



**Afghan Transitional Islamic Administration  
Ministry of Communications**



## *Computer Locale Requirements for Afghanistan*

**Background.** In December 2002 the United Nations Development Programme Country Office for Afghanistan commissioned a report on the language support required for Pashto and Dari, Afghanistan's official languages. In addition, the Project Team (Michael Everson of Everson Typography, Dublin, and Roozbeh Pournader of the FarsiWeb Project, Tehran) was successful in collecting similar information for a number of other languages of Afghanistan, namely, Southern Uzbek and Brahui. (The survey sought information on other languages, such as Hazaragi, Aimaq, Southwest Pashai, Western Balochi, and Turkmen, but conclusive responses were not available.)

**Need for support.** The reason this report has been commissioned is that *none* of the major computer software providers currently supports *any* of these languages – adequately or at all – and this causes serious constraints and problems for all aspects of information technology for the entire country. Language support involves inputting, display and printing, processing, and provision of a user's locale format information.

**Fonts: alphabets and glyph shapes.** This document presents information on the letters required to support the languages of Afghanistan, and the minimum shaping behaviour required for correct rendering. We urge software manufacturers in the strongest possible terms to update their Arabic and Persian fonts to include the letters and combinations which are missing. A summary of these forms is found at the end of the document.

**Input: keyboard layouts.** This document also proposes keyboard layouts suitable for inputting the languages of Afghanistan. Responses to our questionnaire did not address keyboard layout issues; however, we did receive word-of-mouth comments during the interviews we held at the University of Kabul, the Afghanistan Academy of Sciences, and the Afghanistan Assistance Coordination Authority. Most people were using old Iranian-made mechanical typewriters based on the Iranian standard ISIRI 820:1973 (*Character arrangement on keyboards of Persian typewriters*), and because they were used to them, they said that they preferred these to anything else. Because, however, that layout is glyph-based and makes use of two to four keys to form each letter, it does not suit modern computer-based entry mechanisms which use one key per letter. Accordingly, we determined that it was necessary to develop practical keyboard layouts for Afghanistan. The ISIRI 2901:1994 (*Keyboard layout for Farsi: Characters in Computer*) layout, which we used in the Farsi sample in the questionnaires, resembles the general layout of ISIRI 820 quite closely, and is modernized for current usage. Therefore, we used the general layout of ISIRI 2901 to develop the Dari layout, and then modified that to derive the Pashto and Uzbek layouts to minimize the learning curve for users as much as possible. Each of the keyboard layouts enables the user to input characters used in each of the languages supported by this report. The basic letters of the language for which the keyboard layout is designed are always on the plain and shifted keys. In some software environments the AltGr keys may not be supported.

## *Computer Locale Requirements for Afghanistan*

**Locales: ICU data.** Additional data we collected has been formatted in locale specifications via the ICU (International Components for Unicode) website hosted at “<http://oss.software.ibm.com/icu/>”. In this way, the Afghan preferred formats for language, date format (including month and day names and calendar), time format, number format, and currency format will be available to vendors who can provide software which will be of use to the people of Afghanistan, at governmental, business, academic, and local level. Whenever ICU locale data format is adequate for describing a part of the locale information, the data in the ICU registry should be consulted for normative and accurate data. The specifications shown in this report are only informative examples in comparison with the ICU data files we have provided. However, in areas such as the Afghan calendar, which ICU locales cannot specify, the text in this report should be considered normative. (Note that the ICU files we have prepared are not standalone. They must be used in combination with other locale data, most importantly the “fa.txt” and “root.txt” files from ICU itself.)

**Ordering.** Sorting order specification is also provided with the ICU locales. This deserves some additional discussion here. In our analysis of the Pashto, Dari, and Uzbek alphabets provided to us, together with our analysis of authoritative Pashto-English and Pashto-Pashto dictionaries, as well as comparison of a Dari-English dictionary with a number of bilingual and monolingual Persian dictionaries, we discovered that, with the exception of the treatment of *hamza*, the alphabets were all mutually compatible with respect to ordering. Because the ordering of “singular” *hamza* (ء) is of little consequence – it is rare and typically comes toward the end of a word – we determined that placing it between *alef* and *beh* would have two advantages: first, it would allow all of the languages of Afghanistan to avail of a *single unified sort order*, and second, it would enable the closely-related Persian languages, Dari and Farsi, to be sorted in the same way. It should be mentioned, however, that the order shown in this report is only the general ordering for letters, and does not show the ordering of less significant elements like the ordering of vowel marks or the handling of punctuation or Unicode control characters. The complete and exact ordering is the one specified in the ICU locale data files.

We commend this report to the software vendors of the world, and trust that its specifications will be implemented as a matter of urgency, so that the people Afghanistan can avail of the essential tools which many of us take for granted.



Michael Everson  
Everson Typography, Dublin  
[www.evertype.com](http://www.evertype.com)



Roozbeh Pournader  
FarsiWeb Project, Tehran  
[www.farsiweb.info](http://www.farsiweb.info)

The latest version of this report, updates, and other resources (such as keyboard layout specifications in data files conformant to ISO/IEC 9995, *Information Technology – Keyboard layouts for text and office systems*, and ICU data files based on this report) are available at “<http://www.evertype.com/standards/af/>”.

© Copyright 2003 United Nations Development Programme, Michael Everson, and Roozbeh Pournader. Everyone is permitted to copy and distribute verbatim copies of this document, but changing it is not allowed. The authors welcome feedback from people who download and use the report and data files.

***Latest update: 2003-05-16***

## *Computer Locale Requirements for Afghanistan*

### **The Pashto Alphabet**

The table below lists the letters used to write Pashto, including loanwords from Arabic. Our survey indicated a number of different sorting orders for Pashto. *All respondents* said that the alphabetic order was (reading from left to right)  $\text{چ} > \text{خ} > \text{ځ} > \text{ح} > \text{څ}$ . This is the order agreed on at a 1991 meeting of Pashto experts in Peshawar. It should be noted that the most authoritative dictionary, the *Pas̄tó-Pas̄tó Descriptive Dictionary*, 4 volumes (Department of Linguistics, Institute of Languages and Literature, Academy of Sciences of Afghanistan, 1979–1987) uses the order  $\text{چ} > \text{ځ} > \text{خ} > \text{څ} > \text{ح}$ , as does Rahimi and Rohi's *Pashto-English Dictionary* (1979). Materials for schoolchildren collected in Kabul, on the other hand, gave this same sequence as  $\text{چ} > \text{ځ} > \text{ح} > \text{خ} > \text{څ}$ . Nevertheless, we suggest that the order agreed at the 1991 Peshawar meeting is the best to follow.

In the list below, the Unicode letters which sort as separate letters at the first level are given flush to the left margin, and letters which sort at the second level are indented. We have also given the Urdu letters which are used in older Pashto orthographies to show how Pashto data in that orthography will sort.

Unicode	Name	isolated	final	medial	initial	Arabic	Dari	Pashto	Uzbek	Brahui
0622	ā	آ	اَ	—	—	•	•	•	•	
0627	alif	ا	اَ	—	—	•	•	•	•	•
0623	alif-hamza	أ	اْ	—	—	•	•	•	•	•
0671	alif-wasla	آ	اِ	—	—	•				
0625	al.-h.-below	آ	آِ	—	—	•				
0621	hamza	ء	—	—	—	•	•	•	•	•
0628	be	ب	بَ	بِ	بِ	•	•	•	•	•
067E	pe	پ	پَ	پِ	پِ		•	•	•	•
062A	te	ت	تَ	تِ	تِ	•	•	•	•	•
067C	ṭe	ٹ	ٹَ	ٹِ	ٹِ			•		
0679	Urdu ṭe	ٲ	ٲَ	ٲِ	ٲِ					•
062B	se	س	سَ	سِ	سِ	•	•	•	•	•
062C	jim	ج	جَ	جِ	جِ	•	•	•	•	•
0681	ḷe	ځ	ځَ	ځِ	ځِ			•		
0686	če	چ	چَ	چِ	چِ		•	•	•	•
0685	ce	چ	چَ	چِ	چِ			•		
062D	he	ح	حَ	حِ	حِ	•	•	•	•	•

## Computer Locale Requirements for Afghanistan

Unicode	Name	isolated	final	medial	initial	Arabic	Dari	Pashto	Uzbek	Brahui
062E	xe	خ	خ	خ	خ	•	•	•	•	•
062F	dāl	د	د	—	—	•	•	•	•	•
0689	ḍāl	ډ	ډ	—	—			•		
0688	Urdu ḍāl	ڈ	ڈ	—	—					•
0630	zāl	ذ	ذ	—	—	•	•	•	•	•
0631	re	ر	ر	—	—	•	•	•	•	•
0693	ṛe	ړ	ړ	—	—			•		
0691	Urdu ṛe	ر	ر	—	—					•
0632	ze	ز	ز	—	—	•	•	•	•	•
0698	že	ژ	ژ	—	—		•	•	•	•
0696	ẓe	ږ	ږ	—	—			•		
0633	sin	س	س	س	س	•	•	•	•	•
0634	šin	ش	ش	ش	ش	•	•	•	•	•
069A	ṣin	ښ	ښ	ښ	ښ			•		
0635	swāt	ص	ص	ص	ص	•	•	•	•	•
0636	zād	ض	ض	ض	ض	•	•	•	•	•
0637	twe	ط	ط	ط	ط	•	•	•	•	•
0638	zwe	ظ	ظ	ظ	ظ	•	•	•	•	•
0639	ʿayn	ع	ع	ع	ع	•	•	•	•	•
063A	ğayn	غ	غ	غ	غ	•	•	•	•	•
0641	fe	ف	ف	ف	ف	•	•	•	•	•
0642	qāf	ق	ق	ق	ق	•	•	•	•	•
06A9	kāf	ک	ک	ک	ک		•	•	•	•
0643	Arabic kāf	ك	ك	ك	ك	•				
06AB	gāf	گ	گ	گ	گ			•		
06AF	Persian gāf	گ	گ	گ	گ		•		•	•
0644	lām	ل	ل	ل	ل	•	•	•	•	•
06B7	Brahui lhām	ل	ل	ل	ل					•
0645	mim	م	م	م	م	•	•	•	•	•
0646	nun	ن	ن	ن	ن	•	•	•	•	•

## Computer Locale Requirements for Afghanistan

Unicode	Name	isolated	final	medial	initial	Arabic	Dari	Pashto	Uzbek	Brahui
06BA	nun ghunna	ن	ن	—	—					•
06BC	ṅun	ښ	ښ	ښ	ښ			•		
0648	wāw	و	و	—	—	•	•	•	•	•
0624	wāw hamza	ؤ	ؤ	—	—	•	•	•	•	•
06C7	Uzbek u	ۇ	ۇ	—	—				•	
06C9	Uzbek o	ۋ	ۋ	—	—				•	
0647	he	ه	ه	ه	ه	•	•	•	•	•
0629	gəɾda te	ۀ	ۀ	—	—	•	•	•	•	•
06CC	ye	ی	ی	ی	ی		•	•	•	•
0649	alef maksura	—	—	ا	ا	•				
06D2	Urdu ye	ے	ے	—	—					•
064A	saxta ye	ي	ي	—	—	•		•	•	
06D0	pasta ye	ې	ې	ې	ې			•	•	
06CD	şəjīna ye	ى	ى	—	—			•		
0626	fe'li ye	ئ	ئ	ئ	ئ	•	•	•	•	

**Note.** Since the shapes of the **ي** initial and **ي** medial forms of the Pashto letters **ی** *ye* (U+06CC) and **ي** *saxta ye* (U+064A) are exactly the same, to avoid encoding ambiguities in Pashto data – and its known implications in security-related issues including possible Pashto domain names – we recommend that the Unicode character for *saxta ye*, namely **ي** U+064A, *never* be used in initial and medial forms in Pashto data. Where input data is explicitly known to be in Pashto, applications *should* automatically replace such usage in the input data (for example, in keyboard input) with the normal *ye* character, namely **ی** U+06CC. Applications *may not* automatically change U+064A to U+06CC in Pashto input if it is used in final or isolated forms.

## *Computer Locale Requirements for Afghanistan*

### **The Dari Alphabet**

The table below lists the letters used to write Dari, including loanwords from Arabic. In the list below, the Unicode letters which sort as separate letters at the first level are given flush to the left margin, and letters which sort at the second level are indented. We have also given the Urdu letters which are used in older Pashto orthographies to show how Pashto data in that orthography will sort.

<b>Unicode</b>	<b>Name</b>	<i>isolated</i>	<i>final</i>	<i>medial</i>	<i>initial</i>	Arabic	Dari	Pashto	Uzbek	Brahui
0622	ā	آ	آ	—	—	•	•	•	•	
0627	alef	ا	ا	—	—	•	•	•	•	•
0623	alef-hamza	أ	أ	—	—	•	•	•	•	•
0671	alif-wasla	آ	آ	—	—	•				
0625	al.-h.-below	ا	ا	—	—	•				
0621	hamza	ء	—	—	—	•	•	•	•	•
0628	be	ب	ب	ب	ب	•	•	•	•	•
067E	pe	پ	پ	پ	پ		•	•	•	•
062A	te	ت	ت	ت	ت	•	•	•	•	•
067C	Pashto ʔe	ټ	ټ	ټ	ټ			•		
0679	Urdu ʔe	ٹ	ٹ	ٹ	ٹ					•
062B	se	س	س	س	س	•	•	•	•	•
062C	jim	ج	ج	ج	ج	•	•	•	•	•
0681	Pashto ʒe	ځ	ځ	ځ	ځ			•		
0686	če	چ	چ	چ	چ		•	•	•	•
0685	Pashto ce	ټچ	ټچ	ټچ	ټچ			•		
062D	he	ه	ه	ه	ه	•	•	•	•	•
062E	xe	خ	خ	خ	خ	•	•	•	•	•
062F	dāl	د	د	—	—	•	•	•	•	•
0689	Pashto ɖāl	ډ	ډ	—	—			•		
0688	Urdu ɖāl	ڍ	ڍ	—	—					•
0630	zāl	ذ	ذ	—	—	•	•	•	•	•
0631	re	ر	ر	—	—	•	•	•	•	•

## Computer Locale Requirements for Afghanistan

Unicode	Name	isolated	final	medial	initial	Arabic	Dari	Pashto	Uzbek	Brahui
0693	Pashto ʀe	ر	ر	—	—			•		
0691	Urdu ʀe	ر	ر	—	—					•
0632	ze	ز	ز	—	—	•	•	•	•	•
0698	že	ژ	ژ	—	—		•	•	•	•
0696	Pashto ʒe	ږ	ږ	—	—			•		
0633	sin	س	س	س	س	•	•	•	•	•
0634	šin	ش	ش	ش	ش	•	•	•	•	•
069A	Pashto šin	ښ	ښ	ښ	ښ			•		
0635	sād	ص	ص	ص	ص	•	•	•	•	•
0636	zād	ض	ض	ض	ض	•	•	•	•	•
0637	tā	ط	ط	ط	ط	•	•	•	•	•
0638	zā	ظ	ظ	ظ	ظ	•	•	•	•	•
0639	ayn	ع	ع	ع	ع	•	•	•	•	•
063A	qayn	غ	غ	غ	غ	•	•	•	•	•
0641	fe	ف	ف	ف	ف	•	•	•	•	•
0642	qāf	ق	ق	ق	ق	•	•	•	•	•
06A9	kāf	ک	ک	ک	ک		•	•	•	•
0643	Arabic kāf	ك	ك	ك	ك	•				
06AB	Pashto gāf	گ	گ	گ	گ			•		
06AF	gāf	گا	گا	گا	گا		•		•	•
0644	lām	ل	ل	ل	ل	•	•	•	•	•
06B7	Brahui lhām	ل	ل	ل	ل					•
0645	mīm	م	م	م	م	•	•	•	•	•
0646	nūn	ن	ن	ن	ن	•	•	•	•	•
06BA	nūn ghunna	ں	ں	—	—					•
06BC	Pashto nūn	ڼ	ڼ	ڼ	ڼ			•		
0648	vāv	و	و	—	—	•	•	•	•	•
0624	vāv hamza	و	و	—	—	•	•	•	•	•
06C7	Uzbek u	ۇ	ۇ	—	—				•	
06C9	Uzbek o	ۈ	ۈ	—	—					•

## *Computer Locale Requirements for Afghanistan*

<b>Unicode</b>	<b>Name</b>	<i>isolated</i>	<i>final</i>	<i>medial</i>	<i>initial</i>	Arabic	Dari	Pashto	Uzbek	Brahui
0647	he	ه	ه	ه	ه	•	•	•	•	•
0629	te-ye gerd	ة	ة	—	—	•	•	•	•	•
06CC	ye	ی	ی	ی	ی		•	•	•	•
0649	alef-e maqsura	—	—	ا	ا	•				
06D2	Urdu ye	ے	ے	—	—					•
064A	saxta ye	ي	ي	—	—	•		•	•	
06D0	pasta ye	پي	پي	پي	پي			•	•	
06CD	şəjīna ye	ئ	ئ	—	—			•		
0626	ye hamza	ء	ء	أ	أ	•	•	•	•	



# Computer Locale Requirements for Afghanistan

## The Southern Uzbek Alphabet

The table below lists the letters used to write Southern Uzbek, including loanwords from Arabic. This is a new official orthography the development of which was inspired by the UNDP project. In the list below, the Unicode letters which sort as separate letters at the first level are given flush to the left margin, and letters which sort at the second level are indented. We have also given the Urdu letters which are used in older Pashto orthographies to show how Pashto data in that orthography will sort.

Unicode	Name	<i>isolated</i>	<i>final</i>	<i>medial</i>	<i>initial</i>	Arabic	Dari	Pashto	Uzbek	Brahui
0622	alef-madha	آ	آ	—	—	•	•	•	•	
0627	alef	ا	ا	—	—	•	•	•	•	•
0623	alef-hamza	أ	أ	—	—	•	•	•	•	•
0671	alif-wasla	آ	آ	—	—	•				
0625	al.-h.-below	ا	ا	—	—	•				
0621	hamza	ء	—	—	—	•	•	•	•	•
0628	beh	ب	ب	ب	ب	•	•	•	•	•
067E	peh	پ	پ	پ	پ		•	•	•	•
062A	teh	ت	ت	ت	ت	•	•	•	•	•
067C	Pashto țeh	ت	ت	ت	ت			•		
0679	Urdu țeh	ٹ	ٹ	ٹ	ٹ					•
062B	seh	ث	ث	ث	ث	•	•	•	•	•
062C	jeem	ج	ج	ج	ج	•	•	•	•	•
0681	Pashto țeh	چ	چ	چ	چ			•		
0686	cheh	چ	چ	چ	چ		•	•	•	•
0685	Pashto ceh	چ	چ	چ	چ			•		
062D	heh	ح	ح	ح	ح	•	•	•	•	•
062E	kheh	ح	ح	ح	ح	•	•	•	•	•
062F	dal	د	د	—	—	•	•	•	•	•
0689	Pashto đal	ډ	ډ	—	—			•		
0688	Urdu đal	ڈ	ڈ	—	—					•
0630	zal	ذ	ذ	—	—	•	•	•	•	•
0631	reh	ر	ر	—	—	•	•	•	•	•

## Computer Locale Requirements for Afghanistan

Unicode	Name	isolated	final	medial	initial	Arabic	Dari	Pashto	Uzbek	Brahui
0693	Pashto reh	ر	ړ	—	—			•		
0691	Urdu reh	ر	ر	—	—					•
0632	zeh	ز	ز	—	—	•	•	•	•	•
0698	zheh	ژ	ژ	—	—		•	•	•	•
0696	Pashto zeh	ږ	ږ	—	—			•		
0633	sin	س	س	س	س	•	•	•	•	•
0634	shin	ش	ش	ش	ش	•	•	•	•	•
069A	Pashto shin	ښ	ښ	ښ	ښ			•		
0635	swad	ص	ص	ص	ص	•	•	•	•	•
0636	zwad	ض	ض	ض	ض	•	•	•	•	•
0637	tuy	ط	ط	ط	ط	•	•	•	•	•
0638	zuy	ظ	ظ	ظ	ظ	•	•	•	•	•
0639	'ain	ع	ع	ع	ع	•	•	•	•	•
063A	ghain	غ	غ	غ	غ	•	•	•	•	•
0641	feh	ف	ف	ف	ف	•	•	•	•	•
0642	qaf	ق	ق	ق	ق	•	•	•	•	•
06A9	kaf	ک	ک	ک	ک		•	•	•	•
0643	Arabic kaf	ك	ك	ك	ك	•				
06AB	Pashto gaf	ګ	ګ	ګ	ګ			•		
06AF	gaf	ګ	ګ	ګ	ګ		•		•	•
0644	lam	ل	ل	ل	ل	•	•	•	•	•
06B7	Brahui lham	ل	ل	ل	ل					•
0645	meem	م	م	م	م	•	•	•	•	•
0646	nuun	ن	ن	ن	ن	•	•	•	•	•
06BA	nuun ghunna	ں	ں	—	—					•
06BC	Pashto nuun	ڼ	ڼ	ڼ	ڼ			•		
0648	waw	و	و	—	—	•	•	•	•	•
0624	waw hamza	وْ	وْ	—	—	•	•	•	•	•
06C7	u	وُ	وُ	—	—				•	
06C9	o'	وُ	وُ	—	—				•	

## *Computer Locale Requirements for Afghanistan*

<b>Unicode</b>	<b>Name</b>	<i>isolated</i>	<i>final</i>	<i>medial</i>	<i>initial</i>	Arabic	Dari	Pashto	Uzbek	Brahui
0647	heh	ه	ه	ه	ه	•	•	•	•	•
0629	teh marbuta	ة	ة	—	—	•	•	•	•	•
06CC	yih	ی	ی	ی	ی		•	•	•	•
0649	alef maksura	—	—	ا	ا	•				
06D2	Urdu yih	ے	ے	—	—					•
064A	yeeh	ي	ي	—	—	•		•	•	
06D0	yeh	ې	ې	ې	ې			•	•	
06CD	şəjina yih	ی	ی	—	—			•		
0626	yih hamza	ئ	ئ	ئ	ئ	•	•	•	•	

**Note.** To avoid encoding ambiguities in Uzbek data, and its known implications in security-related issues including possible Uzbek Internet domain names, we recommend that the Uzbek letter *u* *always* be encoded as **ۇ** U+06C7 and *never* as a sequence of *waw* and a *damma*. Where input data is explicitly known to be in Uzbek, applications *should* automatically replace a *waw* followed by a *damma* in the input data (for example, in keyboard input) with the U+06C7 character.

## *Computer Locale Requirements for Afghanistan*

### **The Brahui Alphabet**

The table below lists the letters used to write Brahui, including loanwords from Arabic. Brahui is not widely written either in Pakistan, where most Brahui speakers live, or in Afghanistan. An expert respondent informed us that in Pakistan, Brahui speakers use Urdu orthography with the addition of one letter, **ڄ** *lhām*. If Brahui is written in Afghanistan, it may prefer the Pashto letters such as **ټ** *ṭe* to the corresponding Urdu **ٹ** *ṭe*. We have no evidence of this, but have given the Brahui alphabet here in order to ensure the support of its own unique letter.

In the list below, the Unicode letters which sort as separate letters at the first level are given flush to the left margin, and letters which sort at the second level are indented. We have also given the Urdu letters which are used in older Pashto orthographies to show how Pashto data in that orthography will sort.

<b>Unicode</b>	<b>Name</b>	<i>isolated</i>	<i>final</i>	<i>medial</i>	<i>initial</i>	Arabic	Dari	Pashto	Uzbek	Brahui
0622	ā	آ	اَ	—	—	•	•	•	•	•
0627	alif	ا	ا	—	—	•	•	•	•	•
0623	alif-hamza	أ	اْ	—	—	•	•	•	•	•
0671	alif-wasla	آ	اَ	—	—	•				
0625	al.-h.-below	اِ	اِ	—	—	•				
0621	hamza	ء	—	—	—	•	•	•	•	•
0628	be	ب	ب	ب	ب	•	•	•	•	•
067E	pe	پ	پ	پ	پ		•	•	•	•
062A	te	ت	ت	ت	ت	•	•	•	•	•
067C	ṭe	ټ	ټ	ټ	ټ			•		
0679	Urdu ṭe	ٹ	ٹ	ٹ	ٹ					•
062B	se	ث	ث	ث	ث	•	•	•	•	•
062C	jim	ج	ج	ج	ج	•	•	•	•	•
0681	je	چ	چ	چ	چ			•		
0686	če	چھ	چھ	چھ	چھ		•	•	•	•
0685	ce	چھ	چھ	چھ	چھ			•		
062D	he	ح	ح	ح	ح	•	•	•	•	•
062E	xe	خ	خ	خ	خ	•	•	•	•	•
062F	dāl	د	د	—	—	•	•	•	•	•

## Computer Locale Requirements for Afghanistan

Unicode	Name	isolated	final	medial	initial	Arabic	Dari	Pashto	Uzbek	Brahui
0689	ḍāl	ڊ	ڊ	—	—			•		
0688	Urdu ḍāl	ڙ	ڙ	—	—					•
0630	zāl	ذ	ذ	—	—	•	•	•	•	•
0631	re	ر	ر	—	—	•	•	•	•	•
0693	ṛe	ړ	ړ	—	—			•		
0691	Urdu ṛe	ر	ر	—	—					•
0632	ze	ز	ز	—	—	•	•	•	•	•
0698	že	ژ	ژ	—	—		•	•	•	•
0696	ẓe	ږ	ږ	—	—			•		
0633	sin	س	س	س	س	•	•	•	•	•
0634	šin	ش	ش	ش	ش	•	•	•	•	•
069A	ṣin	ښ	ښ	ښ	ښ			•		
0635	swāt	ص	ص	ص	ص	•	•	•	•	•
0636	zād	ض	ض	ض	ض	•	•	•	•	•
0637	twe	ط	ط	ط	ط	•	•	•	•	•
0638	zwe	ظ	ظ	ظ	ظ	•	•	•	•	•
0639	ʿayn	ع	ع	ع	ع	•	•	•	•	•
063A	ğayn	غ	غ	غ	غ	•	•	•	•	•
0641	fe	ف	ف	ف	ف	•	•	•	•	•
0642	qāf	ق	ق	ق	ق	•	•	•	•	•
06A9	kāf	ک	ک	ک	ک		•	•	•	•
0643	Arabic kāf	ك	ك	ك	ك	•				
06AB	Pashto gāf	ګ	ګ	ګ	ګ			•		
06AF	gāf	ګ	ګ	ګ	ګ		•		•	•
0644	lām	ل	ل	ل	ل	•	•	•	•	•
06B7	lhām	ل	ل	ل	ل					•
0645	mim	م	م	م	م	•	•	•	•	•
0646	nun	ن	ن	ن	ن	•	•	•	•	•
06BA	nun ghunna	ن	ن	—	—					•
06BC	ṇun	ڼ	ڼ	ڼ	ڼ			•		

## Computer Locale Requirements for Afghanistan

Unicode	Name	isolated	final	medial	initial	Arabic	Dari	Pashto	Uzbek	Brahui
0648	wāw	و	و	—	—	•	•	•	•	•
0624	wāw hamza	وْ	وْ	—	—	•	•	•	•	•
06C7	Uzbek u	ۇ	ۇ	—	—				•	
06C9	Uzbek o	ۋ	ۋ	—	—				•	
0647	he	ه	ه	ه	ه	•	•	•	•	•
0629	te marbuta	ة	ة	—	—	•	•	•	•	•
06CC	ye	ی	ی	ی	ی		•	•	•	•
0649	alef maksura	ـ	ـ	ـ	ـ	•				
06D2	Urdu ye	ے	ے	—	—					•
064A	saxta ye	ي	ي	—	—	•		•	•	
06D0	pasta ye	ې	ې	ـ	ـ			•	•	
06CD	şəjīna ye	ئ	ئ	—	—			•		
0626	ye hamza	ئْ	ئْ	ـ	ـ	•	•	•	•	

## Computer Locale Requirements for Afghanistan

### Glyph-shape Requirements for Afghanistan

The table below lists the glyphs not usually found in most Arabic fonts, which it is essential to support in order to enable the people of Afghanistan to write and process their languages.

Unicode	Name	<i>isolated</i>	<i>final</i>	<i>medial</i>	<i>initial</i>	Dari	Pashto	Uzbek	Brahui
067E	pe	پ	پا	پا	پا	•	•	•	•
067C	Pashto ɽe	ږ	ږا	ږا	ږا		•		
0679	Urdu ɽe	ط	طا	طا	طا				•
0681	je	ج	جا	جا	جا		•		
0686	če	چ	چا	چا	چا	•	•	•	•
0685	ce	س	سا	سا	سا		•		
0689	ḍāl	ڍ	ڍا	—	—		•		
0688	Urdu ḍāl	ڍا	ڍا	—	—				•
0693	ɽe	ر	را	—	—		•		
0691	Urdu ɽe	ر	ر	—	—				•
0698	že	ز	زا	—	—	•	•	•	•
0696	ze	ز	ز	—	—		•		
069A	šin	ش	شا	شا	شا		•		
06A9	Persian kāf	ک	کا	کا	کا	•	•	•	•
06AB	Pashto gāf	گ	گا	گا	گا		•		
06AF	gāf	گ	گا	گا	گا	•		•	•
06B7	lhām	ل	لا	لا	لا				•
06BA	nun ghunna	ن	ن	—	—				•
06BC	ṇun	ڼ	ڼا	ڼا	ڼا		•		
06C7	Uzbek u	ۇ	ۇ	—	—			•	
06C9	Uzbek o	ۈ	ۈ	—	—			•	
06CC	ye	ي	يا	يا	يا	•	•	•	•
06D2	Urdu ye	ی	ی	—	—				•
06D0	pasta ye	پي	پي	پا	پا		•	•	
06CD	şəjīna ye	ی	ی	—	—		•		

## *Computer Locale Requirements for Afghanistan*

Almost all of the languages used in Afghanistan require the use of diacritical marks which follow a base letter, to make pronunciation or grammar usage clear. These marks are known by different names in the different languages of Afghanistan, but are listed here with their Unicode character names. Font and application developers should ensure that they support these marks in their products. The *zwarakay* at the bottom of the list is used in some Pashto educational materials, and as a result of this survey is now being proposed for addition to the Unicode Standard. Until such time as it has been formally encoded, we recommend the Private Use Area code position U+E659 for the *zwarakay*.

<b>Unicode</b>	<b>Name</b>		<b>Unicode</b>	<b>Name</b>	
064B	fathatan	◌َ	0651	shadda	◌̣
064C	dammatan	◌ُ	0652	sukun	◌◌̣
064D	kasratan	◌ِ	0653	maddah above	◌̣̣
064E	fatha	◌َ	0654	hamza above	◌̣̣̣
064F	damma	◌ُ	0670	superscript alef	◌̣̣̣̣
0650	kasra	◌ِ	(E659)	zwarakay	◌̣̣̣̣̣

One additional character appears to be in wide use by local banks and accountants in Afghanistan, namely the *afghani sign*, ﷲ. This character is also being proposed for addition to the Unicode Standard. Until such time as it has been formally encoded, we recommend the Private Use Area code position U+E0B4 for the *afghani sign*. It should be noted that glyph variants for this character exist, for example ﷲ and ﷲ, but in our discussions several experts indicated that the first one shown above is the most original and may therefore be more appropriate. Research on the best glyph shape is still ongoing. As of this writing, neither the Ministry of Finance nor Da Afghanistan Bank had formally adopted this symbol for use, although it is in use in both the public and private sectors.

It should also be mentioned that certain punctuation marks and control characters are used in Afghan computing. We have listed all of these characters on the collectively-equivalent keyboard layouts. Font developers are expected to include *all* of these additional characters in fonts which are intended to support Afghan languages.





# *Computer Locale Requirements for Afghanistan*

## **Keyboard layouts for Afghanistan**

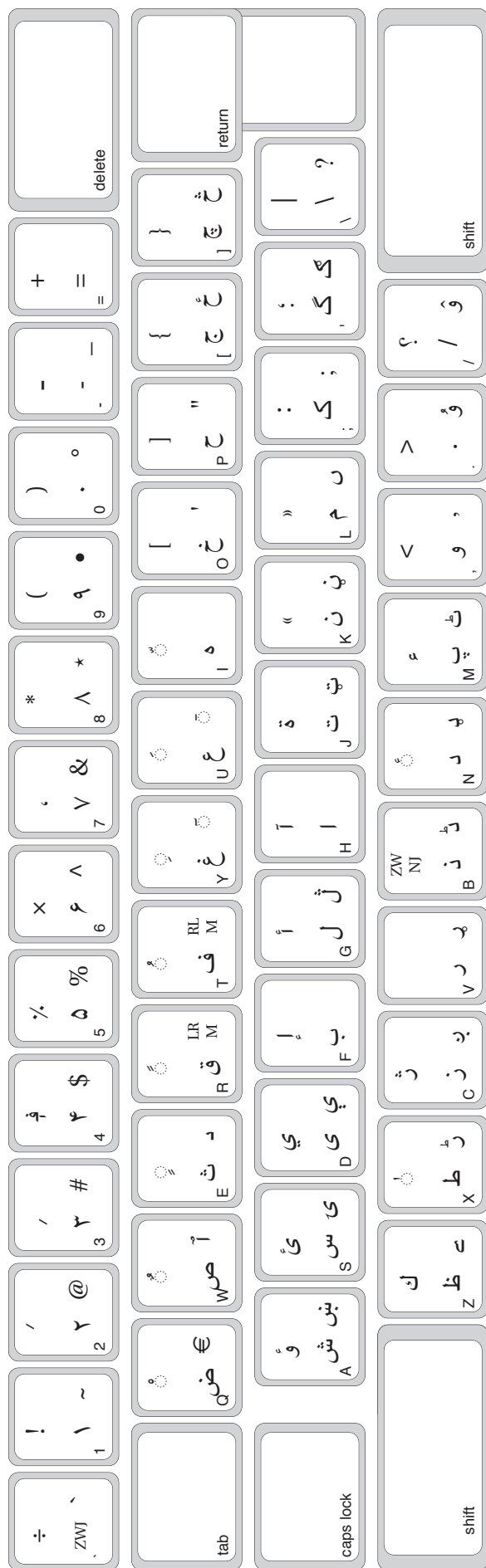
Key assignments for Pashto, using ISO/IEC 9995 notation:

	<b>Unshifted</b>	<b>Shifted</b>	<b>AltGr</b>
TLDE	200D	0654	0060
E01	06F1	0021	007E
E02	06F2	066C	0040
E03	06F3	066B	0023
E04	06F4	AFGHANI	0024
E05	06F5	066A	0025
E06	06F6	00D7	005E
E07	06F7	00BB	0026
E08	06F8	00AB	066D
E09	06F9	0029	2022
E10	06F0	0028	00B0
E11	002D	0640	005F
E12	003D	002B	00F7
D01	0636	0652	20AC
D02	0635	064C	0671
D03	062B	064D	0649
D04	0642	064B	200E
D05	0641	064F	200F
D06	063A	0650	0653
D07	0639	064E	ZWARAKAY
D08	0647	0651	0670
D09	062E	0681	0027
D10	062D	0685	0022
D11	062C	005D	007D
D12	0686	005B	007B
C01	0634	069A	<FREE>
C02	0633	06CD	<FREE>
C03	06CC	064A	06D2
C04	0628	067E	06BA
C05	0644	0623	06B7
C06	0627	0622	0625
C07	062A	067C	0679
C08	0646	06BC	003E
C09	0645	0629	003C
C10	06A9	003A	0643
C11	06AB	061B	06AF
B01	0638	0626	003F
B02	0637	06D0	003B
B03	0632	0698	<FREE>
B04	0631	0621	<FREE>
B05	0630	200C	<FREE>
B06	062F	0689	0688
B07	0693	0624	0691
B08	0648	060C	002C
B09	0696	002E	06C7
B10	002F	061F	06C9
BKSL	005C	002A	007C
SPCE	0020	200C	00A0

# Computer Locale Requirements for Afghanistan

## Dari keyboard layout for Afghanistan

by Michael Everson & Roozbeh Pournader



shift	altgr shift
plain	altgr

Guide to the key assignments, reading top-down and left-right

### Unshifted:

ZWJ, one, two, three, four, five, six, seven, eight, nine, zero, hyphen, equals  
 zād, sād, se, qāf, fe, qayn, ayn, he, xe, he, jim, če  
 šin, sin, ye, be, lām, alef, te, nūn, mūm, kāf, gāf, backslash  
 zā, tā, ze, re, zāl, dāl, pe, vāv, full-stop, slash

### Shifted:

division, exclamation, thousands, decimals, afghani, percent, multiplication, comma, asterisk, parenthesis, kashida, plus  
 sukun, dammatan, kasratan, fathatan, damma, kasra, fatha, shadda, square-bracket, curly-bracket, curly-bracket  
 vāv-hamza, ye-hamza, saxta-ye, alef-hamza-below, alef-hamza, ā, te-ye-gerd, guillemet, colon, semicolon, vertical-bar  
 arabic-kāf, alef-above, že, —, ZWNJ, hamza-above, hamza, less-than, greater-than, question

### Alt-Gr:

grave, tilde, at, number, dollar, percent, circumflex, ampersand, 5-star, bullet, degree, underscore, —  
 euro, alef-wasla, alef-e-maqsurā, LRM, RLM, maddah, zwarakay, —, apostrophe, quote, pashto-je, pashto-ce  
 pashto-šin, šejina-ye, pasta-ye, —, brahui-lhām, —, pashto-fe, pashto-nūn, nūn-ghunna, semicolon, pashto-gāf, question  
 urdu-ye, urdu-fe, pashto-že, pashto-fer, urdu-dāl, pashto-dāl, urdu-že, comma, uzbek-u, uzbek-o

# *Computer Locale Requirements for Afghanistan*

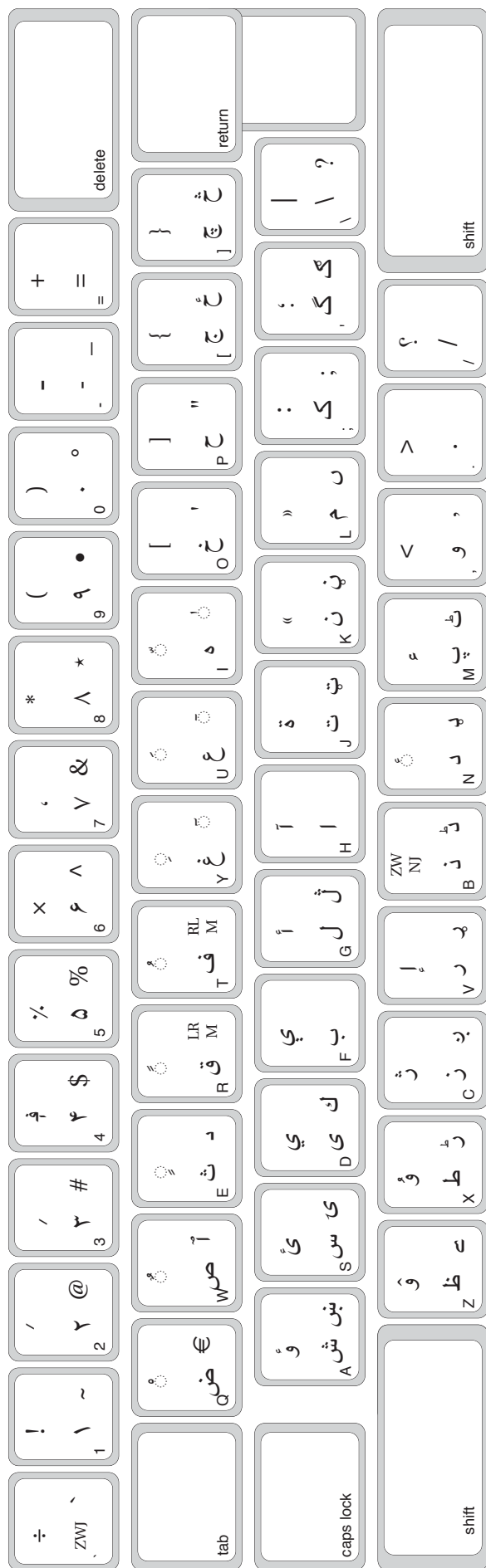
## **Keyboard layouts for Afghanistan**

Key assignments for Dari, using ISO/IEC 9995 notation:

	<b>Unshifted</b>	<b>Shifted</b>	<b>AltGr</b>
TLDE	200D	00F7	0060
E01	06F1	0021	007E
E02	06F2	066C	0040
E03	06F3	066B	0023
E04	06F4	AFGHANI	0024
E05	06F5	066A	0025
E06	06F6	00D7	005E
E07	06F7	060C	0026
E08	06F8	002A	066D
E09	06F9	0029	2022
E10	06F0	0028	00B0
E11	002D	0640	005F
E12	003D	002B	<FREE>
D01	0636	0652	20AC
D02	0635	064C	0671
D03	062B	064D	0649
D04	0642	064B	200E
D05	0641	064F	200F
D06	063A	0650	0653
D07	0639	064E	ZWARAKAY
D08	0647	0651	<FREE>
D09	062E	005D	0027
D10	062D	005B	0022
D11	062C	007D	0681
D12	0686	007B	0685
C01	0634	0624	069A
C02	0633	0626	06CD
C03	06CC	064A	06D0
C04	0628	0625	<FREE>
C05	0644	0623	06B7
C06	0627	0622	<FREE>
C07	062A	0629	067C
C08	0646	00BB	06BC
C09	0645	00AB	06BA
C10	06A9	003A	003B
C11	06AF	061B	06AB
B01	0638	0643	06D2
B02	0637	0670	0691
B03	0632	0698	0696
B04	0631	<FREE>	0693
B05	0630	200C	0688
B06	062F	0654	0689
B07	067E	0621	0679
B08	0648	003E	002C
B09	002E	003C	06C7
B10	002F	061F	06C9
BKSL	005C	007C	003F
SPCE	0020	200C	00A0

## Southern Uzbek keyboard layout for Afghanistan

by Michael Everson & Roozbeh Pournader



shift	altgr shift
plain	altgr

Guide to the key assignments, reading top-down and left-right

### Unshifted:

- ZWJ, one, two, three, four, five, six, seven, eight, nine, zero, hyphen, equals
- zwad, swad, seh, qaf, feh, ghain, 'ain, heh, kheh, heh, jeem, cheh
- shin, sin, yih, beh, lam, alef, teh, nuun, meem, kaf, gaf, backslash
- zuy, tuy, zeh, reh, zal, dal, peh, waw, full-stop, slash

### Shifted:

- division, exclamation, thousands, decimals, afghani, percent, multiplication, comma, asterisk, parenthesis, kashida, plus
- sukun, dammatan, kasratan, fathatan, damma, kasra, fatha, shadda, square-bracket, curly-bracket, curly-bracket, curly-bracket
- waw-hamza, yih-hamza, yeeh, yeh, alef-hamza, alef-madha, teh marbuta, guillemet, colon, semicolon, vertical-bar
- o', u, zheh, alef-hamza-below, ZWNJ, hamza-above, hamza, less-than, greater-than, question

### Alt-Gr:

- grave, tilde, at, number, dollar, percent, circumflex, ampersand, 5-star, bullet, degree, underscore, —
- euro, alef-wasla, alef-maksura, LRM, RLM, maddah, zwarakay, —, apostrophe, quote, pashto-jeh, pashto-ceh
- pashto-shin, shajina-yih, arabic-kaf, —, brahui-lham, —, pashto-teh, pashto-nuun, nuun-ghunna, semicolon, pashto-gaf, question
- urdu-yih, urdu-teh, pashto-zeh, pashto-feh, urdu-dal, pashto-dal, urdu-zeh, comma, —, —

# *Computer Locale Requirements for Afghanistan*

## **Keyboard layouts for Afghanistan**

Key assignments for Southern Uzbek, using ISO/IEC 9995 notation:

	<b>Unshifted</b>	<b>Shifted</b>	<b>AltGr</b>
TLDE	200D	00F7	0060
E01	06F1	0021	007E
E02	06F2	066C	0040
E03	06F3	066B	0023
E04	06F4	AFGHANI	0024
E05	06F5	066A	0025
E06	06F6	00D7	005E
E07	06F7	060C	0026
E08	06F8	002A	066D
E09	06F9	0029	2022
E10	06F0	0028	00B0
E11	002D	0640	005F
E12	003D	002B	<FREE>
D01	0636	0652	20AC
D02	0635	064C	0671
D03	062B	064D	0649
D04	0642	064B	200E
D05	0641	064F	200F
D06	063A	0650	0653
D07	0639	064E	ZWARAKAY
D08	0647	0651	0670
D09	062E	005D	0027
D10	062D	005B	0022
D11	062C	007D	0681
D12	0686	007B	0685
C01	0634	0624	069A
C02	0633	0626	06CD
C03	06CC	064A	0643
C04	0628	06D0	<FREE>
C05	0644	0623	06B7
C06	0627	0622	<FREE>
C07	062A	0629	067C
C08	0646	00BB	06BC
C09	0645	00AB	06BA
C10	06A9	003A	003B
C11	06AF	061B	06AB
B01	0638	06C9	06D2
B02	0637	06C7	0691
B03	0632	0698	0696
B04	0631	0625	0693
B05	0630	200C	0688
B06	062F	0654	0689
B07	067E	0621	0679
B08	0648	003E	002C
B09	002E	003C	<FREE>
B10	002F	061F	<FREE>
BKSL	005C	007C	003F
SPCE	0020	200C	00A0

## Computer Locale Requirements for Afghanistan

### The official calendar of Afghanistan

The official calendar of Afghanistan is the solar Islamic calendar (also known as the Persian calendar or the *hejrī-e šamsī* calendar). This calendar counts up from the year of the Hegira of Muhammad. The Gregorian calendar is used in international activities or occasions, and the Lunar Islamic calendar (also known as *hejrī-e qamarī*) is used for religious ceremonies and some of the holidays.

The Persian calendar of Afghanistan, although very similar to the Persian calendar of Iran, differs in the algorithm it uses to calculate leap years, which may lead to a one-day difference for some years. There are plans to synchronize these calendars, but the present report only describes the current calendar of Afghanistan. The Persian calendar has 12 months, consisting of 29 to 31 days. The names of the months in Dari and Pashto, together with the number of the days in each month, is given below. Other languages of Afghanistan usually use the Dari names of the months.

Dari	Pashto	Days
حمل Hamal	وری Wray	31
ثور Sawr	غویی Ğwayay	31
جوزا Jawzā	غبرگولی Ğbargolay	31
سرطان Saratān	چنگاښ Čungāš	31
اسد Asad	زمری Zmaray	31
سنبله Sonbola	وږی Waṛay	31
میزان Mīzān	تله Təla	30
عقرب Aqrab	لړم Laṛam	30
قوس Qaws	لیندی Lindəy	30
جدی Jady	مرغومی Margūmay	30
دلو Dalv	سلواغه Salwāğə	30
حوت Hūt	کب Kab	29 (30 in leap years)

Iranian leap years are computed astronomically, but the Afghan leap years have an arithmetic formula. The formula is for synchronizing with the Gregorian calendar, and is rather simple in that regard: the Persian year  $x$  is a leap year if and only if the Gregorian year  $x + 621$  is a leap year. For example, the Persian year 1383 will be a leap year in the Afghan Persian calendar, since the Gregorian year 2004 will be a leap year. For the sake of conversion between Gregorian dates and Persian years, one can use the following date for a reference point: **1 Hamal 1382 = 21 March 2003**.

The Lunar Islamic calendar used in Afghanistan is a civil Lunar Islamic calendar, based on pre-computation of the months. The number of the days of the months are 29 and 30 days alternatively. Leap years occur every two or three years, and make the month *Dhu l-Qa'da* a day longer. The exact algorithm used for leap years is reported to be complicated and was not made available to us. It is worth mentioning that because of the importance of Islamic months *Ramadan* and *Shawwal* in Muslim fasting and other ceremonies, the starting day of these two months is as observed by religious or governmental authorities rather than pre-computed. There have been (and will be) frequent cases of one day difference between the pre-computed calendar and the observed one. This difference is usually adjusted at the end of *Shawwal*, and the pre-computed calendar is used from the next month.

# Computer Locale Requirements for Afghanistan

## Acknowledgements

*This report could not have been produced without the kind and expert assistance of a great many people. Our heartfelt thanks are especially given to Habibullah Rafi and to Said Marjan Zazai, who shared their expertise with great kindness and patience. We apologize to any person whose name we have inadvertently omitted here. – ME & RP*

**Muhammad Afzal**, University of Arid Agriculture, Rawalpindi  
**Nurollah Altay**, Institute of Uzbek and Turkmen Languages and Literature, Kabul  
**Muhammad Assam**, Ministry of Communications, Kabul  
**Christian Cooke**, fishamble.net, Dublin  
**Noorullah Delawari**, Ministry of Finance, Kabul  
**Rahim Ebrahim**, Institute of Uzbek and Turkmen Languages and Literature, Kabul  
**Abbas Edalat**, Science and Arts Foundation, London  
**Joseph Elfenbein**, Ancient India and Iran Trust, Cambridge  
**Behdad Esfahbod**, FarsiWeb Project, Tehran  
**Asmus Freytag**, Asmus, Inc., Seattle  
**Roman Gehring**, SIL International  
**Jost Gippert**, TITUS project, Frankfurt  
**Bob Hallissy**, SIL International, Horsleys Green, England  
**Javid Hamdard**, Ministry of Finance, Kabul  
**Horia Samira Hamidi**, UNDP Afghanistan, Kabul  
**Fazl Ahmad Hedayat**, Head Office of Geodesy and Cartography, Kabul  
**Sarmad Hussain**, Foundation for Advancement of Science and Technology, Lahore  
**Shirzad Kamawal**, shirzad.de, Berlin  
**Jonathan Kew**, SIL International, Horsleys Green, England  
**Raj Wali Shah Khattak**, Pushto Academy, University of Peshawar  
**Ghutay Khawre**, Afghanistan Academy of Sciences, Kabul  
**Alireza Kheirkhahan**, FarsiWeb Project, Tehran  
**Mohammad Alem Kuhkan**, Radio Liberty, Kabul  
**Sayyed Mohammad Alem Labib**, Institute of Uzbek and Turkmen Languages and Literature, Kabul  
**Marc Lepage**, UNDP Afghanistan, Kabul  
**Abdur Rauf Mahsood**, Khpala Pashto, Peshawar  
**Nasrollah Shubman Mangol**, Afghanistan Academy of Sciences, Kabul  
**Muhammad Aimal Marjan**, Ministry of Communications, Kabul  
**Ron Meyer-Hagen**, SIL International  
**Yari Ata Mohammad**, UNDP Afghanistan, Kabul  
**Ercan Murat**, UNDP Afghanistan, Kabul  
**Mohammad Tamim Noori**, Afghanistan Assistance Coordination Authority, Kabul  
**Mohammad Amin Oriya**, UNDP Afghanistan, Kabul  
**Knut Ostby**, UNDP Afghanistan, Kabul  
**Ah. Farhad Popal**, Ministry of Finance, Kabul  
**Habibullah Rafi**, *Kelid Magazine*, Kabul

**Elnaz Sarbar**, FarsiWeb Project, Tehran  
**James Seng**, Infocomm Development Authority, Singapore  
**Mohammad Sepehry Rad**, High Council of Informatics of Iran, Tehran  
**Nazar Ahmad Shah**, UNDP Afghanistan, Kabul  
**Mr Shams**, Department of Linguistics, University of Kabul  
**Nicholas Sims-Williams**, School of Oriental and African Studies, London  
**Masoum Stanekzai**, Ministry of Communications, Kabul  
**Yahya Tabesh**, Sharif University of Technology, Tehran  
**Massoumeh Torfeh**, Afghanistan Assistance Coordination Authority, Kabul  
**Mohammad Ya'ghub Vahedi**, Institute of Uzbek and Turkmen Languages and Literature, Kabul  
**Haron Wardak**, hewad.com, Göteborg  
**Ken Whistler**, Sybase, Dublin, California  
**Cathy Wissink**, Microsoft, Redmond  
**Ismail Yoon**, University of Kabul  
**Ghulam Rasoul Yosoufzai**, Ministry of Information and Culture, Kabul  
**Mahmood Zahir**, UNDP Afghanistan, Kabul  
**Abdul Zahir Gulistani**, Ministry of Education, Kabul  
**Said Marjan Zazai**, Afghanistan Assistance Coordination Authority, Kabul

*We are also grateful to the staff of the following organizations:*

Afghanistan Academy of Sciences, Kabul  
Afghan National Bank, Kabul  
Centre for Dari Language and Literature, University of Kabul  
Da Afghanistan Bank, Kabul  
Department of Literature and Linguistics, University of Kabul  
Department of Turkmen and Uzbek Language and Literature, School of Literature and Human Sciences, Balkh University  
Institute of Uzbek and Turkmen Languages & Literature, Afghanistan Academy of Sciences, Kabul  
Ministry of Communication, Kabul  
Ministry of Education, Kabul  
Ministry of Finance, Kabul  
Ministry of Foreign Affairs, Kabul  
Ministry of Information and Culture, Kabul  
Office for Public Libraries, Kabul  
UNDP Afghanistan, Kabul