ICANN's open, participative, bottom-up approach is critical because a stable and secure network requires global solutions that include local languages, characters, and cultural conventions.

To achieve a fully internationalized Internet, we need the collaborative involvement of technical and linguistic experts, policy-makers, application developers, end users and other stakeholders. ICANN looks forward to continuing to facilitate open forums for all interested parties.

Additional information about the IDN Program can be found at *http://icann.org/topics/idn/*. For more information on how to participate in IDN activities, contact idn@icann.org.

Internationalized Domain Name Glossary

Historically, domain names on the Internet were restricted to using ASCII characters (a–z, 0–9, and "-"). However, with the increasing use of the Internet in all regions and by diverse linguistic groups around the world, the need for multilingual domain names has become more intense. Various acronyms are used widely in communications around internationalizing the domain name space. Explanations for these are provided below to make this topic simpler to understand. The glossary here is based on a broader, more dynamic online glossary of IDN-related terms available at *http://www.icann.org/en/topics/idn/idn-glossary.htm*.

ASCII (American Standard Code for ASCII (American Standard Code for Information Interchange) ASCII is a common numerical code for computers and other devices that work with text. Computers can only understand numbers, so an ASCII code is the numerical representation of a character such as "a" or "@". When used in relation to ASCII TLDs or ASCII domain names, this refers to the fact that before internationalization only the letters a–z, digits 0–9, and the hyphen "-" were allowed in domain names.

DNS (Domain Name System) The DNS makes using the Internet easier by allowing a familiar string of letters (the domain name) to be used instead of the arcane IP address. So instead of typing 207.151.159.3, you can type www.internic.net.

IDNA (Internationalized Domain Names in Application) protocol IDNA is a protocol defined in RFC3490 by the Internet Engineering Task Force (*http://www.ietf.org*) that makes it possible for applications to handle domain names with non-ASCII characters. IDNA converts domain name strings with non-ASCII characters to ASCII domain name labels that applications that use the DNS can accurately understand. Not all characters used in the world's languages will be available for use in domain names. Hence IDNA is not able to convert all such characters into ASCII labels.

IDN (Internationalized Domain Name) IDNs are domain names represented by local language characters (non-ASCII characters). Such domain names could contain characters with diacritical marks as required by many European languages, or characters from non-Latin scripts (for example, Arabic or Chinese)

IDN SLDs Usually a reference for domain names with local characters at the second level, while the top level remains in ASCII-only characters. For example: [*παράδειγμα.test*] ("example.test" in Greek).

IDN TLDs Usually the short reference for internationalized top level labels referring to the entire domain name being represented by local characters. For example: [실례.테스트] ("example.test" in Hangul script).

Languages | Scripts | Alphabets Languages are used by speech communities. Scripts are used to write the various languages and this is done by using the corresponding alphabets or alternative writing systems.

LDH (Letter, Digit, Hyphen) A subset of the ASCII characters that only contains letters a-z, digits 0-9 and the hyphen "-", and the term "LDH code points" usually refers to this subset. Originally, domain name labels were restricted to this subset of characters.

Punycode This is the sequence of ASCII characters all IDNs will be encoded into in order for the Domain Name System (DNS) to understand and manage the names. The intention is that domain name registrants and users will never see this decoded form of a domain name. The sole purpose is for the DNS to be able to resolve for example a web address containing local characters. The DNS is only capable of handling ASCII characters. For example, the punycode version of [실례.테스트] (this is the Hindi, in Devanagari script, version of "example.test") is: xn--p1b6ci4b4b3a. xn--11b5bs3a9aj6g.

The prefix for the punycode version of the domain names is always "xn—". Hence this prefix is often reserved at the registry level to avoid confusion in registration of IDNs.

The Unicode Consortium is a nonprofit organization founded to develop, extend and promote use of the Unicode standard. For membership and more information, please visit *http://www.unicode.org*. Unicode is a commonly used single encoding scheme that provides a unique number for each character across a wide variety of languages and scripts. The Unicode Tables hold the code pints for each local character identified. These tables continue to expand as more and more characters are digitalized.

About ICANN

ICANN was formed in 1998 to coordinate the Internet's unique identifiers around the world. Without that coordination we wouldn't have one global Internet. It is a not-for-profit public-benefit corporation with participants from all over the world dedicated to keeping the Internet secure, stable and interoperable. It promotes competition and develops policy on the Internet's unique identifiers. ICANN doesn't control content on the Internet. It cannot stop spam and it doesn't deal with access to the Internet. But through its coordination role of the Internet's naming system, it does have an important impact on the expansion and evolution of the Internet. For more information please visit: *www.icann.org*.

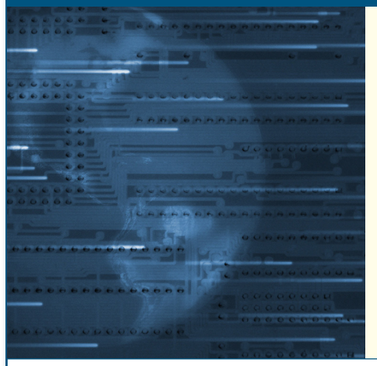
ICANN

Brussels	6 Rond Point Schuman, Bt. 5	B-1040 Brussels	Belgium	T +32 2 234 7870	F +32 2 234 7848
Marina del Rey	4676 Admiralty Way, Suite 330	Marina del Rey, CA 90292	USA	T +1 310 823 9358	F +1 310 823 8649
Washington, DC	1875 I Street NW, 5th Floor	Washington, DC 20006, USA	USA	T +1 202 429 2407	F +1 202 429 2714
Sydney	Level 2, 48 Hunter Street	Sydney NSW 2000	Australia	T +61 2 8236 7900	F +61 2 8236 7913

ICANN

INTERNET CORPORATION FOR ASSIGNED NAMES AND NUMBERS

IDNs: Internationalized Domain Names



Did you know?

- Internationalized Domain Names (IDNs) are domain names with characters other than a, b…, z; 0, 1,…, 9; and "-".
- IDNs on the second and third levels exist in some generic top-level domains (gTLDs) and in some country code top-level domains (ccTLDs). TLD registries determine the choice of characters available under these TLDs.
- Some languages, like Arabic, are written from right to left. IDNs at the second level are not very useful for the communities that use these languages because users would need to change the type direction and script in the middle of entering a domain name.
- IDN TLDs will be made available through two separate processes, initially through the New gTLD Program and the IDN ccTLD Fast Track Process.
- 60 percent of Internet users are non-English speakers, while the dominant language used on the Internet is English (see Global Reach at www.glreach.com).

One of the most significant innovations in the Internet since its inception is the introduction of Internationalized Domain Names (IDNs) in the Internet's address space. You may be familiar with some of the larger top-level domains: dot-com, dot-net, and dot-org are just three examples. IDN top-level domain names will offer many new opportunities and benefits for Internet users around the world by allowing them to establish and use domains in their native languages and scripts.

ICANN, the Internet Corporation for Assigned Names and Numbers, is responsible for managing and coordinating the Domain Name System (DNS) to ensure that every address is unique and that all of us who use the Internet can find all valid Internet addresses. It does this by overseeing the distribution of unique IP (Internet Protocol) addresses and domain names. It also ensures that each domain name maps to the correct IP address.

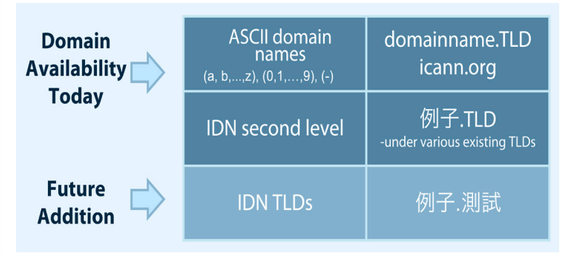
One of ICANN's top priorities is its IDN Program. This program is made up of several initiatives that involve a significant cross-section of the global Internet community and are aimed at providing a secure method for implementing internationalized top level domains in the root zone. Through this program, users will be able to register and use domain names based on their local language scripts. This includes users of languages based on right-to-left scripts such as Arabic and users of languages based on non-alphabetic scripts such as Mandarin Chinese. End users will be able to more easily identify localized online content by using fully localized identifiers. One example is *http://*실례.테스트 ("example.test" in Korean Hangul script).

Most of the program's objectives are related to:

- The secure and stable technical and operational deployment of IDNs in the Domain Name System (DNS) at the root level.
- Engaging the global Internet community in the policy processes that will guide the deployment, allocation and secure management of IDNs.

What are IDNs?

IDN is the short name for an Internationalized Domain Name. These domain names are represented by local language characters. Such domain names could contain characters with diacritical marks as required by many European languages, or characters from non-Latin scripts (for example, Arabic or Chinese).



Although many websites contain multilingual content, if you look closely at the site's web address on your browser, you will find IDNs represented only on the second level or lower levels. This representation happens in some gTLDs and some ccTLDs, and each registry decides which characters are available for registration in domain names. A map showing the IDN implementation across TLDs is available at *http://www.icann.org/en/maps/idntld.htm*.

Soon users around the globe will have access to IDN TLDs made up of non-ASCII characters to the right of the dot. ICANN is carrying out three initiatives to bring this about. Two are being implemented and one is at the policy development stage. These three processes are described briefly in the graphic on page 2.

IDN usage in emails—or, more specifically, in the domain name part of the email address—is not yet available. The technical standard that will make this possible is being developed by the Internet Engineering Task Force, or IETF. And some implementations are now available for users to try out. More information and test environments can be found at *http://cslab.kr/eai-test/index.php/Main_Page*, *http://global-email.info/intro.html*, and *http://www.ietf.org/html.charters/eai-charter.html*.

Why are IDNs important?

For many reasons. The most important is the growing number of Internet users around the world for whom it is difficult to use ASCII characters. The fact is, the Internet is accessed by more people who do not use Latin languages and scripts than those who do. This means that it is difficult for them to recognize ASCII characters and reproduce them on keyboards or use software to enter website addresses in browsers. For example:

- If you read an online newspaper in Russian and there is a link to more information, then it can be a challenge to reproduce an ASCII web address. It also makes more sense to have a Cyrillic-based address for a site with Russian content.
- If you see a billboard advertisement that contains a link or email address, it is more useful if the website or email address is based on characters that you can reproduce. In other words, in many countries and territories this link must be in scripts other than the basic Latin script.
- Suppose your child comes home from school with an assignment that requires obtaining information online on a specified website. What if the website address is in a different script than your native language? It may be hard to recognize the characters and reproduce the address in a browser. IDNs in your native language make navigating the Internet much easier.
- Sometimes a website address is not a name that makes sense or has any meaning attached. In such cases it is even more important to use the script or alphabet that intended users can recognize and be able to reproduce on a computer keyboard.

The ICANN community is addressing the needs of this growing community. Certainly, IDNs have clear advantages in targeting local markets and non-English speaking Internet users, and for local promotions and advertisements. So if you are a user not familiar with the basic Latin script, or if you wish to reach such users (with information, products or services), then IDNs will be very important for your ability to reach these markets.

How can I register an IDN?

First you might consider why you need an IDN.

If you want an IDN at the second level, or lower levels, you should select a TLD that meets your purpose. Check with your registrar and registry whether IDNs are available, and in what languages. For a good overview, go to *http://www.icann.org/en/maps/idntld.htm*. This map shows where IDNs are available. Currently, more than 47 TLD registries and registry managers support IDNs.

If you want to register a domain name with the entire string in local characters, or if you want to apply to become an IDN TLD operator, you can choose from three possible programs. Each program has its own specific qualification requirements, rules and launch time.

The diagram here gives an overview of the three programs available. As these initiatives continuously progress, please view the latest status at:

- IDN ccTLD Fast Track Process: *http://www.icann.org/en/topics/idn/fast-track/* The shorthand term IDN ccTLDs refers to new top-level domains associated with entries in the ISO 3166-1list.
- New gTLD Program (including IDN gTLDs): *http://www.icann.org/en/topics/new-gtld-program.htm*
- IDN ccNSO PDP: *http://ccnso.icann.org/*

In Implementation: IDN ccTLD Fast Track Process

- To introduce a limited number of IDN TLDs in non-Latin scripts
- Requires strings to be meaningful representations of country or territory names
- Only for countries or territories represented in the ISO 3166-1 list

In Policy Development: IDN ccTLDs – Long Term

- Full policy that caters for all
- Follows the ccNSO policy development process
- Was initiated formally on 13 April 2009 by the ccNSO Council *http://www.icann.org/en/announcements/announcement-13apr09-en.htm*

In Implementation: New Generic TLD Program

- New process for applying for new gTLDs
- Includes IDN TLDs
- Technical string requirements are the same for all IDN TLDs
- Includes a mechanism for avoiding confusingly similar strings

What is the IDN ccTLD Fast Track Process?

The IDN ccTLD Fast Track Process focuses on meeting a near-term demand by allocating a limited number of new ccTLDs based on non-Latin scripts while a full IDN ccTLD policy is being developed. The process is further limited to countries and territories represented in the ISO 3166-1 list. The Fast Track Process is expected to be launched in Q4 2009, and a separate factsheet for this process is available at www.icann.org/topics/idn.

What is the IDN ccTLD Policy Development?

Along with the Fast Track Process, with its goal of meeting near-term demand for ccTLDs to selected countries and regions, ICANN is develop a long-term process for implementing IDN ccTLDs. Once the long-term process is implemented, the Fast Track Process will become unnecessary. However, experience gained in the IDN ccTLD Fast Track Process will be a valuable addition to completing the process and implementing the final policy.

What is the New gTLD Program Applicant Guidebook?

The New gTLD Program focuses on expanding choices for generic top-level domains, or gTLDs. This expansion includes IDNs at the top level in order to meet growing diversity and encourage competition for more innovation, choice and change to the Internet's addressing system.

ICANN is developing both a process for registries to apply for new gTLDs and an Applicant Guidebook that takes applicants through the process and explains the implications of the many complicated issues surrounding these new top-level domains. The Internet community is currently wresting with resolving string contention, protecting intellectual property rights, handling internationally recognized issues of morality and public order, and the geographical names process, among other issues. After going through several rounds of public comment, the process and the guidebook are expected to be fully implemented early in 2010.

For more information about the Applicant Guidebook and the New gTLD Program, go to *http://icann.org/en/topics/new-gtld-program.htm*.

Where can I try IDNs?

ICANN created a dedicated IDNwiki facility to introduce users to the testing of IDN top-level domains. The test is based on eleven IDN TLDs representing "example.test" in non-Latin based scripts. The table below shows the languages and scripts initially selected for the IDNwiki. You can enter the Wiki by typing in any of the illustrated domains, or by going to *http://idn.icann.org*.

Public participation in the evaluation of these domains is an important aspect of the project. To participate, access one or more of the addresses in the table and report your experience on the discussion pages.

More languages and scripts are being added as requests come in from the various language communities.

Script	Language	SLD.TLD U-labels	SLD A-label	TLD A-label
Arabic	Arabic	مثال.اختبار	xn--mgbh0fb	xn--kgbechtv
Arabic	Persian	مثال.ازمایشی	xn--mgbh0fb	xn--hgbk6aj7f53bba
Chinese, simplified	Chinese	例子.测试	xn--fsqu00a	xn--0zwm56d
Chinese, traditional	Chinese	例子.測試	xn--fsqu00a	xn--g6w251d
Cyrillic	Russian	пример.испытание	xn--e1afmkfd	xn--80akhbyknj4f
Devanagari	Hindi	उदाहरण.परीक्षा	xn--p1b6ci4b4b3a	xn--11b5bs3a9aj6g
Greek	Greek	παράδειγμα.δοκιμή	xn--hxajbheg2az3al	xn--jxalpdlp
Hangul	Korean	실례.테스트	xn--9n2bp8q	xn--9t4b11yi5a
Hebrew	Yiddish	בײַשפּיל.טעסט	xn--fdbk5d8ap9b8a8d	xn--deba0ad
Kanji Hirigana, and Katakana	Japanese	例え.テスト	xn--r8jz45g	xn--zckzah
Tamil	Tamil	உதாரணம்.பரிட்சை	xn--zkc6ce5bi7f6e	xn--hlcj6aya9esc7a

What is the IDNA protocol revision?

The IETF is leading in creating standards for using non-ASCII characters in the Domain Name System. Early in this process, an international IDN working group developed suggested standards for review and comment. Participating organizations developed technologies to the standards to test their application. Through revision and comment, participants reached agreement and issued standards in 2003. Since then, the IETF IDN working group has been evaluating the experience gained in implementing the IDNA protocol. They have identified several key areas of work now under way. The core components still being revised include:

- Eliminating Unicode version dependencies, thereby permitting more characters to be used automatically in IDNs now and in the future.
- Clear defining valid IDN labels.
- Fixing the complications and errors in the existing protocol for many right-to-left domains. These domains will become valid when the revised protocol is implemented.

The issues with the current IDN model that led to the revision work are discussed in RFC4690. The Internet-drafts with proposed revisions are available at *http://icann.org/topics/idn/rfc.htm*.

The next IETF meeting where the IDNA protocol revision will be discussed is scheduled for Stockholm, Sweden, on 26-31 July 2009. *http://www.ietf.org*

What are the IDN Guidelines?

These guidelines list the standards for IDN registration. The policies and practices in the guidelines will minimize the risk of cybersquatting and consumer confusion, and ensure that the interests of local languages and character sets are respected. Registries that deploy IDNs will be required to follow these guidelines.

In addition all new TLD managers, either through the IDN ccTLD Fast Track Process or the New gTLD Program, will be required to fulfill all existing or future versions of the IDN Guidelines. The latest version of the guidelines is available at *http://www.icann.org/en/topics/idn/implementation-guidelines.htm*.

The IDN Guidelines will be revised again right after the IDNA protocol revision to ensure they are consistent. All changes in the protocol will be reflected in the guidelines as well.

Brief history of internationalization and the Internet's Domain Name System

Internationalization measures, including IDNs, might be the largest change in Internet operation since the TCP/IP (Transmission Control Protocol/Internet Protocol) was introduced. Furthermore, IDN deployment might be more significant and complex than the original DNS introduction. To fully understand the difficulties of internationalization of the domain name space, it is important and useful to give a brief history of IDN technology development.

The idea of internationalizing domain names goes back to the developmental stages of the Internet in the 1970s. Discussions centered on the usability of languages and scripts. However, the technology deployed in the Domain Name System today, which allows practically any character in the registration of a domain name, was not developed at that time. Therefore, the characters available for registering domain names was limited to ASCII characters: the letters a-z, the digits 0-9, and the hyphen "-".

In 2003, technical bodies raised the topic again, and the Internet Engineering Task Force released the standards (RFCs 3490, 3491, and 3492), also often referred to as IDNA, or IDNs in Application. These standards provide technical guidelines for successful deployment of IDNs; that is, a mechanism to handle non-ASCII characters in domain names in a standard fashion. It works by converting names with non-ASCII characters to ASCII labels that the Domain Name System will understand.

Later, ICANN and several top-level domain registries released version 1.0 of the IDN Guidelines. These guidelines were created for generic TLD (gTLD) registries to follow when implementing IDNs at the second level only. The current guidelines, version 2.1, is available at *http://icann.org/topics/idn/implementation-guidelines.htm*.

Work is under way to revise the IDNA protocol, based on experience since the 2003 introduction. Additional work is under way by an IDN TLD registry working group to amend the guidelines further to ensure that the guideline directions will be used deeper into the DNS hierarchy, particularly as a set of principles for implementing internationalized top-level labels.

For a more detailed history of technical development see *http://www.isoc.org/pubpolpillar/docs/i18n-dns-chronology.pdf*.