DOC TYPE:	Working Group Document
TITLE:	SEI Liaison Report
SOURCE:	Anushah Hossain and Deborah Anderson Script Encoding Initiative, UC Berkeley
DATE:	June 10, 2025
STATUS:	Liaison contribution
ACTION:	For consideration by WG2
DISTRIBUTION:	ISO/IEC JTC1/SC2/WG2

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

I. Submitted Proposals

Proposals incorporated in CD2 include:

- **Tolong Siki** [L2/23-024], Anshuman Pandey
- **Beria Erfe** [L2/24-004], Siddick Adam Issa et al.
- **Sidetic** [L2/23-019]**,** Anshuman Pandey

Proposals or documents currently submitted to WG2:

- **Small Seal** [WG2 N5306], Michel Suignard
- **Sidetic Naming Conventions** [WG2 N5308], Debbie Anderson, Robin Leroy, and Ken Whistler

Proposals submitted to the Unicode Script Encoding Working Group:

- **Minim Dag Noore** [L2/25-136], Oreen Yousuf
- **Mwangwego** [L2/25-135, L2/25-039, L2/24-241], Oreen Yousuf and Daniel Yacob
- **Sirmauri** [L2/25-134, L2/25-038], Anshuman Pandey and Biswajit Mandal
- **Masaba** [L2/24-248], Oreen Yousuf

II. Ongoing Proposals

The following script proposals are in active development, and are expected to be submitted to the Script Encoding Working Group for initial review in the next six months:

- Egyptian Hieroglyphs (Database Improvements)
- Proto-Cuneiform (Revised)
- Lampung (Revised)
- Zou (Zolai) (Revised)
- Kulitan (Revised)
- Sheek Bakrii Saphaloo (Revised)

- Odùduwà (Preliminary)
- N'Ti (Preliminary)
- Luo Lakeside (Preliminary)

Other ongoing projects include: Book Pahlavi, Proto-Elamite, Proto-Sinaitic, Vatteluttu, Ranjana, Linear Elamite, Byblos, Kerinci, Maya Hieroglyphs (Codical and Classical).

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

III. Technical Support

Alongside our preparation of script proposals for unencoded scripts from the Southern Sumatra region of Indonesia, SEI intern Ariq Syauqi has been conducting a review of the existing Rejang code block (a script family from the same region).

SEI worked with the design studio Jamra Patel to prepare a Noto Sunuwar font. This font has also been supplied to the Common Locale Data Repository (CLDR) hosted by Unicode Consortium for the upcoming release.

With the help of collaborators from Jamra Patel, Translation Commons, and Yale University, we have developed a messaging application for the Bété script using Private Use Area code points. SEI Research Fellow Adam Yeo has been user testing this application across Côte d'Ivoire. In the future, we hope that a general application based on this prototype might serve as a strategy for aiding new/experimental scripts gain traction in digital environments prior to encoding.

Debbie Anderson has been helping to identify language experts to supply and verify information for CLDR.

IV. Research Activities

SEI's latest research efforts include preparing a "Script Readiness Rubric" and a set of "Principles Underlying Script Readiness" to help communicate a script's viability for Unicode inclusion. These are published on the SEI website, alongside evaluations of 150+ unencoded scripts (see https://sei.berkeley.edu/scripts-to-encode/).

Assisted by student volunteer Ahad Bashir from Brown University, SEI has also begun compiling an open data set that aims to illustrate the length of time it takes for a script to move from the first proposal to full stack support. Anushah Hossain and Ahad Bashir presented preliminary findings at the Unicode Technical Workshop 2 in Sunnyvale, USA in October 2024.

We are planning to work with four interns this summer, on a range of projects including the history of the *ZWJ* and *ZWNJ*, typewriter reforms in South Asia, a guide for formatting script proposals, and research on Indonesian scripts.

We also expanded our team for the first time, with the hire of the new Program Manager, Helena Kansa. Helena expands our administrative capacities and oversees programs such as our new Research Fellowship.

V. Public communication

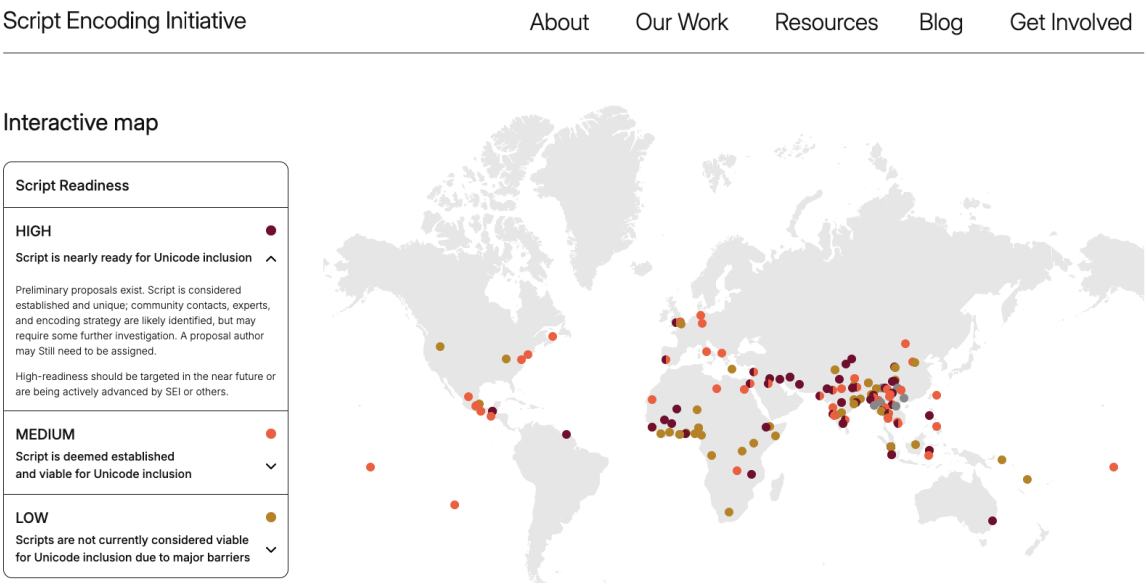
Presentations

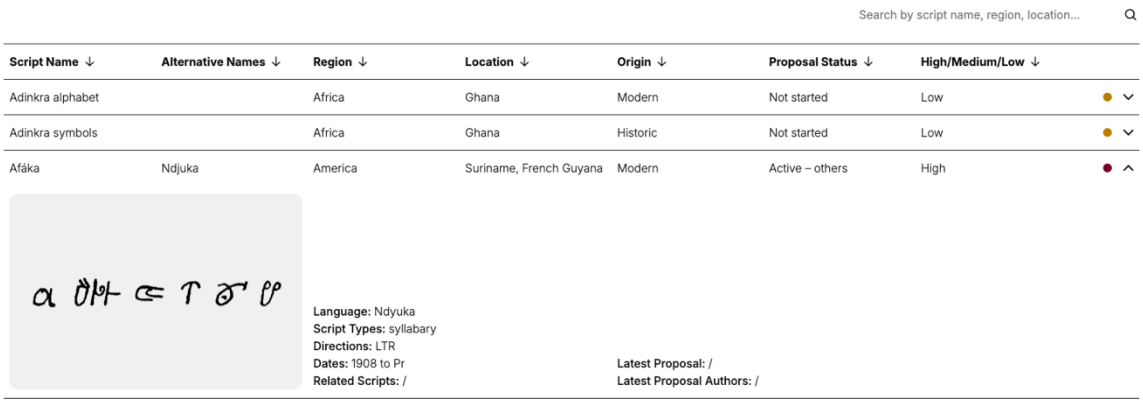
This year, Anushah Hossain gave talks at Face/Interface at Stanford University, World Endangered Writing Day, Automatic Type Design 3 in Nancy, France, and UNESCO's Language Technologies for All conference in Paris, France. She also participated in a panel at the Computer History Museum in Mountain View, USA and spoke at a monthly event for the International Multilingual User Group in San Jose, USA.

Logan Simpson delivered a talk entitled, "Behind the Scenes of Unicode: Navigating the Proposal Process for Indigenous Scripts" at the International Conference on Language Documentation & Conservation at the University of Hawaii at Manoa. Gabrielle Vail, Andrew Glass, and Alexandre Bassi presented "New Digital Tools for Engaging Audiences with Maya Text" at the Tulane University's Latin American Library.

Website + Newsletter

SEI's new website launched in May 2025 and features an updated script database, blog, and resources for learning and teaching about script digitization. Visit us at **sei.berkeley.edu** to learn more and sign up for the monthly newsletter.





Preview of Scripts to Encode page

VI. Funding

SEI is formally housed in the Department of Linguistics at the University of California, Berkeley, but is entirely sustained by external grants. Basic operations and script research funding currently comes from the Mellon Foundation's Public Knowledge program, which will provide support from Sept 2024 – Aug 2028. Additional support for public communications projects from Sept 2024 - Aug 2026 comes from the Internet Society Foundation. Proposal work by Anshuman Pandey is also being supported through the Unicode Adopt-a-Character (AAC) program. Carlos Pallán's work on Codical Maya hieroglyphs is being supported jointly by the Unicode Consortium and SEI.

SEI accepts donations from private donors via our website (https://sei.berkeley.edu/donate/). We are continually seeking funding to be able to expand our team and our work.