

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, Panchavati, 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
Support Contact Details	18

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 multilingual in Python Django CMS for the Indian government websites. The information given below differ for different websites based on its logic implementation applied.

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.

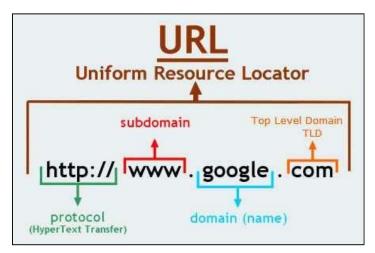


Figure 1: URL structure for reference only

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.
- 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.



*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
bhashanet.in	भाषानेट.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
भाषानेट.भारत	xnc2bro4b8ab8d.xnh2brj9c

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 IDNA2008¹ 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.
- 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.



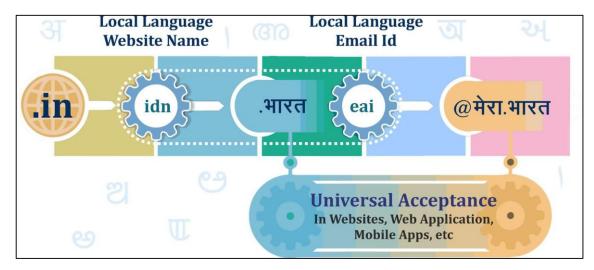


Figure 2: Universal Acceptance

IDN Registrar

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

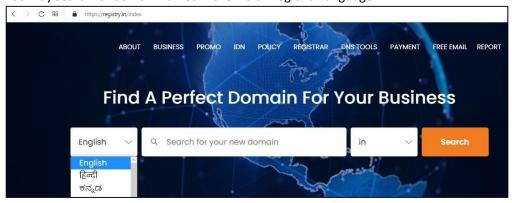


Figure 4: Domain Search in the registry.in



Availability search result

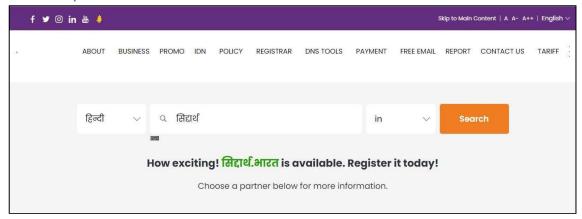


Figure 5: Domain availability in Hindi

Appointed Registrars

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

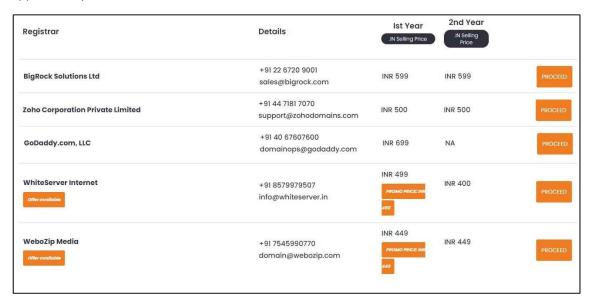


Figure 6: Appointed Registrar

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



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

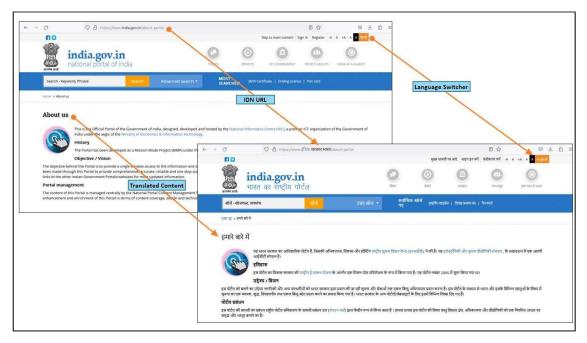


Figure 20: www.india.gov.in



Indian Domain Website (.IN/.भारत)

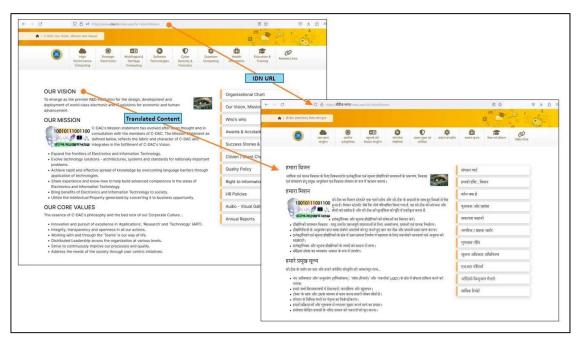


Figure 21: www.cdac.in

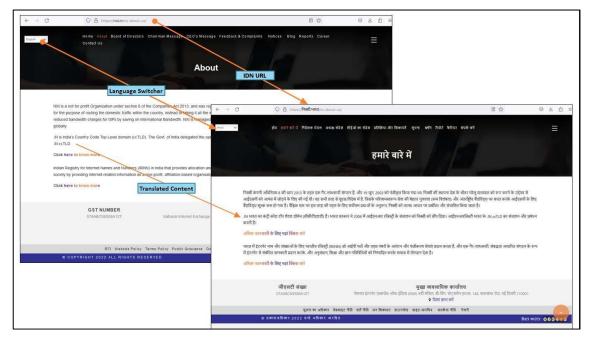


Figure 22: www.nixi.in



Support Contact Details



Thank You

Centre for Development of Advanced Computing

C-DAC Innovation Park,

Panchavati, Pashan,

Pune - 411 008, Maharashtra (India)

