SOP OF IDN-UA COMPLIANCE FOR INDIAN GOVERNMENT WEBSITES

Support document to enable URLs from English to Multilingual Indian languages

MARCH 28, 2023

CENTRE FOR DEVELOPMENT OF ADVANCED COMPUTING
C-DAC Innovation Park, Panchvati, Pashan, Pune - 411008

Table of Contents

Objective 2
URL (Uniform Resource Locator) Structure 2
IDN enabled Domain registration 2
Punycode 3
Universal Acceptance 3
IDN Registrar 4
IDN Domain Availability 4
Availability search result 5
Appointed Registrars 5
Drupal 7 Configuration 6
Drupal Version 8 and above IDN compliance 8
Content Translation Module 8
Enabling Multilingual Content 9
Content in bilingual/multilingual 9
Frontend Language Switcher [Dropdown/Link] 12
IDN support 13
IDN-UA Compliance Script 15
GoI Content Management Framework Compliance 15
IDN domain-compliant examples 16
Government Website (.gov.in/.सरकार.भारत) 16
Indian Domain Website (.IN/.भारत) 17
Support Contact Details 18

Copyright © 2023 C-DAC

IDN-UA Compliance

Enabling multilingual internet

Objective

This document is a reference document to enable IDN (internationalize domain name) based URL support for Indian regional bilingual/multilingual in Drupal CMS for the Indian government websites.

URL (Uniform Resource Locator) Structure

A URL (Uniform Resource Locator) is a reference to a web resource. Most browser display URLs in their address bar. In general, a URL consists of the following structure.

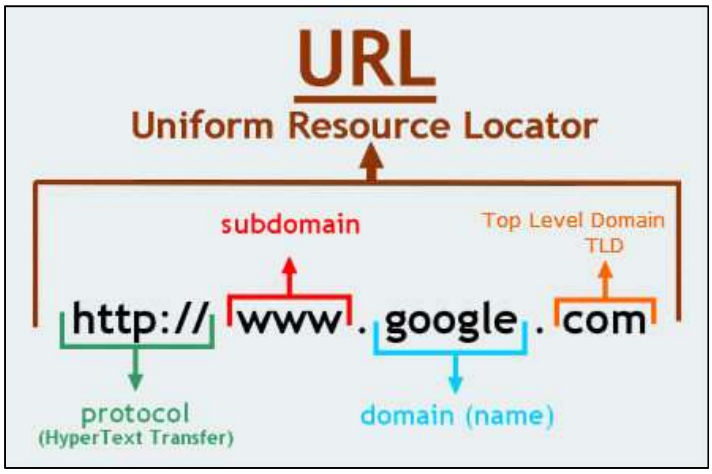


Figure 1: URL structure for reference only

IDN enabled Domain registration

To change a domain name to the Hindi language, you would need to use Internationalized Domain Names (IDN), Internationalized Domain Names (IDNs) enable people around the world to use domain names in local languages and scripts. So, domain names are to be written using characters from non-Latin scripts such as Hindi.

Steps to register a domain name in a regional language:

1. Choose a registrar that supports IDNs. Not all registrars support IDNs, so make sure to choose one that does.
2. Search for the Hindi domain name you want. You can use a tool like Punycode Converter to convert your desired regional language domain name into its Punycode equivalent, which can then be used to search for the domain.
3. Register the domain name. Once you've found an available domain name, you can register it like any other domain name. The registration process will be the same as for any other domain name.
4. Set up the domain name. Once you've registered the domain name, you can set it up like any other domain name. You will need to set up DNS records and configure any necessary settings.

Copyright © 2023 C-DAC

**Note that while IDNs allow domain names to be written in non-Latin scripts, they may not be supported by all web browsers and email clients. Additionally, some search engines may not be able to properly index IDN domain names. So, it's important to carefully consider the potential implications before changing your domain name to a Hindi IDN.*

ASCII Domain Name	Unicode Domain Name
cdac.in	सीडैक.भारत

Punycode

Punycode is a representation of Unicode with the limited ASCII character subset used for Internet hostnames. Using Punycode, host names containing Unicode characters are transcoded to a subset of ASCII consisting of letters, digits, and hyphens, which is called the letter–digit–hyphen (LDH) subset.

The Punycode syntax is a method of encoding strings containing Unicode characters, such as internationalized domain names (IDNA), into the LDH subset of ASCII favored by DNS.

It is specified in IETF Request for Comments 3492.

Unicode Domain Name	Punycode
सीडैक.भारत	xn--11bx2e6a3b.xn--h2brj9c

Universal Acceptance

Universal Acceptance (UA) is a set of best practices that ensures that all domain names are accepted, validated, stored, processed, and displayed correctly by all software applications. By making your website Universal Acceptance (UA) compliant, you can ensure that users with non-Latin script domain names can access your website without any issues.

All Domain names should be validated against the Internationalized Domain names in applications IDNA2008[1] standard.

Pre-requisite to enable website UA compliance:

1. **Use Internationalized Domain Names (IDNs):** IDNs allow domain names to be written in non-Latin scripts, such as Arabic, Chinese, Cyrillic, Devanagari, etc. You can register IDNs with your domain registrar and use them on your website as provided steps above.
2. **Ensure your website and applications can handle Unicode:** Unicode is a character encoding standard that supports almost all language scripts used in the world. Make sure your website and applications can handle Unicode to correctly display non-Latin script domain names.

With these prerequisites, you can make your website Universal Acceptance compliant and ensure that it can be accessed by users with non-Latin script domain names without any issues.

Copyright © 2023 C-DAC

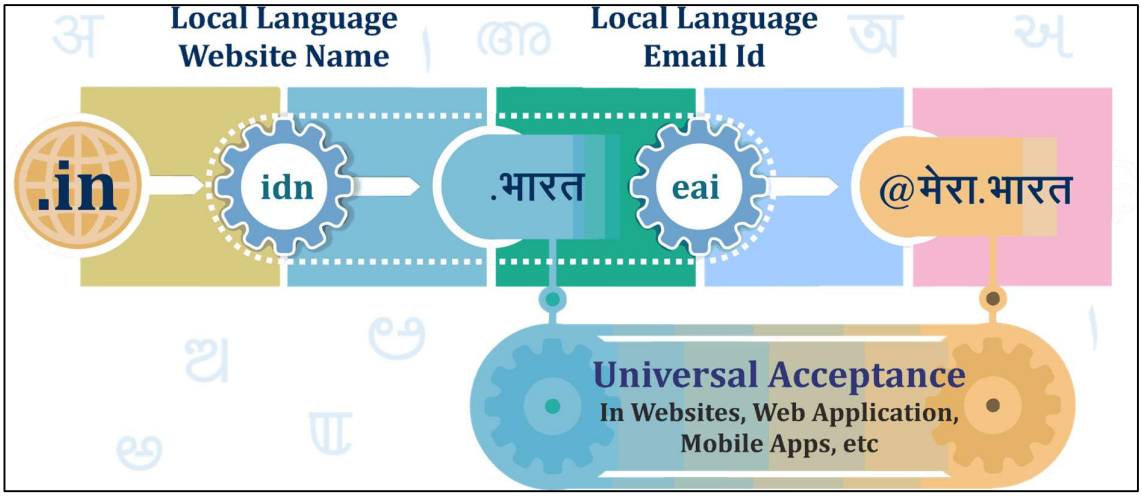


Figure 2: Universal Acceptance

IDN Registrar

The government of India has authorized NIXI (National Internet Exchange of India: https://nixi.in/).IN/.भारत registry through https://registry.in/ and .gov.in/ .सरकार.भारत through https://registry.gov.in/.

Figure 3: Registry.in process

IDN Domain Availability

You may search for domain names in the Indian regional language.

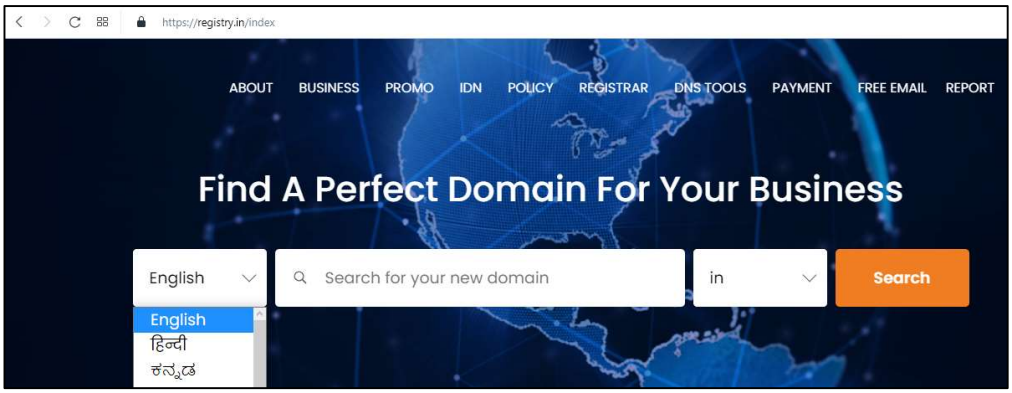


Figure 4: Domain Search in the registry.in

Copyright © 2023 C-DAC

Availability search result

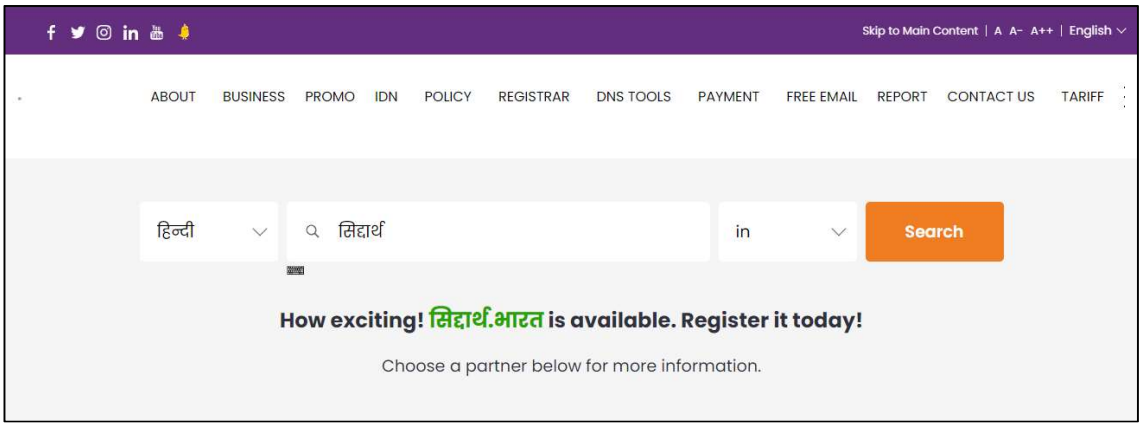
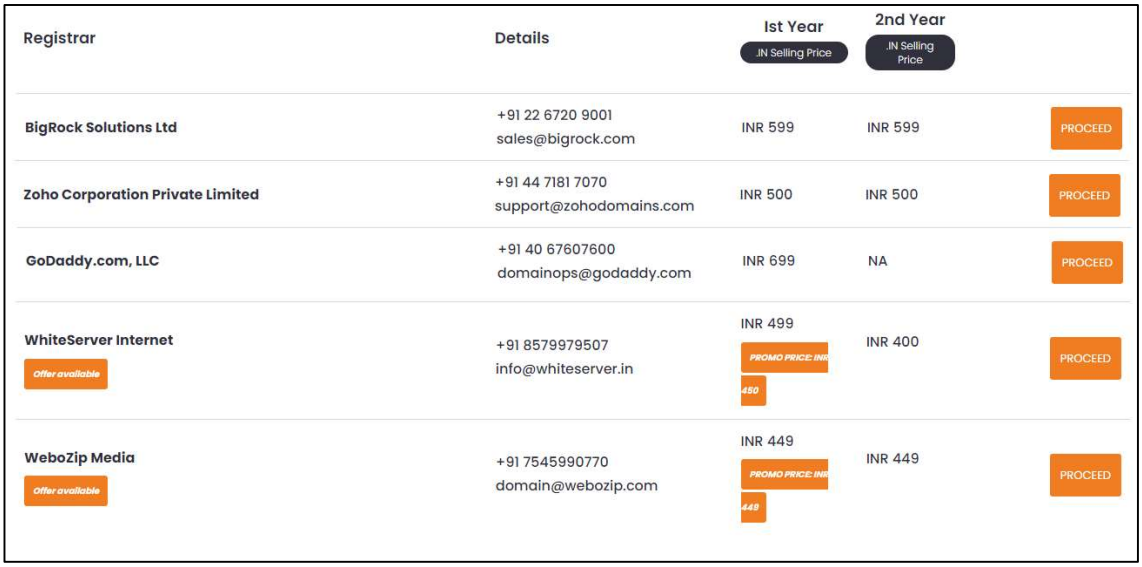


Figure 5: Domain availability in Hindi

Appointed Registrars

The.IN Registry by itself will not carry out registrations. It will do so through several Registrars to be appointed by it.



Registrar	Details	Ist Year .IN Selling Price	2nd Year .IN Selling Price	
BigRock Solutions Ltd	+91 22 6720 9001 sales@bigrock.com	INR 599	INR 599	PROCEED
Zoho Corporation Private Limited	+91 44 7181 7070 support@zohodomains.com	INR 500	INR 500	PROCEED
GoDaddy.com, LLC	+91 40 67607600 domainops@godaddy.com	INR 699	NA	PROCEED
WhiteServer Internet Offer available	+91 8579979507 info@whiteserver.in	INR 499 PROMO PRICE: INR 450	INR 400	PROCEED
WeboZip Media Offer available	+91 7545990770 domain@webozip.com	INR 449 PROMO PRICE: INR 449	INR 449	PROCEED

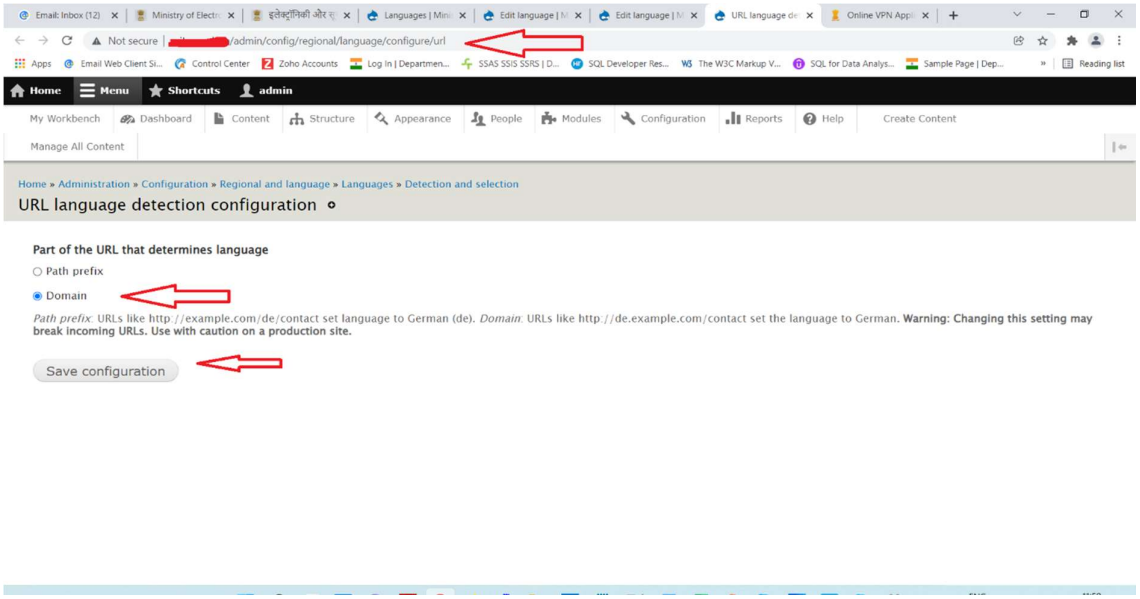
Figure 6: Appointed Registrar

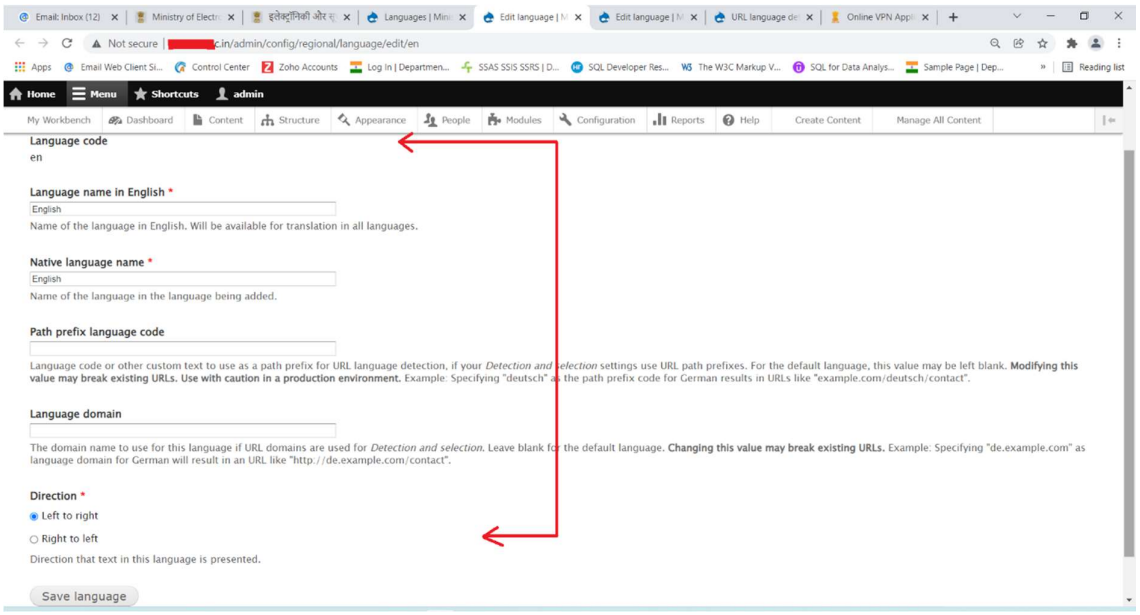
You may select any of the registrars to register the domain & hosting purposes.

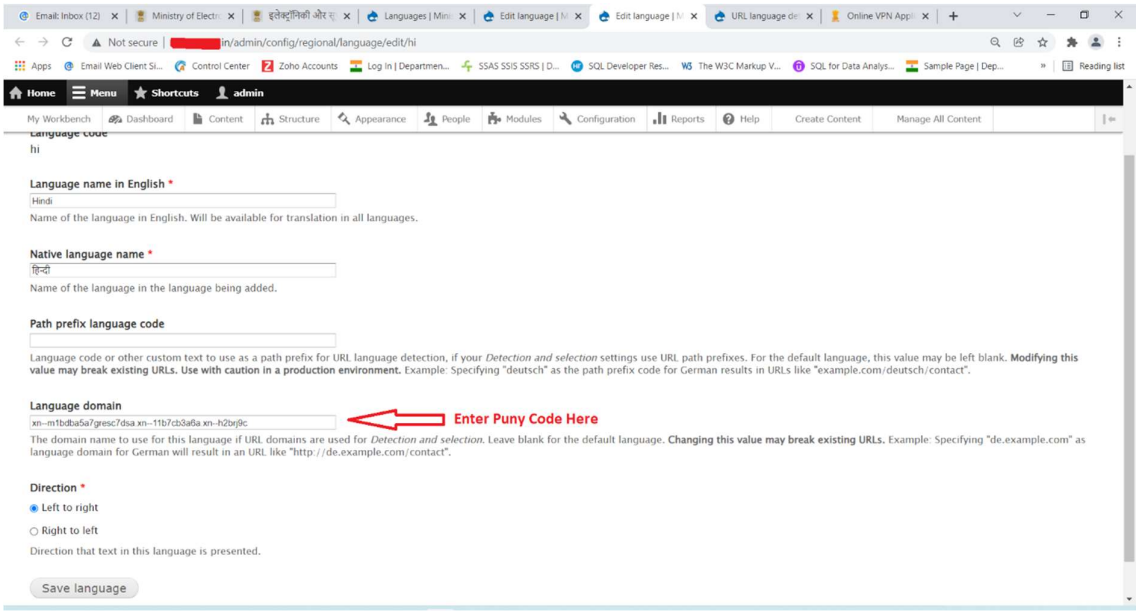
Copyright © 2023 C-DAC

Drupal 7 Configuration

Go to the path to configure the URL for






Not secure
[redacted]in/admin/config/regional/language/edit/hi
Home
Menu
Shortcuts
admin
My Workbench
Dashboard
Content
Structure
Appearance
People
Modules
Configuration
Reports
Help
Create Content
Manage All Content
hi
Language name in English *
Hindi
Name of the language in English. Will be available for translation in all languages.
Native language name *
हिन्दी
Name of the language in the language being added.
Path prefix language code
Language code or other custom text to use as a path prefix for URL language detection, if your Detection and selection settings use URL path prefixes. For the default language, this value may be left blank. Modifying this value may break existing URLs. Use with caution in a production environment. Example: Specifying "deutsch" as the path prefix code for German results in URLs like "example.com/deutsch/contact".
Language domain
xn--m1bdba5a7gresc7dsa.xn--11b7cb3a6a.xn--h2brj9c
Enter Puny Code Here
The domain name to use for this language if URL domains are used for Detection and selection. Leave blank for the default language. Changing this value may break existing URLs. Example: Specifying "de.example.com" as language domain for German will result in an URL like "http://de.example.com/contact".
Direction *
Left to right
Right to left
Direction that text in this language is presented.
Save language

Drupal Version 8 and above IDN compliance

Drupal is an open-source content management system (CMS) used for personal, corporate, political & government sites.

Most of the Indian government uses a standardized framework (Content Management Framework: https://cmf.gov.in/) using Drupal CMS.

Content Translation Module

In Drupal, the main method for translating content is to use the core Content Translation module, which has an approach very similar to the Drupal 7 and above "Entity Translation" module. Only a single node or entity is created. The entity is language-independent. Only the associated fields are flagged with a language.

Note: The content translation module is inbuilt with Drupal 8 and above.

To enable content translation in Drupal, follow the following steps:

1. Enable the core "Content Translation" module.
2. Navigate to configure: Configuration > Content language and translation (path: admin/config/regional/content-language)

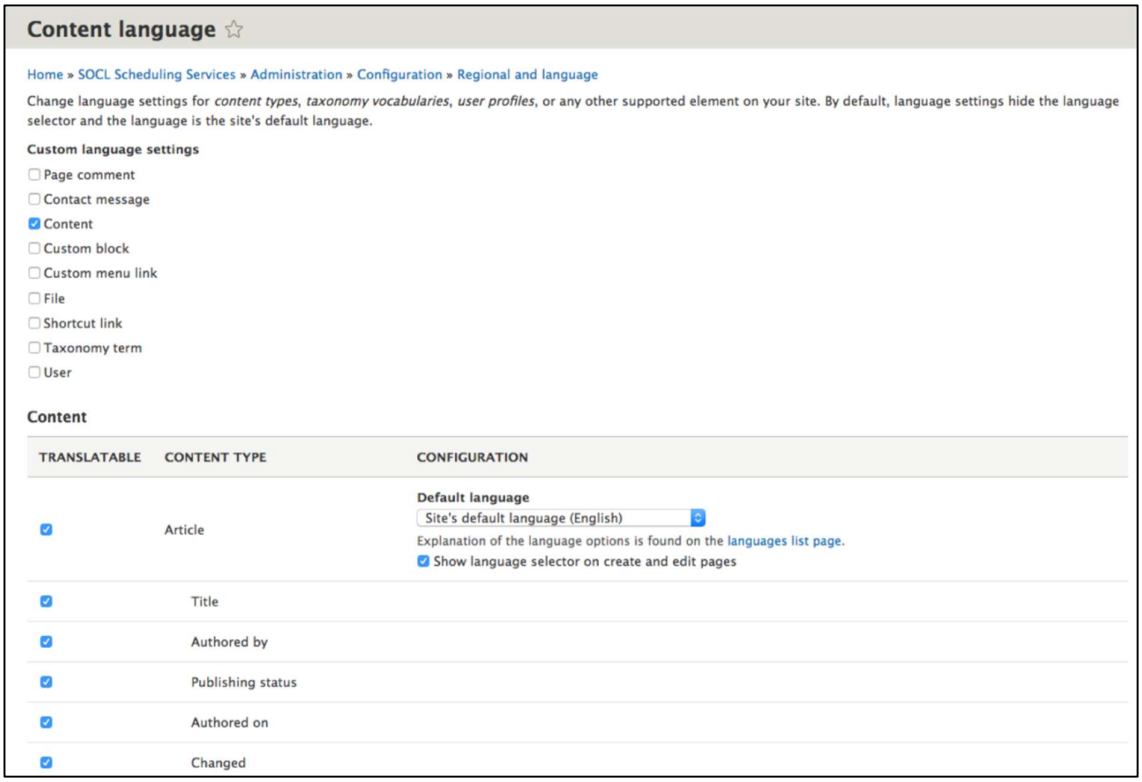


Figure 7: Content Translation Module

This module allows you to decide whether each type of content entity (node pages, comments, custom blocks, taxonomy terms, user accounts, etc.) should be translatable or not. Then, within each entity type, you can decide whether the sub-types (content types for node page content, terms in

particular vocabularies for taxonomy, etc.) should be translated. And within each translatable entity sub-type, you can decide which fields should be translatable.

Enabling Multilingual Content

Once you have designated certain fields of certain entity sub-types to be translatable, and you have installed at least two languages on your site, you can translate content items.

Users with translate permissions will see links to "Translate" alongside where you'd normally see "Edit" links, and you'll be able to add translations for each configured language.

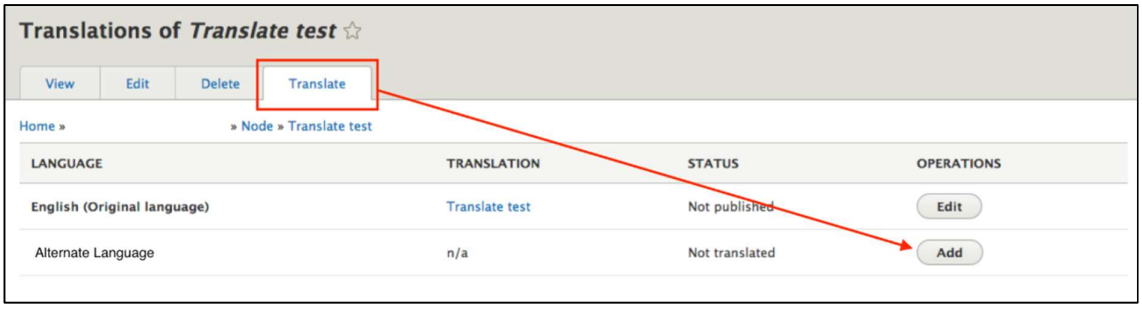


Figure 8: Table for Translation

Behind the scenes, you'll see that all the translations share the same ID (such as the node ID), and if you look at the Drupal database, you'll see that there is a table that stores information about each translation (such as the language and the node ID), and then each field stores its translated values in a separate table.

Content in bilingual/multilingual

To create content in bilingual/multilingual please follow the steps:

1. After admin login goes to content. Here we have a list of all created page content. Here, in the “Operations” column, select the “Translate” option for existing content. For creating new content click on the “Add Content” button.

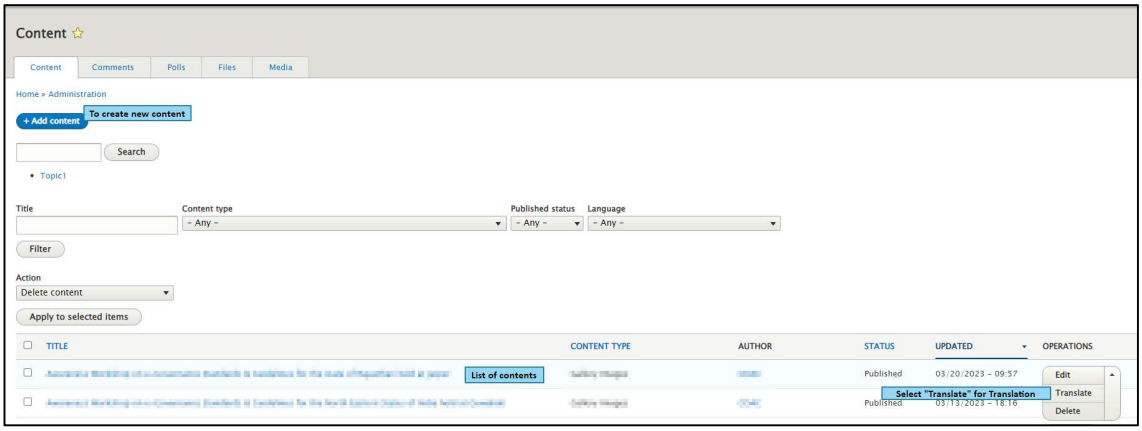


Figure 9: Translate in Multilingual

2. For creating new content, provides an option to choose a language for adding content.

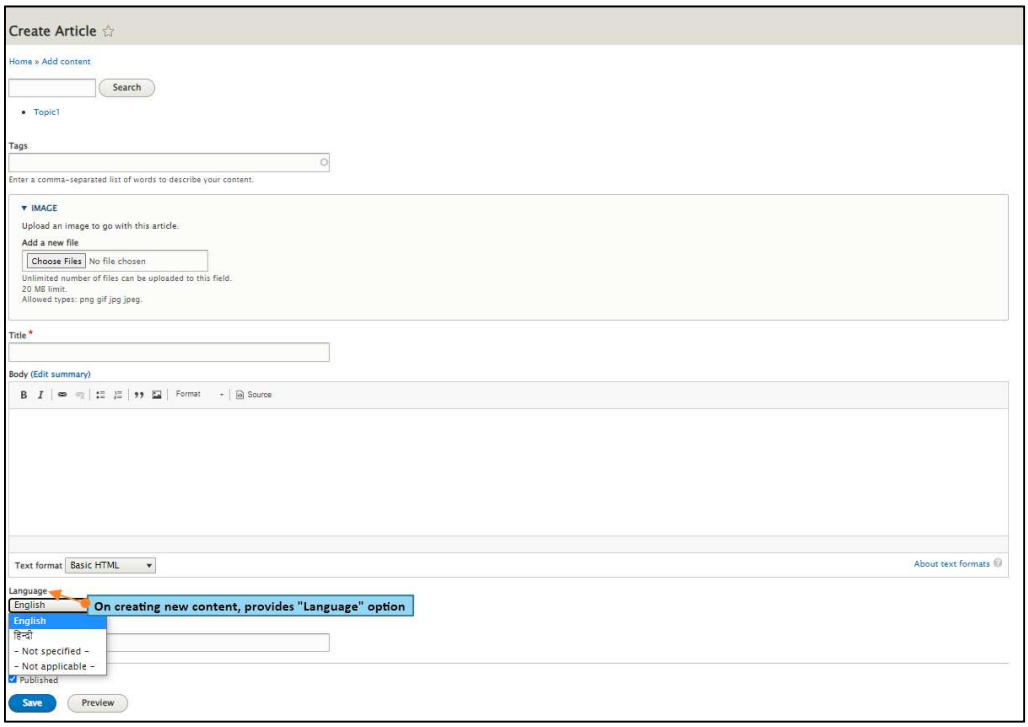


Figure 10: New content language selection

3. For existing content, shows available content in languages. You can edit content based on the translate tab shows an option.

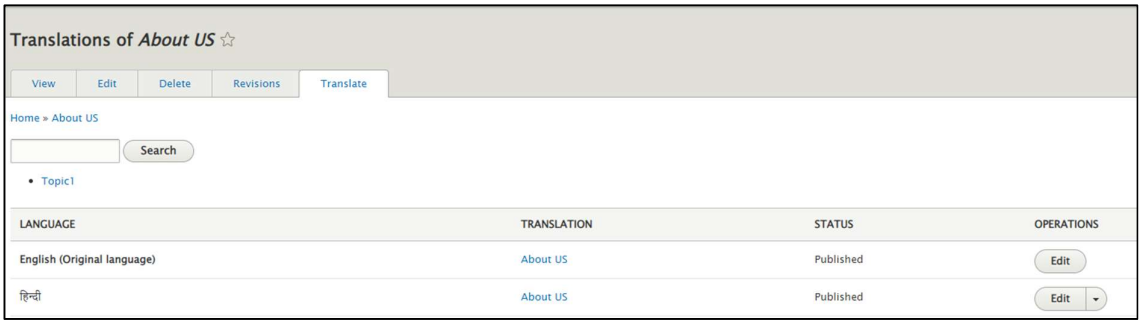


Figure 11: Content to translate

Copyright © 2023 C-DAC

4. Drupal article content in the English language.

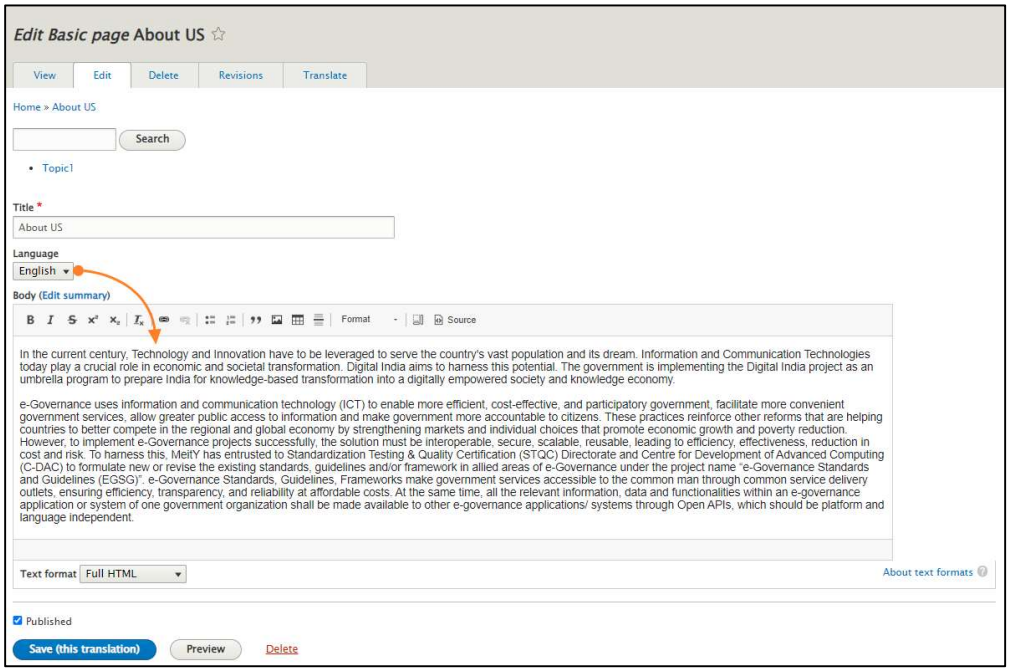


Figure 12: Content in the English language

5. Same article in the Hindi language.

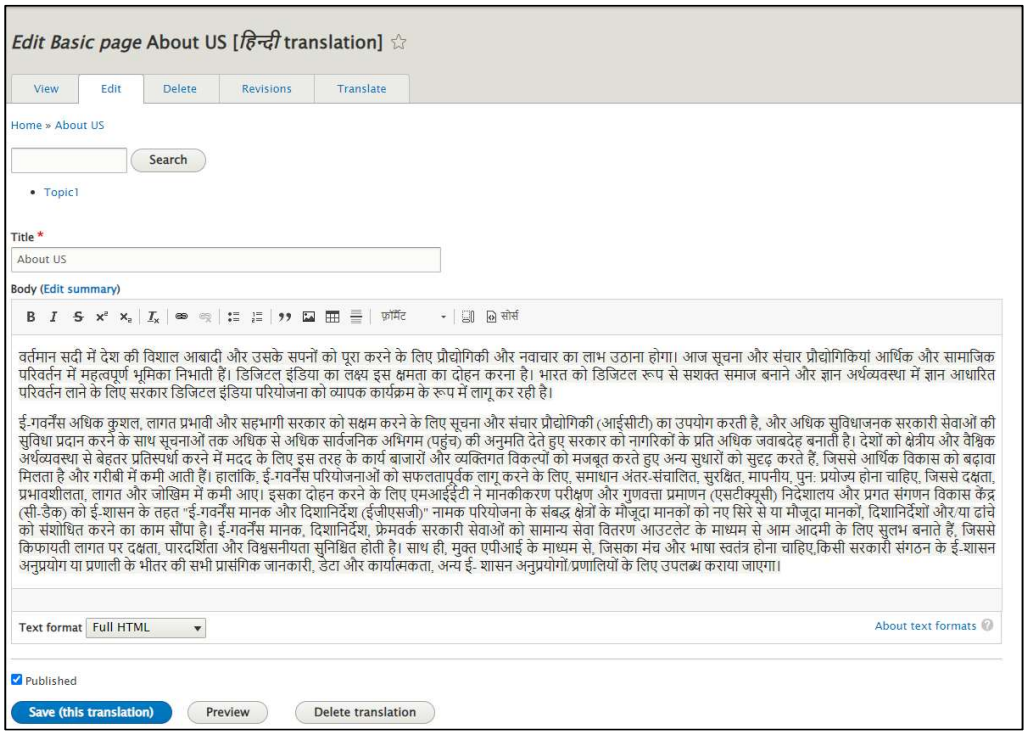


Figure 13: Content in Hindi

Frontend Language Switcher [Dropdown/Link]

After enabling the module, in the front by default a language dropdown appears, which allows changing content based on the selected language.

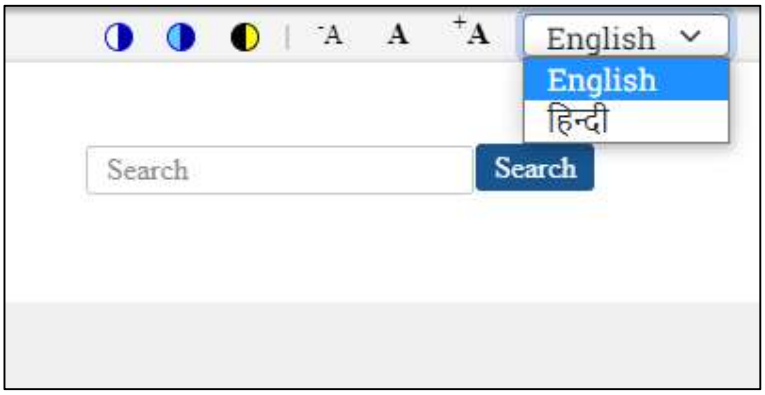


Figure 14: Frontend language selector

Note*: To be effective use, your content must be available in bi/multilingual.

Figure 15 Drupal default language selector functionality

IDN support

To enable domain names in regional languages like ".in" to ".भारत" in Drupal CMS, follow the steps:

1. **Install the Internationalization (i18n) module:** This module provides multilingual features to Drupal, including language negotiation and translation tools.
2. **Enable the Hindi language:** After installing the i18n module, go to the "Languages" page (Configuration > Regional and language > Languages) and enable the Hindi language.

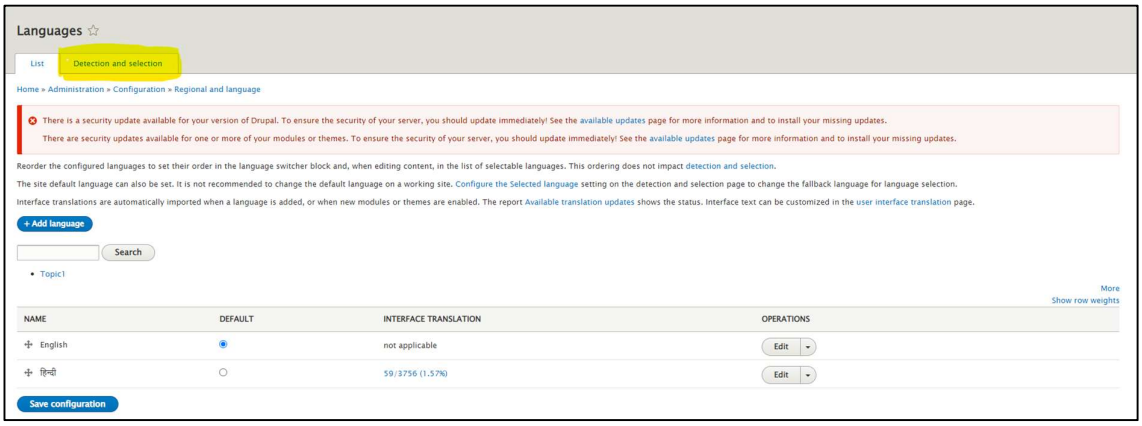


Figure 16: Drupal adds languages

Keep the setting as given in the Screenshot

3. **Configure language negotiation:** Language negotiation determines which language to use for a particular request. Go to the "Detection and selection" tab on the "Languages" page and set up the language negotiation method to use the "URL" method. This means that the language will be determined by the URL, so you'll need to create a new URL for the Hindi version of your site.

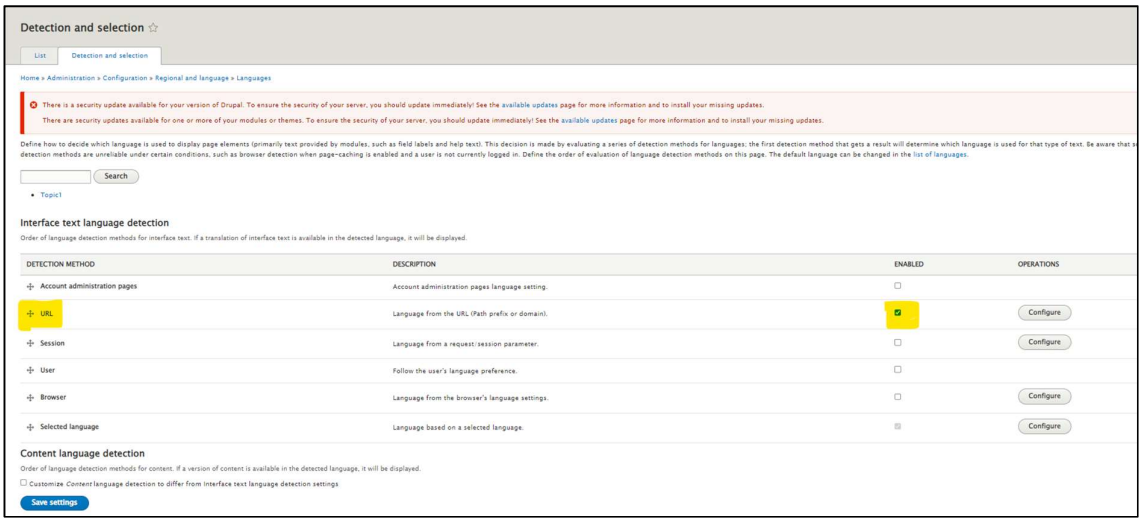


Figure 17: Drupal language negotiation configuration

4. **Create a new URL for the Hindi version of your site:** You can create a new URL for the Hindi version of your site by adding the language code to the URL. For example, if your current URL is "example.com", the Hindi version of your site could be "example.com/hi". To create this URL, go to the "URL aliases" page (Configuration > Search and metadata > URL aliases) and create a new alias for the "hi" language code.
5. **Create a new URL for the Hindi version of your site:** You can create a new URL for the Hindi version of your site by adding the language code to the URL. For example, if your current URL is "example.com", the Hindi version of your site could be "example.com/hi". To create this URL, go to the "URL aliases" page (Configuration > Search and metadata > URL aliases) and create a new alias for the "hi" language code.
6. Go to the "Configuration" tab on the top menu and select "Regional and language".
7. Click on the "Languages" tab.
8. Click the "Add Language" button and select "Hindi" from the drop-down list.
9. Set the "Path prefix" field to "hi" (which is the language code for Hindi).

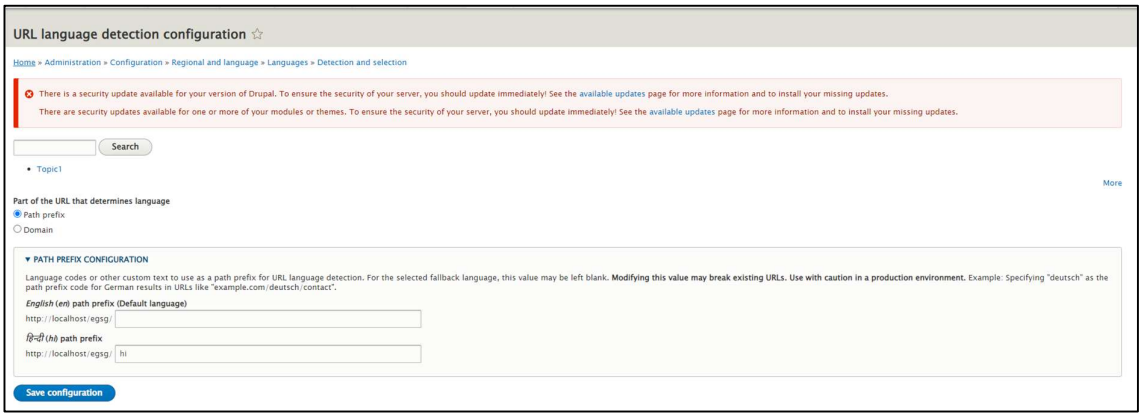


Figure 18: Drupal language detection configuration

Keep the setting as given in the Screenshot

IDN-UA Compliance Script

To comply with the regional language domain name to render the same language content a generic compliance script needs to be added to the Drupal CMS default index page or footer section.

```
window.onload = function() {
    var langChecker = document.getElementById('edit-lang-dropdown-select').value;
    if ( langChecker != localStorage.getItem('lang')){

    if (langChecker == 'hi' ){
    localStorage.setItem('lang', 'hi')
    location.assign("http://ईगवमानक.भारत/hi")
    }
    else {
    localStorage.setItem('lang', 'en')
    location.assign("http://egovstd.in/")
    }

    }
}
```

Please Note the Screenshot is just for example kindly use your domain name

GoI Content Management Framework Compliance

To make GoI CMF application for IDN-UA compliance, need to follow following steps:

1. As per GoI CMF .htaccess file (mostly available at root directory of website) having rewrite rule for domain name need to be disable by comment.

```
RewriteCond %{HTTP_HOST} !^dot.gov.in$ [NC]
 RewriteRule ^(.*?)$ https://dot.gov.in/$1 [L,R=301,NC,NE]
```

2. Add a new rule for punycode in .htaccess file, replace punycode with equivalent punycode for non-ASCII URLs.

```
RewriteCond %{HTTP_HOST} !^punycode$ [NC]
RewriteRule ^(.*?)$ https://punycode/hi/$1 [L,R=301,NC,NE]
```

3. .module file is CMF module responsible for dropdown language selection.

```
if($strLan == "hi"){
        $new_path="https://non-ASCII Domain name/hi";
}else{
        $new_path="https://ASCII Domain name/";
}
```

Copyright © 2023 C-DAC

IDN domain-compliant examples

Some of the Indian government websites used Drupal as CMS and IDN compliant.

Government Website (.gov.in/.सरकार.भारत)

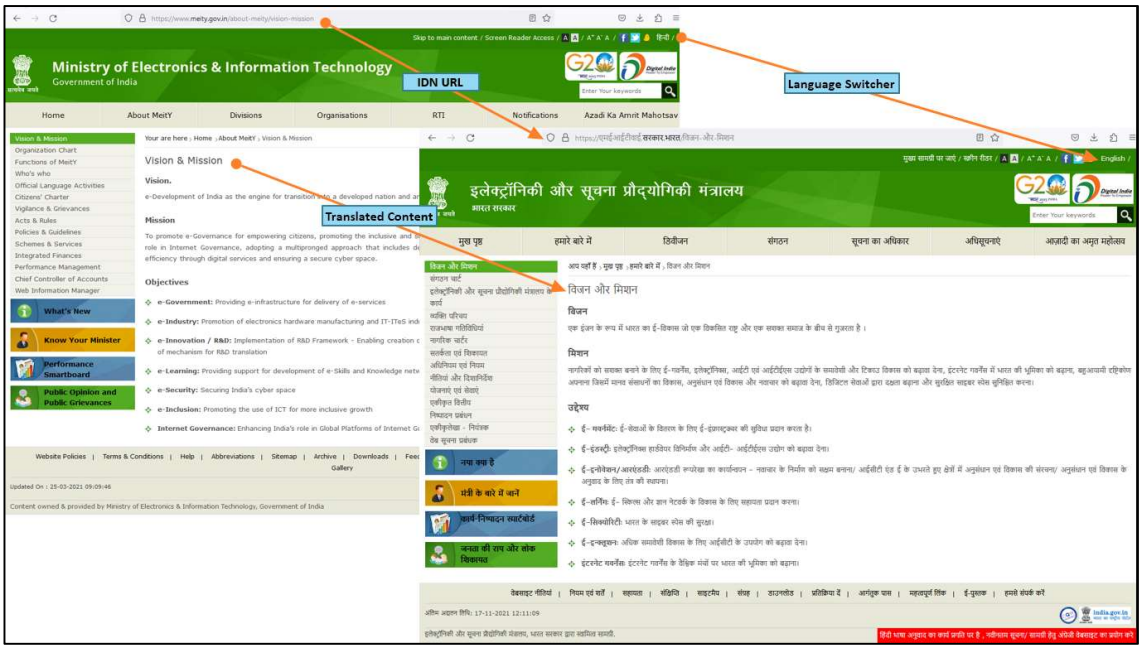


Figure 19: www.meity.gov.in

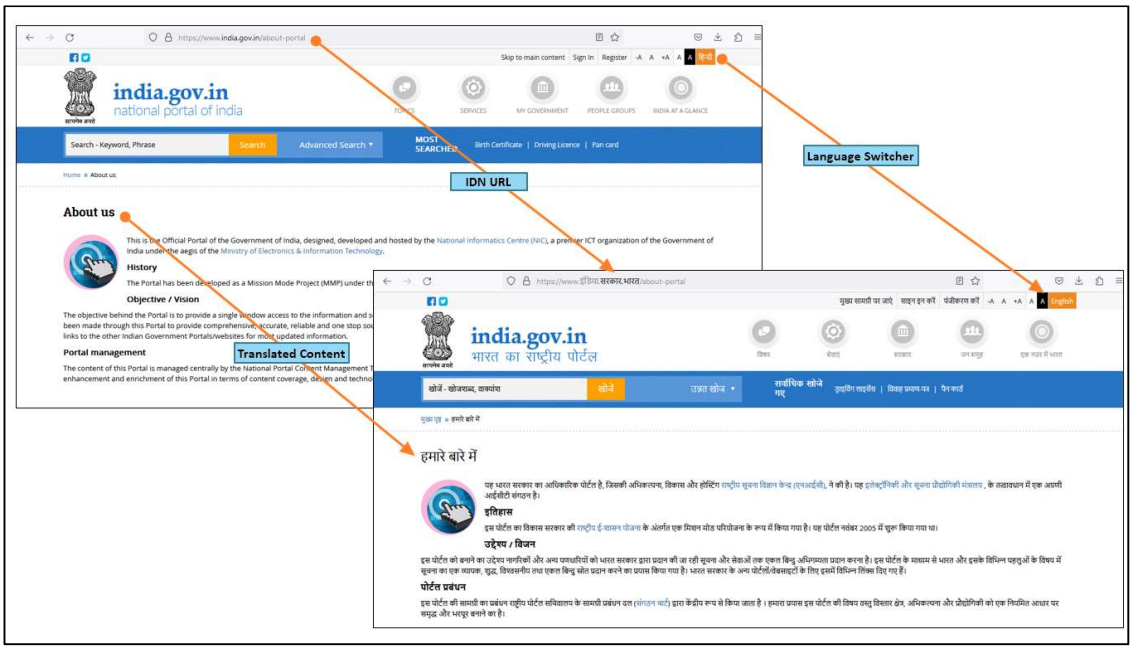


Figure 20: www.india.gov.in

Indian Domain Website (.IN/.भारत)

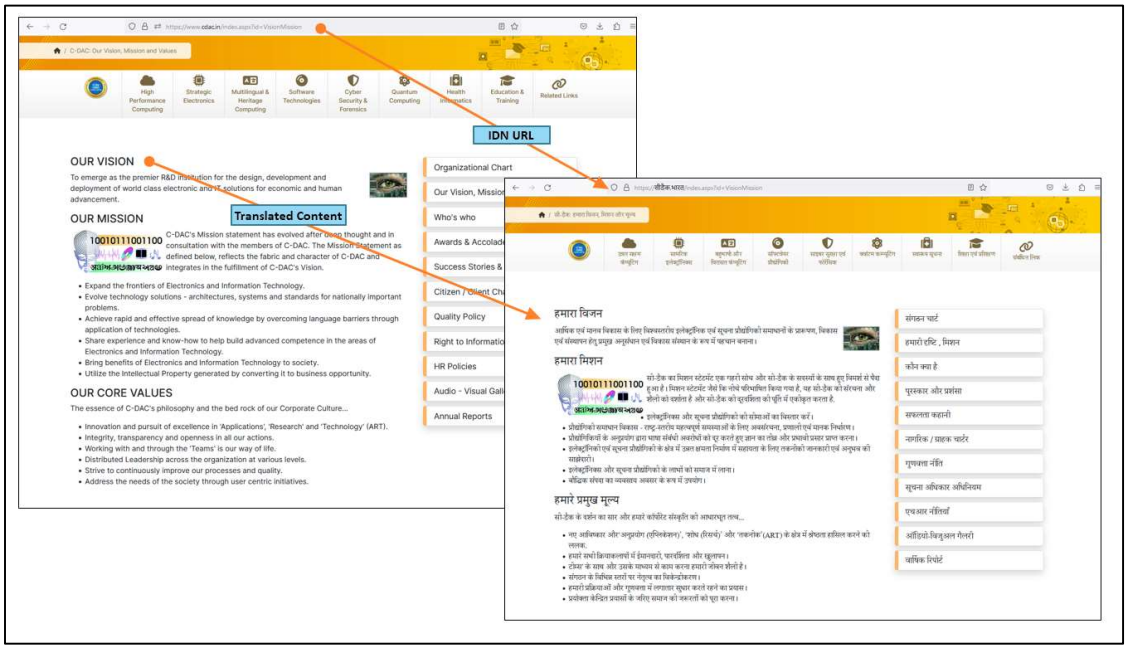


Figure 21: www.cdac.in

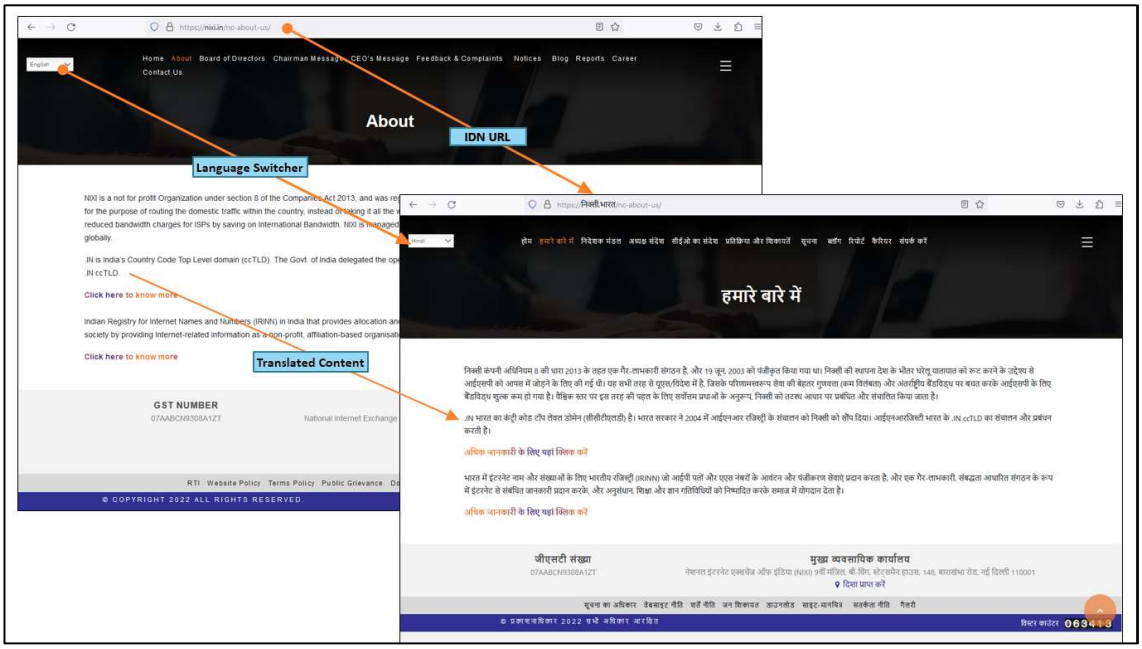


Figure 22: www.nixi.in

Support Contact Details

Copyright © 2023 C-DAC

Thank You

Centre for Development of Advanced Computing

C-DAC Innovation Park,

Panchavati, Pashan,

Pune - 411 008, Maharashtra (India)

Copyright © 2023 C-DAC