Non-Unicode to Unicode conversion

John Stephen Mangam | johnmangam@gmail.com

Contents

[Introduction 1](#_Toc521590519)

[Technology 1](#_Toc521590520)

[SILBulkWordConverter 1](#_Toc521590521)

[Visual Studio and C# 2](#_Toc521590522)

[Code base 2](#_Toc521590523)

[Pre-requisites 2](#_Toc521590524)

[Web-Server 2](#_Toc521590525)

[Non-Unicode fonts 2](#_Toc521590526)

[MS-Word 2](#_Toc521590527)

[SIL Converters and TEC maps 2](#_Toc521590528)

[Usage 3](#_Toc521590529)

[File Conversion 3](#_Toc521590530)

[Text Conversion 5](#_Toc521590531)

[Adding conversion ability 6](#_Toc521590532)

[TECk file generation 6](#_Toc521590533)

[TEC file addition to web-application 6](#_Toc521590534)

# Introduction

**File and Text conversion** for **non-Unicode fonts using TEC maps** is made possible through this web-application using the **SIL Converters**.

# Technology

**SILBulkWordConverter windows application** has been modified and made as a **library** and **TECkit maps** are created for the legacy fonts to facilitate the **Unicode conversion.**

## SILBulkWordConverter

The **C#** source code for it is available from: <https://github.com/bobeaton/SilConverters> (for the solution that includes the project “SILConvertersWordML” – the Bulk Word Document Converter project).

## Visual Studio and C#

# Code base

It can be found at <https://github.com/Bridgeconn/DOTNET-UnicodeConversion-WebApp>

# Pre-requisites

## Web-Server

### Non-Unicode fonts

Fonts folder in the web app code base has the non-Unicode fonts

### MS-Word

SIL Conversion uses MS-Word for conversion and we use SIL Converter platform so MS-Word is required.

### SIL Converters and TEC maps

Step (1) Install SIL Converters from <http://downloads.sil.org/EncodingConverters/SEC%204.0%20Web/SEC_FullInstall.exe>

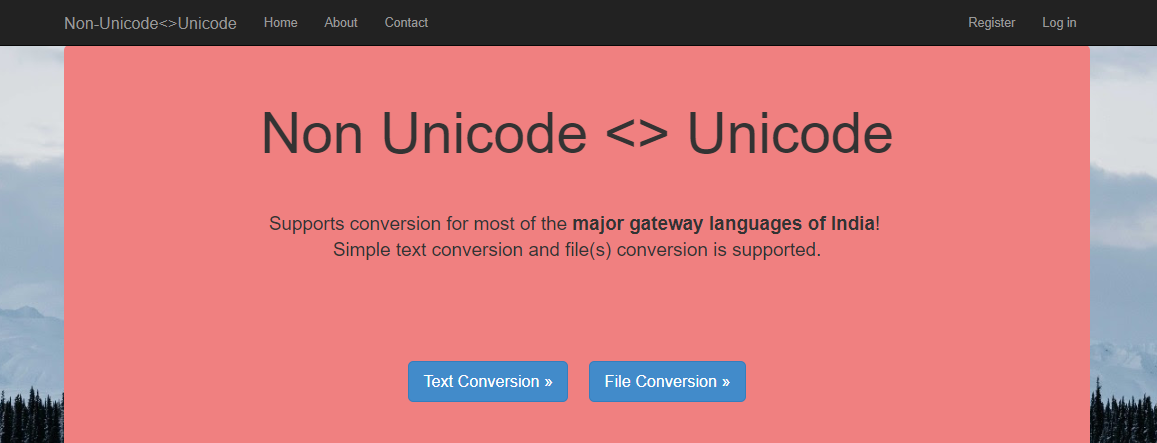
Step (2) After installing SIL Converters, it would ask if we need to install .NET 2.0 shared support. Please click **yes**.

Step (3) Finally, the below-shown installer would come up, **select all and click OK**.

A screenshot of a cell phone

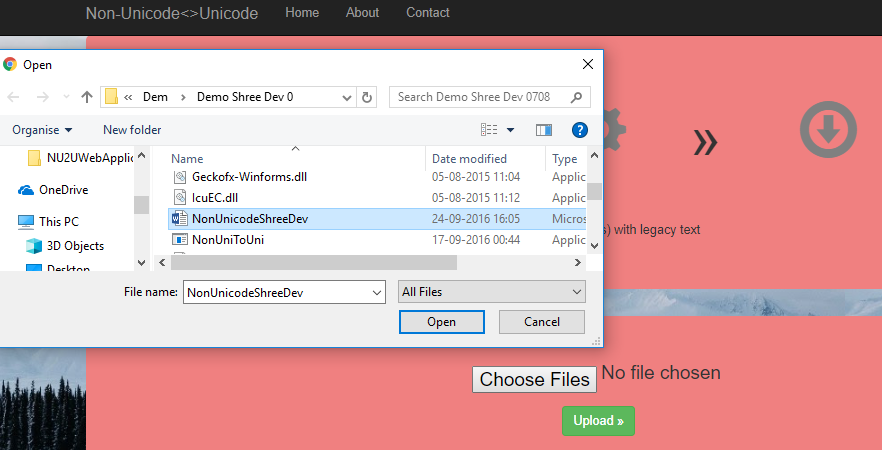
Description generated with very high confidence

# Usage



## File Conversion

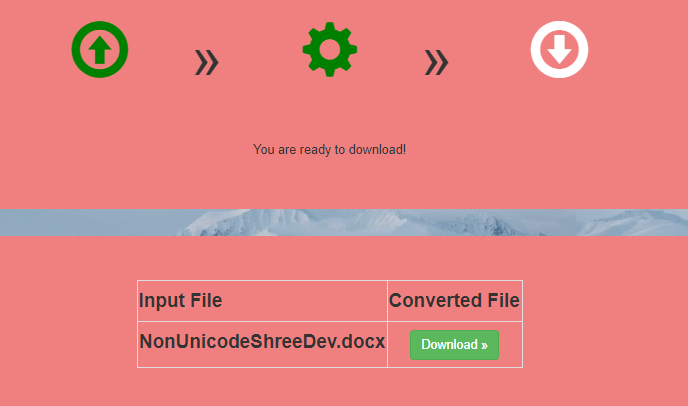
Step 1: **Choose a file** and click **the upload button**



Step 2: **Select Unicode fonts** for the conversion and click on **convert button**



Step 3: Click on **download button**



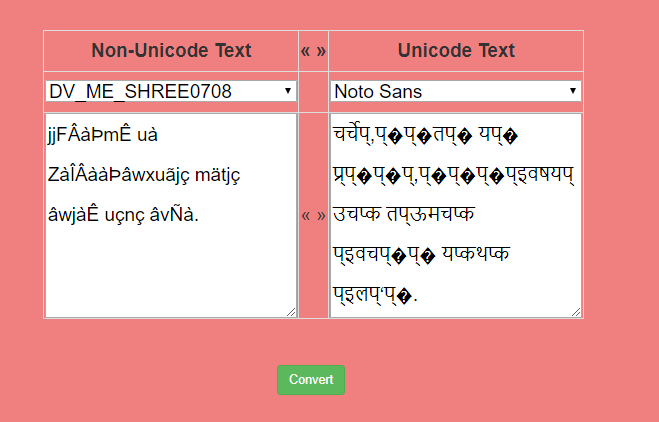
## Text Conversion

Step 1: Select the non-Unicode font

Step 2: Paste the text that needs to be converted

Step 3: Select Unicode font

Step 4: Click convert



# Adding conversion ability

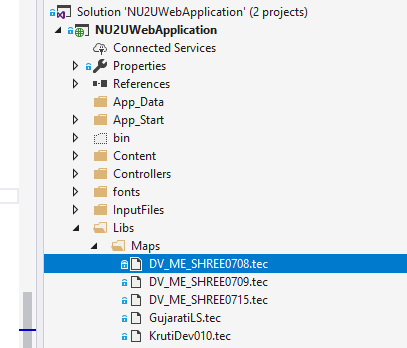
## TECk file generation

To add the ability to convert non-Unicode to Unicode, for that legacy non-Unicode font, a TEC file needs to be created. The process is best explained by SIL: <https://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=TECkitMapping01>

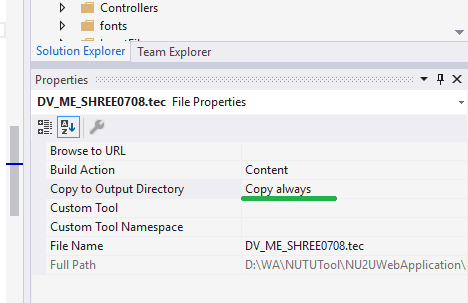
The non-Unicode’s equivalent Unicode and the rules can be stated in the TEC file. Once the TEC file works well, it needs to be added to the web application.

## TEC file addition to web-application

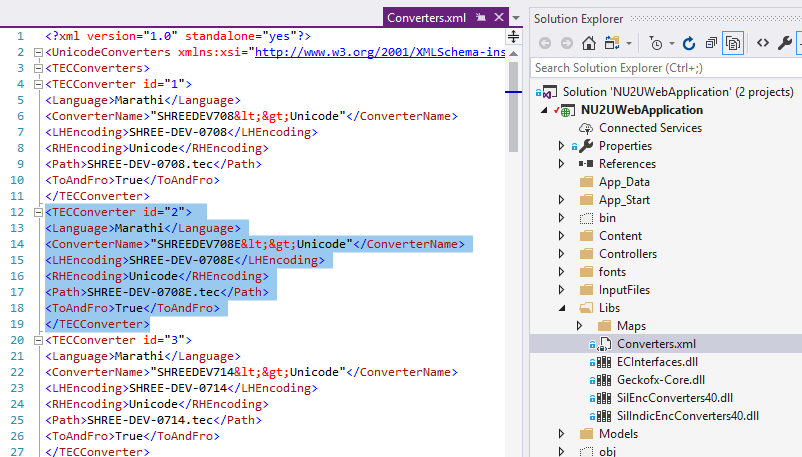
Step 1: Add the TEC file as shown below



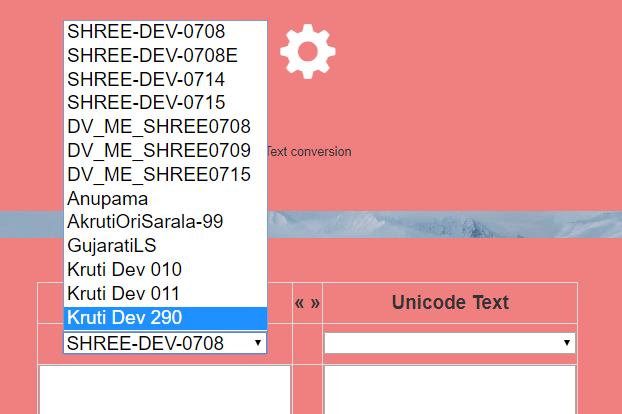
Step 2: In the properties of the TEC file, copy to output needs to be selected:



Step 3: In the **Converters.xml**, add an <**TECConverter** id=<next number>”> element with respective details as shown below:



Step 4: To verify that it is working, in the text conversion of the web-application, the non-Unicode font needs to be seen.



Step 5: Once the conversion works, the web application can be republished so that the new capability is made available for use by others.