

# INCITS: Character Encoding Enables Electronic Access to Text in Languages Across the Globe

## Background

Computers store letters and other characters by assigning a number for each one. In the 1970s and 1980s, a variety of custom-built fonts and platform-specific character sets were developed in an attempt to achieve this requirement but no single encoding system was adequate to cover all the letters, punctuation, and technical symbols in common use in all the languages world-wide. By the mid-1980s, industry experts began investigating the practicalities of creating a universal character set.

## Problem

Because antiquated standards and fonts handled letters and symbols in different ways, users couldn't be sure the text they sent was being properly received. And encoding systems also conflicted with one another. For example, two encodings could use the same number for two different characters or use different numbers for the same character, and data passed between different encodings or platforms was at risk of corruption. The result was chaotic for businesses, governments, librarians, academics, and others, and led to the idea of a single standard for character encoding.

## Approach

In the late 1980s, work on a single character encoding standard was launched by two separate entities: the [International Organization for Standardization](#) (ISO), and a group that coined the term "unicode" and would later develop into the [Unicode Consortium](#). The efforts of the two groups eventually resulted in the creation of [ISO/IEC JTC1/SC2/WG2](#), *Universal Coded Character Set*, and the merging of their character repertoires in 1993, which appeared as ISO/IEC 10646-1 and Unicode Standard 1.1, *Universal Coded Character Set (UCS)*.

## Outcome

Today, both [ISO/IEC 10646:2014](#), *Information technology - Universal Coded Character Set (UCS)*, and the [Unicode Standard](#) are used for encoding multilingual text for the exchange of data internationally, are used on over 80% of webpages, and are found on virtually all modern computer systems and devices. These standards are the result of an extremely successful partnership between ISO, the Unicode Consortium, and [INCITS](#) (the InterNational Committee for Information Technology Standards), who serves as the ANSI-accredited U.S. Technical Advisory Group (TAG) Administrator to JTC 1/SC 2, *Coded Character Sets*.

The groups work in tandem to approve new character additions so the standards remain synchronized. Both the ISO committee and the Unicode Consortium regularly receive proposals from their members and the public identifying characters that are not in the international character encoding standard, but are needed by users. The proposals are carefully

reviewed by both groups to ensure that the characters are eligible and enough information is provided in the proposals to allow implementation in software and fonts. For the Unicode Consortium, having character proposals reviewed and accepted by ISO members provides a useful check on the accuracy of such proposals, and helps to ensure involvement from countries where such characters are used. For ISO, review of proposals by Unicode members provides assurance that technical aspects have been adequately addressed. After the proposals have been approved by the Unicode technical committee and have successfully completed the ISO balloting process, the characters are published and can be implemented by industry, governments, and others.