

# Proposal for the Generation Panel for the Thai Script Label

## Generation Ruleset for the Root Zone

### 1 General information

In general, the Thai script could be tracked back more than seven centuries. It has derived from Khmer scripts at 12th century. However, Thai script today has been gradually changed from time to time both in the script's shape and writing system.

Thai script is primarily used for writing Thai language which has been used by 66 million people throughout Thailand. Moreover, there are a number of regional dialects such as northern and southern dialects which also use the Thai script.

Thai script has 46 consonants that do not have a notion of uppercase and lowercase characters. Forming words usually comes up from a combination of one or more consonants. In addition, Thai character including 4 tone marks and 18 vowels. While some of vowels are placed before or after the main consonant in order to form a word. The Thai script also includes Thai punctuation signs and digits. In terms of writing order, Thai is a left-to-right systems, without space between words.

#### 1.1 Script for which the panel is to be established

ISO 15924 – Code: Thai

ISO 15924 – Number: 352

ISO 15924 - English Name: Thai

Native Name: ไทย

Thai Unicode range: U+0E00 – U+0E7F

The Thai Generation Panel will consider the subset of code points within this range which have been shortlisted in Maximal Starting Repertoire version 2 (MSR-2) as shown in Appendix 1.

#### 1.2 Principal languages using that script

The Thai script is principally used for writing the Thai language, which is the official language of Thailand and widely used across the country. In addition, there is also a number of other languages in Thailand, Laos, Myanmar and China written in Thai script; however, with moderate to small number of population. Therefore, the panel considers mainly working on the Thai script for the languages that are still actively used and has been reported having high-usage population as listed in the Table 1 [1]-[2]. Besides the listed languages, the panel may also consider other languages written in the Thai script during the study.

Table 1: The list of selected languages written in Thai script.

Language	ISO 639-3 Code	Locations	Population	Status
Thai	tha	Thailand (official language of Thailand)	20,200,000 (2000)	1
Northeastern Thai	tts	Widespread in Northeast Thailand	15,000,000 (1983 SIL)	6a
Northern Thai	nod	Northern region of Thailand	6,000,000 (1983 SIL)	5
Southern Thai	sou	Southern region of Thailand	4,500,000 (2006 Mahidol University)	5
Northern Khmer	kxm	Northeastern and Eastern regions of Thailand along the border with Cambodia	1,400,000 (2006 Mahidol University)	5
Pattani Malay	mfa	Southern region of Thailand near the border with Malaysia	1,000,000 (2006 Mahidol University)	5

### 1.3 Geographic territories or countries with significant user communities for the script

Thai script has been used to write Thai language, the national language of Thailand, and also other languages of ethnic groups in Thailand and neighboring countries of Thailand which having different writing systems. It is also used by the Thai expatriates in other countries such as United States of America, United Kingdom, Germany, France, Australia, Southeast Asia, East Asia, Middle East and Scandinavian countries.

### 1.4 Are there related scripts? If so, which ones?

Some Thai scripts are similar with Lao and Khmer scripts; however, the panel will conduct more detailed analysis ensuring that the Thai script is differentiated from those scripts.

## 2 Proposed initial composition of panel

### 2.1 List prospective chair and members

#### Advisory Committee

##### Role of the Advisory committee

- This committee will overview, give advice and facilitate to the Thai Script Generation Panel ensuring it is able to complete this task effectively and timely.
- This committee will also provide advice to panel on public policy, especially when collaborating with other panels for the national laws and international agreements.

Table 2: List of advisory committees.

Name	Designation	Organization
Dr. Thaweesak Koanantakool	President of the National Science and Technology Development Agency (NSTDA)	National Science and Technology Development Agency (NSTDA)
Mrs. Surangkana Wayuparb	Executive Director, CEO	Electronic Transactions Development Agency (Public Organization)

### Panel chair and members

Table 3: List of panel members.

Name	Role	Designation	Organization	Relevant experience
Mr. Wanawit Ahkuputra	Chair	Deputy Executive Director of ETDA, Vice-Chair of ICANN Government Advisory Committee	Electronic Transactions Development Agency (Public Organization)	Develop on ICT policy and software standard
Ms. Pitinan Kooarmornpatana	Secretary	Director of office of Information Technology Infrastructure	Electronic Transactions Development Agency (Public Organization)	DNS/IDNS/UNICODE expert
Ms. Ubolthip Sethakaset	Member	Specialist	Electronic Transactions Development Agency (Public Organization)	Current research in Thai-script writing system for Thai language
Mr. Panus Na Nakorn	Member	Specialist	Electronic Transactions Development Agency (Public Organization)	Policy expert
Mr. Pakpoom Tripatana	Member	Director	Thai Name Server Co., Ltd.	ccTLD/IDN ccTLD registry
Mrs. Pensri Arunwatanamongkol	Member	Director	Dot Arai Co., Ltd.	ICANN accredited registrar
Mr. Arthit Suriyawongkul	Member	Director	Foundation for Internet and Civic Culture	Representative of community on internet governance

Name	Role	Designation	Organization	Relevant experience
Dr. Yunyong Teng-amnuay	Member	Board of Director	Thai Network Information Center Foundation	DNS/IDNS/ccTLD expert
Mr. Ome Sivadith	Member	National Technology Officer	Microsoft (Thailand) Limited	Policy expert
Mr. Worapon Pitayaphongpat	Member	Senior Software Engineer Manager	Microsoft (Thailand) Limited	Software development domain and related to Thai language in Windows operating system
Ms. Monthika Boriboon	Member	Senior Researcher	National Electronics and Computer Technology Center	Thai Linguistics
Dr. Thepchai Supnithi	Member	Head of Language and Semantic Technology Laboratory	National Electronics and Computer Technology Center	Computational Linguistics
Mr. Wichai Termwuttipreecha	Member	Founding member of Mozilla Thai localization team	Mozilla Thailand Community (Mhafai.com)	
Mr. Theppitak Karoonboonyanan	Member	Freelance developer	None	Thai language standards for computer

Note: The panel will invite additional member as required.

### 3 Work plan

#### 3.1 Suggested Timeline, listing of significant milestones

The generation panel will start the work from October 2015 and intends to finalize the proposal by February 2016.

The panel will be working via mailing-list as a normal basis, and it will hold teleconference after reaching each milestone. As the majority of the panel members are in Bangkok, any face to face meetings will be organized by ETDA.

The panel will hold an open public consultation before submitting the proposal to ICANN. The purpose of public consultation is to get feedback on the work from the larger community and experts. Experts will be invited from Thai Internet Governance Forum; such as, but not limited to, internet users, IT companies, internet service providers, universities and media. The feedback from this consultation will be used to finalize of the proposal for submission.

Table 4: Schedule of project milestones.

Activity	Description	Start date	Duration
<b>1) Develop Principles</b>	Principles to be used to determine valid code points, variants and labels.	05/10/2015	2 weeks
<b>2) Determine Code Points</b>	Select the code points from MSR which are needed for Root Zone LGR.	19/10/2015	2 weeks
<b>3) Determine (any) Variants</b>	From the codes points selected, determine in case end-user may confuse two code points: <ul style="list-style-type: none"> <li>- Within script</li> <li>- Cross scripts</li> </ul>	02/11/2015	4 weeks
<b>4) Determine Label Rules</b>	Determine in case there are any label level constraints on the use of selected code points.	07/12/2015	2 weeks
<b>5) Hold Public Consultation</b>	Hold a workshop on the work accomplished by the generation panel to get feedback from the experts and community: <ul style="list-style-type: none"> <li>- Online (1month)</li> <li>- Face to Face (focus on experts but open for others) (21/12/2015)</li> </ul>	21/12/2015	4 weeks +1 new year holiday week
<b>6) Write Proposal and Create XML</b>	Write up the Root Zone LGR proposal, including references to each code point included, why variants needed and details of label rules developed + XML file.  Coordinate with Lao and Khmer scripts.	25/01/2016	2 weeks
<b>7) Submit</b>	Get public comments then finalize and submit (tentatively submit the final proposal on 22/02/2016).	08/02/2016	2 weeks

### 3.2 Proposed schedule of meeting and teleconferences

Table 5: Schedule of teleconferences.

Meeting Point	Description	Meeting Date
<b>1) Before Develop Principles</b>	Output: Development Principles	05/10/2015
<b>2) After Determine Code Points</b>	Output: Defined Code Points	02/11/2015
<b>3) After Determine (any) Variants</b>	Output: Defined Variants	08/12/2015 (07/12/2015 is public holiday.)
<b>4) After Determine Label Rules and before Hold Public Consultation</b>	Output: <ul style="list-style-type: none"> <li>• Public comments report</li> </ul>	21/12/2015

	<ul style="list-style-type: none"> <li>• Revised Label Generation Ruleset</li> </ul>	
<b>5) Before Submit</b>	Output: <ul style="list-style-type: none"> <li>• Label Generation Ruleset Report</li> <li>• Machine readable Label Generation Ruleset (XML)</li> </ul>	08/02/2016

### 3.3 Anticipated need for logistic support, such as mailing list?

The panel requests for confluence space and [thaigp@icann.org](mailto:thaigp@icann.org) mailing list creation for the project and setting up for adobe-connect for teleconferences.

### 3.4 Does the panel expect to call on ICANN provided advisors?

The panel anticipates that it will require a DNS/IDN expert to interact with the Panel and assist arranging collaboration with Lao-script and Khmer-script panel members.

## Appendix 1: Thai script [3]

The code points of Thai script shortlisted in MSR-2 is shown in Figure 1. The code cells with yellow highlighting are part of the MSR, whereas those with pink highlighting are excluded from the MSR and those without highlighting are for code points that are not PVALID in IDNA 2008.

	0E0	0E1	0E2	0E3	0E4	0E5	0E6	0E7
0		จ ๐๕10	ภ ๐๕20	๕๕ ๐๕30	เ ๐๕40	๐ ๐๕50		
1	ก ๐๕01	ท ๐๕11	ม ๐๕21	เ ๐๕31	แ ๐๕41	๑ ๐๕51		
2	ข ๐๕02	ฃ ๐๕12	ย ๐๕22	า ๐๕32	โ ๐๕42	๒ ๐๕52		
3	ช ๐๕03	ฅ ๐๕13	ร ๐๕23	ำ ๐๕33	ใ ๐๕43	๓ ๐๕53		
4	ค ๐๕04	ด ๐๕14	ถ ๐๕24	อ ๐๕34	ไ ๐๕44	๔ ๐๕54		
5	ฆ ๐๕05	ต ๐๕15	ล ๐๕25	ั ๐๕35	า ๐๕45	๕ ๐๕55		
6	ฆ ๐๕06	ถ ๐๕16	ภ ๐๕26	อ ๐๕36	ง ๐๕46	๖ ๐๕56		
7	ง ๐๕07	ท ๐๕17	ว ๐๕27	อ ๐๕37	อ ๐๕47	๗ ๐๕57		
8	จ ๐๕08	ธ ๐๕18	ศ ๐๕28	อ ๐๕38	อ ๐๕48	๘ ๐๕58		
9	ฉ ๐๕09	น ๐๕19	ษ ๐๕29	อ ๐๕39	อ ๐๕49	๙ ๐๕59		
A	ช ๐๕0A	บ ๐๕1A	ส ๐๕2A	อ ๐๕3A	อ ๐๕4A	๑๑ ๐๕5A		
B	ช ๐๕0B	ป ๐๕1B	ห ๐๕2B		๑๒ ๐๕4B	๑๓ ๐๕5B		
C	ฌ ๐๕0C	ผ ๐๕1C	ฬ ๐๕2C		๑๔ ๐๕4C			
D	ญ ๐๕0D	ฝ ๐๕1D	อ ๐๕2D		๑๕ ๐๕4D			
E	ฎ ๐๕0E	พ ๐๕1E	ฮ ๐๕2E		๑๖ ๐๕4E			
F	ฏ ๐๕0F	ฬ ๐๕1F	๑๗ ๐๕2F	฿ ๐๕3F	๑๘ ๐๕4F			

Figure 1: Thai script Unicode shortlisted in MSR-2.

**Based on TIS 620-2533****Consonants**

0E01	ก	THAI CHARACTER KO KAI
0E02	ข	THAI CHARACTER KHO KHAI
0E03	ฃ	THAI CHARACTER KHO KHUAT
0E04	ฅ	THAI CHARACTER KHO KHWAI
0E05	ฆ	THAI CHARACTER KHO KHON
0E06	ง	THAI CHARACTER KHO RAKHANG
0E07	จ	THAI CHARACTER NGO NGU
0E08	จ	THAI CHARACTER CHO CHAN
0E09	ฉ	THAI CHARACTER CHO CHING
0E0A	ช	THAI CHARACTER CHO CHANG
0E0B	ซ	THAI CHARACTER SO SO
0E0C	ฌ	THAI CHARACTER CHO CHOE
0E0D	ญ	THAI CHARACTER YO YING
0E0E	ฎ	THAI CHARACTER DO CHADA
0E0F	ฏ	THAI CHARACTER TO PATAK
0E10	ฐ	THAI CHARACTER THO THAN
0E11	ฑ	THAI CHARACTER THO NANGMONTHO
0E12	ฒ	THAI CHARACTER THO PHUTHAO
0E13	ณ	THAI CHARACTER NO NEN
0E14	ด	THAI CHARACTER DO DEK
0E15	ต	THAI CHARACTER TO TAO
0E16	ถ	THAI CHARACTER THO THUNG
0E17	ฑ	THAI CHARACTER THO THAHAN
0E18	ธ	THAI CHARACTER THO THONG
0E19	น	THAI CHARACTER NO NU
0E1A	บ	THAI CHARACTER BO BAIMAI
0E1B	ป	THAI CHARACTER PO PLA
0E1C	ผ	THAI CHARACTER PHO PHUNG
0E1D	ฝ	THAI CHARACTER FO FA
0E1E	พ	THAI CHARACTER PHO PHAN
0E1F	ฟ	THAI CHARACTER FO FAN
0E20	ภ	THAI CHARACTER PHO SAMPHAO
0E21	ม	THAI CHARACTER MO MA
0E22	ย	THAI CHARACTER YO YAK
0E23	ร	THAI CHARACTER RO RUA
0E24	ฤ	THAI CHARACTER RU
0E25	ล	THAI CHARACTER LO LING
0E26	ฌ	THAI CHARACTER LU
0E27	ว	THAI CHARACTER WO WAEN
0E28	ศ	THAI CHARACTER SO SALA
0E29	ษ	THAI CHARACTER SO RUSI
0E2A	ส	THAI CHARACTER SO SUA
0E2B	ห	THAI CHARACTER HO HIP
0E2C	ฬ	THAI CHARACTER LO CHULA
0E2D	อ	THAI CHARACTER O ANG
0E2E	ฮ	THAI CHARACTER HO NOKHUK

**Sign**

0E2F	๙	THAI CHARACTER PAIYANNOI • symbol
------	---	--------------------------------------

**Vowels**

0E30	๐	THAI CHARACTER SARA A
0E31	๑	THAI CHARACTER MAI HAN-AKAT
0E32	๒	THAI CHARACTER SARA AA
0E33	๓	THAI CHARACTER SARA AM
0E34	๔	THAI CHARACTER SARA I
0E35	๕	THAI CHARACTER SARA II
0E36	๖	THAI CHARACTER SARA UE
0E37	๗	THAI CHARACTER SARA UEE

0E38	๘	THAI CHARACTER SARA U
0E39	๙	THAI CHARACTER SARA UU
0E3A	๐	THAI CHARACTER PHINTHU

**Currency symbol**

0E3F	฿	THAI CURRENCY SYMBOL BAHT
------	---	---------------------------

**Vowels**

0E40	ิ	THAI CHARACTER SARA E
0E41	ึ	THAI CHARACTER SARA AE
0E42	๓	THAI CHARACTER SARA O
0E43	๔	THAI CHARACTER SARA AI MAIMUAN
0E44	๕	THAI CHARACTER SARA AI MAIMALAI

**Vowel length sign**

0E45	๖	THAI CHARACTER LAKKHANGYAO
------	---	----------------------------

**Repetition mark**

0E46	๗	THAI CHARACTER MAIYAMOK
------	---	-------------------------

**Vowel**

0E47	๘	THAI CHARACTER MAITAIKHU
------	---	--------------------------

**Tone marks**

0E48	๙	THAI CHARACTER MAI EK
0E49	๐	THAI CHARACTER MAI THO
0E4A	๑	THAI CHARACTER MAI TRI
0E4B	๒	THAI CHARACTER MAI CHATTAWA

**Signs**

0E4C	๓	THAI CHARACTER THANTHAKHAT
0E4D	๔	THAI CHARACTER NIKHAHIT
0E4E	๕	THAI CHARACTER YAMAKKAN
0E4F	๖	THAI CHARACTER FONGMAN

**Digits**

0E50	๐	THAI DIGIT ZERO
0E51	๑	THAI DIGIT ONE
0E52	๒	THAI DIGIT TWO
0E53	๓	THAI DIGIT THREE
0E54	๔	THAI DIGIT FOUR
0E55	๕	THAI DIGIT FIVE
0E56	๖	THAI DIGIT SIX
0E57	๗	THAI DIGIT SEVEN
0E58	๘	THAI DIGIT EIGHT
0E59	๙	THAI DIGIT NINE

**Signs**

0E5A	๐๑	THAI CHARACTER ANGKHANKHU
0E5B	๐๒	THAI CHARACTER KHOMUT



## References:

- [1] <http://www.ethnologue.com/country/TH/languages>, September 9, 2015.
- [2] [http://scriptsource.org/cms/scripts/page.php?item\\_id=script\\_detail&key=Thai](http://scriptsource.org/cms/scripts/page.php?item_id=script_detail&key=Thai), September 9, 2015.
- [3] Maximal Starting Repertoire – MSR-2, April 13, 2015.