB</td>
</tr>
<tr>
<td>Total</td>
<td>4,307</td>
<td>10,929</td>
<td>277,040</td>
<td>323,695</td>
<td>2.71 GB</td>
</tr>
</tbody>
</table>

Figure 4. Usage data for NECA, 2017.

From 2016 through 2018, the United States and Ukraine generated most of the site's hundreds of thousands of page views, 75% of which lasted thirty seconds or less. Presumably, a large portion of these were bots. The next largest proportions of visits, however, lasted for over one hour (around 8%), thirty minutes to an hour (around 6%), and fifteen to thirty minutes (around 4%), suggesting that a significant minority of users were seriously engaging the site. Notably, when we ceased to actively promote the site in 2019 we saw a drop in unique visitors, a consistent narrowing of the pages-to-hits ratio indicating shallower exploration of the site, and 88% of visits lasting less than thirty seconds. (For the first time, a large number of such visits in 2019 came from the Netherlands, bumping Ukraine to third place in the “probably a bot” category). Nevertheless, in 2019 the most intensive users — those spending thirty minutes to over an hour on the site at a time — still constituted our next largest user group, or 7% of the total number of visits.

As a digital archive, therefore, NECA is doing reasonably well even when we do not take advantage of conferences, social media, and other means to publicize and promote it. As an online platform for collaborative transcription, it has been less successful. A few people have used the "Add a Document" feature to provide new citations and high-quality images, but most of the 19 new documents added since the site's inception have come from our own research or direct outreach. The same is true of the Discussion area, where invited essays by Janine Gasco on Nahuan agricultural terms in the Soconusco and by Adriana Alvarez on Nahuatl instruction at the Universidad de San Carlos in Guatemala have generated a handful of comments from community members mostly from El Salvador or of Salvadoran descent in the United States, but no serious scholarly engagement,
without which we cannot move forward to better understand why, how, or to whom these documents might be important.

The document teams and Scripto's transcription tool have attracted no users at all since beta testing in March 2017. This may be a design issue. This first iteration of NECA is based on a pre-designed Omeka platform and utilizes only the Scripto features made available through the plug-in. Certainly we could improve the transcription and translation tool to be more appealing and effective, including a simpler user interface, better versioning, an improved commenting feature that identifies the user and is always visible, side-by-side images and workspace, progress bars, and the ability to toggle between transcriptions, translations, and versioning on a single page. The features and functionality of the transcription tool at the Codex Aubin project, hosted on software developer Ben Brumfield's transcription platform FromThePage based on Ruby on Rails (https://fromthepage.com), are exemplary (https://fromthepage.ace.fordham.edu/benwbrum/codex-aubin), as is the transcription and search tool created for the Freedom on the Move project (https://freedomonthemove.org/index.html). Other projects with more user-friendly transcription workspaces than NECA include the Newberry Library's Newberry Transcribes (https://publications.newberry.org/digital/mms-transcribe/index) and Maynooth University's Letters 1916-1923 (http://letters1916.maynoothuniversity.ie/learn), both of which are based on Omeka and Scripto.

How the digital archive's form can encourage engagement with its content is not, however, only a design issue. The most successful transcription projects come from outward-facing institutions digitizing items from their physical archives and making them available to "citizen humanists" with the clear goal of public engagement — for instance (among many other examples), the Smithsonian Institution's Digital Volunteers initiative (https://transcription.si.edu) and the Library of Virginia's Making History project (http://www.virginiamemory.com/transcribe). Often, featured collections are chosen with audience interest and capabilities in mind. The Stanford University Archives (https://library.stanford.edu/spc/university-archives), for example, invites online transcription of manuscripts related to the university's history, in English. Broad or targeted appeal of the subject matter, readability of the documents, and language accessibility seem equally relevant to the success of the aforementioned Freedom on the Move Project, which crowdsources transcriptions of mostly English, printed newspaper announcements of rewards for runaway African American slaves; Newberry Transcribes, which presents mostly English-language diaries and letters about family life in the Midwest; and the narrower but commemorative Letters 1916-1923, which invites visitors to submit and transcribe their own family's documents for upcoming anniversaries of the Easter Rising, World War I, and the Irish War of Independence and Civil War.

A search through the transcription platform FromThePage's various collections suggests that more academic projects often involve fewer participants, especially where handwritten manuscripts of earlier time periods with idiosyncratic paleography in languages other than English are concerned. Online transcription in these circumstances seems to work best as a collaboration tool between professors and students, or between small groups of colleagues with similar skills. This is the case of the Codex Aubin and French from Outremer (https://fromthepage.ace.fordham.edu /collection/show?collection_id=1) projects from Fordham University, which deal with medieval and early modern manuscripts in Nahuatl and French requiring highly specialized transcribers. Many digital archives of similarly challenging material rely entirely on professional teams and do not make collaborative online transcription tools available, for instance the BFM - Base de François Médéval (txm.bfm-corpus.org), the 1641 Depositions Project from Trinity College, Ireland (http://1641.tcd.ie/project.php), and the Native Northeast Research Collaborative (https://www.thenativenevertheast.org).

Comparing these projects, and NECA, to the Ticha project described in this issue by Broadwell et al. makes clear that the challenges faced by creators of digital archives are highly contingent. As a digital archive and online transcription platform for colonial-era texts in Zapotec languages from Oaxaca, Ticha encountered some of the same design limitations as NECA when using software such as Scripto and the Fieldworks Language Explorer (FLE) [Broadwell et al. 2020] [Broadwell and Lillehaugen 2013]. Ticha was aided by the fact that the interests of Zapotec speakers and scholars, ethnographers, and linguists converged on the same period and language, as opposed to the criss-crossing and sometimes conflicting interests faced by NECA. However, Ticha is also a powerful example of what sustained attention to the human side of digital projects — conferences and workshops, acceptance and accommodation of a wide range of user communities, and outreach especially to non-academic stakeholders, in this case native speakers of Indigenous and minority languages — can achieve.

To re-design the weakest link of NECA, its transcription tool, would require at minimum a switch from the current pre-designed website and/or outsourcing of the tool, and possibly changing from WikiMedia to a standalone database. It is not clear that, at this stage of the project, the effort would be worth it. While some have expressed interest in using the site as a teaching tool for advanced students who are simultaneously learning Nahuatl and paleography, there is no way to know whether this is happening. Likewise, if more established scholars are working with documents from NECA, they are doing outside the context of the site. At a practical level, scholars may find online transcription and translation, which requires working within the confines of the program and/or between multiple formats, less efficient than traditional methods. They may also appreciate opportunities for face-to-face discussion prior to performing their work online. Scholarship is risky and takes time. Sturdy, creative collaborations between people who have not traditionally worked together — such as the local, national, disciplinary, and academic networks that have expressed interest in NECA yet remain siloed from each other — may initially develop better in person. Rather than immediately overhauling the site or the transcription tool, a better next step for NECA may be more old-fashioned: to convene scholars and community members in different combinations and venues, with the goal of creating collaborative teams and identifying viable research questions and interests in common.

Digital humanities promises more than a new marriage between mathematical, qualitative, and design methodologies and tools. It also proposes a paradigmatic change in how scholars collaborate, flattening research and/or learning communities and vaunting an idealized, non-hierarchical community where people willingly share their research, promote interdisciplinarity, and work in teams of members with complementary skills sets, none of which is seen as more important than another. Despite the ways in which this mimics Silicon Valley-ese (rightly criticized for its hypocrisy), there is much to hold onto here: the potential of digital humanities to communicate with broader publics, to democratize the production of knowledge, to make the fruits of scholarship more accessible, and to make us all more flexible thinkers. As NECA argues, digital archives also have the potential to push scholarship in certain directions by calling attention to understudied texts or problems and by making the materials for studying them available.

But the digital humanities’ optimistic, even utopian view of the scholarly workplace is tinged with disciplinary, financial, and intergenerational
anxieties. In the United States, humanities scholars of all stripes fear the devaluation of their work in the information age. The younger generation faces an increasingly freelance economy and shrinking humanities job market from the peculiar position of being simultaneously valued for their digital savvy (writing code, understanding algorithms, managing project teams, marketing their work), expected to be innovators and jacks-of-all-trades, and suspected of not doing the kinds of specialized research that got their professorly elders tenure. Established scholars are suspected of lagging behind the digital turn, but have more freedom to experiment with digital tools — or not — with far less risk to their future careers. They are also the gatekeepers of the academy.

It is therefore incumbent upon senior scholars, especially, to ponder the lessons of creative failure in digital humanities projects. NECA shows the potential for digital archiving to turn a wide range of people’s attention towards a particular corpus of historical documentation and set of questions. NECA also highlights the difficulty of attracting scholars to skills-intensive transcription and translation online in collaborative projects without prior commitments, goals, and relationships in common. While we maintain the first iteration of the NECA digital archive, our next best step for transcription and translation — the necessary building blocks of any future database — will involve human, not digital, development: recruiting and funding new team members, acquiring grant money to pay for skilled transcriptions and translations, and organizing conferences. With data in hand and new ideas on the table, we can start to contemplate smaller, more limited digital tools — what Rockwell and Sinclair [2016] call "embeddable toys" — for scholars to play with, exploring what value computation might bring to the analysis of the entire NECA corpus. To move forward we must forcefully argue for the funding of both methods of scholarship, digital and traditional, most especially for those who will be the generators, guardians, and teachers of Nahua right and Nawat in the future.

Notes

[1] NECA is a joint effort between the authors that has also depended on the generous support of others. We thank Marquette University's College of Arts and Sciences for start-up funding, Sergio Romero for helping assemble the core document list and translations of sample text, Héctor Conococh Chet for photographing documents in the Archivo General de Centroamérica in Guatemala City, Rafael Lara Martinez for Spanish translations of the primary pages, and Jorge Lemus for his collaboration in El Salvador. Ann Hanlon at the Digital Humanities Lab, University of Wisconsin-Milwaukee, Adriana Alvarez, David Dominguez Herbón, and Miriam Peña Pimentel of the Red de Humanidades Digitales, Universidad Autónoma de México, Sharon Leon at the Roy Rosenzweig Center for History and New Media, George Mason University, and David Bodenhamer of the Polis Center, Indiana University-Purdue University were invaluable sounding boards, as were advisory board members Michael Swanton, Janine Gasco, Matilde lvic de Monterosso, and Karl Offen. Julia Madajczak and Agnieszka Brylak tested the transcription tool, and graduate assistants Ben Nestor and Cory Haala explored mapping applications. We especially thank the many contributors of documents to the site, listed at http://nahua-nawat.org/aboutus.

[2] The terms "Náhuatl" and "Náhuatl" in Spanish, "Nahuatl" and "Nawat" in English, are currently the most conventional ways of referencing these related but distinct languages. As Hansen 2016 explains, the orthographic conventions of Nahua languages are fluid and we do not intend any definitive statement by selecting these particular ones. On the politics of orthography and revitalization see also Olko and Sullivan 2013 esp. pp. 201-11, and van Zantwijk 2011.

[3] For instance, documents of only a few pages each in many different Indigenous languages and genres — cofradía documents, letters, tribute records, etc., — are archived together as standalone documents, removed from their original documentary context with no paper trail, within the folder A1 legajo 6074 in the Archivo General de Centro América in Guatemala City. Similarly, see Quiroa 2017 on the recent decision by the Newberry Library in Chicago to physically separate the Popol Wuj from the rest of the clerical text by Dominican Fr. Francisco Ximénez to which it once belonged, while preserving a record of its provenance and the state in which it arrived at the Newberry. The removal of texts from their place of origin by antiquarian collectors and scholars, with permission or not, represents yet another kind of decontextualization.

[4] But see the cautionary example provided by Madajczak and Hansen 2016, who show that even the characteristics assigned to these generally accepted linguistic labels may be combined or modified and fail to precisely identify the language of any given document.

[5] The literature on this process of linguistic, religious, and intellectual exchange is vast and varied, resting on the shoulders of scholars such as Angel María Garibay, Fernando Horcasitas, and Miguel León-Portilla in Mexico, and Dennis Tedlock, James Lockhart, Louise Burkhart, and Judith Maxwell in the United States.


[7] CHNS has since released a new version Omeka S with a built-in multilingual option, but as of this writing it is not compatible with Scripto. Rafael Lara Martinez generously translated all our simple pages into Spanish.

Works Cited


This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.