

Release Management Group Report to UTC #175

April 19, 2023

Peter Constable & UTC Release Management Group

The Release Management Group (“RMG”) is composed of UTC working group leaders and other volunteers working on the process for development of the Unicode 15.1 release and for evolving processes for future Unicode Standard releases.

Unicode 15.1 timeline

A timeline for the development and release of Unicode 15.1 was proposed in [L2/22-270](#) and confirmed by UTC #173 (cf. [173-C1](#)). The following is a summary of key dates:

- 2023-4-27, UTC #175: Finalize beta content
- 2023-5-23: start of beta review
- 2023-7-4: end of beta public review
- 2023-7-27: UTC #176: Finalize 15.1 content
- 2023-9-12: Unicode 15.1 release

Unicode 15.1 alpha and beta

With help from EDC, PAG and others, RMG published the alpha release on schedule. There were few 15.1-related tasks for RMG during the public review period, which allowed time to work on release-process transition.

The beta review period will begin four weeks from the start of UTC #175. Since there is more content to be published for the beta than for the alpha, there will be more work required to prepare the beta for publication. RMG has begun reviewing the beta-prep task list in preparation for this phase.

RMG recommends that UTC take the following actions:

Consensus: The UTC authorizes starting the beta review for Unicode version 15.1.

Action Item for Ken Whistler, PAG: Prepare an updated NamesList.txt for Unicode 15.1 repertoire, as finalized during UTC #175.

Action Item for Michel Suignard, Rick McGowan, RMG: Prepare a set of Unicode version 15.1 beta review code charts for posting.

Action Item for Ken Whistler, PAG: Prepare an updated and complete set of data files for the UCD, and the data directories for UTS #10, UTS #39, and UTS #46, for the beta review of Unicode version 15.1.

Action Item for Ned Holbrook, Emoji Subcommittee: Prepare an updated and complete set of data files for the data directory of UTS #51, and a complete set of emoji beta review charts for Emoji 15.1.

Action Item for Peter Constable, RMG: Coordinate the publication of the Unicode 15.1 beta review.

Action Item for Ken Whistler, RMG: Prepare a background document for a PRI on the Unicode 15.1 beta review.

Action Item for Rick McGowan, RMG: Post the PRI for the Unicode 15.1 beta review, to close July 4, 2023.

CJK Extension I Considerations

RMG discussed Ken Lunde's proposal ([L2/23-106](#)) for a set of urgently needed characters (UNC) to be encoded as *CJK Extension I*, which aims to provide China with an alternative to their current plan to encode characters in GB18030 that use reserved code points in plane 10. A couple of open questions are whether approval of the proposal by UTC will influence China to change their plans, and whether Unicode timing—15.1 versus 16.0—would be a factor in convincing China to change their plans.

RMG discussed tradeoffs of a 15.1 versus 16.0 release timing in relation to release management issues. From that perspective, the less risky alternative would be to progress CJK Extension I as part of the proposed repertoire for Unicode 16.0. That would provide plenty of time for final review and would enable normal process for the eventual update of all the related Unihan properties for the addition.

RMG recognizes, however, that there might be considerations that favor accelerating the publication of CJK Extension I into Unicode 15.1. L2/23-106 explains how this would be possible. An important consideration, from the release management perspective, would be to aim to get CJK Extension I immediately into the 15.1 beta review. However, we would advise the UTC to establish special contingency expectations for the status of CJK Extension I in such a case. Ordinarily, once we hit the beta review stage, we fix all the code points and character names in the repertoire, and do not add or withdraw from that repertoire before publication. But given the speed of processing and the potential risk of adding 603 CJK characters to the repertoire this late in the 15.1 cycle, it would be expedient to allow for the potential postponement of the CJK Extension I to 16.0, even if we have already included it in the 15.1

beta review, if the additional speed up in publication schedule turns out not to be required in order to help resolve the status of these additions for GB 18030.

Processes: progression of new script and character repertoire

Tracking tasks

UTC #174 agreed with RMG’s recommendation to adopt a practice of *provisional assignment of code points for new characters*. The intent has been to establish (relatively) stable code point assignments without yet committing to a release version for the new characters, providing flexibility to distribute tasks to prepare final content across available volunteers.

With that decision in place, RMG recognized the need to track the various tasks and deliverables needed to prepare content for release—charts, data, core spec, etc. To that end, a *Script Checklist* has been prepared. This is a GitHub issue template with a boilerplate checklist for tasks and deliverables, starting with SAH (or CJK) review of an encoding proposal through to integration in a release. For each encoding proposal brought to UTC, a separate issue is created using the Script Checklist template, allowing progress on each unit of new repertoire to be tracked.

<input type="checkbox"/>	<input checked="" type="radio"/>	Script checklist: ARABIC BIBLICAL END OF VERSE	scripts	#47 opened last week by dwanders-A	3 of 22 tasks
<input type="checkbox"/>	<input checked="" type="radio"/>	Script checklist: Latin: S with Diagonal Stroke	scripts	#45 opened last month by dwanders-A	7 of 22 tasks
<input type="checkbox"/>	<input checked="" type="radio"/>	Script checklist: Cyrillic letter tje for Khanty	scripts	#44 opened last month by dwanders-A	8 of 22 tasks
<input type="checkbox"/>	<input checked="" type="radio"/>	Script checklist: Latin Capital Rams Horn	scripts	#43 opened last month by dwanders-A	7 of 22 tasks
<input type="checkbox"/>	<input checked="" type="radio"/>	Script checklist: Kawi NUKTA	scripts	#42 opened last month by dwanders-A	10 of 22 tasks
<input type="checkbox"/>	<input checked="" type="radio"/>	Script checklist: Balinese punctuation characters	scripts	#41 opened last month by dwanders-A	8 of 22 tasks
<input type="checkbox"/>	<input checked="" type="radio"/>	Script checklist: Kannada and Telugu Archaic Ligature Shrii	scripts		

A similar checklist is also being considered for tracking tasks related to glyph changes for already-encoded characters.

Seeding data

Once code points have been provisionally assigned to new characters, there are separate work threads that need to begin and that will involve different working groups. These include the following:

- Update pipeline (RMG)
- Chart development (Charts Group)
- UCD development (PAG)
- Core spec (SAH, EDC)

One thing needed in common across these threads is a minimum set of data that includes (i) character names, (ii) code points, and (iii) encoding status.

RMG has begun to explore possibilities for maintaining this base data in a common way that might serve to seed different work threads without duplication of data. As an initial exploration, Liang Hai and Ken Whistler have begun prototyping new tooling that can be used to maintain the Pipeline page in a more automated manner using base character data. If any new characters are provisionally assigned code points during UTC #175, that will provide an early test case for exercising the prototype process.

If this Pipeline prototype is promising, we will consider integration of any tools or data into other work threads that might have potential for gains in efficiency.

Processes: tracking UTC action items

For many years, Ken Whistler has maintained the data for tracking UTC action items, with help from Rick McGowan to publish reports. A year or more ago, Markus Scherer, Josh Hadley and other contributors did work in parallel to put UTC AI data into an online spreadsheet (shared with RMG members) with automation that enables different UTC working groups to get reports of AIs they own. For many months now, these two process threads have been going in parallel. While the spreadsheet automation has provided value for (at least some) working groups, it has also required a small increase in the work to maintain the AI data: Whenever the AI data changes, Ken also copies it into the new online spreadsheet.

On completion of UTC #175, Ken will work with other RMG members on creating tracking data for new action items, copy them into the online spreadsheet, and then stop updating the AI data in the old location. This will reduce the maintenance burden by eliminating some or all of the older workflows and replacing them with more automation.

Going forward, data for new AIs and updates to existing ones will only be done in the online spreadsheet. Rather than Ken managing all UTC AIs, the chairs and vice chairs of the UTC working groups (EDC, PAG, ...) will be expected to maintain the AIs relevant to their groups.