Diaspora and Digital Discourse: Papiamento/u Research in the Digital Archives of Coleccion Aruba

Jan Bant
Eric Mijts
Peter Scholing

This article describes the development, scope and possibilities of the Aruban collaborative digital archive initiative Coleccion Aruba. Coleccion Aruba is an online Aruban digital hub where users can consult historical texts, videos, and images. Among other things, it comprises the digitalized databases of the Aruban National Library, the Aruban National Archive, and the Aruban Archeological Museum. In addition to Aruban documents, it hosts sources from European and regional institutes that touch on the cultural, social, and political history of Aruba and the former Netherlands Antilles. A unique feature is that many of the digitized sources are in Papiamento/u. This article first describes the history and creation of this collection. Second, it demonstrates how Coleccion Aruba has opened up new research possibilities in and about Papiamento/u, and sketches directions for future research on the basis of this collection. Lastly, it explains how the digital collection enables new and decolonial opportunities for the development of Artificial Intelligence solutions for the Papiamento/u language as well as regional and international collaborations between archives, libraries, and research institutions. Because of the scattered and diasporic nature of the Papiamento/u language community, the authors conclude that online digital collections such as this one are essential in ensuring a sustainable future for the language of Papiamento/u as well as research in and about this language.

In April 2024, Aruba appeared on the international news several times: the island was featured and discussed by different large news media.[1] While Aruba is no stranger to

[1] The sequence of authors of this article is alphabetical, and the three authors contributed equally to this article.

international coverage – it is often praised as a tourist destination – the fact that the island was covered by several tech-websites was indeed unprecedented. Why did websites like *Wired* and *The Verge* suddenly write about Aruba? They reported that the digital library of Internet Archive became "the official custodian of an entire nation's history for the first time."[2] Indeed, the Internet Archive hosts the digital collections of *Coleccion Aruba*, the online hub of several national Aruban historical and cultural organizations. Users can freely consult over 150,000 digital items from Aruban archives, museums, and libraries, thus enabling a diverse range of people to study the history and culture of this small Caribbean island.

This article delves deeper into Coleccion Aruba and its digital possibilities, especially with regards to Papiamento. As this article focuses on an Aruban digital hub, the authors will speak of Papiamento – the Aruban variation of the language – instead of Papiamentu – the variation used in Bonaire and Curaçao. Through this deep dive into Aruba's digital heritage hub, this article argues that Coleccion Aruba is vital for sustaining the future of the Papiamento language, and especially promoting research related to it. To make this point, this article first describes the history and coming into existence of Coleccion Aruba. It then discusses the benefits the collection has for enabling knowledge production about *and* in Papiamento. Lastly, it describes how Coleccion Aruba can create decolonial opportunities with regards to Artificial Intelligence (AI) and Machine Learning (ML) possibilities for Papiamento.

Coleccion Aruba: A short history

The *Plataforma Coleccion Aruba* was established in October of 2022 by its three founding partners: Biblioteca Nacional Aruba (National Library of Aruba, est. 1949), Archivo Nacional Aruba (National Archives of Aruba, est. 1994), and UNOCA (Union of Aruban Cultural Organizations, est. 1986), all with a long history in preserving and providing access to the island's knowledge, to its documentary heritage, and to its cultural and historical materials, with physical collections spanning from the 16th century to the present day.

These three founding partners decided to formalize their existing collaboration in 2022, and joined forces in order to mitigate the bottlenecks usually associated with documentary and cultural heritage preservation work in small island development states (SIDS). As of April 2024 the Plataforma Collection Aruba consists of a total of 12 heritage partners from across the GLAM (galleries, libraries, archives, and museums) and cultural spectrum (encompassing government/public and private institutions and NGOs), all contributing with their collections, their expertise and connected by their willingness to collaborate, share, and to open up access. Collaboratively and jointly, these partners are the custodians of Aruba's current documentary and cultural heritage, building upon decades of preservation and digitization work, and dedication to public access to their collections and services.

[2] Kate Knibbs, "The Internet Archive Just Backed Up an Entire Caribbean Island," Wired.com, April 8, 2024, https://www.wired.com/story/internet-archive-backed-up-aruba-caribbean-island/; Wes Davis, "The Internet Archive is now hosting Aruba's history," Theverge.com, April 10, 2024, https://www.theverge.com/2024/4/10/24126091/the-internet-archive-aruba-historic-materials-digital-preservation.

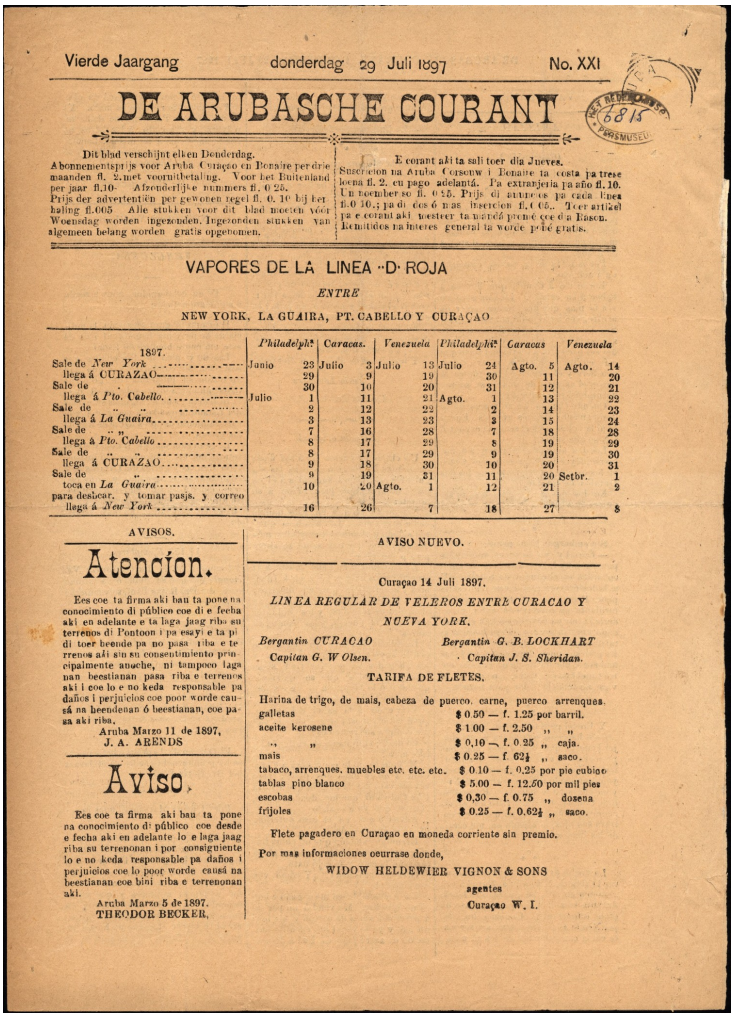
The online Coleccion Aruba is the main product of this collaboration. The objective of the online collection is to actively mitigate the lack of accessibility of relevant collections, a major limiting factor for research and education, through digital means. The digital presence enables users, who otherwise might not be able to visit an archive or a library, to consult Aruba's historical and cultural collections. It makes accessible and available materials that touch the identity, history and language of Aruba. As such, Coleccion Aruba aims to take this scattered nature of historical and cultural documents of Aruba, including documents about the Papiamento, and bring all the puzzle pieces together on one single online platform. All materials relevant to Aruba, regardless of their physical location, wherever they are in the worldwide "diaspora of collections," became available in one online database, instead of dozens. In order to accomplish this, collaboration not only takes place on the island itself (between local partners), but also with institutions or individuals holding valuable documentary heritage materials anywhere else in the world. Recent examples are digital collection exchanges and linked data projects with national libraries, archives and museums in both Europe and in the Latin-American and Caribbean region, with missionary organizations (formerly) active in the region, and even with the private owners, like with the archives of the royal family of the Kingdom of the Netherlands, the Koninklijk Huisarchief.

The Digital Collections hold over 150,000 digital items, almost as many items as the island has inhabitants (110,000). The webpage of the Coleccion, `https://coleccion.aw/`, was started as a continuation of – and expansion on – the Aruba National Library's online National Digital Library Collection, which was established in 2019. During the COVID-19 pandemic and its periods of lockdowns, this webpage proved to be an important resource for both local and international audiences interested in Aruba's history and culture. It includes over 60,000 textual resources (ranging from full-text searchable handwritten/archival materials to scanned publications and digital-born materials, in total covering over one million pages), 80,000 images, over a thousand videos, and 3D-scanned versions of archeological artifacts from the colonial and pre-Columbian era currently held by European institutions. By providing digital access along with the (pre-existing but scattered and limited) physical access, these small island collections are open to the world, with no barrier or bottleneck other than access to the internet.

The large corpus of materials from a wide variety of sources/partners, and on a wide range of topics, makes Coleccion Aruba interesting both as a research object and as a resource. An example of this is the research conducted and facilitated on Transatlantic slavery and the (forced and voluntary) migration within the region, research on the use of linked data and AI-powered Handwritten Text Recognition technology to further access, research and understanding of these impactful historical events and forces.[3] This research has resulted in a nomination to the International Memory of the World Register, submitted in November of 2023 (addition of Aruba to the already registered "Documentary heritage of the enslaved people of the Dutch Caribbean and their descendants (1816-1969)" which previously did not include Aruba), and also in an online database on the People Enslaved

[3] Peter Scholing and Raymond Hernandez, "Coleccion Aruba, Inter-sectoral Collaboration on Aruba as a Model for the (Dutch) Caribbean: A collaborative approach for preservation and access of collections in Small Island States," in *Sailing in Salt Water: Advancing the Dutch Caribbean Reparations Agenda and Establishing an Inter-Island Knowledge and Expertise Network on the Slavery Past and its Heritage*, ed. Luc Alofs (2024), 81-88, `https://coleccion.aw/show/?UA-DIG-SAILING-SALT-WATER/page/81`; See also the Database of People Enslaved on Aruba 1840-1863, `https://coleccion.aw/database/en/`; "History of slavery on Aruba and St. Eustasius online," Radboud University Nijmegen, `https://www.ru.nl/en/research/research-news/history-of-slavery-on-aruba-and-st-eustasius-online`.

on Aruba 1840-1863, in thematic collections on the Coleccion Aruba, and in various research publications and presentations.



Vierde Jaargang donderdag 29 Juli 1897 No. XXI

DE ARUBASCHE COURANT

Dit blad verschijnt elken Donderdag.
Abonnementsprijs voor Aruba Curaçao en Bonaire per drie maanden fl. 2.met vooruitbetaling. Voor het Buitenland per jaar fl.10. Afzonderlijke nummers fl. 0.25.
Prijs der advertentiën per gewonen regel fl. 0. 10 bij herhaling fl.005 Alle stukken voor dit blad moeten vóór Woensdag worden ingezonden. Ingezonden stukken van algemeen belang worden gratis opgenomen.

E corant aki ta sali toer diu Jueves.
Suscricion na Aruba Corsouw i Bonaire ta costa ja trese loena fl. 2. cu pago adelantá. Pa extranjeria ja año fl. 10.
Un nomber so fl. 0.25. Prijs di anuncios ja cada linea fl. 0 10.; ja di cos ó mas insercion fl. 0 05.. Toer artikel ya e corant aki bestrer ta nanéd prome cu di Bosch. Remitidos na interes general ta worde prie gratis.

VAPORES DE LA LINEA D. ROJA

ENTRE

NEW YORK, LA GUAIRA, PT. CABELLO Y CURAÇAO

1897.	*Philadelphia*		*Caracas.*		*Venezuela*		*Philadelphia*		*Caracas*		*Venezuela*	
Sale de *New York*	Junio	23	Julio	3	Julio	13	Julio	24	Agto.	5	Agto.	14
llega á CURAZAO		29		9		19		30		11		20
Sale de		30		10		20		31		12		21
llega á *Pto. Cabello*	Julio	1		11		21	Agto.	1		12		22
Sale de " "		2		12		22		2		14		23
llega á *La Guaira*		3		13		23		3		15		24
Sale de " "		7		16		28		7		18		28
llega á *Pto. Cabello*		8		17		29		8		19		29
Sale de " "		8		17		29		9		19		30
llega á CURAZAO		9		18		30		10		20		31
Sale de "		9		19		31		11		20	Setbr.	1
toca en *La Guaira*		10		20	Agto.	1		12		21		2
para descar. y tomar pasjs. y correo												
llega á *New York*		16		26		7		18		27		8

AVISOS.

Atencion.

Ees coe ta firma aki hau ta pone na conocimiento di público coe di e fecha aki en adelante e ta loga jaag riba su terrenos di Pontoon i pa esayi e ta pi di toer beende pa no pasa riba e terrenos ahi sin su consentimiento principalmente anoche, ni tampoco laga nan beestianan pasa riba e terrenos aki i coe lo e no keda responsable pa daños i perjuicios coe poor worde causá na beendenan ó beestianan, coe pasa aki riba.

Aruba Marzo 11 de 1897,
J. A. ARENDS

Aviso.

Ees coe ta firma aki hau ta pone na conocimiento di público coe desde e fecha aki en adelante lo e laga jaag riba su terrenonan i por consiguiente lo e no keda responsable pa daños i perjuicios coe lo poor worde causá na beestianan coe bini riba e terrenonan aki.

Aruba Marzo 5 de 1897.
THEODOR BECKER.

AVISO NUEVO.

Curaçao 14 Juli 1897.

LINEA REGULAR DE VELEROS ENTRE CURACAO Y NUEVA YORK.

Bergantin CURACAO — *Capitan G. W Olsen.*
Bergantin G. B. LOCKHART — *Capitan J. S. Sheridan.*

TARIFA DE FLETES.

Harina de trigo, de mais, cabeza de puerco, carne, puerco arrenques.	
galletas	$ 0.50 — f. 1.25 por barril.
aceite kerosene	$ 1.00 — f. 2.50 " "
" "	$ 0,10 — f. 0.25 " caja.
mais	$ 0.25 — f. 62½ " saco.
tabaco, arrenques, muebles etc. etc. etc.	$ 0.10 — f. 0.25 por pie cubico
tablas pino blanco	$ 5.00 — f. 12.50 por mil pies
escobas	$ 0,30 — f. 0.75 " dosena
frijoles	$ 0.25 — f. 0.62½ " saco.

Flete pagadero en Curaçao en moneda corriente sin premio.

Por mas informaciones ocurrase donde,

WIDOW HELDEWIER VIGNON & SONS
agentes
Curaçao W. I.

Figure 1 *De Arubasche Courant* was the first newspaper printed on Aruba. This edition, from July 29, 1897, is the only known copy still in existence, and is held by the IISG in Amsterdam, The Netherlands. A digital copy was digitally donated to form part of Biblioteca Nacional Aruba's digital collection on the occasion of BNA's 70th anniversary in 2019.

Coleccion Aruba and knowledge production about and in Papiamento

A key role of coleccion.aw is its contribution to digital scholarship by the creation and maintenance of the platform that allows for accessibility and study of datasets on a wide range of topics and in a wide range of disciplines. An important part of the collection of Coleccion Aruba deals with the history and nature of Papiamento, is written in (older variants of) Papiamento, or is both. As such, the collection enables many possibilities for Papiamento, when it comes to research and knowledge creation. This section explores the different regions where Coleccion Aruba can do this. First of all, Coleccion Aruba hosts

a number of documents that contain information about the history, spelling, grammar, and syntax of the language. The publications by the Aruban department of education, Departamento di Enseñansa, are good examples of this. In 2009, for instance, the Departamento di Enseñansa published *Vocabulario Ortografico di Papiamento*, an overview of more than 17,000 Papiamento words spelled the Aruban way.[4] This is a relevant publication, since the difference between Papiamento and Papiamentu lies most visibly in their different ways of spelling. Other publications by the Departamento di Enseñansa on Coleccion Aruba go further than merely spelling: between 2010 and 2015, they published a manual on the morphology of Papiamento (2010), a manual of syntax (2015), and a manual on prepositions and interjections (2016).[5] According to the department, these books were published because there was yet little knowledge and information publicly available about the grammar, syntax, and vocabulary of Papiamento. The goal of these books was therefore to get people familiar with the rules and words of the language. As such, it was geared towards students and teachers of Papiamento, with various levels of knowledge and understanding of the language. These publications thus offer a way for people to learn and familiarize themselves with the rules of Papiamento. Moreover, these publications can play an important role in programming and in the building of computational models. On the basis of books by the Departamento di Enseñansa one can, for instance, create automatic syntax-checking tools that take a step beyond a traditional (word-for-word) spell checker. Additionally, one can create all kinds of "learned machines" (and large language models) based on these "rule books" and the entire corpus of Papiamento texts available there.

Besides these official and more rule-based publications, Coleccion Aruba hosts publications which critically engage with the history and politics of Papiamento. These documents add context and additional information to the publications of the Departamento di Enseñansa. There is a large variety of these types of documents. This article discusses two sorts: speeches and presentations on the one hand, and academic theses and dissertations on the other. Over the years, during many different occasions, and in different locations, scholars and educators have given talks about the history and present of Papiamento. While many of these talks have gone unrecorded, the text files of some of them can be found on Coleccion Aruba. While these texts are often short and miss references to literature, they do give a clear overview of the history and state of the language. On the 11th of September 2015, for instance, Ramon Todd Dandaré, Aruban teacher and language-activist, gave a talk at an event of SPLIKA, an organization that aims to enhance the status and development of Papiamento/u. In his talk, Todd Dandaré focused on the historical developments of the Aruban spelling of Papiamento.[6] A decade earlier, Aruban Papiamento-scholar Joyce Pereira gave a talk in honor of International Mother Language Day. In her speech, Pereira shared recent developments that had been made with the institutionalization of Papiamento and analyzed the steps that still needed

[4] Departamento di Enseñansa, *Vocabulario Ortografico di Papiamento* (Aruba: Departamento di Enseñansa, 2009), `https://coleccion.aw/show/?BNA-DIG-PAP-VOCABULARIO-2009`.

[5] Departamento di Enseñansa, *Manual di gramatica di Papiamento: morfologia* (Aruba: Departamento di Enseñansa, 2010), `https://coleccion.aw/show/?BNA-DIG-PAP-GRAMATICA-MORFOLOGIA`; Departamento di Enseñansa, *Manual di gramatica di Papiamento: sintaxis* (Aruba: Departamento di Enseñansa, 2015), `https://coleccion.aw/show/?BNA-DIG-PAP-GRAMATICA-SINTAXIS`; Departamento di Enseñansa, *Preposicion y interheccion di Papiamento* (Aruba: Departamento di Enseñansa, 2016), `https://coleccion.aw/show/?BNA-DIG-PAP-GRAMATICA-PREPOSICION`.

[6] Ramon Todd Dandaré, "Ontwikkeling van de spelling van het Papiamento op Aruba," lecture delivered at SPLIKA symposium, September 11, 2015, `https://coleccion.aw/show/?BNA-DIG-CHARLA-TODDDANDARE-2016`.

to be taken.[7] The findings of these talks often represent a specific moment in time that are, more often than not, not found in books and articles. While primarily enabling users on Coleccion Aruba to expand their knowledge of the language, these documents also open up research possibilities. On the basis of these speeches, one could trace the historical development of language activism and advocacy about Papiamento.

The possibilities to conduct research are expanded in the case of the theses and dissertations that Coleccion Aruba hosts. It is particularly important that these documents are digitally stored, since they contain rich source material and analysis, but are seldom saved for later use. These documents reflect the dynamic and diverse possibilities that research into Papiamento has taken. In order to attain her master's degree in 1995, for instance, Régine Croes analyzed if and how Aruban education meets the language needs of the Aruban working population. The thesis shows that many people on Aruba use not only Papiamento, but also English, Spanish and Dutch on a near-daily basis. To safeguard the status of Papiamento within this multilingual situation, Croes concluded that strengthening teacher training and producing more Papiamento reading materials were essential.[8] In 2008, Vanessa Pietersz studied the role of Papiamento in Aruban secondary education, from the perspective of teachers. It analyzed how Papiamento can be most fruitfully put to use in schools. While some teachers support the use of Papiamento, policy decisions would need much additional research before being put into practice.[9] And in 2021, Jan Bant studied critical Aruban journals between the 1960s and 1980s. One of the journals studied was the monthly publication of the Aruban teacher's union, who were very involved with language policy. Bant showed that debate about instruction language in Aruban education already existed from the 1960s onwards.[10] These theses can form the basis for important follow-up research about the multifaceted history and future of Papiamento. How did the institutionalization of Papiamento in Aruban society and education come about? What are best practices with regards to introducing Papiamento as the language of instruction in schools? And how does Aruba's diverse and multilingual society influence all of this?

Coleccion Aruba is also the digital host of a number of PhD dissertations on Papiamento. For instance, the abovementioned Joyce Pereira defended her dissertation, *Valorization of Papiamento in Aruban society and education* at the University of Curaçao in 2018.[11] Her dissertation is a valuable resource for anyone interested in the history of Papiamento, and creole languages more generally. It offers a wealth of historical, educational, and political information about the position of Papiamento locally, regionally, and internationally. Shortly thereafter, in 2021, Eric Mijts defended his dissertation, *The situated construction of language ideologies in Aruba*.[12] Mijts demonstrates that

---

[7] Joyce Pereira, "UNESCO Y ARUBA: Hunto pa un maneho di idioma bon balansa!" Lecture delivered at Celebracion di dia di lenga materno, March 5, 2004, `https://coleccion.aw/show/?BNA-DIG-CHARLA-PEREIRA-2004`.

[8] Régine Croes, "Multilingualismo i enseñansa: Language needs in Aruban education." (MA Thesis, Catholic University of Nijmegen, 1995), `https://coleccion.aw/show/?BNA-DIG-TESIS-CROES-REGINE-1995`.

[9] Vanessa A.M. Pietersz, "How to open the door to Papiamento? Aruban language policy in secondary education." (MA Thesis, University of Groningen, 2008), `https://coleccion.aw/show/?BNADIGTESISPIETERSZVANESSA2008`.

[10] Jan Bant, "*Un ramo sosyal izkyerdista*: De visie van *Feneta* en *Skol i Komunidat* op dekolonisatie, 1969-1989." (MA Thesis, Leiden University, 2021), `https://coleccion.aw/show/?BNA-DIG-TESIS-BANT-JAN-2021`.

[11] Joyce Pereira, *Valorization of Papiamento in Aruban society and education, in historical, contemporary and future perspectives* (PhD diss., University of Curaçao, 2018), `https://coleccion.aw/show/?BNA-DIG-ARUB-9789990469417`.

[12] Eric Mijts, *The situated construction of language ideologies in Aruba: a study among participants in the language planning and policy process* (PhD diss., University of Antwerp & Ghent University, 2021), `https://coleccion.aw/show/?BNA-DIG-DISSERTATION-MIJTS-2021`.

the strongly ideological and colonial nature of the discussions on the development of inclusive multilingual education in the Caribbean is one of the main obstacles for the implementation of educational reform in these regions. He argues that educational reform will have little chance of success if these ideological underpinnings of the debate are not addressed. These theses and dissertations not only give those interested in the language information and tools to work with, it also shows that Papiamento has been an inspiration for many different generations of students, from the Caribbean and beyond.

While Coleccion Aruba thus showcases information about Papiamento, it also hosts documents that can enable innovative research avenues in diverse disciplines. For instance, early books in Papiamento can enable analyses of the origins of the language. Note, however, that the distinction between Papiamento and Papiamentu was still many years away when these books were published: these texts can therefore be used to understand the interinsular history of the language. Publications like *Ewanhelie di San Matheo* (1844), *Predikaasie den paapiamente* (1847), and *Boeki di orasjon pa moetcha Katolieka nan di Curaçao* (1849) are among the earliest official publications in Papiamento.[13] Users can consult these works and add a comparative layer to their analysis by studying them next to later publications in Papiamento, such as *Rekenboek pa muchanan di Curaçao* (1882), *Ciento cuenta corticoe* (1899), *Giambo bieuw a bolbe na wea* (1928), and *Cuentanan di Nanzi* (1952).[14] As such, scholars of Papiamento, and of creole languages more generally, can study the way the language developed over decades, even up to the present.

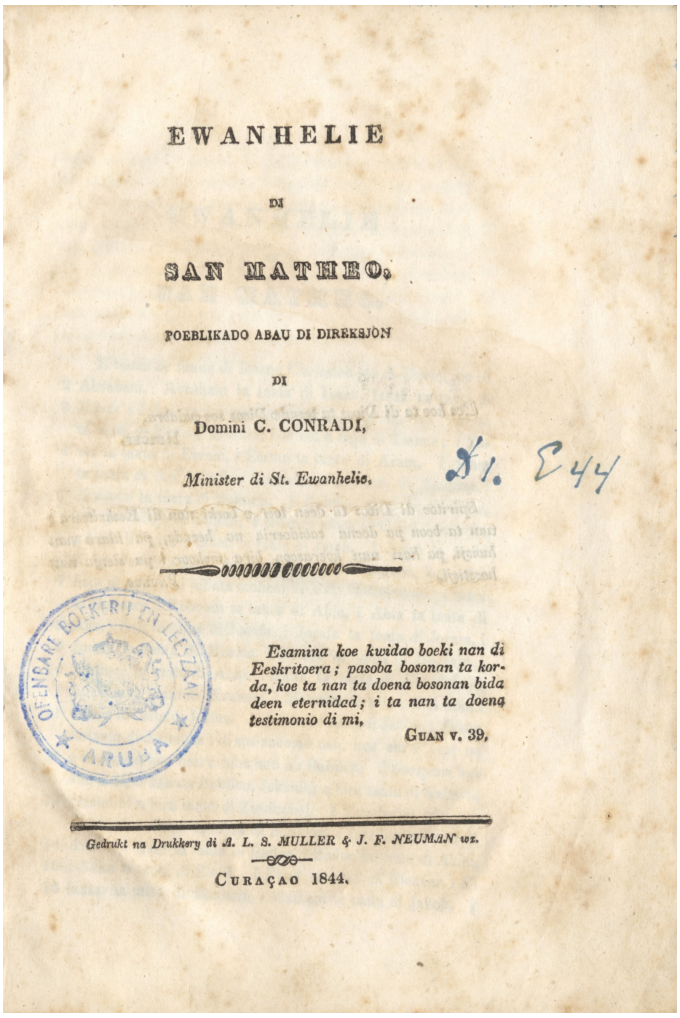
Another discipline for which Coleccion Aruba offers scholarly opportunities in the field of media studies: the database holds several digitized and digital-born Aruban newspapers in Papiamento, such as *Bon Dia Aruba*, *Awe Mainta*, and *Solo di Pueblo*.[15] Furthermore, the collection hosts archival material from several TV-programs, such as *Showtime* and *TeleAruba*.[16] Scholars interested in the daily written and spoken Papiamento, as well as scholars interested in Caribbean media studies more generally, can use these sources freely and fruitfully. These collections thus enable research *about* Papiamento, but also ideally empower research to write *in* Papiamento. The latter is especially important, given the relative lack of research carried out in that language.

[13] Cornelis Conradi, *Ewanhelie di San Matheo* (Curaçao: Drukkery di A.L.S. Muller & J.F. Neuman wz, 1844), `https://coleccion.aw/show/?BNA-DIG-KOSTBARE-0385`; Abraham van Dragt, *Predikaasie den paapiamente* (Aruba, 1847), `https://coleccion.aw/show/?BNA-DIG-MICROFILM-VANDRAGT-1847`. Jacobus. Putman, *Boeki di orasjon pa moetcha Katolieka nan di Curaçao* (Curaçao: Drukkery di A.L.S. Muller & J.F. Neuman wz, 1894), `https://coleccion.aw/show/?UNIVERSITEITLEIDEN-DIG-388056`.

[14] *Rekenboek pa Muchanan di Curaçao* (Curaçao: Imprenta de la libreria, 1882), https://coleccion.aw/show/?UNIVERSITEITLEIDEN-DIG-38611*; Ciento Cuenta Corticoe. Boeki di Leza Pa Uso di School* (Nijmegen: L.C.G. Malmberg, 1899), `https://coleccion.aw/show/?BNADIGITOTROMP1899CIENTOCUENTACORTICOE`; Willem E. Kroon, *Giambo bieuw a bolbe na wea. Novela intima Curazoleña* (Curaçao: Imprenta di Vicariato, 1928), `https://coleccion.aw/show/?UNIVERSITEITLEIDEN-DIG-444994`; N.M. Geerdink and Jesurun Pinto, *Cuentanan di Nanzi* (Curaçao: Fundasjon Nilda Pinto, 1952), `https://coleccion.aw/show/?BNA-DIG-PINTO-NANZI-1952`.

[15] "Corant – Newspapers," Coleccion Aruba. `https://coleccion.aw/browse/?col=arubanewspapers&sort=date+desc`.

[16] "Coleccion Showtime," Coleccion Aruba, `https://coleccion.aw/browse/?col=ana-dig-showtime&sort=addeddate+desc`; "Telearuba," Coleccion Aruba, `https://coleccion.aw/browse/?col=telearuba&sort=date+desc`.



EWANHELIE

DI

SAN MATHEO,

POEBLIKADO ABAU DI DIREKSJON

DI

Domini C. CONRADI,

Minister di St. Ewanhelie.

Esamina koe kwidao boeki nan di Eeskritoera; pasoba bosonan ta korda, koe ta nan ta doena bosonan bida deen eternidad; i ta nan ta doena testimonio di mi,

GUAN v. 39.

Gedrukt na Drukkery di A. L. S. MULLER & J. F. NEUMAN wz.

CURAÇAO 1844.

Figure 2 Title page of *Ewanhelie di San Matheo, poeblicado abau di direksjon di Domini C. Conradi* (1844)

Coleccion Aruba and decolonial digital creole language studies

While Coleccion Aruba thus offers many possibilities with regards to Papiamento and knowledge production, its potential also exceeds this and can open up new forms to understand the dynamics of the Papiamento language, as well as the richness of *languaging* and linguistic practices in the multilingual realities of the ABC islands where Papiamento/u has developed into its present day shape. Over the past decades, a wealth of new developments in AI have opened up new methods and opportunities for corpus based research to discover underlying rules and relations. The primary focus of such forms of research has been on corpora in the institutes of the Global North and - most often - in the dominant languages from the Global North. Dictionaries and grammars have rigidly codified Western, colonial languages, reflecting a prescriptive approach that reinforced the ideology of monolingual conformity – one language, one people. This endeavor was accompanied in colonial times by Christian missionaries' attempts to codify

languages in the colonial world, aiming to make these languages countable, identifiable, territorially bound, and controllable – not by their native speakers but by Western scholars who described the languages, and subsequently claimed and prescribed these language varieties.[17] The study of the historically disadvantaged languages of the Global South has primarily been done in Global Northern centers of study, and little effort or funding has been invested in the study of these languages in Global Southern institutes.[18]

In the Dutch Caribbean, few efforts were made to describe the local languages, with notable exceptions by Catholic missionaries like Jacobus Putman.[19] Papiamento, as one of the languages that emerged from the colonial project, originated from the interactions between various groups affected by the African enslavement and forced migration, the colonizers, and plantation workers. Despite efforts to codify the language, its exclusion from formal education allowed Papiamento to develop and expand organically. Its vocabulary and grammar have uniquely evolved in Aruba, Curaçao, and Bonaire, leading to a pluricentric nature characterized by variations in vocabulary, grammar, and pronunciation both between and within these islands.

The digital collection of documents across time and across regions allows for a comprehensive understanding of the richness and diversity of the Papiamento language, and the vibrant dynamics of the language in the multilingual ABC-islands. If a sufficiently large corpus is developed – as is currently being done – comparative Artificial Intelligence (AI) studies between digitized texts in different languages can allow for the development of a vast number of new possibilities for the use and study of the language. As Papiamento has not been described and prescribed as rigidly as the Global Northern languages, this allows for new approaches to understanding languaging through AI and Machine Learning (ML) with the support of Coleccion Aruba. It is clear however, that in order to achieve such goals, the right tools, skills, budget and collaborations are to be in place. The Universities and the Language Academies of the Dutch Caribbean have an important role to play in the achievement of these goals as the primary actors in developing these forms of research, and in order to ensure contextual understanding, local knowledge development and dissemination of research findings.

Present day development of AI and ML offer significant new pathways of study for Creole languages like Papiamento, that were thus far only reserved for larger and well-studied Global Northern languages.[20] In the first place AI technologies can aid in the preservation, documentation and promotion of Creole languages, which often face the risk of being marginalized and/or lost over time. ML algorithms can be used to transcribe oral histories, digitize written texts, and develop language learning tools tailored to Creole-speaking communities. This contributes to the preservation and revitalization of cultural heritage and linguistic diversity. Also in more sociolinguistic forms of study, ML can play an

[17] Rachael Gilmour, "'A Nice Derangement of Epitaphs': Missionary Language-Learning in Mid-Nineteenth Century Natal," *Journal of Southern African Studies* 33, no. 3 (September 2007): 521-538; Michael Meeuwis, "Involvement in language: The role of the Congregatio Immaculati Cordis Mariae in the history of Lingala," *The Catholic Historical Review* 95, no. 2. (2009): 240-260.

[18] Eric Mijts, "Home languages and higher education: challenges for inclusive higher education language policies for postcolonial small island states in the twenty-first century," in *Internationalization of Universities and the National Language: Language Policy Interventions and Case Studies*, eds. Danny Pieters & Thijs Keersmaekers, (Berlin: Peter Lang, 2020), 155-173, https://doi.org/10.3726/b17283.

[19] Wim Rutgers et al., *Jacobus Putman: Godsdienst, taal en onderwijs op Curaçao in de negentiende eeuw.* (Curaçao: Fundashon pa Planifikashon di Idioma & University of Curaçao, 2016).

[20] Nathaniel R. Robinson et al., "Kreyòl-MT: Building MT for Latin American, Caribbean and Colonial African Creole Languages," Arxiv.org, 2024, https://arxiv.org/abs/2405.05376.

important role, think for example about in depth linguistic landscape studies, in which not only quantification but also insights into the hierarchical relations between languages in the multilingual linguistic landscape can be developed. In doing so, Coleccion Aruba can be crucial in a decolonial shift in linguistics in which the expansion of fields of study can lead to a deeper understanding of the inclusive potential of multilingual and translingual practices.

ML algorithms can be trained to create and curate content in Creole languages. This is particularly valuable for educational purposes in contexts where several varieties of a Creole language are present, as is the context in the Papiamento/u speaking islands Aruba, Bonaire and Curaçao where there is a formal distinction between the etymological orthography which is used in Aruba, and the phonetic orthography which is used in Bonaire and Curaçao. Also in government communication, legal proceedings, but also in online platforms, social media, and digital content where personalized and relevant content is crucial for user engagement, the use of creole languages can be facilitated using AI. By generating content in Creole languages, AI can help bridge the digital divide and ensure inclusivity for speakers of these languages.

Not only will AI and ML techniques facilitate the development of (socio-)linguistic solutions in written form, but they can also be used to develop Natural Language Processing (NLP) models that understand and process Creole languages. This includes tasks such as speech to text application that allow for advances in elderly care, medical care and psychological care, but will also open up digital interfaces such as Alexa or Siri to speakers of Creole Languages. Other applications can include text classification, named entity recognition, sentiment analysis, and machine translation.

Through these developments, AI and ML will empower the Papiamento speaking communities by providing them with the tools to participate in multilingual education that includes the Papiamento language without burdening the education system with undue pressures to continuously develop new materials and tools for multilingual inclusive education for a small population. It will also facilitate Papiamento-inclusive multilingual governance and law enforcement, while increasing participation in the digital world and knowledge-sharing networks.

The development of Coleccion Aruba will further allow for the development of solutions for automatic transcription and translation, both for written and spoken language, which would also further allow for the study of the complex dynamics of Papiamento. In doing so, we may also be able to develop further insight into the disputed origins of Papiamento, not only by bringing to light the development of the language throughout the periods in which the corpus has been established, but also by using ML to track the multiple origins and roots of the language.[21] Finally, the corpus in combination with AI will allow for new forms of discourse analysis that will further support understanding of the reception of historic events in the past of the Papiamento speaking world, which in combination with the fields of anthropology, sociology and history can lead to a comprehensive understanding of the development of the society over time.

The potential of AI and ML can open up new horizons in the study of Creole languages such as Papiamento. The development of Coleccion Aruba is a first and fundamental building block of the realization of this potential, however, it is important that the research

[21] Efraim Frank Martinus, *The Kiss of a Slave. Papiamentu's West-African connections (PhD diss., University of Amsterdam, 1996); Bart Jacobs, _Origins of a Creole: The History of Papiamentu and Its African Ties* (Berlin: De Gruyter, 2012).

centers in the Global South are provided with the right AI and ML tools and support to further build upon that potential. The knowledge and insight that can be developed in doing so will only truly benefit the Global Southern communities if they are rooted in the communities and institutes of the Global South.



Figure 3 Little girls in Parochial School in Santa Cruz (April 1944). Photographer: Nelson Morris. Part of the LAGO Picture Collection, Biblioteca Nacional Aruba.

Future Development

Coleccion.aw is a dynamic project that is intended to further develop and grow. In the years ahead, Coleccion Aruba will focus on expanding its partnerships, collections and services in multiple ways: the audiovisual collections of both ANA, BNA and Telearuba are currently being digitized, and will add thousands of hours of Papiamento language content to the existing offerings. All current TeleAruba broadcasts are also being added to the digital archives on a daily basis. Combined with ML- and/or AI-based automatic transcriptions becoming available as well in the near future this will add immensely to

the expanding Papiamento text corpus of the Coleccion[22]. In early 2025, a high capacity book scanner will be donated to BNA and ANA by the Internet Archive and will be tasked to digitize the following: first all works in the Papiamento language held by BNA, second: all (government and research) reports held by BNA and ANA. A different line of expansion is into the region: expanding collaboration with memory institutions within the (Dutch) Caribbean, joining collections, and offering assistance to institutions and islands interested in embarking on a journey like the one undertaken in Aruba[23].

Concluding Remarks

To conclude, this piece introduced and described the history, function, and uses of Coleccion Aruba. It discussed the possibilities the Coleccion offers with regards to research about Papiamento, and explored the uses of this digital database with regards to the decolonial possibilities of AI and ML. As such, it has positioned Coleccion Aruba as a central digital resource in the cultural and historical context of Aruba, the wider Caribbean, and, indeed, all over the world. The collaboration between the different project partners has demonstrated to be the solid foundation that enabled the development of the collection, and the international partnership with archive.org allows for the structural safety and integrity of the collection. Through the accessibility that is created by coleccion.aw, a wealth of opportunities for digital humanities has opened up, ranging from historical studies to media studies, linguistics, ethnography and anthropology. Seen more broadly, Coleccion Aruba might show the way for other small creole languages to survive and thrive. Because many creole languages have a scattered language community, and because many documents related to these languages are also scattered over the world, a digital solution like the Coleccion can be an answer to the risks endemic to small creole languages.

Author Bios

Notes

[22] Peter Scholing, "Preserving Papiamento—Safeguarding Aruba's Language and Cultural Heritage" in: Messara et al. (eds.), *Vanishing Culture: A Report on Our Fragile Cultural Record* (2024), pp. 40-43,. https://blog.archive.org/2024/10/31/vanishing-culture-preserving-papiamento-safeguarding-arubas-language-and-cultural-heritage/

[23] Peter Scholing and Raymond Hernandez, "Coleccion Aruba, Inter-sectoral Collaboration on Aruba as a Model for the (Dutch) Caribbean: A collaborative approach for preservation and access of collections in Small Island States," in *Sailing in Salt Water: Advancing the Dutch Caribbean Reparations Agenda and Establishing an Inter-Island Knowledge and Expertise Network on the Slavery Past and its Heritage*, ed. Luc Alofs (2024), 81-88, https://coleccion.aw/show/?UA-DIG-SAILING-SALT-WATER/page/81

Jan Bant

Jan Bant (jan.bant@ru.nl) is a PhD Candidate in History at Radboud University Nijmegen. He moved from Aruba to the Netherlands in 2014, where he studied Philosophy and American Studies in Groningen, and Colonial History in Leiden. In his dissertation, he analyzes the history of sport in the Kingdom of the Netherlands. In particular, he looks at the role of baseball and softball in processes of belonging and identity formation among Dutch Caribbean people on both sides of the Atlantic since the 1950s. Through a combination of oral and written sources, he locates and interrogates the historical position of sport within the Kingdom. His PhD project is part of the Re/Presenting Europe research consortium, funded by the Dutch Research Council (NWO). In Jan's research, he is guided by the aim to forge meaningful connections between the Caribbean and Europe.

Eric Mijts

Dr. Eric Mijts (eric.mijts@ua.aw) is a researcher, educator and manager at the University of Aruba with a keen interest in program design for sustainable development in small island states. He leads the University of Aruba Research Center, co-founded and implemented the Sustainable Island Solutions through Science, Technology, Engineering and Mathematics (SISSTEM) project cluster (www.ua.aw/SISSTEM) as well as the Academic Foundation Year. He facilitates and initiates international students research exchange programs. As a researcher Eric is also affiliated to the University of Antwerp and is an active partner in the Global Resilience Research Network, in the United Nations Sustainable Development Solutions Network and in the Metabolism of Islands think tank. His research and publications focus on multilingualism, language policy and planning, education for sustainable development and inclusion/exclusion processes. Driven by his passion for the balance between humanity and the environment, Eric supports several international research and education projects on marine and terrestrial ecosystems health, governance and restoration.

Peter Scholing

Drs. Peter Scholing (p.scholing@bibliotecanacional.aw) is an information scientist and researcher at Biblioteca Nacional Aruba, Aruba's National Library, responsible for digitization, digital collections and research. As curator of the Aruba National Library's digital collections, his current main project is Coleccion Aruba, the "Aruba Collection," a digital collection platform and information portal that aims to digitally unite and make available all of Aruba's documentary and cultural heritage that is physically scattered across different islands, countries and continents. Peter is a member of the Aruba National Commission for UNESCO and is the immediate past President of MoWLAC, UNESCO's Memory of the World Programme for Latin America and the Caribbean (2022-2024). He also currently serves as an Editorial Board Member for the International Journal of Documentary Heritage (IJDH), established in 2024, and as an Executive Board Member of dLOC, the

Digital Library of the Caribbean. In May of 2024 he received the ACURILEANA Medal of Excellence as ACURIL's Caribbean Information Professional of the Year 2024. In October of 2024 Coleccion Aruba was awarded Internet Archive's 2024 Internet Hero Award.