PAG assessment: Proposal for Reclassifying the Balinese Script

Markus Scherer & other PAG members
2025-oct-06

The Unicode Technical Committee's Properties and Algorithm working group (PAG) has received and reviewed document L2/25-218 "Proposal for Reclassifying the Balinese Script".

First, we recognize that the Balinese language is actively spoken by several million people, and that the Balinese writing system is historically valuable and being promoted again for contemporary use in public signage and public education.

Second, we would like to point out that the Balinese script has been encoded in Unicode since version 5.0 (2006), which forms the basis for fonts, keyboards, and general widespread support for the script.

Third, the classification under discussion is one in Unicode Standard Annex #31 (UAX #31), "Unicode Identifiers and Syntax". This is not a judgment call on the general viability or value of a script or language; this annex is not about "normal text". Instead, this part of the annex makes recommendations for use of characters and scripts in technical contexts, for example in names of variables in computer source code, or in internet domain names. The technical jargon for such names is "identifiers". The Default Identifier Syntax in UAX #31 allows letters of all scripts, including Balinese.

The annex discusses possible "profiles" which expand or reduce the set of characters that are used in identifiers. The section "Specific Character Adjustments" focuses on "more natural-language identifiers", adding some characters, but also suggesting three groups of scripts based on (a) whether a script is in widespread common use, and (b) whether there are issues with the use of a script in a security-sensitive context where it should be easy for a user to uniquely recognize an identifier.

The scripts that are "recommended" for this specific type of use are scripts that are in widespread common everyday use and that have also been fully analyzed for whether letters can be confused with each other or present other security concerns. Scripts that are not in customary modern use or that have architectural issues are recommended for exclusion from identifiers. In addition, even "recommended" scripts include letters that are not in common everyday use or have security concerns, and are not recommended for default use in identifiers.

In between those two classes of scripts are currently about 35 scripts which "are in more limited use"; UAX #31 suggests that "to avoid security issues, some implementations may wish to disallow the limited-use scripts in identifiers". In other words, in a context of technical use in "identifiers" which are also subject to security considerations, Unicode lists some scripts that implementers may allow or disallow based on their own assessment. The Balinese script is one of these "limited-use scripts in identifiers".

One such implementer is ICANN, which supports only "recommended" scripts for top level domain names (e.g., .org or .みんな ["everyone"]), but formally endorses selected "limited-use" scripts in Reference LGRs

(rules for valid domain names) for the Second level (e.g., **fußball**.de or ตามใจตามจิ่ม.id). The set of scripts endorsed for second-level domain names is growing and includes Balinese.

The "Proposal for Reclassifying the Balinese Script" provides evidence for use of the script in historical texts, public signage, research and teaching materials, and cases similar to those. This is a good example of what in the context of UAX #31 "Unicode Identifiers and Syntax" is a script "in more limited use".

Document L2/24-019 "Requirements and Process for Changing Script Status for Identifier Use" provides details for what evidence we expect to see for this kind of classification.

For example, we would be looking for the *body text* of publications written and commonly read in the Balinese script, rather than merely use of alternate versions of Latin-script names and titles. We would be looking for evidence of an active community of users authoring and reading online forum content such as a Wikipedia in Balinese script. We would be looking for evidence of widespread usage in social media, such as Facebook posts, and YouTube descriptions, captions and comments. We would welcome evidence of commonly used school materials in Balinese script for topics other than teaching the script.