



வா.மு.சே.கவிஅரசன்
சேது கணிணி மையம்.

Va.Mu.Se. Kaviarasan
Sethuinc
Annai Sethu Classical Tamil Institute

वा. मु. से. कवि अरसन
सेतु इन्क्

Subject: Response to Govt.of India Proposal L2/10-426

To

Date: January 7, 2011

Dr. Lisa Moore

Chair,

Unicode Tech Committee,

Unicode Consortium,

U.S.A.

Dear Dr. Lisa Moore

I am writing this letter to you as our response to Government of India proposal L2/10-426, dated 18th October 2010.

While we are in agreement to represent only the unrepresented archaic grantha characters in Unicode SMP, all other disputed characters of Govt. of India proposal need to be deferred, pending wider and in-depth considerations among various scholars and stakeholders.

Government of Tamilnadu (L2/10-464) dated 6th November 2010, has requested Government of India to defer consideration of this proposal, pending wider and in-depth considerations among various scholars and stakeholders. In line with above request we are requesting additional time from Unicode Consortium to hash out few details including possible security risks from phishing that can arise from the encoding found in the proposal L2/10-426, when SMP support is introduced in iDNS. We also need to understand the necessity and implication of entirely adding the letters from Tamil block in the new extended Grantha proposed in L2/10-426. When we can use the characters from different block in the same application, I am failing to understand the need of repeating the same characters, which may add additional complexities in sorting and text processing applications based on indic languages. In spite of the UCA algorithms available most of the products based on Unicode is not sorting properly (Refer Annexure III) and we are still trying to fix these issues in indic language. Adding a complexity at this stage will further delay the fixing of the existing issues in indic languages.

When SMP support is added in iDNS, If all the Tamil letters are represented in extended grantha block, சேது.வணி (0BC70B9A 0BA40BC1..... Tamil Block), In grantha block , if some writes சேது.வணி (11AD111AD211AD311AD4..... Grantha Block) This introduces



வா.மு.சே.கவிஅரசன்
சேது கணினி மையம்.

Va.Mu.Se. Kaviarasan
Sethuinc
Annai Sethu Classical Tamil Institute

वा. मु. से. कवि अरसन
सेतु इन्क्

(Refer UTR#36 for issues)

We have learned that there are quite a few proposals about encoding of Tamil characters in a new archaic Grantha block and encoding unused archaic block and hence we are writing this letter requesting the Unicode Consortium to delay taking a formal decision on proposals **L2/10-426 in order** to give ourselves sufficient time to understand and digest the proposals and their impact on the Tamil Block and on the users of Tamil Unicode. We have also reached out to the Government of India also to reconsider their recommendations on the above proposal.

A few points which seem to have been misrepresented in these proposals are as follows:

1. In Proposal L2/10-426, Item B-3 of the proposal summary form seems to misrepresent the category as "contemporary" whereas the true nature of the script is 'E-Minor extinct', since barring the few characters which are already encoded in Tamil block in BMP the rest of the grantha characters are not in use any more.
2. In Proposal L2/10-426, the responses to Item C-2a and C-2b regarding contacts made with the user community are of questionable value. In the list of contacts, we do not find a single major Tamil institution or its official representative even though Tamil language has a plethora of official bodies to contact on this serious matter such as the Tamilnadu government's branches, Tamilnadu universities and institutions such as the Central Institute of Classical Tamil or the International Institute of Tamil Studies. On the computing side, reputable institutions such as INFITT should have been contacted. But not so. Almost all the scholars listed in C-2b are scholars of Sanskrit, another Classical language of India, and of Devanagari script which is used to write Sanskrit. While some of them may have used Grantha script here and there, the major focus of these scholars seems to be Sanskrit and Devanagari. In summary the contacts claimed to have been made therein are not in the true spirit of diligent discovery. It can be seen by the complete surprise with which the Tamil user community at large has been taken by the news of proposal and the attendant political and social controversy it has created since then.
3. In Item C-4b, "the context of use for the proposed characters" has not at all been answered directly as one of "common" or "rare". It should have been answered as "rare". It is used as an alternative script to Devanagari by very few to write Sanskrit. All the Grantha characters necessary to write almost all of the "Tamil Manipravala style (மணிப்பிரவாளம்)", are already found in Tamil Unicode 0B9C, 0BB5, 0BB6, 0BB7, 0BB8 and 0BB9. Most of the Tamils who studied Tamil at elementary school level will understand that these are not Tamil characters. Respecting Unicode's stability policy, INFITT is working on a proposal to add an annotation to these characters to clarify them as Grantha Characters, although these characters truly do not belong to Tamil block as per Tamil linguists. This annotation should



வா.மு.சே.கவிஅரசன்
சேது கணிணி மையம்.

Va.Mu.Se. Kaviarasan
Sethuinc
Annai Sethu Classical Tamil Institute

वा. मु. से. कवि अरसन
सेतु इन्क्

help to remove these six characters from being again duplicated in the new proposal. In any case, Manipravalam, from the script point of view, is the product of switching between two different scripts and is not the product of a single integrated script and we can accomplish the same by keeping the full-blown Grantha characters as a separate script in Unicode from Tamil Unicode and let the user switch in between these two scripts in her text processor just as this very document has between Roman and Tamil scripts in composing this paragraph. Anyone can write an editor that allows easy switching between Unicode Tamil and Grantha. The editor might have text entered in Roman and then converted. It is easy to switch from one Unicode character set to another and this would also make it simple to use Grantha glyphs for Sanskrit sounds."

4. Item C-5, asking whether the proposed characters are in current use by the user community, is best answered by a clear "NO", since a subset of the proposed characters in use by the Tamil user community is already encoded in the Tamil block (Refer to our explanation above to Item C-4b). Item C-5b seeking where the proposed characters are used seems to have been misrepresented. You can reach out to [Classical Tamil Apex Implementation Committee : http://en.cict.in/node/3](http://en.cict.in/node/3). Representing an inscription letter or character should not be termed as a contemporary usage of the script it should be treated as "historic and archaic".

5. Item C-6a has been answered with an inaccurate statement that the grantha is not a historic script. Barring most of the Tamil characters duplicate encoded and perhaps a few Malayalam characters included in the proposed block, all other characters are historic and archaic.

6. Items C-10a, 10b and 10c admits of similarities of some Grantha characters with Tamil and Malayalam characters. Whatever similarity is there between Grantha and Tamil is sufficient to cause significant security issues. Such duplicate representation of characters already represented in Tamil and perhaps in Malayalam block may lead to serious phishing issues when used in domain names. Sufficient study needs to be made before accepting/adding these characters again in the proposed new block, which may seriously impair the domain name users of Tamil and Malayalam, resulting in additional security problems. Please read further for a list of duplicate representation of characters.

7. List of some of the repeated Characters

- 11309 உ Grantha Letter U , Already represented in 0B29
- 1130A ஊ Grantha Letter UU , Already Represented in 0B89
- 1131C ஜ Grantha Letter JA, Already represented in 0B9C



வா.மு.சே.கவிஅரசன்
சேது கணினி மையம்.

Va.Mu.Se. Kaviarasan
Sethuinc
Annai Sethu Classical Tamil Institute

वा. मु. से. कवि अरसन
सेतु इन्क्

- 1131F ஸ Grantha Letter TTA – Font Style change , Already represented in 0BB1
- 11323 ண Grantha Letter NNA, Already represented 0BA3
- 11324 த Grantha Letter TA – Font style change, Already represented in 0BA4
- 11329 ன Grantha Letter NNNA, Already represented in 0BA9
- 1132A Grantha Letter PA,Already represented in 0BB5
- 1132F ய Grantha Letter YA, Already represented in 0BAF
- 11331 ற Grantha Letter RRA – Font style change, Already represented in 0BB1
- 11334 ழ Grantha Letter LLLA, Already represented in 0BB4
- 11335 வ Grantha Letter VA, Already represented in 0BB5
- 11336 ஶ Grantha Letter SHA, Already represented in 0BB6
- 11337 ஷ Grantha Letter SSA, Already represented in 0BB7
- 11338 ஸ Grantha Letter SA, Already represented in 0BB8
- 11339 ஹ Grantha Letter HA, Already represented in 0BB9
- 1133E Grantha Vowel Sign AA, Already represented in 0BBE
- 1133F Grantha Vowel Sign I, Already represented in 0BBF
- 11340 Grantha Vowel Sign II, Already represented in 0BC0
- 11341 Grantha Vowel Sign U, Already represented in 0BC1
- 11342 Grantha Vowel Sign UU, Already represented in 0BC2
- 11347 Grantha Vowel Sign EE, Already represented in 0BC6



வா.மு.சே.கவிஅரசன்
சேது கணிணி மையம்.

Va.Mu.Se. Kaviarasan
Sethuinc
Annai Sethu Classical Tamil Institute

वा. मु. से. कवि अरसन
सेतु इन्क्

- 1134B Grantha Vowel Sign OO, Represented as modifies for Grantha JA
- 1134C Grantha Vowel Sign AU, Represented as 0BCC
- 11357 Grantha AU Length Mark, Represented as 0BD7

8. Grantha is a script used in the ancient days to precisely write Sanskrit sounds. The new extended grantha seems to have "invented characters" which were not part of the ancient Grantha script

- 1130E Grantha Letter E
- 11312 Grantha Letter O
- 11329 Grantha Letter NNNA
- 11331 Grantha Letter RRA
- 11334 Grantha Letter LLLA
- 11346 Grantha Vowel Sign E
- 1134A Grantha Vowel Sign O

While I understand and support the need for representing old archaic forms of Grantha script which are not in contemporary use to understand old inscriptions, we need to follow the same principle used for all other world historic and archaic scripts. Instead of representing the actual Grantha found in historical documents and inscriptions, the proposal seems to have been written to introduce a new Grantha, based on Modern Tamil and ancient Grantha usage for representing Sanskrit, which could have been very easily accomplished with Devanagari script or by adding a few unrepresented archaic Grantha forms in SMP.

To avoid mixing the Tamil modern script with the Grantha character set with the potential for technical/security problems, it is best if the Grantha block is formally recognized by the Unicode Consortium as archaic and it had only hitherto-unrepresented archaic Grantha letters found in inscriptions/historical documents encoded in its independent block with no connection to any Tamil characters.

Should you need more clarifications feel free to reach me on my cell or via email.



வா.மு.சே.கவிஅரசன்
சேது கணிணி மையம்.

Va.Mu.Se. Kaviarasan
Sethuinc
Annai Sethu Classical Tamil Institute

वा. मु. से. कवि अरसन
सेतु इन्क्

Thanking you in anticipation.

Best Regards,

Kaviarasan Va.Mu.Se.

Cell: 1-614-517-4267,

Email : kavi@sethuinc.com

Web: www.sethuinc.com

CC

1. Honorable Chief Minister, Government of Tamilnadu, India
2. Honorable Minister of I.T., Government of India.
3. Dr. M. Rajendran, Vice-Chancellor Tamil University
4. Executive committee, INFITT
5. Dr. V.C. Kulandaisami Chairman, Tamilnadu Virtual Academy
6. Dr. M. AnandaKrishnan, Advisor, INFITT; Chairman IIT Kanpur

Annexure I: About Me

I am an I.T. Architect by profession and I am passionate about implementing various language tools, My involvement in language computing tools goes back to mid 1980s and involved in development of such tools from late 1980's. I am currently in an voluntary position in INFITT and in that capacity, I have implemented Unicode using standard IT tools. For a sample, Pl. refer <http://www.infitt.org/ed2008/ListMemberAll.php>. Since I have been employed in the past in various posts in Indian Government, I had gained knowledge of Hindi and the Devanagari script used to write Hindi, Sanskrit and other Indo-Aryan Languages. I use my language skills with my computing background to advise and develop products that suits Indian language computing market and how they can take advantage of Unicode without additional re-coding efforts. Kind of preventing the reinventing the wheel efforts. Eg., <http://www.sethuinc.com/bht>, <http://www.sethuinc.com/ta/ShowDet.php?id=149>, a page designed to use any Indic language, prototype in Tamil language.



வா.மு.சே.கவிஅரசன்
சேது கணிணி மையம்.

Va.Mu.Se. Kaviarasan
Sethuinc
Annai Sethu Classical Tamil Institute

वा. मु. से. कवि अरसन
सेतु इन्क्

Annexure II: Indic Database Application - Prototype in Tamil.

சேது தமிழ் பயிற்சிப் பட்டறை - 2 - Windows Internet Explorer

http://www.sethuinc.com/ta/ShowDet1.php?id=197

File Edit View Favorites Tools Help

Web Slice Gallery

சேது தமிழ் பயிற்சிப் பட்டறை - 2

Page Safety Tools

அன்னை சேது செந்தமிழ் நிறுவனம், அமெரிக்கா - ழ

Annai Sethu Classical Tamil Institute, America - ழ

Total Words Found : = 70
Requested by at Tuesday, November 02, 2010 18:14

சொல்	Letter	Picture
அன்னாசிப் பழம்	ழ	Picture
அழைப்பிதழ்	ழ	Picture
ஆழாக்கு	ழ	Picture
ஈச்சம் பழம்	ழ	Picture
உழவர்	ழ	Picture
உழவு	ழ	Picture
உழிதை	ழ	Picture
ஊழோடு	ழ	Picture

Done

Internet | Protected Mode: On

110%

7:14 PM



வா.மு.சே.கவிஅரசன்
சேது கணிணி மையம்.

Va.Mu.Se. Kaviarasan
Sethuinc
Annai Sethu Classical Tamil Institute

वा. मु. से. कवि अरसन
सेतु इन्क्

sethuinc elearning - Windows Internet Explorer

http://www.sethuinc.com/lan/ShowLet.php?id=108&id=1002

File Edit View Favorites Tools Help

sethuinc elearning X Google Google

Page Safety Tgols

Bharatiya Hindu Temple - Gurukul -

Total Letters Found : = 10
Requested by at Sunday, January 09, 2011 14:23

Click	Letter
	அ
	ஆ
	இ
	ஈ
	உ
	ஊ
	எ
	ஏ

Contains commands for working with the selected items.

Downloads Kavi_Response to G... http://www.unicod... Tamil Virtual Acade... sethuinc elearning ...

3:23 PM



வா.மு.சே.கவிஅரசன்
சேது கணிணி மையம்.

Va.Mu.Se. Kaviarasan
Sethuinc
Annai Sethu Classical Tamil Institute

वा. मु. से. कवि अरसन
सेतु इन्क्



Anneuxure III. Sorting Issues in Indic Languages : Case Study Tamil



வா.மு.சே.கவிஅரசன்
சேது கணினி மையம்.

Va.Mu.Se. Kaviarasan
Sethuinc
Annai Sethu Classical Tamil Institute

वा. मु. से. कवि अरसन
सेतु इन्क्

List of Registered Members of INFITT from INFITT database - Windows Internet Explorer

http://www.infitt.org/ed2008/ListMemberAll.php

File Edit View Favorites Tools Help

☆ Favorites ☆ Web Slice Gallery

List of Registered Members of INFITT from INFITT...

உத்தமம் INFITT International Forum for Information Technology in Tamil
உலகத் தமிழ்த் தகவல் தொழில்நுட்ப மன்றம்
International Forum for Information Technology in Tamil

Home: INFITT Members Database: List of INFITT Members 2009

#Search date: Tuesday, November 02, 2010 18:17
#Number of Records in this INFITT Membership Database as of search date = 352

Note:

Payment ID	Category	Name	Role	PaidMember Since PaymentExpiry	2008 Membership Payment Details	2009 Membership Payment Details	2010 Membership Renewal Status	2011 Membership Renewal Status	2012 Membership Renewal Status	Country	City
1199-N	Individual	Farook Fahath Ahmed புகத் அகமது அபுருக்		0000-00-00	...	...	உறுப்பினராக	உறுப்பினராக	உறுப்பினராக	LK	Kinniya
357-E	Individual	Fredric C. Gey பிரெட்ரிக்		2003-04-11 2003-12-31	...	...	உறுப்பினராக	உறுப்பினராக	உறுப்பினராக		
1100-E	Individual	Ashokan, Vijayshankar அசோகன், விஜயசங்கர்	Student Member மாணவர் கணினி	0000-00-00	...	...	உறுப்பினராக	உறுப்பினராக	உறுப்பினராக	INDIA இந்தியா	Dharmapuri
1303-Y	Individual	Achuthan G. Menon அச்சுதன் மேனன்	Member GB உறுப்பினர் மன்றம்	2009-10-24 2010-12-31	...	US \$ 5 Paid on 2009-10-24	US \$ 10 Paid on 2009-10-24	புதிப்பிக்க	புதிப்பிக்க	INDIA இந்தியா	
405-E	Individual	Annamalai Natarajan அண்ணாமலை நடராசன்		2002-12-31 2002-12-31	...	...	உறுப்பினராக	உறுப்பினராக	உறுப்பினராக		

Done

Internet | Protected Mode: On

7:18 PM