Root Zone Evolution Review Committee (RZERC)

Internationalized Domain Name ,IDN,"IDNs are domain names that include characters used in the local representation of languages that are not written with the twenty-six letters of the basic Latin alphabet ""a-z"". An IDN can contain Latin letters with diacritical marks, as required by many European languages, or may consist of characters from non-Latin scripts such as Arabic or Chinese. Many languages also use other types of digits than the European ""0-9"". The basic Latin alphabet together with the European-Arabic digits are, for the purpose of domain names, termed ""ASCII characters"" (ASCII = American Standard Code for Information Interchange). These are also included in the broader range of ""Unicode characters"" that provides the basis for IDNs. The ""hostname rule"" requires that all domain names of the type under consideration here are stored in the DNS using only the ASCII characters listed above, with the one further addition of the hyphen ""-"". The Unicode form of an IDN therefore requires special encoding before it is entered into the DNS. The following terminology is used when distinguishing between these forms: A domain name consists of a series of ""labels"" (separated by ""dots""). The ASCII form of an IDN label is termed an ""A-label"". All operations defined in the DNS protocol use A-labels exclusively. The Unicode form, which a user expects to be displayed, is termed a ""U-label"". The difference may be illustrated with the Hindi word for ""test"" — परीका — appearing here as a U-label would (in the Devanagari script). A special form of ""ASCII compatible encoding"" (abbreviated ACE) is applied to this to produce the corresponding A-label: xn--11b5bs1di. A domain name that only includes ASCII letters, digits, and hyphens is termed an ""LDH label"". Although the definitions of A-labels and LDH-labels overlap, a name consisting exclusively of LDH labels, such as""icann.org"" is not an IDN."