

TO: UTC
FROM: Anushah Hossain and Deborah Anderson
 Script Encoding Initiative, UC Berkeley
TITLE: SEI Liaison Report
DATE: October 24, 2025

This document provides a summary of the UC Berkeley Script Encoding Initiative's recent activities.

Proposals/documents

The following SEI-funded proposals were recently submitted to the Script Encoding Working Group:

- Revised proposal to encode **Small Seal** in UCS, Michel Suignard [[WG2 N5344R](#)]
- Final Proposal for Encoding the **Mwangwego** script in the UCS, Oreen Yousuf and Daniel Yacob [[L2/25-135R](#)]
- Revised proposal to encode **Proto-Cuneiform** in Unicode, Steve Tinney, Anshuman Pandey, and Robin Leroy [[L2/25-211](#)]
- Revised proposal to encode **Sirmauri** in Unicode, Anshuman Pandey and Biswajit Mandal [[L2/25-224](#)]
- Updated proposal to encode **Lampung** in Unicode, Febri Muhammad Nasrullah [[L2/25-247](#)]
- Proposal to Encode **Codical Maya** in Unicode, Carlos Pallán, Andrew Glass, and Céline Tamignaux [not yet posted]
- Revised Proposal for Encoding the **Shaaldaa** Script in the UCS, Oreen Yousuf, Daniel Yacob, Mohammed Hassen Ali, and Aneso Mohammed [not yet posted]*

*Submitted in 2025 Q3 but not yet reviewed by the Script Encoding Working Group

Ongoing Script Work

Alongside the scripts mentioned above, the following projects are in active development, and are expected to be submitted to SEW for review in the next six months:

- Zou (Zolai) (Revised)
- Kulitan (Revised)
- Masaba (Revised)
- Odùduwà (Preliminary)
- N'Ti (Preliminary)
- Luo Lakeside (Preliminary)

Other ongoing projects include: Book Pahlavi, Minim Dag Noore, Proto-Elamite, Proto-Sinaitic, Vatteluttu, Ranjana, Linear Elamite, Byblos, Kerinci, Maya Hieroglyphs (Classic).

The full list of SEI-funded projects, including those currently inactive or awaiting more information, is available on the SEI website (sei.berkeley.edu/scripts-to-encode).

Research and Technical Activities

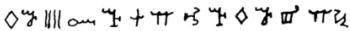
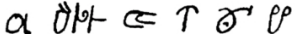
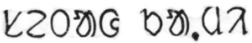
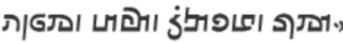
This year we launched the SEI Research Fellows program, a new fieldwork-oriented fellowship designed to support early-stage research on unencoded scripts. We are currently funding three projects, including script research on **Kulitan** in the Philippines, **Nsibidi** in Nigeria, and **Kerinci** and **Buginese-derived scripts** (Bima, Sumbawa, and Ende) in Indonesia. The blog post introducing this year’s research fellows can be read [here](#). Information about last year’s pilot run of the program is available on the SEI website [here](#).

SEI took on two research interns this summer, who began projects on the history of the *zwoj* and *zwnj* and on typewriter reforms in South Asia. The interns wrote blog posts to reflect on their respective experiences [here](#) and [here](#).

This fall, we are working with UC Berkeley PhD student Julián Vargo on an ongoing project estimating the time it takes to achieve “full stack support” for encoded scripts.

SEI is beginning a new stream of research on modern “neographies,” or recently invented scripts. This work will involve compiling data on past encoding decisions with respect to neographies and engaging with neography communities to understand how digital tools interact with script stability and adoption. A preliminary article by Anushah Hossain discussing the challenges of encoding neographies will appear in *Modern Languages Open* this fall.

This research is primarily funded by a new grant from the U.S. National Science Foundation, under the Science and Technology Studies program. We will be recruiting a two-year postdoctoral scholar to support this project. Recruitment details are available [here](#).

 Vai 1833 West Africa Vai Encoded in Unicode 5.1	 Pahawh Hmong 1959 Southeast Asia & diaspora Hmong Encoded in Unicode 7.0
 Afaka 1910 Suriname Ndyuka Not yet encoded in Unicode	 Tani 2001 Northeast India Tani languages Not yet encoded in Unicode
 Lisu 1915 Southwest China & Northern Myanmar Lisu Encoded in Unicode 5.2	 Ditema tsa Dinoko 2010s Southern Africa Bantu languages Not yet encoded in Unicode
 Ol Chiki 1929 Eastern India Santali Encoded in Unicode 5.1	 Rma 2017 Southwest China Qiang Not yet encoded in Unicode

Examples of neographies, indicating year and place of origin and target language(s)
(From “Unsettled Scripts: Script Invention and the Bind of Digital Recognition,” forthcoming)

Public Communications

SEI hosted its first ever open house at UC Berkeley in September, introducing our efforts to students, staff, and researchers across campus.

Also in September, Anushah Hossain spoke at the annual *Society for the Social Studies of Science (4S)* conference in Seattle, WA on a panel on text technologies. In October, Helena Kansa attended *LocWorld54* in Monterey, CA. Anushah is also giving a talk this month at *Alphabetica II: Entering Otherworlds* at the West Den Haag Museum in the Netherlands.

Several SEI affiliates will be giving talks at the Unicode Technology Workshop in November on topics such as: encoding undeciphered and unstandardized scripts, analyzing Indic script variation, and assessing a script's Unicode readiness.

SEI has continued sharing its work via a monthly [newsletter](#) and its [blog](#). A recent blog post described "[The Making of a Unicode Release](#)" in honor of Unicode 17.0.

Administrative Updates & Funding

SEI is pleased to host artist and type designer Arthur Francietta this fall under the prestigious Villa Albertine Artist Residency. Arthur will be developing typographic experiments that blend Afro-Caribbean writing traditions with California's cultural and technological influences.

SEI is looking forward to hosting UTC #187 at UC Berkeley in April 2026. We are also working closely with Unicode's Community Engagement Team to take UTW abroad next year.



Nancy / ARTEM Campus Connecting Art, Technology and Management

Overhead view of the Atelier National de Recherche Typographique in Nancy, France

SEI is formally housed in the Department of Linguistics at the University of California, Berkeley, but is entirely sustained by external grants and donations. Basic operations and script research funding currently comes from the Mellon Foundation's Public Knowledge program. Funding for research on neographies comes from the National Science Foundation (award number 2521875). Additional support for public communications projects comes from the Internet Society Foundation. Proposal work by Anshuman Pandey is also being supported through the Unicode Adopt-a-Character (AAC) program. In addition, proposal work by Carlos Pallán on Maya Hieroglyphs (Codal) is supported through SEI and the Unicode Consortium.

SEI accepts donations from private donors via our website (<https://sei.berkeley.edu/donate/>); employer matching is available. We are continually seeking funding to be able to expand our team and support more projects.