THE DOMINATION OF THE ENGLISH LANGUAGE IN THE GLOBAL VILLAGE: EFFORTS TO FURTHER DEVELOP THE INTERNET BY POPULATING IT WITH NON-LATIN-BASED LANGUAGES

Molly Torsen*

*University of Washington School of Law, J.D., 2005. The author would like to thank Christian Degueldre, Bryan Stech and Jane Winn. Ms. Torsen is currently at the AHRC Research Centre for Studies in Intellectual Property and Technology at the University of Edinburgh School of Law performing intellectual property research.

[1] At the top of the homepage of the Internet Corporation for Assigned Names and Numbers (ICANN) is a choice of nine different languages in which to read information about the organization; four of them are languages written in non-Latin script (Arabic, Chinese, Japanese and Russian).[1] Clicking upon any of the language options brings the reader to a new website in that language and, presuming the reader has a computer and screen that can handle non-Latin script, the webpage reads legibly and clearly.[2] The Uniform Resource Locator (URL),[3] however, still reads in

---

[1] Internet Corporation for Assigned Names and Numbers, http://www.icann.org (last visited Sept. 11, 2005) (hereinafter ICANN).

[2] *See, e.g.*, ICANN, http://www.icann.org/tr/chinese.html (last visited Sept. 11, 2005).

[3] A URL

> [I]s the address of a specific Web site or file on the Internet. A URL cannot have spaces or certain other characters and uses forward slashes to denote different directories. Some examples of URLs are http://www.sharpened.net/glossary/index.html,

Latin script despite the webpage's content being in a foreign script. This article will examine the current state of affairs in policy-oriented Internet realms and suggest that the cohesive development of Internationalized Domain Names (IDNs), which are domain names in character sets other than American Standard Code for Information Interchange (ASCII), is a critical step to take in order to more fully utilize the potential the Internet offers for international communication.

[2] As technology enables more and more people to connect with each other, so must the policies about the platforms for communication respond to international needs. Indeed, "two-thirds of content on the Internet is in English, but only one-third of users speak English as a native language."[4] Furthermore, the inhabitants of the world's 228 countries speak an approximate 6,700 languages and enabling everyone to participate in the global village that the world has become is essential to promoting true international cooperation.[5] Not surprisingly, non-English speaking Internet users complain "that Latin-script domain names, web addresses and e-mail addresses are difficult to remember and easy to misspell."[6] The internationalization of the Internet should not be surprising, considering that 92% of the world's population speaks a primary language other than English.[7] Forecasters believe that Chinese will be the number one language used on the World Wide Web by 2007.[8] The Internet is a

---

> http://www.wheaton.edu/, and ftp://info.apple.com/. As you can see, not all URLs begin with 'http'. The first part of a URL indicates what kind of resource it is addressing.

Definition of URL, http://www.sharpened.net/glossary/definition.php?url (last visited Sept. 11, 2005).

[4] VeriSign, Inc., Internationalized Domain Names, http://www.verisign.com/products-services/naming-and-directory-services/naming-services/internationalized-domain-names/ (last visited Sept. 11, 2005).

[5] Hong Xue, *The Voice of China: A Story of Chinese-Character Domain Names*, 12 CARDOZO J. INT'L & COMP. L. 559, 560 (2004).

[6] *Id.* at 563.

[7] Int'l Telecomm. Union (ITU) and World Intell. Prop. Org. (WIPO), *Multilingual Domain Names: Joint ITU/WIPO Symposium*, at 4-5 (Dec. 6-7, 2001) (citing statistics from Walid, Inc., http://www.walid.com), *available at* http://www.itu.int/mlds/briefingpaper/wipo/wipofinal3.doc (hereinafter Joint ITU/WIPO Symposium).

[8] *Id.* at 5.

truly global medium; it would be foolish not to use it to the most expansive and inclusive extent possible.[9]

THESIS

[3] *The several international governmental and non-governmental organizations involved in the development of IDNs should agree to cooperate on technical and policy issues. Standardization in this realm – such as universal use of Unicode – will promote the adoption and implementation of IDNs by registrars, increase their sales and put market pressure on other application providers, such as browsers, to support IDNs. It is imperative to avoid fragmentation of this process so that the Internet remains a globally-useful platform. Not only will standardization be more convenient for Internet users, but also for policymakers in realms such as international intellectual property dispute resolution.*

INTRODUCTION

[4] The following discussion focuses on international policies and problems resulting from experimentation and implementation of IDNs; it does not attempt to provide a technical or semantically-refined analysis of IDNs, as the discussion and literature varies depending on context and some of the technical considerations are beyond the scope of this overview of legal policy issues.[10] This paper does, however, endorse the

---

[9] *Id.* at 5.

[10] An IDN, for example, can have several different meanings depending on the context in which it is used.

> The term 'IDN' has a number of different uses: (a) as an abbreviation for "Internationalized Domain Name"; (b) as a fully qualified domain name that contains at least one label that contains characters not appearing in ASCII, specifically not in the subset of ASCII recommended for domain names (the so-called 'hostname' or 'LDH' subset, see RFC1035 [STD13]); (c) as a label of a domain name that contains at least one character beyond ASCII; (d) as a Unicode string to be processed by Nameprep; (e) as a string that is an output from Nameprep; (f) as a string that is the result of processing through both Nameprep and conversion into Punycode; (g) as the abbreviation of an IDN (more properly, IDL) Package, in the terminology of this document; (h) as the abbreviation of the IETF IDN Working Group; (g)

recommendations of the United Nations Educational, Scientific and Cultural Organization (UNESCO), which encourage the local adaptation of operating systems, search engines and web browsers with extensive multilingual capabilities, online dictionaries and terminologies.[11] Aside from sanctioning the general concept of multilingualism, this writing encourages international cooperation in the development of a single policy to implement this goal; there are currently several initiatives "underway to explore the means by which this internationalization of the DNS" should or could expand.[12] Alternate addressing systems were created that use different IP number-to-name mapping systems.[13] This writing reinforces the Internet Society's findings that individualized efforts to create separate Internet systems will ultimately undermine the global capabilities of the World Wide Web; international cooperation is imperative right now. Finally, this writing highlights some recent international trademark cases that demonstrate some new concerns that IDNs have introduced in the intellectual property context. While there are an array of "alternative" domain name registries, such as .club, .sport and .church,[14] this writing will focus on generic top level domain names (gTLDs), such as .com, .gov and .org, assuming that their widespread use and recognition will continue to maintain their popularity and value. The new issues with language are not unexpected, and provide an example of the inevitable awkward adjustments that will occur as the world tries to use a necessarily singular platform – the Internet – for a gamut of languages and characters.

I. BACKGROUND AND KEY ORGANIZATIONS

A. ICANN

---

as the abbreviation of the ICANN IDN Committee; and (h) as standing for other IDN activities in other companies/organizations.
K. KONISHI ET AL., JOINT ENGINEERING TEAM (JET) GUIDELINES FOR INTERNATIONALIZED DOMAIN NAMES (IDN) REGISTRATION AND ADMINISTRATION FOR CHINESE, JAPANESE, AND KOREAN 4 (2004), http://www.rfc-editor.org/rfc/rfc3743.txt.

[11] UNESCO, Recommendation Concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace, (Nov. 21, 2003), http://portal.unesco.org/ci/en/file_download.php/41e32bf91c3d30c7855cefe4251cba6fRecommendation-Eng.pdf, at 2.

[12] Joint ITU/WIPO Symposium, *supra* note 7, at 5.

[13] *Id.,* at 7.

[14] *E.g.,* Exciting and Descriptive Alternative Domains, http://www.easyspace.com/domains/newdotnet.html (last visited Sept. 11, 2005).

[5] ICANN was founded as a not-for-profit corporation based in California.[15] As an international organization set on American soil, it has sought to "legitimate itself" as an open and representative working group of international players, but several critics believe that the intricate and obscure structure ICANN has installed for its governance has prevented any significant contribution from many interested parties.[16] The Internet is an invention of the United States, and ICANN's mechanisms, as well as the majority of domain names on the Internet, are primarily proliferated in rules and, of course, a language that are best-suited to the United States.[17] Another problem identified by critics is that ICANN is backed by the United States Department of Commerce due to the Department's significant power over the domain name system, which is based on its control over the "A" root name server.[18] This is significant because there are thirteen root servers (assigned letters A-M) and only the A root is able to refer "inquiring computers to the Internet address of the computer that has the authoritative list of the registered domain names" in the relevant top-level domain (e.g., .com, .info, etc).[19] Nine of the secondary servers are located in the United States; seven are owned by the United States government.[20]

[6] ICANN is reportedly seeking full independence from the U.S. Department of Commerce by 2006,[21] however, and on ICANN's webpage, it invites global participation. "Participation in ICANN is open to all who have an interest in global Internet policy as it relates to ICANN's mission of technical coordination. ICANN provides many

---

T[15] ICANN, Articles of Incorporation of Internet Corporation for Assigned Names and Numbers (Nov. 21, 1998), http://www.icann.org/general/articles.htm.

[16] *See, e.g.,* John Palfrey, *The End of the Experiment: How ICANN's Foray into Global Internet Democracy Failed*, 17 HARV. J. L. & TECH. 409, 409 (2004).

[17] Brent T. Yonehara, *Landoftherisingsun.co.jp: A Review of Japan's Protection of Domain Names Against Cybersquatting*, 43 IDEA 207, 231 (2003).

[18] Xue, *supra* note 5, at 580–81.

[19] Kim G. von Arx and Gregory R. Hagen, *Sovereign Domains: A Declaration of Independence of ccTLDs from Foreign Control*, 9 RICH. J.L. & TECH. 4, ¶ 15 (Fall 2002), http://www.law.richmond.edu/jolt/v9i1/article4.html.

[20] *Id.* For a thorough discussion of the United States government's relationship to ICANN and the Internet, *see generally id.*

[21] William New, *Net Governance: ICANN Meeting to Address Global Issues*, NAT'L J. TECH. DAILY, Nov. 29, 2004.

online forums which are accessible through ICANN's website, and the Supporting Organizations and Advisory Committees have active mailing lists for participants."[22] Likewise, on ICANN's IDN webpage, it introduces the topic by inviting participation:

> This area is designed to document the progress of the implementation of IDNs as well as *allow for discussion of issues* encountered in implementation. As part of this effort, ICANN hosts a publicly archived mailing list for discussion of IDN implementation issues. All [generic top-level domains] and [country-code top-level domains] registries are *encouraged to participate in the list*.[23]

ICANN also hosted a workshop in July of 2004 in which it aimed to concentrate on the more practical aspects of implementing IDNs.[24]

[7] In October 2002, ICANN's Internet Engineering Steering Group (IESG) approved a means by which to implement non-ASCII IDNs in the Internet's domain-name system.[25] Subsequent to this initial effort, the IESG drew up a set of rules titled *Guidelines for the Implementation of Internationalized Domain Names*; Version 1.0 of these Guidelines was published in June of 2003.[26] ICANN arguably wields great authority in this realm and, as such, it is evident that the development of non-ASCII domain names is still very new, both conceptually and technically.

B. THE MULTILINGUAL INTERNET NAMES CONSORTIUM (MINC)

[8] MINC is a non-profit, non-governmental, international organization that focuses on developing and promoting multilingual Internet domain

---

[22] ICANN, ICANN Information, http://www.icann.org/general/ (last visited Sept. 12, 2005).

[23] ICANN, Internationalized Domain Names, http://www.icann.org/topics/idn.html (last visited Sept. 12, 2005) (emphasis added).

[24] ICANN, Workshop: Internationalized Domain Name, http://www.icann.org/meetings/kualalumpur/idn-workshop-08jul04.htm (Nov. 29, 2004).

[25] ICANN, Internationalized Domain Names, http://www.icann.org/topics/idn.html (last visited Sept. 12, 2005).

[26] *Id.*

names and keywords.[27]   It aims to internationalize Internet names
standards and protocols, provide technical coordination, and to liaison
with other international bodies.[28] MINC "has organizational and
individual members from all continents of the world" and from various
backgrounds including industry, academia, research, government,
investors and international organizations.[29] MINC's beginnings are based
on a research project regarding multilingual Internet domain names begun
by Dr. Tan Tin Wee, the outgoing Head of the Internet Research and
Development Unit at the National University of Singapore in early 1998.[30]
An IDN prototype was developed in mid-1998.[31] In order to avoid any
disruption of the domain name operations as they were currently
functioning, and to demonstrate that IDNs could be implemented, the
researchers created a proxy system which intercepted multilingual
character strings sent out from multilingual enabled client applications
such as Web browsers, and converted these characters into ASCII
compatible encodings.[32] These forms of ASCII domain names can reside
on any DNS records on any DNS servers regardless of their ability to host
non-ASCII script.[33]

[9] The proxy system prototype demonstrated that multilingual forms of
domain names could work with the then-current existing DNS servers
without breaking the system.[34] "It provided the impetus for an integrated
approach towards the gradual internationalization of the DNS system
worldwide, laid the basis and paved the way forward for further work on
the Internationalization of Domain Names, which has now led to the
formation of the Multilingual Internet Names Consortium."[35] ICANN's
interest in IDNs paralleled the test-bed stage. In June of 2002, following

---

[27] MINC, Introduction, http://www.minc.org (last visited Sept. 13, 2005).
[28] *Id.*
[29] *Id.*
[30] MINC, History of MINC, http://www.minc.org/about/history/earlyhistory.shtml (last
visited Sept. 13, 2005); *see also* Wikipedia, Internationalized Domain Name,
http://en.wikipedia.org/wiki/Internationalized_domain_name (last visited Aug. 23, 2005)
(providing a timeline of the history of Internationalized Domain Names).
[31] MINC, History of MINC, http://www.minc.org/about/history/earlyhistory.shtml (last
visited Sept. 13, 2005).
[32] *Id.*
[33] *Id.*
[34] *Id.*
[35] *Id.*

the initial work of MINC, ICANN published a paper on non-ASCII TLD registry selection considerations and generally found that procedures for ASCII and non-ASCII registration should be harmonized.[36] ICANN also reported on its study of technical-related issues, finding, for example, that nothing within any current or future non-ASCII TLD space "inherently constrains names or labels to any language character set….[N]othing in the protocols would prevent a domain label from being created…that consists of a Chinese character, followed by a Roman-derived character, followed by a Thai character, followed by an Arabic character, followed by a Cyrillic character, etc."[37]

C. THE PUBLIC INTEREST REGISTRY AND THE INTERNET SOCIETY

[10] The Internet Society [ISOC] is a professional membership society that consists of more than 100 organizations and over 20,000 individual members in at least 180 countries.[38] ISOC provides leadership in addressing issues that confront the future of the Internet, and is the umbrella organization for the groups responsible for Internet infrastructure standards, including the Internet Engineering Task Force [IETF].[39] "ISOC is an international non-profit organization whose mission is to assure the open development, evolution and use of the Internet for the benefit of all people throughout the world. ISOC focuses on Internet standards development, educational initiatives around the world and technically sound policy formation."[40] In April of 2004, ISOC published guidelines for the immediate future of IDNs. It specified that the development and test-bed process of IDNs:

> [M]ade it clear that use of characters with similar
> appearances and/or interpretations created potential for
> confusion, as well as difficulties in deployment and
> transition. The conclusion was that, while those issues

---

[36] ICANN INT'L DOMAIN NAMES COMM., FINAL REPORT TO THE ICANN BOARD (2002), http://www.icann.org/committees/idn/final-report-27jun02.htm.
[37] Id.
[38] Internet Society, All About the Internet Society, http://www.isoc.org/isoc/ (last visited Sept. 13, 2005).
[39] Id.
[40] PIR, Internet Society, http://www.pir.org/InternetSociety/InternetSociety.aspx (last visited Sept. 13, 2005).

were important, they could best be addressed
administratively rather than through restrictions embedded
in the protocols. This [publication from the Joint
Engineering Team] defines a set of guidelines for applying
restrictions of that type for Chinese, Japanese and Korean
(CJK) scripts and the zones that use them and, perhaps, the
beginning of a framework for thinking about other zones,
languages, and scripts.[41]

[11] The Public Interest Registry (PIR), also an offshoot of ISOC, is a
not-for-profit corporation. PIR's central task is to operate the .ORG
domain; this role is a result of a competitive bid process conducted by
ICANN.[42] PIR's technical partner in its administration of the .ORG
domain is Afilias, a global provider of advanced domain name registry
services based in Ireland that makes available a variety of "capabilities
essential to the smooth and efficient operation of any Internet domain
name registry."[43] PIR is dedicated to providing a domain that is global in
scope and sensitivity, by providing outreach and resources to
noncommercial and nonprofit organizations worldwide.[44]

[12] ISOC suggests that, while "it is important to be able to write the
names of TLDs, especially country-associated TLDs, in languages and
scripts associated with those countries," care must be taken to address
some technical issues.[45] From the perspective of an Internet user, a
reference to a web site which is located in Greece, whose content is in
Greek, and which uses Greek characters for its domain name, should be in
Greek.[46] The position of ISOC, however, is that ICANN and other policy-
oriented bodies need to understand that this logical aspiration is *not
possible* at a protocol level. For example, the "http" in a URL is the name

---

[41] KONISHI, *supra* note 10, at 1.

[42] PIR, Public Interest Registry Launches New .ORG Web Site to Global Audience (Aug.
23, 2005), http://www.pir.org/PDFs/Press/NewWebSite08_23_05.pdf.

[43] Afilias, About Afilias, http://www.afilias.info/about_afilias/ (last visited Sept 13,
2005).

[44] *See* PIR, About PIR, http://www.pir.org/AboutPir/AboutPir.aspx (last visited Sept. 13,
2005).

[45] JOHN KLENSIN, INTERNATIONALIZING TOP-LEVEL DOMAIN NAMES: ANOTHER LOOK
(2004), http://www.isoc.org/briefings/018/briefing18.pdf, at 2.

[46] *See id.*

of a protocol and if it were translated into a different language, it would be an entirely different protocol.[47] In other words, it would be unwise to look into creating different formats for "http," the standard language by which computers connected to the World Wide Web communicate with one another.[48] Furthermore:

> As with any attempt to localize, or otherwise optimize a system for use within a specific community, the technique proposed makes global interoperability more difficult. Just as is the case with IDNs themselves, the user sees strings that are not the ones being passed across the network and that may not be globally comprehensive. If a user of one language passes a domain name containing IDNs that are expressed in their native script to another user, the second user may not be able to read them or key them back into a computer and, at least with the state of the technology today, a cut-and-paste operation on the characters from, say, an email message, may or may not work as intended.[49]

D. WORLD WIDE WEB CONSORTIUM

[13] Tim Berners-Lee founded the World Wide Web Consortium (W3C) in 1994 at the Massachusetts Institute of Technology, Laboratory for Computer Science (MIT/LCS) in conjunction with Centre Européen de Recherche Nucléaire (CERN), with support from the Defense Advanced Research Projects Agency (DARPA) and the European Commission.[50] In April 1995, the Institut National de Recherche en Informatique et Automatique (INRIA) became the first European W3C host. Keio University of Japan in Asia followed suit in 1996.[51] In 2003, the European Research Consortium in Informatics and Mathematics became the European W3C Host.[52] W3C currently pursues an international

---

[47] *Id.*
[48] GetNetWise, Guide to Internet Terms: A Glossary, http://www.getnetwise.org/glossary.php (last visited Sept 13, 2005).
[49] KLENSIN, *supra* note 45, at 4.
[50] W3C, History, http://www.w3.org/Consortium/history (last visited Sept. 14, 2005).
[51] *Id.*
[52] *Id.*

audience through offices in fifteen countries.[53] Insofar as its efforts in the realm of IDNs, W3C is aiming to coordinate any techniques, conventions, guidelines and activities both within its own framework and "together with other organizations that allow and make it easy to use W3C technology worldwide, with different languages, scripts, and cultures."[54] Inherent in its name and development, W3C demonstrates the global nature of the Internet and provides a spotlight for ISOC's concern that internationalization is a technically challenging feat because there is only one Internet for the world's hundreds of languages and scripts.[55]

[14] The W3C Internationalization Activity's Working Group recently published a Recommendation called the *Character Model for the World Wide Web 1.0: Fundamentals*.[56] It presents a "well-defined and well-understood way for Web applications to transmit and process the characters of the world's languages."[57] The Working Group has members from various entities including BBC, Boeing, Ecole Mohammadia d'Ingénieurs, IBM, Microsoft, Siemens, Sun Microsystems, and webMethods.[58] Their recommendation gives a common reference to authors of specifications, software developers, and content developers, "enabling interoperable text manipulation on the World Wide Web."[59] It expands the Universal Character Set that was defined by the Unicode Standard. The primary goal of the *Character Model for the World Wide Web* is to "facilitate use of the Web by all people, regardless of their language, script, writing system, and cultural conventions, in accordance

---

[53] W3C, W3C Offices, http://www.w3.org/Consortium/Offices/ (last visited Sept. 14, 2005).

[54] W3C, W3C Internationalization Activity, http://www.w3.org/International/ (last visited Sept. 14, 2005).

[55] *See* KLENSIN, supra note 45, at 1-5.

[56] W3C, http://www.w3.org/TR/2005/REC-charmod-20050215/ (last visited Sept. 14, 2005); W3C, World Wide Web Consortium Issues Critical Internationalization Recommendation, http://www.w3.org/2005/02/charmod-pressrelease (last visited Sept. 14, 2005).

[57] W3C, World Wide Web Consortium Issues Critical Internationalization Recommendation, http://www.w3.org/2005/02/charmod-pressrelease (last visited Sept. 14, 2005).

[58] *Id.*

[59] *Id.*

with the W3C goal of universal access."[60] As more Web applications develop:

> [T]he need for a shared character model [becomes] more critical. Unicode is the natural choice as the basis for that shared model, especially as applications developers begin to consolidate their encoding options. However, applying Unicode to the Web requires additional specifications; this is the purpose of the W3C Character Model series.[61]

The Recommendation put out by W3C is the first in a series of three documents. In progress are *Character Model for the World Wide Web 1.0: Normalization*, "specifying uniform normalization and string identity matching for text manipulation," and *Character Model for the World Wide Web 1.0: Resource Identifiers*, specifying IRI conventions.[62]

E. THE CHINA SITUATION

[15] In terms of fitting onto the Internet, not only does China have special difficulties with the difference between its traditional and simplified characters, but also with its government's historic restrictions on – and suspicion of – the Internet. Internet Service Providers [ISPs] are approved by the Chinese State Council; as of 2001, there were only six.[63] The Chinese government has always been concerned with and involved in the Chinese-language domain system. In June of 1997, the Chinese management regime under the Computer Network Information Center of the Chinese Academy of Science [CNNIC], a government entity, was established to provide registration services for the ".cn" domain. And on November 7, 2000, a few days before United States companies introduced IDN registration, CNNIC exhibited its own Chinese character domain name registration services with the suffixes of .corporation (translated) and .net (translated).[64] Furthermore, "China's domain name authorities dramatically opposed foreign firms registering Chinese language domain

---

[60] *Id.*

[61] *Id.*

[62] *Id.*

[63] Jiang-yu Wang, *The Internet and E-Commerce in China: Regulations, Judicial Views, and Government Policies*, COMPUTER & INTERNET LAW., Jan. 2001, at 12.

[64] *Id.*

names. They sent a complaint to [ICANN]… arguing that the U.S. government had no right to authorize any company to manage domain names with Chinese characters because Chinese character domain names 'have unique… cultural and historical implication.'"[65]

[16] In 2001, "no one [was] allowed to conduct a Chinese-language domain name service within China" without first obtaining approval by Chinese authorities.[66] As of the end of January 2005, however, China's government had loosened its hold on the .cn domain, which is now administered by NeuLevel, a registrar based in Virginia. NeuLevel is currently interested in ICANN's bidding for the ".net" domain; NeuLevel's bid emphasizes new technology and services, including support for internationalized domain names.[67] CNNIC's Director General commented, "We are very pleased that Chinese .cn domain names are now available to the international market."[68] The country's change in policy may be attributable to the fact that, since 2001, CNNIC has actively participated in the standard-making process led by the IETF IDN Working Group.[69] Indeed, "China's Internet authority has become fully aware of the importance of international collaborations. Prospectively, only under the cooperation of all the stakeholders could a reasonable and effective management system of Chinese-character domain names be established."[70]

II. (SIMPLIFIED) TECHNICAL CONSIDERATIONS: WHAT AN IDN REALLY IS

[17] The Internet domain name system (DNS) facilitates end users' ability to navigate the Internet by mapping a given domain name to its

---

[65] *Id. (quoting* Jamila Zhou, *Beijing Lodges Complaint Over Domain Names*, S. CHINA MORNING POST, Nov. 4, 2000).

[66] Wang, *supra* note 63.

[67] Carolyn Duffy Marsan, ICANN Accepts Bids to Operate .Net Registry, PC World (Jan. 21, 2005), http://www.pcworld.idg.com.au/index.php?id=397612571. NeuLevel teamed with Japan Registry Services, which operates Japan's .jp domain, on its .net bid. The two companies created a joint company called Sentan Registry Services. *Id.*

[68] NeuLevel Introduces Chinese Language Internationalized Domain Names (IDNs) in China's .CN Domain, NeuLevel (Jan. 18, 2005), http://www.neulevel.biz/press/press_release/IDN.CNrelease1-18-05.pdf.

[69] Xue, *supra* note 5, at 586.

[70] *Id.*

corresponding numeric Internet Protocol Number.[71] On a conceptual level, one writer makes an important point about IDNs: "It might be more precise to call these non-ASCII domain names 'localized' domain names than calling them 'internationalized' domain names"[72] because their aim is largely to serve for the local Internet communities whose native language is not English. American Internet users are familiar with the URL format http://www.*somedomainname*.com. An IDN in a different character set would, for example, retain the "http://" but not the domain name or the ".com" (or whichever appropriate designation is applicable, such as ".info" or ".org") in ASCII script.[73] This is done by using the established Latin letters, digits and hyphen, to encode the new IDN characters, which are any characters not restricted to the twenty-six letters of the Latin alphabet, the ten digits, and the hyphen.[74]

> Software that understands this system displays these characters as a user expects to see them, transparently encoding and decoding them as required. This is necessary, for example, when an IDN is entered into the address line of a Web browser. Recent versions of many Web browsers can perform the requisite conversions, and plug-ins are readily available for others."[75]

[18] Because there is one unique root system for the Internet, however, the ASCII script is still the basis of all non-ASCII domain names; they are in effect filtered through a type of translator called punycode. Punycode is a "simple and efficient transfer encoding syntax" created for use with IDNs in applications.[76] "It uniquely and reversibly transforms a Unicode string into an ASCII string. ASCII characters in the Unicode string are represented literally, and non-ASCII characters are represented by ASCII

---

[71] Joint ITU/WIPO Symposium, *supra* note 7, at 6.

[72] Xue, *supra* note 5, at n.4.

[73] *See, e.g.,* VeriSign, Web Addresses in Your Own Language, http://www.idnnow.com/index.jsp?lang=no (last visited Sept. 14, 2005).

[74] *See* .Museum, Internationalized Domain Names (IDNs) in .Museum, http://about.museum/idn (last visited Sept. 14, 2005).

[75] *Id.*

[76] A. COSTELLO, PUNYCODE: A BOOTSTRING ENCODING OF UNICODE FOR INTERNATIONALIZED DOMAIN NAMES IN APPLICATIONS (IDNA) (2003), http://www.rfc-editor.org/rfc/rfc3492.txt.

characters that are allowed in host name labels (letters, digits, and hyphens)."[77] Punycode is a way of representing Unicode codepoints using only ASCII characters; for comparative purposes, punycode is the same as ASCII for most strings, but changes the interpretation of some special strings (which are not in use anywhere) to allow characters such as the Greek αβγ.[78] The encoding is applied to each aspect of a domain name which is unable to be represented within the ASCII character set, and a reserved prefix "xn—" is added to the translated Punycode string. For example, *bücher* becomes *bcher-kva* in Punycode, and therefore the domain name "bücher.ch" would be represented as "xn–bcher-kva.ch" in IDN application language.[79] A problem with this system could arise if the resulting character strings are copied to programs that do not read punycode. Subsequent domain-name lookups will then fail unless all relevant software has been upgraded and tested for interoperability.[80] This is why name preparation or "Nameprep" exists. Nameprep is the process of preparing IDN labels "in order to increase the likelihood that name input and name comparison work in ways that make sense for typical users throughout the world."[81] In a more technical light, Nameprep is the process of Unicode normalization, mapping look-alike characters together, and eliminating restricted codepoints applied to text so that it is suitable to represent a domain name.[82] A given domain name string is assumed to have been normalized using Nameprep and filtered against an officially registered language table before being *punycoded*.[83] The DNS protocol

---

[77] *Id.*

[78] D.J. Bernstein, Internationalized Domain Names, http://cr.yp.to/djbdns/idn.html (last visited Sept. 17, 2005).

[79] Wikipedia, Punycode, http://www.answers.com/main/ntquery;jsessionid=12h2b59qdswph?tname=punycode&curtab=2222_1&hl=nameprep&sbid=lc04a (last visited Sept. 17, 2005).

[80] D.J. Bernstein, Internationalized Domain Names, http://cr.yp.to/djbdns/idn.html (last visited Sept. 17, 2005).

[81] *See* P. Hoffman & M. Blanchet, Nameprep: A Stringprep Profile for Internationalized Domain Names (IDN) (2003), http://www.ietf.org/internet-drafts/draft-ietf-idn-nameprep-11.txt. This definition is, of course, an oversimplification of the term, but the technical aspects of the process are beyond the scope of this paper.

[82] *See* Wikipedia, Nameprep, http://www.answers.com/topic/nameprep (last visited Sept. 17, 2005).

[83] *See* Wikipedia, Punycode, http://www.answers.com/main/ntquery;jsessionid=12h2b59qdswph?tname=punycode&curtab=2222_1&hl=nameprep&sbid=lc04a (last visited Sept. 17, 2005).

will have set limits on the acceptable lengths of the output Punycode string and the IDN will then presumably exist.

[19] At this point, a more specific understanding of Unicode is relevant. Unicode is the universal character encoding maintained by the Unicode Consortium that "provides the basis for processing, storage and interchange of text data in any language in all modern software and information technology protocols."[84] Punycode, therefore, is a mechanism for retaining the uniqueness of Unicode while allowing for IDNs and, so long as a user's Internet browser supports IDN standards, it would be able to reach the website with the non-ASCII URL. Different IDN registries have different designations and different language availability: WALID, for example, an IDN registry based in Michigan, uses a Java-based data entry application that enables direct input of multilingual characters in over 40 languages.[85] Companies and individuals all over the world can register for an IDN from any ICANN-accredited registrars that are IDN-certified.[86] The most prominent IDN registrars are located in China, Japan, Korea, Europe and the United States.[87] It is important to note here that non-ASCII characters are not just those with entirely different alphabets; they also include those languages with Latin scripts that have diacritics,[88] like the German "o" with an umlaut: "ö" and the French "c" with the cedilla: "ç," which are available on most United States' keyboards.[89] From a technical perspective and disregarding policy considerations such as cybersquatting and Nameprep, for example, there could be a difference now between the website www.chateau.com and www.château.com.   A punycode conversion

---

[84] Unicode Consortium, Glossary, http://www.unicode.org/glossary/#U (last visited Sept. 18, 2005).

[85] WALID, Technical Solutions, http://www.walid.com/en/docs/world_solution.shtml (last visited Sept. 18, 2005).

[86] VeriSign, VeriSign Internationalized Domain Names, http://www.verisign.com/static/002277.pdf (last visited Sept. 18, 2005).

[87] *See* Verisign, Selected Internationalized Domain Name Partners, http://www.verisign.com/static/002268.pdf (last visited Sept. 18, 2005).

[88] A diacritic is "[A]n accent near or through an orthographic or phonetic character or combination of characters indicating a phonetic value different from that given the unmarked or otherwise marked element." MERRIAM-WEBSTER'S COLLEGIATE DICTIONARY 318 (Frederick C. Mish et al. eds., 10th ed. 1999).

[89] *See, e.g.*, Wikipedia, Internationalized Domain Names, http://en.wikipedia.org/wiki/Internationalized_domain_name (last visited Sept. 18, 2005).

website hosted by an ICANN-accredited registrar shows that
"www.château.com" would look like "www.xn--chteau-xta.com" in
punycode.[90]

III. INTELLECTUAL PROPERTY: A NEW FORMAT FOR OLD PROBLEMS

[20] The actuality of a difference in domain names between such URLs as
www.chateau.com and www.château.com inevitably incurs domain name
and trademark disputes. The introduction and dissemination of IDNs "will
introduce added dimensions to the problem of conflicting identifiers
across different languages and language scripts which are in addition to
those already experienced in respect of conflicting trademark
registrations."[91] Such issues as phonetic similarity in the context of a
largely visual or textual medium are expected. The prevalent practice for
dealing with trademarked names on the international platform is to grant a
non-ASCII domain name to the owner of that trademark who has
registered that domain name in ASCII. One example is the World
Intellectual Property Organization's [WIPO] recent arbitration of
Getränke Holding AG v. De Souza.[92] The dispute concerned the domain
name www.rhäzünser.ch – .ch is the country code for Switzerland. The
plaintiff in the case owns four trademarks in Switzerland: Rhäzünser,
Rhäzünser+, Rhäzünser Aquaplus and Rhäzünser Plus, all being used for
sparkling and still mineral waters. From 1997 to 2004, the plaintiff
operated a website to market its products: www.rhaezuenser.ch. The letter
combinations "ae" and "ue" are recognized ways for representing "ë" and
"ü," respectively. On March 1, 2004, IDNs with diacritics were made
available for .ch domain names, and www.rhäzünser.ch was registered on
March 2, 2004 by the defendant. The WIPO arbiter found there to be a
clear infringement of the plaintiff's intellectual property right, as the
domain name was clearly identical to the plaintiff's trademark under
which it was nationally known and the respondent did not provide any

---

[90] *See* NameISP, http://www.nameisp.com/puny.asp (last visited Sept. 18, 2005). This
website allows users to enter the domain name (i.e., "www.château.com") and convert it
into punycode.
[91] Joint ITU/WIPO Symposium, *supra* note 7, at 8.
[92] Getränke v. De Souza, WIPO Case No. DCH2004-0012 (2004),
http://arbiter.wipo.int/domains/decisions/html/2004/dch2004-0012.html (last visited Sept.
18, 2005).

reasonable explanation for the registration of the disputed domain name.[93]
The WIPO Joint Recommendation Concerning Well-Known Marks
provides guidelines designed to uphold general tenets of international
cohesion: first, to avoid conflicts between trade and service marks across
different languages and, second, to provide specifically for registration
and protection of a mark, its translation and its transliteration.[94]

[21] A very similar fact pattern and outcome occurred in the case
CHERIE FM v. Sablon-Dauberton, wherein the respondent registered a
domain name identical to the plaintiff's except without the French
accent.[95] The trademark was almost exactly copied in the domain name,
the only difference being the égu accent mark on the 'e' of 'chérie.'[96]

> It is well established that, in determining identity or
> confusing similarity under Paragraph 4(a)(i) of the
> [Uniform Dispute Resolution] Policy, the generic top-level
> domain must be excluded from consideration (WIPO Case
> No. D2001-0868) and minor omissions of punctuation do
> not sufficiently alter the trademarked word to negate a
> finding of identity of confusing similarity (WIPO Case No.
> D2000-0059). This must also apply to the omission of an
> adjunct part of a letter, as with the French "égu" accent
> mark on an "e" letter, as in "chérie."[97]

[22] Cases are not always so clear, of course, and traditional
international intellectual property treaties come into play. "It
remains to be seen what significance will attach to the phonetic
similarity of trademarks and domain names, as internationalization
develops in the context of a largely visual or textual medium….
Several domain name cases have already addressed the issue, and

---

[93] *Id.*

[94] *See* WIPO, Joint Recommendation Concerning Provisions on the Protection of Well-
Known Marks (Sept. 20-29, 1999), http://www.wipo.int/about-
ip/en/development_iplaw/pub833-03.htm#P143_8519, at Articles 3-6.

[95] CHERIE FM v. Sablon-Dauberton, WIPO Case No. D2003-0486 (2003),
http://arbiter.wipo.int/domains/decisions/html/2003/d2003-0486.html (last visited Sept.
18, 2005).

[96] *Id.*

[97] *Id.*

hinted at its complexities."[98]   The Paris Convention for the
Protection of Industrial Property of 1883,[99] to which 169 States are
party,[100] applies to trademark and unfair competition law. As
alluded to above, WIPO's Uniform Domain Name Dispute
Resolution Policy [UDRP], setting out the legal framework for the
resolution of disputes between a domain name registrant and a
third party regarding the abusive registration and use of an Internet
domain name, governs disputes for IDNs as well as generic Top
Level Domains and country code Top Level Domains.[101]   The
WIPO Policy generally involves the weighing of three factors: 1.
Whether the domain name is identical or confusingly similar to a
trademark or service mark in which the plaintiff has rights; 2.
Whether the respondent has rights or legitimate interests in respect
of such domain name; and 3. Whether the domain name was
registered in bad faith and is being used in bad faith.[102]   These
rules are weighed and balanced against each other and as compared
to prior cases; the addition of the script and language factors add
another layer of complexity.

A. Special Problems with Foreign Language and Trademark Law

[23] Courts in the United States, following the lead of the United States
Patent and Trademark Office (USPTO), handle translation and
transliteration in a specific manner. Transliteration is the phonetic
spelling in corresponding Latin characters of word(s) that are in non-Latin

---

[98] Joint ITU/WIPO Symposium, *supra* note 7, at 37.

[99] WIPO, Paris Convention for the Protection of Industrial Property, March 20, 1883,
http://www.wipo.int/treaties/en/ip/paris/pdf/trtdocs_wo020.pdf.

[100] WIPO, Treaties Database Contracting Parties, Paris Convention,
http://www.wipo.int/treaties/en/ShowResults.jsp?country_id=ALL&start_year=ANY&en
d_year=ANY&search_what=C&treaty_id=2 (last visited Sept. 18, 2005).

[101] *See, e.g.*, John W. Bagby and John C. Ruhnka, *Protecting Domain Name Assets*, THE
CPA JOURNAL, Apr. 1, 2004, *available at*
http://www.nysscpa.org/cpajournal/2004/404/essentials/p64.htm (last visited Sept. 18,
2005).

[102] *E.g.*, Bella I. Safro and Thomas S. Keaty, *What's in a Name? Protection of Well-
Known Trademarks Under International and National Law*, 6 TUL. J. TECH & INTELL.
PROP. 33, 45 (2004).

characters.[103] For example, a Japanese symbol that is not possible to replicate on this keyboard transliterates to "asahi," which translates to "rising sun."[104] As any professional translator or interpreter knows, the nuances and peculiarities of any given language incur many complications beyond what any set of rules could identify.[105] It is therefore often imperative that anyone adjudicating trademark or domain name disputes between different languages be fluent in both languages at issue.[106] One of the most illustrative cases concerning the majority of issues mentioned in this writing is *Sankyo Co., Ltd. v. Zhu Jiajun*,[107] which was filed with the WIPO Arbitration and Mediation Center on December 21, 2000. At issue was a Chinese-character domain name (translated into "sankyo") which was registered with OnlineNic, Inc. in the United States., doing business as China-channel.com in China.[108] Sankyo, a large Japanese pharmaceutical corporation, owned the trademark "Sankyo" in Japan, in the United States and in China. The respondent, Zhu Jiajun, the owner of Sankyo Art Salon, registered the website "www.sangong.com" in China, which is the Chinese pronunciation of "sankyo."[109] The arbitration panel acknowledged that the respondent's use of the domain name was

---

[103] The United States Patent and Trademark Office, *Translation and Transliteration of Non-English Wording in Mark*, *in* TRADEMARK MANUAL OF EXAMINING PROCEDURE (TMEP) § 809 (Mary E. Hannon, ed., 4th ed. 2005), *available at* http://www.uspto.gov/web/offices/tac/tmep/pdf/0800.pdf.

[104] *Id.*

[105] *See* The Baheyeldin Dynasty, Arabic on the Internet: Translation Pitfalls (Dec. 20, 2004), http://baheyeldin.com/arabization/translation-pitfalls.html (providing illustrative examples of this sort of translation complication in the context of Arabic to English translation).

[106] *See* WIPO Guide to the Uniform Domain Name Dispute Resolution Policy (UDRP), The Role of the Administrative Panel, http://arbiter.wipo.int/domains/guide/index.html (last visited Sept. 18, 2005).

> The persons appearing on the WIPO Center's list of Domain Name Panelists have been selected on the basis of their well-established reputations for their impartiality, sound judgment and experience as decision-makers, as well as their substantive experience in the areas of international trademark law, electronic commerce and Internet-related issues. The WIPO Center's list is truly international, consisting of more than 250 Panelists from 42 countries, many of whom are multi-lingual.
> *Id.*

[107] Sankyo Co., Ltd. V. Zhu Jiajun, WIPO Case No. D2000-1791 (Mar. 23, 2001), http://arbiter.wipo.int/domains/decisions/pdf/2000/d2000-1791.pdf.

[108] *Id.* at 1.

[109] *Id.* at 3.

confusingly similar to that of the plaintiff's trademark and that the respondent's use of the mark violated both Japanese trademark law and Japan's Unfair Competition Prevention Law. The panel concluded that the respondent had no legitimate interest in the domain name and registered in bad faith, and therefore that the domain name should be transferred to the complainant.[110]

[24] Specifically, in parsing through the international legal and phonetic issues, the panel found that the "sankyo" mark is famous in Japan because is composed of two kanji characters, meaning "three together."[111] Japanese and Chinese languages and script have been described as ill-suited to Internet and IDN adoption in general, and this case illustrates something of a primer for the adjudication of non-ASCII domain names.[112] Japanese language IDNs are further complicated by its use of both kanji and kana characters.

[25] Likewise, the Chinese population uses both simplified Chinese characters – mostly in Mainland China – and traditional Chinese characters – mostly in Taiwan, Hong Kong and Macao.[113] This complicates policy and technical considerations at the domestic level as well as the international scale. As one French commentator writes,

> To many, IDNs represent a technological challenge to overcome: We need to make the Internet function in a different way than that in which it was conceived… imagine a Japanese businessman working with a European. He gives the European the URL to his website so that he can better understand his products. More likely than not, the Japanese businessman uses an IDN in Kanji…but the European keyboard does not provide a means to inputting those characters. Even if the European could identify or

---

[110] *Id.* at 8-9.
[111] Yonehara, *supra* note 17, at 228-29.
[112] *Id.* at 229, 231.
[113] Xue, *supra* note 5, at 575.

understand them, which is not likely, he would be unable to type them![114]

This same writer goes on to argue that IDNs do not pose solely technical problems – multilingual domain names also pose functional problems. A Japanese domain name is not likely to be exploited by someone who cannot read the language even if he or she has the technical capacity to do so. The power of the Internet lies in its universality. IDNs reintroduce the limitations of true universality with situations such as the one above, but they also provide the means for its eventuality, in that the ability of a European to type the Japanese-language domain name in the URL bar will at least make him familiar with a few Japanese characters and how to type them.

[26] Returning to the "asahi" example, it is important to note that domestic trademark laws still sometimes handle translation and transliteration differently. Whereas trademarks in the brick-and-mortar world may exist with a modicum of overlap, domain names generally cannot and arguably should not, despite different TLDs. For example, there is an upscale department store on the west bank in Paris called Le Bon Marché. Up until recently, prior to its merger with Macy's, there has also been a Bon Marche department store in the United States and, to further complicate the menu, there is a Bon Marché in the United Kingdom. The country-code domain names differentiate between the companies (.fr, .com and .co.uk, respectively), but the whole of the domain name tangle is not yet resolved. Typing "www.bonmarché.com" in a web browser that supports IDNs brings the end user to a Spanish website called "Polidias," which apparently has nothing to do with any of the Bon Marché entities.[115]

IV. CONCLUSION

[27] Technical glitches are still not entirely overcome in order for IDNs to function seamlessly on the Internet. But it is not only technical problems

---

[114] Stéphane Van Gelder, IDN – Le Miroir aux Alouettes?, http://www.domaines.info/imprimer_chronique.php?chronique_id=27 (last visited Sept. 18, 2005) (translated by the author).

[115] Polidias, http://www.bonmarché.com/ (last visited Sept. 18, 2005).

with which the international Internet community should be concerned. Various policy-making bodies, ranging from localized consortia, such as the Chinese Domain Name Consortium, and international organizations, such as ICANN, should work diligently to solidify policy guidelines and best practices that come as close as possible to providing localized services in native languages while at the same time maintaining the worldwide nature of the Internet. There will doubtless be problems throughout the implementation stage ranging from orthography and phonetics to government involvement and foreign backlash. While it is, of course, easier to advocate than it is to implement, there is no doubt that standardization is essential.

[28] The IETF Working Group on IDNs has the delicate task of specifying the "requirements for internationalized access to domain names and to specify a standards track protocol based on the requirements."[116] A fundamental requirement for their work is to refrain from disturbing the current use and operation of the domain name system, and "for the DNS to continue to allow any system anywhere to resolve any domain name."[117] The Internet has influenced so many things, not the least of which is international communication. Adding language capabilities should only strengthen the viability and worth of the Internet and a single protocol – both technically and ideologically – is required to achieve that ambition.

---

[116] IETF, Internationalized Domain Name Charter, http://www.ietf.org/html.charters/OLD/idn-charter.html (last visited Sept 18, 2005).
[117] *Id.*