

**Universal Multiple-Octet Coded Character Set  
International Organization for Standardization**

**Doc Type:** ISO/IEC JTC1/SC2/WG2/IRG

**Title:** Request for IDS Data

**Source:** Unicode Consortium

**Status:** Liaison Contribution

**Action:** Request to the IRG

**Date:** 2015-11-03

The [Unihan Database](#) is a well-known and broadly used resource that provides to researchers, scholars, software developers, and others a rich set of information that is associated with each CJK Unified Ideograph, such as source references, radical/stroke data, various readings, variant relationships, and so on. The Unihan Database represents ongoing work by a team of experts, in terms of extending its scope to accommodate new extensions, and also to make adjustments and corrections for existing characters.

IDSes (*Ideographic Description Sequences*), which serve to loosely describe the components and structure of CJK Unified Ideographs, and which can also be leveraged to flag potential duplicate characters in new submissions, are curiously absent from the Unihan Database.

The Unicode Consortium requests that the IRG supplies a complete set of IDS data from their master database for inclusion in the Unihan Database. Of course, the IRG would be credited accordingly.

That is all.