ISO/IEC JTC 1 N9698

2009-09-10

Replaces:

ISO/IEC JTC 1
Information Technology

Document Type: business plan

Document Title: SC 2 Business Plan, October 2008-October 2009

Document Source: SC 2 Chairman

Project Number:

Document Status: This document is forwarded to JTC 1 National Bodies for review and consideration at the October 2009 JTC 1 Plenary meeting in Tel Aviv.

Action ID: ACT

Due Date:

No. of Pages: 6

BUSINESS PLAN FOR JTC 1/SC 2

PERIOD COVERED: October 2008 – September 2009

SUBMITTED BY: Tatsuo L. Kobayashi, Chair of JTC 1/ SC 2

1.0   MANAGEMENT SUMMARY

1. 1   CHAIRMAN'S REMARKS

In accordance with the 23rd Plenary Resolution 34 – African Languages and Keyboard Layouts, SC 2 prepared a survey report on African Languages and ISO/IEC 10646, with the kind cooperation of SC 2/WG 2 experts, especially Dr. Deborah Anderson, SEI, UC Berkeley (JTC1/SC2/WG2 N3641). The document is attached at the bottom of this report.

Based on this report, SC 2 recognizes that almost all of the generally known languages are already covered by ISO/IEC 10646. However, SC 2 has to continue to reach out to seek further requirements from minority language communities, especially in the African continent.

Note)

Resolution 34 – African Languages and Keyboard Layouts
Based on the results of the Technology Watch session at the 23rd Plenary, JTC 1
• instructs its Secretariat to distribute the presentation on "Writing African Languages: The Keyboard Issue" to SC 35/WG 1 for further consideration and
• requests SC 2 to check whether all characters needed for the representation of the noted African languages are indeed contained in ISO/IEC 10646 and address any omissions.
Unanimous

After the last JTC 1 plenary, there happened to appear two very important market products in terms of SC 2 standard implementation. One is Adobe Acrobat 9, and the other is Microsoft Windows 7. These two market products silently implemented the Ideographic Variation Sequence (IVS) technology of ISO/IEC 10646 and Unicode.

The core technology of this functionality had been standardized more than ten years ago, i.e. variation selector. The original idea of this mechanism was to harmonize market requirements to distinguish slight differences and Han unification principle in ISO/IEC 10646. This mechanism has been adapted to several scripts and symbols, such as Old Mongolian and Mathematical Symbols. However, adaptation to Han characters had been put on hold. The reason was the difficulty and complexity of shape variations of Han characters. If the user communities are different from each other, the requirements are also different from each other. There has been long discussion on this issue. And there has been no harmonized solution.

Our last solution is to give up on developing consolidated IVS sequences, and to develop a new registration procedure only to guarantee the uniqueness of each user community's IVS sequences to be registered.

Based on this solution, Adobe and MS started to implement this mechanism in their products, and now is the evaluation stage to elaborate the implementations.

This case reminded us again that the standardization is not the goal, but the beginning of the implementation and actual usage, and sometimes it takes a long time to be implemented and start to serve the actual users.

1.2   JTC 1 SC 2 STATEMENT OF SCOPE

Title: Coded Character Sets
Scope: 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.

1.3 PROJECT REPORT

The Programme of Work of SC 2 is available at: http://std.dkuug.dk/jtc1/sc2/open/pow.htm

Ongoing Projects:

- Project 02.10646.00.06 (10646/Amd 6), Information technology -- Universal Multiple-Octet Coded Character Set (UCS) -- AMENDMENT 6: Bamum, Javanese, Lisu, Meetei Mayek, Samaritan, and other characters (Current status: FDAM)

- Project 02.10646.00.07 (10646/Amd 7), Information technology -- Universal Multiple-Octet Coded Character Set (UCS) -- AMENDMENT 7: Mandaic, Batak, Brahmi, and other characters (Current status: FPDAM)

- Project 02.10646.00.08 (10646/Amd 8), Information technology -- Universal Multiple-Octet Coded Character Set (UCS) -- AMENDMENT 8: Additional symbols, Bamum supplement, CJK Unified Ideographs Extension D, and other characters (Current status: PDAM)

- Project 10646.00.00.00.02 (Consolidation of ISO/IEC 10646: 2003 and Amendment 1 – 7 into a new Edition), Information technology -- Universal Multiple-Octet Coded Character Set (UCS) (Current status: CD)

- Project 14651.00.01.00.02, Information technology -- International string ordering and comparison -- Method for comparing character strings and description of the common template tailorable ordering -- Amendment 1 (Current status: FDAM)

- Project 14651.00.02.00.02, Information technology -- International string ordering and comparison -- Method for comparing character strings and description of the common template tailorable ordering -- Amendment 2 (Current status: WD)

Total Number of Projects: 61 (including subprojects)

New project/subproject: 2

Withdrawn project: 0

1.4 CO-OPERATION AND COMPETITION

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.

Internal Liaisons:

ISO/IEC JTC 1/SC 22	Programming Languages, their Environments and System Software Interfaces
ISO/IEC JTC 1/SC 29	Coding of audio, picture, multimedia and hypermedia information
ISO/IEC JTC 1/SC 34	Document description and processing languages
ISO/IEC JTC 1/SC 35	User Interfaces
ISO/TC 37/SC 2	Terminographical and lexicographical working methods
ISO/TC 46/SC 4	Information and documentation – Technical interoperability
ISO/TC 211	Geographic information/Geometrics

External Liaisons

IETF/ISOC	Internet Society	A
ITU-T	International Telecommunication Union - Telecommunication Standardization Sector	A
CEC	Commission of European Communities	B
UNCTAD	United Nations Conference on Trade and Development	B
UN-ECE	United Nations Economic Commission for Europe	B
WIPO	World Intellectual Property Organization	B
WMO	World Meteorological Organization	B
TCA	Taipei Computer Association	C
UC Berkeley	UC Berkeley	C
UNICODE	The Unicode Consortium	C
W3C	World Wide Web Consortium	C

SC 2 has IRG (Ideographic Rapporteur Group) under its WG 2. This Rapporteur 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. Taiwan has been participating through TCA (Category C liaison), and Macao as a guest. At the last IRG meeting, Malaysia (O member) attended for the first time and contributed a couple of its unique Ideographs to be included in UCS.

Recently, IRG began a preliminary research work for encoding of Old Hanzi characters. In this activity, several experts from universities and academic institutes such as Beijing Normal University, East China Normal University, Ibaraki University, National Hualien Teachers College, are involved.

SC 2 has been continuously co-working with the Unicode Consortium from the first stage of the development of ISO/IEC 10646 for more than fifteen years. The Unicode Consortium was recently assigned as an Approved RS Originator Organization (ARO) of JTC 1. This assignment makes normative reference of not only the Unicode Standard, but also ISO/IEC 10646 itself, very easy for other JTC 1 standards.

SC 2 also has the established C liaison between its WG 2 and UC Berkeley to develop particular minority and historic scripts. Besides these official relationships, SC 2 has active and close relationships with several academic institutions, such as Tokyo University of Foreign Studies.

2.0   PERIOD REVIEW

2.1   MARKET REQUIREMENTS

Coded character sets and their orderings are basic infrastructure for all information and communication technologies.

Lately, because of the rapid spread of information technologies, especially Internet technologies, UCS is widely used throughout the entire world. The importance of universal coded character set is acknowledged among governmental sectors, industrial sectors and open source communities.

The number of standards which refer to UCS and the number of actual implementations based on UCS are increasing rapidly. In these circumstances, requests to keep consistencies between UCS and referencing standards are increasing. These requirements come not only from other SCs but also from standardization organizations outside JTC 1.

On the other hand, potential requests from user groups of minority and historic scripts are still 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. Some of them are in danger of extinction. Moreover, the user groups of these languages are mostly poor and have few chances to satisfy their requirements.

Also, even if some of the minority scripts are standardized, it takes very long time to be implemented in operation systems and environment. The standardization activity is not the goal but the start line for

their actual usage.

2.2   ACHIEVEMENTS

The following projects have been published or progressed to the publication stage during this period.

- Project 02.10646.00.05 (10646/Amd 5), Information technology -- Universal Multiple-Octet Coded Character Set (UCS) -- Amendment 5: Tai Tham, Tai Viet, Avestan, Egyptian Hieroglyphs, CJK Unified Ideographs Extension C, and other characters

- Project 02.14651.00.01.00.02 (14651/Amd. 1), Information technology -- International string ordering and comparison -- Method for comparing character strings and description of the common template tailorable ordering -- Amendment 1

2.3   RESOURCES

From the view point 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. In this sense, SC 2 is one of the largest and most active SCs in JTC 1. The number of current P-member National Bodies is 28 and O-member National Bodies is 21. 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 and its WG 2 have assigned officials, and all developing projects also have assigned officials. SC 2 has sufficient resources.

3.0   FOCUS NEXT WORK PERIOD

3.1   DELIVERABLES

Amendment 6 and 7 to ISO/IEC 10646: 2003 and the 2nd Edition of ISO/IEC 10646 will be published.

3.2   STRATEGIES

SC 2 should focus in the following five 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 3921).
2) Synchronization of 14651 to 10646.
Note: 14651 has been developed and maintained by SC 2 directly. Practical editing work is being done by the editing group, OWG-SORT, operated in accordance with JTC 1 directives. Also, Canadian national body kindly has been taking responsibility for its French version, in accordance with the ISO/IEC directives.
3) Maintaining consistency with countries' and areas' standards.
4) Maintaining consistency with related standards which refer to SC 2's standards.
5) Establish relationship with real user group of targeted scripts and characters.

3.2.1 RISKS

1) RISK: Possibility of standardization without feedback from the real user community.
SOLUTION: Effort to establish relationship with the real user community with the cooperation of other international organizations, governments and academic research institutes.

2) RISK: Delay of synchronization of 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: Contradictions between international standards and country standards when referencing each other.
SOLUTIONS:

- Close relationships between SC 2 and each national body.
- Maintenance of the mapping information between international standards and local standards.

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.

3.2.2 OPPORTUNITIES

1. Expansion of usage in technical areas such as XML, 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.

3.3   WORK PROGRAMME PRIORITIES

All working programs have to be developed simultaneously, and ISO/IEC14651 should catch up the modification and additional repertoires of ISO/IEC 10646 as quickly as possible.