

Southern Illinois University Edwardsville

SPARK

SIUE Faculty Research, Scholarship, and Creative Activity

Winter 12-15-2014

Linguistic Diversity on the Internet: Arabic, Chinese and Cyrillic Script Top-Level Domain Names

Undrah Baasanjav
ubaasan@siue.edu

Follow this and additional works at: https://spark.siu.edu/siue_fac



Part of the [Communication Technology and New Media Commons](#), [International and Intercultural Communication Commons](#), and the [Mass Communication Commons](#)

Recommended Citation

Baasanjav, Undrah, "Linguistic Diversity on the Internet: Arabic, Chinese and Cyrillic Script Top-Level Domain Names" (2014). *SIUE Faculty Research, Scholarship, and Creative Activity*. 71.
https://spark.siu.edu/siue_fac/71

This Article is brought to you for free and open access by SPARK. It has been accepted for inclusion in SIUE Faculty Research, Scholarship, and Creative Activity by an authorized administrator of SPARK. For more information, please contact jkohlbu@siue.edu.



Linguistic diversity on the internet: Arabic, Chinese and Cyrillic script top-level domain names



Undrah B. Baasanjav*

Southern Illinois University Edwardsville, USA

ARTICLE INFO

Available online 20 May 2014

Keywords:

International domain names

IDN

Language diversity

Country-code TLD

ICANN

Arabic

Chinese

Cyrillic

ABSTRACT

The deployment of Arabic, Chinese, and Cyrillic top-level domain names is explored in this research by analyzing technical and policy documents of the Internet Corporation for Assigned Names and Numbers (ICANN), as well as newspaper articles in the respective language regions. The tension between English uniformity at the root level of the Internet's domain names system, and language diversity in the global Internet community, has resulted in various technological solutions surrounding Arabic, Chinese, and Cyrillic language domain names. These standards and technological solutions ensure the security and stability of the Internet; however, they do not comprehensively address the linguistic diversity needs of the Internet. ICANN has been transforming into an international policy organization, yet its linguistic diversity policies appear disconnected from the diversity policies of the United Nations, and remain technically oriented. Linguistic diversity in relation to IDNs at this stage mostly focus on the language representation of major languages that are spoken in powerful nation-states, who use the rhetoric of national pride, local business branding, and inclusion of non-English speakers. This situation surfaces the tension between nation-states and the new international governing institution ICANN.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction: International domain names

On the 40th anniversary of the Internet in November 2009, the Internet Corporation for Assigned Names and Numbers (ICANN), the governing body of the Internet, approved a new standard for fully internationalized domain names that use characters outside the range of the capital and small Roman letters from A to Z, the Arabic numbers of 0–9, and the hyphens as used in the English language (Cooper, 2008). Theoretically, international domain names consisting entirely of native character sets tend to benefit local companies and people who only speak their local languages and improve access to the Internet.

Domain names, textual names of web resources on the Internet, are descriptive markers with corresponding numerical addresses called Internet protocol (IP) addresses. When an end user types a web address or an email address, the domain names system (DNS) on the Internet resolves the entered web address into the IP address of the requested web host or email user addresses (Zook, 2000; see also Kleinwachter, 2000; Mueller, 2002; Síthigh, 2010). Special computers on the Internet, called name servers, resolve a web resource address (e.g. www.mol.mn) into an IP address (202.131.0.3 or an IP address

* Correspondence address: Department of Mass Communications, Dunham Hall, Southern Illinois University Edwardsville, Illinois 62026.
E-mail address: undrah.baasanj@gmail.com

block starting with 202.131). The reliability and security of the Internet depend on the effective resolution of domain names. Domain names consist of top-level domain names placed at the very end, and sub-domain names separated by dots. Top-level domain names are divided into generic top level domain names (gTLD) such as .com, .edu, .gov, .mil, .net, .org, .int, .asia, .africa and country code top level domain (ccTLD) names assigned to certain countries and territories such as .cn (China), .in (India), .ru (Russia) and .sa (Saudi Arabia). There were more than 233 million registrations across all top-level domains (TLDs) and 94.9 million registrations across all country code top-level domains (ccTLDs) in 2012 (Verisign, 2012). Prior to the initiation of fully internationalized top-level domain names (IDN ccTLDs) such as .دبي (Arab Emirates), .中国 (China), and .рф (Russia) in 2009, international characters were used in second-level (SLD) or subsequent level domains in both generic (gTLD) and country-code TLD (ccTLD) domain names, and they were referred to as international domain names (IDN) (Cooper, 2008; Xue, 2004).

Internationalized domain names have socio-economic, political and cultural entailments, and the initiation of ccTLD IDN raises a range of issues concerning competition policy, compatibility, name rights and trade protectionism. The full internationalization of domain names evokes also debates surrounding the US government's control of the DNS root, and the multilateral governance of the Internet (see Froomkin, 2011; Mueller, 2002; Sithigh, 2010). Furthermore, the initiation process of IDN ccTLDs shows the tension between the traditional state-centered international governance model and a new transnational institutional governance model personified by the ICANN regime. Even though nascent transnational networks of actors such as the Internet Engineering Task Force (IETF), regional Internet address registries, and ICANN have created radically different institutional arrangements than those of the traditional nation-state-centered international regime, the initiations of IDN in Arabic, Chinese and Cyrillic scripts have been spearheaded or backed up by the states in the respective countries.

This research analyzes the policy implementation of Arabic, Chinese and Cyrillic script domain names by ICANN and examines technical specification and security challenges in tandem with social and cultural issues of multilingual domain names. The paper evaluates ICANN policies by asking whether or not the initial goals of multilingualism are being achieved.

2. Review of the literature

2.1. Linguistic diversity and multilateral governance of the Internet

As the Internet becomes a mature medium, cultural diversity concerns endure. Diversity, “a rich multifaceted principle” has always been at the center of communications policy (Napoli, 2011). The rise of English as a lingua franca on the Internet is a reality (Hobsbawm, 1996; Warschauer, 2003), and despite the various efforts to integrate international languages and different character sets since the earliest stages, the extent to which the Internet is international “remains a matter of deep concern” (Braman, 2012, p. 28). Digital divide scholars have pointed out that the great discrepancy of language representation on the Internet is one of the contributing factors to the global digital divide (Baasanjav, 2012; DiMaggio, Hargittai, Celeste, & Shafer, 2004; Hamelink, 2000; Hargittai, 2003; Warschauer, 2003).

The United Nations (UN) sponsored a series of highly prominent international summits with the primary intent of addressing the global digital divide and to further develop a global vision pertaining to the information society, and these international summits have become known as the World Summits on Information Society (WSIS). WSIS to a certain extent mobilized and involved transnational advocacy groups, which promote global human rights and equality causes in policy making in relation to the development of global communication. The UN in Article 27 of *The Universal Declaration of Human Rights* (UDHR) states “everyone has a right to participate in the cultural life of the community, to enjoy the arts and to share in scientific advancement and its benefits” (UN, 1948). Theoretically, this statement protects the language rights of non-English speakers worldwide concerning their participation on the Internet. Many governments insisted on the need for multilateral, transparent, and democratic governance of the Internet based on the cause of language rights, and challenged the US centric and private-led governing of critical Internet resources (Mueller, 2010). That is how previously obscure technical issues concerning domain names systems (DNS) have become the center of WSIS debates. WSIS involvement in Internet governance resulted in the strengthening of the Governmental Advisory Committee (GAC) within ICANN, through which nation-states other than the U.S. increasingly influenced decision making (Froomkin, 2011; Mueller, 2010). As discussed in the section below, the multi-stakeholder arrangement of ICANN involving governments, private organizations, civil society institutions, and heterogeneous network actors around the globe led to its *sui generis* status (Mueller, 2010; Sithigh, 2010). This somewhat institutionally innovative ICANN regime has gotten into an uneasy relationship with the state-centered international regime when initiating IDNs. Furthermore, IDN initiation at the top level has brought unprecedented challenges to global media governance institutions that have been primarily concerned with the audiovisual exchange in transnational media space and telecommunications compatibility, standards, and resource allocations (see O'Siochru, Girard, & Mahan, 2002; Hamelink, 2000; Wilkinson, 2004).

2.2. International governance of IDNs: ICANN

The Internet Corporation for Assigned Names and Numbers (ICANN), a not-for-profit international organization, regulates domain name systems (DNS) and helps ensure the stability of the Internet. ICANN oversees the top-level domains, accredits domain name registries and registrars, registers and maintains the root zone of DNS, and establishes

policies for resolving trademark and other claims over domain names. At the earlier stages of the Internet, domain names were managed by a three-person ad-hoc committee led by Dr. Jon Postel at the University of California at Los-Angeles. In 1998, ICANN was established as an independent, private, not-for-profit organization to manage the coordination of the Internet's unique identifiers with the approval of the US government (see [Easton, 2012](#); [Kleinwachter, 2000](#); [Mueller, 2002](#); [Mueller & McKnight, 2004](#)) and since then has been transitioning to a multilateral, independent and international institution. UN sponsored WSIS has pressured ICANN and the US Department of Commerce for the multistakeholder management of critical Internet resources that subsequently led to ICANN's willingness to create a fast track process of ccTLDs that privileges state controlled IDN TLDs ([Mueller, 2010](#), pp. 232–234). In 2009, ICANN and the US Department of Commerce signed a policy document “Affirmation of Commitments” (AoC), in which the two parties asserted their commitments to transparency, accountability, resiliency, security, and competitiveness of the DNS. [Froomkin \(2011\)](#) and [Mueller \(2010\)](#) argue that the AoC does not have binding promises by the two parties; neither does it outline an accountability mechanism for ICANN. The AoC was a good political move by the U.S. government that slightly loosened its unilateral oversight of ICANN and deflected domestic and international pressure for ICANN's independence without binding commitments ([Froomkin, 2011](#); [Mueller, 2010](#)). In terms of the IDN initiation, signing the AoC was an important step toward the greater involvement of international participation in DNS coordination. The AoC also endorsed “the rapid introduction of internationalized country code top level domain names (ccTLDs), provided related security, stability and resiliency issues are first addressed” ([ICANN, 2009](#), September 30, para 5).

ICANN's legitimacy as an unprecedented legal and multinational regulator is inherently controversial. On the one hand, as an international governing organization of the Internet, ICANN follows the principles of international law including the United Nations' *Universal Declaration of Human Rights* (UDHR) and World Intellectual Property Organization (WIPO) treaties ([Sithigh, 2010](#)). On the other hand, as a quasi-corporate structure regulating a market worth billions of dollars mostly concentrated in the US, ICANN's position to present the public interest of the global Internet community is weak and is questioned by non-US governments and international civil society organizations. Powerful commercial interest groups such as the Virginia based corporation Verisign (formerly Network Solution Inc.), which made windfall profits selling .com domain names, play an influential role in ICANN's decision making ([Froomkin, 2011](#)). The DNS root zone file, a single and globally consistent list of top-level domain name assignments with pointers to authoritative name servers (NSs), is still somewhat controlled by the U.S. Department of Commerce and is housed by VeriSign; and this situation has been contested by non-US governments (see [Froomkin, 2011](#); [Mueller, 2002, 2010](#)). Taking advantage of its position, Verisign also attempted to get “the first claim” to IDN strings in non-Roman scripts that match the meaning of the strings in English names in the .com domain ([Mueller, 2010](#), p. 233).

Because domain names in multiple languages raise potential problems concerning country names, territories, trademarks and other name rights that are beyond ICANN capabilities, it has been delegating the registry and management of ccTLDs in non-English scripts to national and regional registries, while the management of gTLDs has been delegated to private registries mostly in English speaking countries. This situation creates some tensions among corporations managing gTLDs, the nation-state authorities in charge of ccTLDs, and cultural/linguistic interests of various groups advocating for the accountability and transparency of the DNS in terms of linguistic diversity. In the past, television and radio industries in the US bought .tv and .fm country code domain names from the developing nations of Tuvalu and the Federation of Micronesia and exploited these ccTLDs for businesses unrelated to those developing countries ([Hrynshin, 2008](#)). Less developed countries like Mongolia tend to use ccTLDs almost three times more than gTLDs ([Baasanjav, 2012](#)). In 2012, ICANN fully liberalized top-level domain names (gTLDs) allowing corporations to apply for new suffixes such as .coke, .invest, and .bank for a \$185,000 application fee ([Easton, 2012](#)); and some developing nations criticized this ICANN move for benefiting corporations while offering little value to users around the world ([Ramachandran, 2011](#)). Furthermore, ccTLD registries in many countries are run by state authorities or organizations implementing cultural and economic policies of the states, which often prioritize language unity over language diversity within the nation ([Sithigh, 2010](#)). When this conservative state-centered approach is applied to new IDN ccTLDs, the grassroots and decentralized character of the Internet is in jeopardy ([Mueller, 2010](#)).

2.3. Technical standards and stability concerns of multilingual TLDs

ICANN has been issuing a limited number of country-code TLDs representing a respective country in scripts other than Latin such as .ا.م.ع (United Arab Emirates), .中国 (China), and .рф (Russia) through its fast track process. An eligible country (e.g., the United Arab Emirates) or a territory (e.g., the Palestinian territory) requests a respective IDN ccTLD(s) from ICANN, which evaluates the proposed “string” (e.g., “.рф”) against a set of criteria focusing on the meaningfulness of linguistic representation and security requirements, and delegates the management to an eligible entity ([ICANN, 2012](#), June 4).

One of the major challenges of implementing IDNs has been the introduction of different character sets in the domain name systems. Two relevant technical standards – Unicode and punycode – need to be explained in relation to IDN ccTLDs. International standards for interchange, processing, and the display of multilingual characters sets and diverse written languages have resulted in Unicode standards, which provide “a repertoire of code points used in different scripts, including various classifications of character properties, and normalization rules” ([ICANN, 2012](#), February 20, p. 10). When a user types, for example, a Chinese ideograph for “hill” 山, a browser or search engine software uniquely renders a Unicode code point of (U+5C71) regardless of the differences in platforms, software and language employing the Unicode standard.

However, while country-code domain registries might recognize IDNs, the root file and the authoritative root servers of the DNS do not recognize Unicode characters and still work only in the LDH (letters, digit and hyphen used in the Roman character set) characters. Because of this hierarchical nature of the DNS, a standard called the International Domain Names in Applications (IDNA) was developed which converts Unicode character sets (U-label) to a “punycode” string in LDH characters sets (A-label) prefixed by ‘xn.’ For example, the Cyrillic name of Russia’s IDN ccTLD ‘.рф’ is represented in the string ‘xn-p1ai.’ The IDNA protocol also specifies rules for determining whether a code point can be included in a domain name (ICANN, 2012, February 20, p.10). The latest version is the IDNA2008 standard which incorporates more variances of IDNs.

Providing a consistent and uniform naming system is essential for the stability of the Internet (ICANN, 2012, February 20, p. 8). One of the major stability and security concerns raised by the internationalized domain names in languages like Arabic, Chinese and Russian has been the “spoofing” of domain names (Sithigh, 2010). IDNs make it easier for criminals to impersonate or spoof web sites by mixing different scripts leading to homograph attacks, phishing, and redirects in order to steal money, information, or goods. The citibank.com web site can be impersonated by replacing the Latin letter c (Unicode character U+0107) with the Cyrillic letter c (Unicode character U+0301) thus luring bank customers to a false site. The two letters look alike and are homographs, and redirecting with malicious intends is called a homograph attack.

2.4. Trademark and intellectual property concerns of multilingual TLDs

The increase in values and registrations of domain names in multiple languages might proliferate trademark and other claims over domain names. Name dispute resolution is one of the major functions of ICANN. Since 1999 ICANN has adopted the Uniform Dispute Resolution Policy (UDRP) in order to effectively settle domain name disputes without lengthy and expensive court proceedings (Zainol, Hussain, & Yaakub, 2011). All ICANN accredited TLD registrars, as well as most country-specific registrars have adopted the UDRP (Kaur & Aggarwal, 2011; Leng, 2010; Zainol et al., 2011). According to the UDRP, a panel decides upon a complaint by a trademark and property right holder who shall show (i) a domain name is identical or confusingly similar to a trademark; (ii) a domain name registrant has no right to or legitimate interest in the domain name; and (iii) a domain name registrant has registered it in bad faith (Leng, 2010; Zainol et al., 2011, p. 276). The UDRP has been effective in curtailing cyber squatting or the registration of domain names with the intent to profit from the trademark or intellectual property belonging to someone else. These tests have not been applied to IDNs yet. A UDRP panel’s decision, however, can be challenged and overturned in local and national courts. This paper explores whether or not national authorities pursue IDN ccTLD to protect cultural spaces in the same way that they protect audiovisual imports (see O’Siochru, Girard, & Mahan, 2002; Waisbord & Morris, 2001).

2.5. Research questions and method

The review of the relevant research revealed major international, technical and policy challenges to the ICANN regime. Few current academic articles have analyzed technical and policy issues surrounding the initiation of IDN ccTLDs (EURid-UNESCO, 2012; Sithigh, 2010; Xue, 2004), and Internet governance policies are being implemented on a trial and error basis. This research analyzes the essential debates surrounding the initiation of IDN in Arabic, Chinese and Cyrillic scripts, which unprecedentedly challenged the existing Internet governance regime. Arabic, Chinese and Cyrillic script IDNs were chosen for their language-specific characteristics, as well as for the cultural and political contexts in which these policies are implemented. This research strives to spotlight policy challenges by posing the following questions: (1) What are the technological standards and security issues related to the deployment of Arabic, Chinese and Cyrillic top-level domain names? (2) What are the economic and policy themes that characterize the deployment of Arabic, Chinese and Cyrillic top-level domain names? (3) How has the internationalization of domain names been characterized and defined in the media of respective Arabic, Chinese and Russian language countries?

This research first analyzes how the problems of internationalization of domain names in Arabic, Cyrillic and Chinese scripts are defined and characterized in ICANN documents, as well as in news reports originating in Asian, Eastern European and Middle Eastern newspaper sources. The analysis of ICANN documents in comparison with news reports in the respective regions explicates different sets of interests in the implementation process of IDNs. Though highly technical, ICANN documents present rigorous data relating to IDN policy, procedure and proceedings given the fact that most of the discussions and operations of ICANN are posted online for public comments. Newspaper reports were obtained by searching the Access World News database from NewsBank, one of the most comprehensive full text newspaper collections globally, by searching for the keywords “international domain names.” Due to the highly technical nature of international domain names, newspaper reports immediately relevant to IDN technical and policy issues are rare, and available reports often overlap resulting in the relatively small number of around 60 full text articles.

Furthermore, this research evaluates the processes of ICANN’s delegation of ccTLD IDN to regional registries through fast track processes. The research strives to evaluate ICANN’s policy by asking (1) to what degrees does the delegation process serve the original goal of the IDN initiation; and (2) have ICANN’s interests skewed the implementation process?

3. Case studies and analysis

3.1. Arabic script domain names: Arab identity

Arabic is one of the six most widely spoken languages in the world, and Arabic script is used to write multiple languages ranging from Semitic and Indo-Iranian to African and Asian languages. It is the fastest growing language on the web ("[Web turns to Arabic](#)," 2010), and ICANN and the Internet Engineering Task Force (IETF) with the support from regional economic and political organizations like the League of Arab States (LAS) has worked to meet the "pressing demands" of Arab speaking countries to initiate international domain names.

ICANN policy documents and reports by the ICANN Arabic case study team explain the challenges surrounding Arabic domain names mostly by focusing on linguistic and technical standards and technological implementation with little reference to their cultural and political significance. The Arabic writing system is cursive and bi-directional, and each letter may have four shapes – initial, medial, final and isolated – depending on where it occurs within a word. In the Unicode standard, the Arabic script characters are presented by code points U+0600–U+06FF and U+0750–U+077F, and only a subset of those characters can be used in domain labels. Different writing styles such as the Perso-Arabic Nastaliq and the Arabic Naskh have different renderings and are not mutually comprehensible (ICANN, 2011, October 7). Regional differences and writing styles have resulted in different variants or alternate labels in the IDNA 2008 standard. For example, IDN ccTLD for Qatar (قطر) considers the characters ق and ك as variants. ICANN delegates the management of visually identical domains names (variants) to the same entity and the same regional registry so that the stability issues are resolved within their zones (ICANN, 2013, p. 24). Furthermore, the lack of a standard keyboard for various Arabic languages makes user experiences inconsistent, and web browsers, web hosting sites, and Internet security services still do not consistently accommodate Arabic domain names (ICANN, 2011, October 7).

While ICANN documents are primarily concerned with technical standards and Internet security concerns, news reports originating in Arabic speaking countries laud the initiation of Arabic domain names as the recognition of Arab culture and identity, and focus on opportunities for local businesses. For example, ArabianBusiness.com cites Mohamed Naser Al Ghanim, the director general of UAE's Telecommunications Regulatory Authority (TRA) who said:

Our objective is to support Arabic content over the internet and build a strong identity for UAE and Arab countries across the World Wide Web. We encourage all companies and entities in the UAE to register for .ae and .emarat as it represents their local business identity and brings their country pride (Sambidge, 2009, Oct 20).

By focusing on business opportunities and Arab identity, news reports rarely provide analyses of the technological issues and security concerns of Arabic domain names discussed at lengths in ICANN documents. This situation indicates the limited effort by the government and telecommunications businesses to inform the public and to reach other grass root organizations and interest groups. Diverging emphases in ICANN documents and Arabic regions news reports suggest a need to coalesce ICANN's policies into regional news discourse. Informing the public about these ICANN policy and technical issues is important in order to involve linguistic groups whose interests IDN ccTLDs are meant to serve.

ICANN, under the pressure from Arab states, as well as from other foreign governments to internationalize domain names, expeditiously delegated the first Arabic IDN ccTLD associated with .ae (Arab Emirate) اتارام (A-label xn – mgbaam7a8h) through the fast-track approval process. By initiating Arabic domain names, and delegating the non-contentious IDN ccTLD to government entities, ICANN received support from the Governmental Advisory Committees and met some pressing needs of non-English speaking countries for IDNs. Furthermore, in its effort to transform to a global entity and to better manage Arabic domain policy and technical issues, ICANN nominated Akram Atta Allah, a Beirut native, as the new chief operating officer ("[Atta Allah making a difference](#)," 2010, October 9).

The government of the United Arab Emirates through its Telecommunications Regulatory Authority (TRA) launched the first Arabic IDN ccTLD and received much popular support in the country. The Arab States including the Arab Emirates and Qatar and the League of Arab States (LAS), a regional intergovernmental organization with great financial and technological resources, supported various technical task forces to initiate Arab script IDNs and to overcome the technological constraints. This political and financial support has played an important role in the initiation of Arab IDNs, especially considering the cost of the fast track approval process of ICANN which is approximately \$26,000 US dollars (ICANN, 2009, June 4). Other Arab countries – Algeria (رياز جلا), Egypt (رصم), Morocco (برغول), Oman (نامع), and Qatar (رطق) – followed and initiated their IDN ccTLDs (ICANN, 2013).

The accountability and transparency of ICANN and the Internet's stability and security issues in relation to the initiation of IDN ccTLDs were discussed without specific responsibilities and binding obligations, as a gesture only. In a mutual understanding letter between ICANN and the local domain names registry, UAE's TRA commits to the Internet's openness, security and stability. ICANN recognizes TRA as "a highly professional registry" which was quick in adopting DNSSEC ("[TRA Signs](#)", 2011, October 11). The relationship between ICANN and UAE's TRA is loose and not binding. References to a few disputed cases concerning name rights and cybersquatting are mentioned in regional newspapers ("[Watchdog predicts](#)," 2009, October 13), but procedures for transparent and accountable management of IDNs at registry and registrars levels are clearly left without obligatory responsibilities on either sides. The Arabic case shows how governments use local language

domain names symbolically as emblems of national and cultural pride citing the importance of including Arabic only speakers and the economic benefit to local businesses.

3.2. Chinese domain names: User experience and business opportunities

One-fifth of the world's population speaks Chinese, and the market value of Chinese language domain names has driven technological and public policy debates. Since 2000, Network Solution Inc. (NSI, later VeriSign) tested second level domain names (SLD) within the .com domain causing a vehement protest by the Chinese government claiming that no company should be allowed to provide Chinese character domain name registration in China without an approval of the Chinese government (Xue, 2004, p. 569). The Chinese government preempted the award of Chinese character domains to NSI and created three Chinese language suffixes independently of the global root – 中国 (.zhonggou), 公司 (gongsi or “company”), and 网络 (wangluo or “network”) (Mueller, 2002, 2010; Sithigh, 2010; Xue, 2004). By operating an alternate national root within the country in the local language, the Chinese government showed the feasibility of operating an alternate root independently of the global root, and this action in return helped to highlight the value of the interoperable global DNS root. Despite some initial unilateral actions, the Chinese government and its agencies like CNNIC have played important roles in IDN debates by developing technical standards for Chinese language IDNs and reforming ICANN into a more international organization (Sithigh, 2010; Xue, 2004).

Technological and policy challenges of Chinese character domain names are still being addressed. Chinese written script or Han script consists of thousands of characters, graphic symbols called ideographs often representing ideas. In the 1960s, the Chinese government introduced simplified Hanzi characters in order to increase literacy. This innovation led to two different written systems: simplified Chinese (SC) used in mainland China and Singapore, and traditional Chinese (TC) used in Taiwan, Hong Kong and Macau. The International Standard Organization and the Unicode Standard have implemented ISO/IEC 10646 standard, which unified more than 70,000 Chinese Hanzi, Japanese Kanji and Korean Hanja characters into “CJK Unified Ideographs,” since both Japanese Kanji and Korean Hanja have been imported from the Han script (ICANN, 2011, October 3). Different writing styles have resulted in different variants in IDNs and in the IDNA 2008 standard. Variants in Chinese domain names are defined as “characters with different visual forms but with the same pronunciations and with the same meanings” (ICANN, 2011, October 3). For example, 中国 and 中國 both stand for .zhonggou in Chinese, and have different Unicode code points that translate to different A-labels of xn–fiqs8s and xn–fiqz9s respectively in the IDNA2008. Variants to the same conceptual Chinese domain name labels under 中国 and 中國 in simplified and traditional forms are delegated to the same registrants (ICANN, 2011, October 3).

The initiation of Chinese domain names is indicative of the uneasy relationship between nation-state and language, as well as between national government and Internet content regulation as Sithigh (2010) points out. Given the relatively long experience of the Chinese language community with IDNs, ICANN delegated the management of 中国/中國 (A-labels xn–fiqs8s /xn–fiqz9s) top-level domains to the China Internet Network Information Center (CNNIC), the government-linked domain name registry agency, and the management of 台湾/台湾 (A-labels xn–kpry57d /xn–kprw13d) top-level domains to the Taiwan Network Information Center (TWNIC) through the fast-track approval process in 2010 (ICANN, 2011, October 3; ICANN, 2013). The Chinese state authority CNNIC enforces content regulation on the Internet and requires proof of business ownership from a registrant of a .cn domain name to “further enhance the authenticity, accuracy and integrality” of registration information for the domain name (Chia, 2009, December 14). The Chinese government imposes a tight requirement upon international domain name registrars like Go Daddy for issuing domain names within the .cn zone by requiring photo IDs and other personal information of the registrants (“Concern for the safety of Chinese Users,” 2010, March 27). It appears that by participating actively in technical and policy developments, the Chinese government took control over the Chinese IDN ccTLDs, and is reaching over other registries with its content regulation policies.

The analysis of ICANN policy documents and newspaper reports suggests that most Chinese users rely on search engines to find websites and do not often type in Chinese character domain names (But, 2009, December 14; Yan, 2010, July 14; ICANN, 2013). From the end users' perspective, typing pinyin characters (China's official system of Romanization) and then selecting proper characters might result in more key-strokes (ICANN, 2013). Chinese users are accustomed to ideographs, rather than letters, and Latin alphabet letters, for example “i” and “r,” in domain names and email addresses have been confusing for the majority of the Chinese users, especially for elderly people (Peng, 2009, November 29). As the user base in China expands, more people who do not speak English use the Internet. Only 22.4% of China's 485 million Internet users have a college education (ICANN, 2011, October 3). Despite the usability drawback associated with typing in pinyin, newspaper reports repeatedly emphasize that Chinese IDNs offer better branding for Chinese companies and eliminate confusion and ambiguity among Chinese users (Yan, 2010, July 14; Peng, 2009, November 29; Wanxian, 2008, December 29).

Thorough explanation of the DNS is rare in newspaper reports, and newspapers tend to equate domain names to IP addresses by leaving out the possibility of a presence on the Internet without domain names. The economic growth in China is contrasted to the receding global economy which incentivizes many Chinese and multinational companies to register their Chinese character domain names (Wanxian, 2008, November 29). Registrars and registries in China promoted Chinese IDNs by reducing the price to CNY 69 (around 11 US dollars) per year from CNY 139. By the end of 2011, there were around 320,000 domains registered under of 中国/中國 (ICANN, 2011, October 3). By separating the economic and policy significances of Chinese character domain names from a thorough explanation of the technological aspects of DNS, newspaper reports in China portray IDN debates as narratives of economic prosperity. Reflected in newspaper reports is the

Chinese government's policy, which imposes extensive content regulation on the Internet and employs an economic prosperity rhetoric when discussing Chinese script domain names. The delegation of IDN ccTLDs to a government authority like CNNIC is not a step toward the privatization direction ICANN is heading towards outlined in the Affirmation of Commitments document. The delegation of the management of IDNs to a state authority like CNNIC by the increasingly privatized ICANN is indicative of the prevalence of the partnership between ICANN and powerful states that leaves less room for civil society and other Internet communities to participate in Internet governance.

3.3. Russian ccTLD: nationalism and security

The launch of Cyrillic script IDN TLDs was spearheaded by the Russian government, which emphasizes the historical, cultural and economic significance of full international domain names in Cyrillic. Russian president Medvedev asserted in 2009 that Cyrillic domain names were “an early priority of his administration” and “a symbol of the importance of the Russian language and Cyrillic” (Antonova, 2009, November 3). Many technology experts in Russia then believed that Russia did not need a Cyrillic domain considering the high literacy rate, and the fact that Latin letters have long been used (“Head of Russian Internet Coordinating Centre views,” 2010, March 21). Compared to Arabic and Chinese, user confusion and ambiguity among Cyrillic users is manageable due to Cyrillic's similarity to Latin in that it uses letters in upper and lower cases and has the same left-to-right directionality and word separations (ICANN, 2012, February 20).

The Cyrillic script is a written alphabetic script for over 60 languages spoken by people in Eastern and Central Europe and Central and Northern Asia (ICANN, 2011, October 6). Historically derived from the Greek alphabet, the Cyrillic alphabet script has been reformed into its modern use by Peter I and the Russian Revolution of 1917. The modern repertoire of the script includes letters from languages of both the Turkic and Perso-Arabic origin (both languages are unrelated to Slavic languages) spoken in Central Asian countries and Mongolia (ICANN, 2012, February 20). Cyrillic script characters occupy code points ranging from U+0400 to U+04FF in the Unicode 6.2 standard, and not all Cyrillic alphabets in the extended Cyrillic can be used in domain names. There is no script-wide variant in Cyrillic; however, variants issues arise mostly at the level of language. Since the root cannot use language-sensitive rules in Cyrillic, domain names need to share aggregate defined variant rules (ICANN, 2012, February 20). Furthermore, Cyrillic shares many visually similar glyphs with the Roman, Greek, and the Perso-Arabic alphabets invoking security concerns surrounding spoofing, impersonation, and homograph attacks. That is why ICANN strongly cautions against mixed characters and confusable collisions due to visual similarities (e.g. the word “pear” in English and “pear” in Cyrillic are visually similar) in IDNs (ICANN, 2011, October 6; ICANN, 2012, February 20).

Russian representatives at ICANN meetings have expressed strong security concerns over the US control of the DNS root, and claimed that the new Domain Name System Security Extensions (DNSSEC) technology, which is verified by US-based ICANN, impinges upon national security (Mueller, 2010, pp. 237–239). DNSSEC is an effort to increase security of Internet transactions by authenticating DNS queries and responses, and by verifying root zone resource records authoritatively. Even though the root records are multiplied, and cached by other servers on the Internet, only the authoritative DNS root files are used by most Internet servers. Russia's root security concern, similar to the steps taken by the Chinese government at the beginning of IDNs, is more political rhetoric than a technological concern (see Froomkin, 2011, 2013). Yet, this rhetoric helped Russia to expedite the delegation of Cyrillic IDN ccTLD.

In 2009 the .рф suffix, the Cyrillic version of .ru ccTLD was officially delegated to Russia through the fast track process of ICANN and public registration of new .рф domain names began in 2010. Similar to China, the National Internet Domain Coordination Center controls .рф domain names, and the center enforces content regulation. Initially, registration within the .рф domain was limited to government bodies, major cities and trademark owners, and certain words belonging to top officials and Government members, for example, Medvedev.rf, and Putin.rf were reserved (Antonova, 2009, November 3; Isaykov, 2009, November 29). The center, like Chinese CNNIC, exercises content regulation and developed a list of around 4023 forbidden domain names that contain “words that go against public interest, principles of humanity and morals” (“Russia's Popular Radio Website,” 2012, February 24).

Despite this effort to control Cyrillic domain name registration, Interfax already reported on a domain name abuse case in which six accredited registrars have been accused of conspiring to acquire potentially valuable website names to resell later at auction (“MPs complain”, 2011, February 7). Factors such as the value of domain names, the market size, and the regulatory environment in Russia contribute to the attractiveness of domain names to speculators (EURid-UNESCO, 2012). Security concerns due to the similarity of Cyrillic glyphs to other script glyphs and the regulatory environments by the authoritative government are also factors observed in the deployment of the Cyrillic IDN ccTLD in Russia.

Similar to the Chinese case, the Russian government has initiated and taken control of the .рф domain mostly for political and cultural reasons even though the technical community did not see a pressing need for Cyrillic domain names. The media and ICANN discourse is dominated by security concerns, speculation issues and content control surrounding Cyrillic domain names. Transparency in issuing Cyrillic domain names and the participation in Internet governance by civil society or linguistic minority groups are almost completely omitted in both ICANN's and Russian media discourse surrounding IDNs.

4. Conclusion

The incorporation of non-English characters into domain names at the top level is an unprecedented undertaking by ICANN and the Internet community, and much has been achieved in initiating Arabic, Chinese, and Cyrillic scripts in

Unicode, IDNA2008 standards and top-level domains, as well as in delegating IDN ccTLDs to regional organizations. Yet, the complexity of the initiation process has surfaced several inherent tensions in Internet governance that call for further policy finessing.

The tension between English uniformity at the DNS root level and the language diversity of the global Internet community beyond the root level has resulted in technological solutions relating to punycode, IDNA standards, and variant tables in different languages at the zone levels in Arabic, Chinese, and Cyrillic language domain names. These standards and technological solutions ensure the security and stability of the Internet; however they do not comprehensively meet the linguistic diversity needs of the Internet. ICANN has been transforming into an international policy organization; yet, as an organization with primarily a technical expertise, its linguistic diversity policy should evolve beyond technological solutions. Despite initial efforts between UNESCO and ICANN to coalesce the linguistic diversity polities (see [EURid-UNESCO, 2012](#)), ICANN policies appear disconnected from the diversity policies of the United Nations, and remain technically oriented. The first step toward a solution would be for ICANN to start communicating its technical, and stability concerns to the global Internet community in regional media without overwhelming the language communities with technical terms.

Linguistic diversity in relation to IDN at this stage mostly focus on the language representation of major languages that are spoken in powerful nation-states, which use the rhetoric of national pride, local business branding, and inclusion of non-English speakers. While different goals were emphasized in three case studies – Arab identity in case of Arabic script IDNs, business branding and business opportunities in the case of Chinese IDN, and cultural pride in the Russian case – the state-centric linguistic unity policies were implemented by the states in all three cases. This situation surfaces the tension between nation-states and the new international governing institution ICANN. It is a bit paradoxical that while ICANN's independence from the US government has solidified since the Affirmation of Commitments document ([Froomkin, 2013](#)), the power of registering and administering IDNs at the local language level is increasingly being consolidated by the states. State-centric international telecommunications regimes, which in the past gave full control of radio frequencies and satellite orbital slots to states, now enforce restrictive communication policy and language unity rather than language diversity in their territories. Chinese, Russian and United Arab Emirate governments spearheaded the initiation of Arabic, Chinese and Cyrillic script IDN ccTLDs, and have supported technical groups; however, the ways to include minority languages within powerful states have not yet been addressed by the states. The imbalance between authoritative states and non-state stakeholders such as civil society groups or other linguistic interest groups within Arabic, Chinese, and Russian language communities should be taken into account when delegating new IDNs to states by ICANN.

The situation in which non-US governments fully control internationalized domain names, while ICANN is given permissions by international governments to liberalize domain names, suggests the increasing need for an accountability-enforcing mechanism for both ICANN and nation-states. Even though ICANN is being internationalized to a greater extent by involving non-US governments, this alliance between the states and the more privatized ICANN is not inclusive of the non-state and non-commercial interest groups within the ICANN regime. There is an increasing concern in the wake of full liberalization of top-level domain names that ICANN and powerful states like China, Russia, and the United Arab Emirates benefit from IDNs while providing few benefits to non-state and non-commercial participants on the Internet. More policy research is needed to explore ways to increase transparency and accountability in allocating, taxing and delegating of IDNs. Involving civil society organizations, especially emerging forms of global civil society – grassroots activist tech groups and networked cyberactivists in the process of deployment of new multilingual domain names is crucial as well. In the hindsight of the WikiLeaks case and the revelation of National Security Administration surveillance, networked collective actions of grassroots tech groups and cyberactivists are increasingly acknowledged by scholars as an emerging form of global civil society in transnational policy making ([Hintz & Milan, 2009, 2013](#)). These groups provide technical expertise, create alternative infrastructures, and resist online corporate and government surveillance ([Coleman, 2013; Hintz & Milan, 2013; Mueller, 2010](#)). Since Internet domain name governance intertwines policy making with technical expertise at the international level, the voices of networked activist groups at the transnational level should necessarily be counted for more transparent Internet governance in order to balance corporate and state powers.

References

- Antanova, M. (2009, November 3). Medvedev gets wish for .pp domain. *St. Petersburg Times*. Retrieved 20.04.12 from the Access World News Database.
- Atta Allah making a difference on a global scale with ICANN (2010, October 9). *Gulf news*. Retrieved 20.04.12 from the Access World News Database.
- Baasanjav, U. (2012). Global digital divide: Language gap and post-communism in Mongolia. In A. Manoharan, & M. Holzer (Eds.), *E-Governance and civic engagement: Factors and determinants of e-democracy* (pp. 210–234). Hershey, PA: IGI Global.
- Braman, S. (2012). Internationalization of the Internet by design: The first decade. *Global Media and Communication*, 8(1), 27–45.
- But, J. (2009, December 14). Websites to get characters for .hk. *South Morning China Post*. Retrieved 20.04.12 from the Access World News Database.
- Chia, T. (2009, December 14). China bars individuals from registering .cn domain Internet websites. *Kyodo News International*. Retrieved 20.04.12 from the Access World News Database.
- Coleman, G. (2013, October). Weapons of the geek and anonymous: Sitting at the nextus between expertise and participation. *Paper presented at the annual conference of the association of internet researchers*. Denver, CO.
- Cooper, E. (2008). A guide to internationalized top-level domains. *Managing Intellectual Property*, 178, 30–34.
- DiMaggio, P., Hargittai, E., Celeste, C., & Shafer, S. (2004). From unequal access to differentiated use: A literature review and agenda for research on digital inequality. In K. Neckerman (Ed.), *Social inequality* (pp. 355–400). New York: Russell Sage Foundation.
- Easton, C. R. (2012). ICANN's core principles and the expansion of generic top-level domain names. *International Journal of Law and Information Technology*, 20(4), 273–290.

- EURid-UNESCO. (2012). *World report on internationalized domain name deployment 2012*. (www.eurid.eu/files/publ/insights_2012_idnreport.pdf) Retrieved 20.04.13.
- Froomkin, A. M. (2011). Almost free: An analysis of ICANN's 'Affirmation of Commitments'. *Journal on Telecommunications & High Tech*, 9, 187–233.
- Froomkin, A. M. (2013). ICANN and domain name system after the 'Affirmation of Commitments'. In I. Brown (Ed.), *Research handbook on governance of the Internet* (pp. 27–51). Cheltenham, UK: Edward Elgar.
- Hamelink, C. J. (2000). *The ethics of cyberspace*. London: Sage.
- Hargittai, E. (2003). The digital divide and what to do about it. In D. C. Jones (Ed.), *The new economy handbook* (pp. 822–841). San Diego, CA: Academic Press.
- Head of Russian Internet Coordinating Centre views launch of new Cyrillic domain. (2010, March 21). *Rossiyskaya Gazeta*. Retrieved 20.04.12 from the Access World News Database.
- Hintz, A., & Milan, S. (2009). At the margins of Internet governance: Grassroots tech groups and communication policy. *International Journal of Media and Cultural Politics*, 5(1&2), 23–38.
- Hintz, A., & Milan, S. (2013). Networked collective action and the institutionalized cyberactivism to the policy arena. *Policy & Internet*, 5(1), 7–26.
- Hobsbawm, E. J. (1996). Language, culture, and national identity. *Social Research*, 63, 1065–1081.
- Hrynshyn, D. (2008). Globalization, nationality and commodification: The politics of the social construction of the Internet. *New Media and Society*, 10(5), 751–770.
- ICANN. (2009, June 4). *Cost analysis of IDN ccTLDs focus on program development and processing costs*. (<http://www.icann.org/en/resources/idn/fast-track/analysis-idn-ccTLD-development-processing-costs-04jun09-en.pdf>) Retrieved 01.04.13.
- ICANN. (2009, September 30). *Affirmation of commitments by the United States Department of Commerce and the Internet Corporation for assigned names and numbers*. Retrieved from (<http://www.icann.org/en/about/agreements/aoc/affirmation-of-commitments-30sep09-en.htm>).
- ICANN. (2011, October 3). *Report on Chinese variants in internationalized top-level domains*. Retrieved from (<http://archive.icann.org/en/topics/new-gtlds/chinese-vip-issues-report-03oct11-en.pdf>).
- ICANN. (2011, October 6). *IDN variant TLDs – Cyrillic script issues*. Retrieved from (<http://archive.icann.org/en/topics/new-gtlds/cyrillic-vip-issues-report-06oct11-en.pdf>).
- ICANN. (2011, October 7). *Internationalized domain names variant issues project: Arabic case study team issues report*. Retrieved from (<http://archive.icann.org/en/topics/new-gtlds/arabic-vip-issues-report-07oct11-en.pdf>).
- ICANN. (2012, February 20). *A Study of issues related to the management of IDN variant TLDs (Integrated issues report)*. Retrieved from (<http://www.icann.org/en/resources/idn/idn-vip-integrated-issues-final-clean-20feb12-en.pdf>).
- ICANN. (2012, June 4). *Final implementation plan for IDN ccTLD fast track process*. Retrieved from (<http://www.icann.org/en/resources/idn/fast-track>).
- ICANN (2013). *Examining the user experience implications of active variant TLDs*. Retrieved from (<http://costarica43.icann.org/meetings/sanjose2012/presentation-vip-13mar12-en.pdf>).
- International Herald Tribune-Asahi Shimbun. Concern for the safety of Chinese users was behind the decision of the world's largest provider of Internet domains (2010, March 27). *Asahi Shimbun*. Retrieved 20.04.12 from the Access World News Database.
- Isaykov, V. (2009, November 29). Medvedev.RF to be first. *Rossiyskaya Gazeta*. Retrieved 20.04.12 from the Access World News Database.
- Kaur, R., & Aggarwal, R. (2011). Cyber-squatting: Legal implications and judicial approaches: An Indian perspective. *Computer Law & Security Review*, 27, 653–658.
- Kleinwachter, W. (2000). ICANN between technical mandate and political challenges. *Telecommunications Policy*, 24, 553–563.
- Leng, T. K. (2010). Internet domain name disputes: From the Singapore perspective. *Computer Law and Security Review*, 26, 210–218.
- Mueller, M. L. (2002). *Ruling the root: Internet governance and the taming of cyberspace*. Cambridge, MA: MIT Press.
- Mueller, M. L., & McKnight, L. W. (2004). The post-COM internet: Toward regular and objective procedures for internet governance. *Telecommunications Policy*, 28, 487–502.
- Mueller, M. L. (2010). *Networks and states*. Cambridge, MA: MIT Press.
- MPs complain of violations in registration of Russian internet names. (2011, February 7). *Interfax*. Retrieved 20.04.12 from the Access World News Database.
- Napoli, P. M. (2011). Diminished, enduring, and emergent diversity policy concerns in an evolving media environment. *International Journal of Communication*, 5, 1182–1196.
- O'Siochru, S., Girard, B., & Mahan, A. (2002). *Global media governance: A beginner's guide*. New York: Rowman and Little.
- Peng, N. (2009, November 29). Chinese-character domain names to benefit online community. *Business Daily Update*. Retrieved 20.04.12 from the Access World News Database.
- Ramachandran, T. (2011, June 27). Package for developing nations in top-level domain expansion. *The Hindu*. Retrieved 20.04.12 from the Access World News Database.
- Russia's Popular Radio Website may change address over registrar's new rules. (2012, February 24). *Izvestiya*. Retrieved 20.04.12 from the Access World News Database.
- Sambidge, A. (2009, October 20). TRA set to launch Arabic domain name for UAE. *ArabianBusiness.com*. Retrieved 20.04.12, from the Access World News Database.
- Sithigh, D. M. (2010). More than words: The introduction of internationalised domain names and the reform of generic top-level domains at ICANN. *International Journal of Law and Information Technology*, 18(3), 274–300.
- TRA Signs an Exchange of Letter with ICANN during GITEX. (2011, October 11). *Asia Pulse*. Retrieved 20.04.12 from the Access World News Database.
- United Nations. (1948). *Universal declaration of human rights*. Retrieved 28.01.13 from (<http://www.un.org/en/documents/udhr/index.shtml>).
- Verisign (2012, July). *The domain name industry brief*. Retrieved 01.03.13 from (<http://www.verisigninc.com/assets/domain-name-brief-july2012.pdf>).
- Waisbord S. and Morris N., (Eds.), *Media and globalization: Why the state matters*, 2001, Rowman & Littlefield; Lanham, MD.
- Wanxian, B. (2008, December 29). Domain name reform to open up Internet wider. *Financial Times*. Retrieved 20.04.12 from the Access World News Database.
- Warschauer, M. (2003). *Technology and social inclusion*. Cambridge, MA: MIT Press.
- Watchdog predicts rise in 'cyber-squatting'. (2009, October 13). *Khaleej Times*. Retrieved 20.04.12 from the Access World News Database.
- Web turns to Arabic. (2010, May 8). *Gulf News*. Retrieved 20.04.12 from the Access World News Database.
- Wilkinson, K. T. (2004). Language difference and communication policy in the information age. *The Information Society*, 20, 217–229.
- Xue, H. (2004). The voice of China: A story of Chinese-character domain names. *Cardozo Journal of International and Comparative Law*, 12, 559–592.
- Yan, H. (2010, July 14). Chinese gains int'l recognition on Web. *Business Daily Updates*. Retrieved 20.04.12, from the Access World News Database.
- Zainol, Z. A., Hussain, W. M. H. W., & Yaakub, N. I. (2011). WIPO Panels' interpretation of the Uniform Dispute Resolution Policy (UDRP) three-prong test. *World Patent Information*, 33(3), 275–281.
- Zook, M. (2000). Internet metrics: Using host and domain counts to map the Internet. *Telecommunications Policy*, 24(6), 613–662.