

ISO/IEC JTC 1 "Information technology"

Secretariat: ANSI

Committee manager: Rajchel Lisa Mrs.



Revised SC 2 Business Plan 2023

Document type	Related content	Document date	Expected action
Project / Other	Meeting: Berlin (Germany) 13 Nov 2023	2023-10-18	COMMENT/REPLY by 2023-11-13

Replaces: N 16482 SC 2 Business Plan 2023

Description

This document is circulated for review and consideration at the November 2023 JTC 1 Plenary.

STRATEGIC BUSINESS PLAN FOR JTC 1/SC 2

PERIOD COVERED: Sept. 2022 – Sept. 2023

SUBMITTED BY: Masaru TAKECHI, Chair of JTC 1/SC 2

Note: This draft strategic business plan is prepared in accordance with new template for strategic business plan for JTC 1 subcommittee. Yellow highlighted description shows the intended content.

1.0 Executive Summary

Scope of the SC: Standardization of graphic character sets and their characteristics, including string ordering, associated control functions, their coded representation for information interchange and code extension techniques. Excluded: audio and picture coding.

While the number of characters standardized in ISO/IEC 10646 increases from 32,884 to more than 143,000 between its first edition to the latest, the need to cover more characters is still strong in accordance with the growth of the information technology and its usage on various devices, better representations of information, and coverage of language/dialect/script.

SC 2 experts have been working hard to develop Universal Coded Character Set, ISO/IEC 10646, and related standards which are the fundamental basis of Information systems and being referred from every standard involving character-based information exchange and/or processing.

SC 2 will keep quality and speed of its work to develop standards for character. And close cooperation between SC 2 and other all committees to develop standards involving character codes is important.

2.0 Environment

The areas of application of connected device are rapidly spreading. And in many cases, character codes are used to convey meaning in information. Mobile devices which exist close to human need more characters to convey information and emotion. As the information technology grows, more characters on the systems are required for various purposes and usages.

For recent about 30 years, the number of characters in the standard has been constantly increasing and has reached as much as more than four times of its first edition. 32,884 characters are coded in the first edition of ISO/IEC 10646-1:1993. The latest edition, ISO/IEC 10646:2020 (Ed.6), standardized over 143,000 characters. To provide better communication, information representation over IT implementations, and wider applicability of Information Technologies for various people, industries, and societies, the need to use wider variety of characters is still strong.

We also need to recognize potential requests from user groups of minority and historic scripts are also strong. In these days, almost all scripts for currently used major and national languages are already encoded. However, there are a huge number of dialects and minority languages, which we should work with.

3.0 Benefits

Coded character sets and their orderings are basic infrastructure for all information and communication technologies. Deliverables of SC 2 offer consistent and harmonized basis of text

representation for human activities using information and communication technologies. However, to use deliverables of SC 2 more effectively, we would like to emphasize the importance that every technology/standard involving interaction with human through characters should be developed under awareness of cultural and linguistic diversity on the earth and interoperability over them.

4.0 Participation and cooperation/collaboration.

From the viewpoint of the active work items, SC 2 has a few work items. However, the number of P-members, O-members, and related organizations are great in number. The number of current P-member National Bodies is 24 and O-member National Bodies is 26. These numbers show continuously active participating National Bodies and are unchanged since the last plan was reported. There are also several, but not many, invited guests in WG meetings and plenary meetings from developing countries, which have no official membership, but have script expertise.

SC 2 has IRG (Ideographic Research Group) under the control of WG 2. This Research group focuses its work on Eastern Asia's ideographic characters, i.e., Han-characters. The participating countries are not limited to P and O members of SC 2, but other related countries and areas are also actively participating as liaison members or guests, i.e. TCA ,HKITF(Category C liaison), and Macao as a guest.

SC 2 is the key organization in the area of coded character set standardization, and has official liaisons with the following organizations. There are no competitive international standards or standardization organizations. The Unicode Standard is a related industrial standard for character encoding and that SC 2 and The Unicode Consortium have been working cooperatively for over twenty years to develop ISO/IEC 10646 and The Unicode Standard in a compatible and synchronous manner.

Internal Liaisons:

Reference	Title
ISO/IEC JTC 1/SC 34	Document description and processing languages
ISO/IEC JTC 1/SC 35	User interfaces
ISO/TC 37/SC 2	Terminology workflow and language coding
ISO/TC 46/SC 4	Technical interoperability
ISO/TC 204	Intelligent transport systems

External Liaisons

Acronym	Title	Category
ISOC	Internet Society	A
ITU	International Telecommunication Union	A
UNICODE	The Unicode Consortium	A
CCSDS	Consultative Committee for Space Data Systems	B
EC – European Commission	European Commission	B
UNCTAD	United Nations Conference on Trade and Development	B
UNECE	United Nations Economic Commission for Europe	B
WIPO	World Intellectual Property Organization	B
WMO	World Meteorological Organization	B
HKITF	Hong Kong Information Technology Federation	C
SAT Daizōkyō Text Database	SAT Daizōkyō Text Database	C
TCA	Taipei computer association	C
UC Berkeley	University of California, Berkeley	C
W3C	World Wide Web Consortium	C

5.0 Objectives and strategies to achieve them

The objective of SC 2 is to enable all the intellectual activities that humankind performs with the characters, including cultural and linguistic aspects, on information technology in a harmonized manner. It allows the IT and other industries to support various languages and scripts in a cost-effective way. To achieve this, SC 2 has to work always keeping in mind the direction given in Annex C.4 of Consolidated JTC 1 Supplement 2018, “cultural and linguistic adaptability” that is the point to be clarified on proposing standards. For the development of the standard, close cooperation between SC 2 and committees developing standards involving character code is very important.

In addition, SC 2 should focus in the following issues;

- 1) Quick and precise standardization of newly proposed characters and scripts, especially proposals from developing countries, user groups of minority and historic scripts.

Note: SC 2/WG 2 has its own guideline to accelerate standardization work and make the criteria

of standardization clear to all experts and user communities as “Principles and Procedures for Allocation of New Characters and Scripts and handling of Defect Reports on Character Names” (SC 2 N 4318).

- 2) Maintaining consistency with countries' and areas' standards.
- 3) Maintaining consistency with related standards which refer to standards developed by SC 2.
- 4) Establish relationship with real user group of targeted scripts and characters.
- 5) Seeking efficient standardization process.
- 6) Secure, efficient and user-friendly publication.

Besides, SC 2 also foresees following opportunities.

- 1) Expansion of usage in technical areas such as Bigdata, IoT, Programming and Scripting Languages, Internet, e-Government, etc., and in a very broad global business application environment that positively impacts developed, as well as developing, countries such as the U.S., Japan, China, Cambodia, Ethiopia, and many others.
- 2) Consolidation and harmonization of huge, coded character sets.
- 3) Infrastructure for improvement of information and communication technology in developing countries, areas and minority scripts users.
- 4) Support as ICT environment for vast area of academic research.

6.0 Factors affecting completion and adoption of the work program

- 1) **RISK:** Possible criticism from cultural, political concerns. As the topics of SC 2 closely relates with identity and dignity of user communities, if the standardization process is not represented by appropriate experts from the user community, or if appropriate input is not given to the standardization process, the standards produced may have a risk of criticism from user communities.

SOLUTION: Effort to enhance openness of discussions and to establish relationship with the user communities with the cooperation of other international organizations, governments and academic research institutes.

- 2) **RISK:** Misalignment of synchronization with other standards which closely refer UCS.
STRATEGY: Promote quick publication of standards, together with the information disclosure of newly standardized scripts and characters.

- 3) **RISK:** Possible deviation of related national standards from ISO/IEC 10646. National standards that reference or are referenced by SC 2 standards may change in ways that create incompatibilities.

SOLUTIONS: The mitigation strategy is to encourage NBs to communicate to SC 2 any plans to make changes to national standards that may affect interoperability and to consider input from other SC 2 members how risks to interoperability between national and international standards might best be avoided.

- 4) **RISK:** Ad hoc solution to the requests from other SCs and standardization organizations outside JTC 1 which harm the consistency of ISO/IEC 10646 itself.

SOLUTIONS:

Welcome the requests from other SCs and other standardization organizations.

Establish close relationship with the requesters and strive to recognize the actual requirements. Seek solutions which will not harm the consistency of the standard and satisfy the requesters' needs as the experts.

- 5) **RISK:** Confrontation between different expert/user communities of scripts to be encoded in UCS. Such situations prevent the progression of developing work.

SOLUTIONS:

It is not so easy to let different positions to be compromised. However, the effort to provide the occasion for discussion is very important.

- 6) **RISK:** Keeping consistency with ISO's policy of publishing which doesn't assume hugeness and format required for character code standards.
- 7) **RISK:** SC2 and working groups under it often need long time discussion of impalpable nuances of visual representations of characters with attendants from all over the world. It makes virtual meetings virtually impossible.

7.0 Structure, current projects, and publications

7.1 Structure

Structure of SC 2 is as follows.

- ISO/IEC JTC 1/SC 2/WG 2 Universal coded character set
- ISO/IEC JTC 1/SC 2/WG 2/IRG Ideographic Research Group

7.2 Current projects

Current Projects in SC 2 are listed below.

- ISO/IEC 10646:2020/CD Amd 2.3, Information technology — Universal coded character set (UCS) — Amendment 2
- ISO/IEC CDTR 2375, Information technology – Registered escape sequences and coded character sets
- ISO/IEC CD 14651 (Ed. 7), Information technology — International string ordering and comparison — Method for comparing character strings and description of the common template tailorable ordering

7.3 Current publication

Current publications of international standard are listed on the web site about the standards by [JTC 1/SC 2¹](#). The publication published during the past one year is listed below.

- ISO/IEC 10646:2020/ Amd 1:2023, Information technology — Universal coded character set (UCS) — Amendment 1: Information technology — Universal coded character set (UCS) —

¹ <https://www.iso.org/committee/45050/x/catalogue/p/1/u/0/w/0/d/0>

Amendment 1: CJK Unified Ideographs Extension H, Vithkuqi, Old Uyghur, Cypro-Minoan, and other characters