Establishing a Common Identity in New Niche gTLDs

Curtis M. Kularski

Univeristy of North Carolina at Charlotte

Department of Sociology

Department of Philosophy

12 May 2016

ESTABLISHING A COMMON IDENTITY IN NEW NICHE GTLDS

"It is extremely unlikely that any other TLDs will be created" stated Dr. Jonathan Postel, one of the original architects of the Domain Naming Service (DNS) protocols and policies, in RFC 1591, a document reviewing the status of the DNS system 10 years after its creation (Postel J. B., 1994). Originally, it was feared that the Internet could not handle more than the several top-level domains that had been created because the entire DNS system was operated from thirteen physical servers, each of which had specifications significantly less than that of the original iPhone. Due to the creation of multicasting, wherein multiple physical servers share an IP address and appear to the Internet to be a single server, and other technological improvements those limitations were quickly overcome. In 1998 Postel suggested opening the root zone to allow for many new top-level domains to be created. Postel, who had been a guiding force in the guidance and policy creation of DNS[1], died unexpectedly in October 1998 before having the opportunity to provide his guidance as to how such an expansion could be performed (Weinberg, 2002).

History and the DNS System

The DNS System began as an experimental technology to evaluate the possibility of a centralized naming system for computers on the Internet (Harrenstien, White, & Feinler, 1982).

[1] Dr. Postel was central to resolving policy disputes and ensuring that facets of the Internet's political and technical structures were stable. Until his death, Postel, as an individual, held the title of Internet Assigned Numbers Authority (IANA) (Cerf, 1998). Postel's death had detrimental effects on the evolution and progress of the DNS system.

DNS has become central to the technical operation of the Internet, but has also entered into the social sphere in the way of identity creation and identity maintenance. Initially the identity space was confined to several generic top-level domains (gTLDs), such as COM, NET, ORG, EDU, GOV and MIL, that were very broadly interpreted and not restricted by any formal process. As such, the identities associated with those top-level domains are very weak and there is no more than incidental commonality between the various organizations and individuals (registrants) who register under those domains. In 2000 and 2003 the Internet Corporation for Assigned Names and Numbers (ICANN) experimented with a minor expansion of the top-level domains by adding seven more domains, such as INFO, BIZ, AERO and PRO, which were also broadly defined. Many critics consider the 2000 and 2003 experiments to have been failures (Halvorson, et al., 2012). In 2011 ICANN started a new process for managing applications for new top-level domains. The new process was not given a specific upper-limit on the number of top-level domains that would be allowed, but ICANN projected up to 1300 new top-level domains could be created (Internet Corporation for Assigned Names and Numbers, 2015). In the 2000 and 2003 batches applicants were competing for a limited number of generic top-level domain spaces, which would need to appeal broadly to the Internet community as a whole. The 2011 process being more open encouraged more specific top-level domain strings. The strings applied for in the first batch are still somewhat broad to appeal to enough registrants to keep the domains financially viable, but cover specific niche markets, specific interest areas and names of geographic places. Do these less general top-level domains encourage more clearly defined Internet communities?

This topic is of social interest as the addition of different top-level domains dramatically changes the nature of the creation of an identity on the Internet for organizations and individuals.

COM became the socialized norm for top-level domain in the 1990s, despite the availability of additional top-level choices. Through this normalization the COM top-level domain lost its meaning as a commercial zone to a wide range of uses. In the case of COM, the term "generic top-level domain" became too accurate. It in essence became a generic term to identify the Internet itself, or at least identify the proliferation of the popularized Internet (Halvorson, et al., 2012). With the addition of a larger base of top-level domains to choose from the next generation of businesses, social networks and individuals wishing to establish a web presence have the opportunity to establish a different type of Internet identity than was previously possible due to limitations and assumptions applied to the naming structures.

A common critique each time new top-level domains are added is that it is expanding territory for intellectual property disputes, but is otherwise not adding anything to the Internet (Hering, 2004; Halvorson, et al., 2012). Despite those critiques, the new top-level domains are different in process of creation and necessity of broad scope relative to the legacy top-level domains. The 2011 new gTLD process makes a fundamental change to the concept of what a TLD is or should be when registering a domain name (at the second level). Moving from only about 20 active top-level domains to several hundred disrupts the centrality of COM as the de facto TLD. The addition of only a few new TLDs in the first two rounds did not make a cultural change because the new gTLDs were not visible enough, due to their small numbers. The present round, adding a large number of new gTLDs makes the change in policy more visible and therefore has the potential to change the cultural interpretation of TLDs other than COM, and for that matter change the interpretation of COM as well. COM could eventually be seen as old and be pushed from its cultural role of being synonymous with the Internet (Goodnight & Green, 2010).

In this paper I wish to investigate the variables associated with the use of top-level domains in the establishment of an Internet identity, whether organizational or personal. As a component of this investigation I am also interested in the establishment of community or otherwise collective identities that may emerge as a result of establishing identities in the same top-level domain. A key consideration for the interaction between the areas of personal/organizational and community/collective identity is the problem of lack of shared meaning. As each entity approaches identity creation using pre-defined textual strings, there is likely to be differences in interpretation in the meaning and purpose of each top-level domain. The difference in interpretation could function counter to the process of a community forming or there being a shared collective identity under each top-level domain.

Literature Review

Literature addressing this specific topic is sparse as the expansion of the top-level domain structure is new (Jarassriwilai, Dauber, Brownlee, & Mahanti, 2015). Most existing literature in this area focuses on the intellectual property and marketing implications of the expansion (Prahl & Null, 2011; Brown, 2012; Alramahi, 2008) and gives no consideration for the social or cultural implications of such an expansion, especially with regard to the changing range of specificity or narrowing of breadth of the top-level domains available. As there is limited research specific to this topic it will be necessary to approach the topic using more theoretical literature related to identity formation and use of the Internet as a communication medium. Primarily this literature is drawn from the social areas of identity and community.

Previous Top-Level Domain Expansion

The 2013 round of top-level domain expansion is the third batch of new top-level domains added by ICANN. The first two expansions were limited to seven top-level domains each (Internet Corporation for Assigned Names and Numbers, 2015). One of the top-level domains added in the 2000 round was BIZ. BIZ was intended to be a compliment to COM due to the lack of available names in the legacy generic TLDs. Many registrants in the COM TLD registered their second-level under the BIZ TLD as well and either redirected traffic to their primary website under COM or otherwise had the domain parked[2]. According to research conducted by Halvorson, et. al. in 2012 up to 23.4% of registered domains in BIZ were parked or not resolving[3], another 17.3% redirect to a COM website. Only 27.7% of domains were found to be completely distinct from their COM counterpart. The remaining domains were found to either serve duplicate content or were owned by the same entities as the COM variant for a different functional purpose (Halvorson, et al., 2012). Duplication of content and defensive registrations indicate that BIZ is not successful as a new namespace because it did not attract a majority of distinct registrants. Halvorson, et. al. also utilized Alexa Internet, an Internet analytics system, as an indicator of TLD success. For each BIZ domain found in the Alexa directory there were 140 COM domains, as of 10 years after COM launched. While COM had several years to become established before BIZ, it should be noted that the annual renewal process for domain names requires a continued effort to maintain an active domain and therefore the registrants must remain engaged (Halvorson, et al., 2012).

[2] Parking is defined by Halvorson, et. al. as being a single-page website designed to sell the domain (Halvorson, et al., 2012)

[3] Resolution refers to the ability of a domain name to be matched to a functioning server

Social Functioning of Domain Names

Identity is embedded in communication and communication is the objective of the Internet. Through the DNS system of the Internet there is a clear area where medium and message intersect. DNS domain names function as addresses, essentially providing a mechanism to reach the intended content, but distinct from postal addresses that are arbitrary or systematic DNS names are partially self-selected. The entity registering the domain name is able to select what string appears to the left of the top-level domain in their domain name and can select the top-level domain in which to register (Weinberg, 2002).

Domain names in DNS are treated by their registrants as assets, essentially the identity converted to a commodity. This is the source of the various debates of intellectual property rights in the new gTLD process. Intellectual property rights holders became concerned that their rights may become diluted in a system where there are several thousand possible TLDs. Concerns arise over the possibility of confusion or of one entity being advantaged over another with regard to reaching their audience due to aspects of the new top-level domain process (Prahl & Null, 2011). These objections to the creation of new top-level domains illustrate the value of being able to strategically establish an internet identity. This also highlights concerns over legitimacy. Through a variety of legal processes and the benefit of extensive financial resources large organizations and celebrities were able to secure their brand, trademark or name as a second-level domain in the COM top-level domain (Halvorson, et al., 2012). Due to the perceived legitimacy and cultural acceptance of COM the named identity was perceived as being authentic. Additional top-level domains disrupt that assumption and allow less powerful entities to establish

their own identity, which may be composed of a string that conflicts[4] with a string already registered in COM.

The meaning conveyed by a top-level domain itself is a critical issue surrounding how online identities are constructed. COM has prestige due to its widespread use and the assumptions noted previously (Weinberg, 2002). Some other top-level domains that were created at the same time as COM have their own power and legitimacy due to the nature of how they were established. GOV, MIL and EDU were established for specialized use and are not able to be registered by the general public (Postel J. B., 1994; Postel & Reynolds, 1984). While knowledge of the specifics of the limitations may not be common knowledge to the average Internet user the cultural knowledge that those top-level domains can be trusted has been generally established such that entities who have domains in those top-level domains are assumed to be who they say they are without further validation of their identity. Whereas COM, NET and ORG were established with very broad missions, the specificity of GOV, MIL and EDU and the governmental limitations applied to them create a specific type of identity in each top-level domain (Postel & Reynolds, 1984). EDU is used exclusively by colleges and universities in the United States as restricted by policy and the majority of qualifying institutions hold domain registrations in the EDU TLD (EDUCAUSE, 2009).

The HEALTH TLD Application – Objections Based on Authority Perceptions

[4] It should be noted that trademarks are only issued for specific goods and service classifications and therefore a trademark holder does not necessarily have monopoly rights to the string/wordmark, despite the urging for more comprehensive protections during the early policy process of the new gTLD program (Internet Corporation for Assigned Names and Numbers, 2010)

The new generic top-level domain program does not have any stipulations for additional limited-use domains such as those found in the legacy TLDs, but it is instead left to the top-level domain operator to develop rules for who can register domains and if any type of validation will occur (Internet Corporation for Assigned Names and Numbers, 2015). The World Health Organization (WHO), which did not apply to manage any top-level domains, objected to several top-level domain applications, including HEALTH. Specifically, WHO objected to HEALTH and other top-level domains that could be associated with the medical professions being operated by for-profit companies (Attaran, Kohler, Liang, & Eysenbach, 2014). In essence the concern of WHO is that Internet users seeking health or medical information will rely upon information found in websites and other resources under the HEALTH top-level domain and assume that it is reliable information from a trustable source. There is not yet empirical data to test how various new top-level domains are perceived in terms of their legitimacy or representation of authenticity. It may be reasonable to assume that as new top-level domains become established their own legitimacy will be established in part by the types of entities that register and use domain names in those top-level domains (Goodnight & Green, 2010).

Theories of Identity as They Apply to Top-Level Domains

Theories of identity formation are essential to analyzing the possible impact that the new top-level domains could have on entities that choose to register within them. Central to my presentation of these theories is Erving Goffman's *Presentation of Self in Everyday Life* (Goffman, 1959). Goffman's book was not written with the Internet in mind, but his concepts are still relevant to the way in which identities are established and interpreted. Goffman

discusses information as a primary determiner of the interpretation of the self. In terms of Internet domain names, the information is the domain name itself, but also other information that is known about other domains in the same top-level domain may be a factor in the way the identity of the entity that owns the domain is perceived. Assuming that the entity uses their domain for a website, the person navigating to that website using the domain name may be generating a concept of the identity of the entity before they see the website due to the information contained in the domain name itself or collected with previous experiences with the top-level domain. It is essential to the continuity of the identity that the domain accurately represents the identity that the entity wishes to portray. In some instances the mission of the top-level domain, as specified by the top-level domain operator in their application, may differ from the interpretation that an end-user associates with the chosen text string. For example, the INK TLD application specifies that the TLD is "to create a designated Internet space for the tattoo industry, and other industries that rely on ink, such as the printing and publishing industries" (King, 2012). The text string "ink" refers to a physical product, but the mission of the TLD involves various service industries. The INK TLD mission also specifies two distinct unrelated industries. There is potential for identity misinterpretation for entities with domain names in the TLD because of this multiple use, but that may be unavoidable in all TLDs due to the overloading of terms in the linguistic lexicon. As stated by Goffman, "the individual's initial projection commits him to what he is proposing to be and requires him to drop all pretenses of being other things" (Goffman, 1959). When applied to the context of the new gTLDs this places a greater burden on registrants than in the past due to the less generic status of the top-level domains. The broadly interpreted nature of the original set of top-level domains allowed for great

flexibility in which only the second-level of a domain name was relevant to identity construction, whereas more specific top-level domains become an integral part of domain name selection.

Goffman did not live in a time where the Internet existed or was a factor of identity creation or identity management, so there are limitations to the extent his theoretical perspective can be useful to understanding how new gTLDs may function in identity. Sherry Turkle considers the implications of a connected society on identity from the perspective of technology becoming an incorporated component of identity rather than simply as a means to express it. Technology, including the Internet, has become incorporated into society. With regard to identity the Internet is a space in which to 'play' with identity and to experiment with expression (Turkle, 2011). However, as danah boyd points out, there are limitations to the experimentation because of expectations of a correlated existence between the Internet and real life (boyd, 2012). Identity is primarily interpreted as an individual attribute. In the new gTLDs individuals have more options for how to present themselves, whether they choose to have a single domain name that represents their primary interest or occupation or if they choose to have domains in multiple TLDs that reflect multiple facets of their life. While it has always been possible for an individual to register multiple domain names, the presence of the COM standard did not afford the capability of maintaining different facets of a consistent identity. Also, scarcity of COM domain names could have presented a challenge for experimentation or multiple presentation.

Nancy Baym discusses the concept of community as it applies to individual websites designed to foster communication and relationship building (Baym, 2010, p. 74). There is no concrete sociological definition for community, despite its usefulness as a construct from which to theorize about group behavior (Baym, 2010). For the purposes of this paper I will utilize a wide definition of community, which essentially means that I am interested in group associations

that share a common value or common elements of identity. For individual identity this make take the form of shared hobbies, demographic similarities or similarity of preferences. From the organizational perspective one of these elements is almost by necessity industry affiliation, but also includes other attributes related to the individual identities of the organization's audience. Individuals and organizations may belong to the same community.

Baym critiques describing social media services such as YouTube and Twitter as communities as being inaccurate. Twitter is a service that can be used for community building, but is composed of many distinct communities built on insider understanding and groups of connections (Baym, 2010). The scope of applying the term community to new gTLDs is considerably wider than the case of social media, but the same types of limitations to the terminology apply. Each of the new gTLDs is a different text string with a different meaning, so inter-TLD community is not an expectation, beyond the shared meaning in the short-lived novelty of being an early adopter, which is no more significant than the shared meaning derived from being restricted to 140 characters inside Twitter. Community is more likely to be found inside individual TLDs where individuals and organizations choose to identify themselves inside the scope of the same text string.

How is that the addition of a large number of new gTLDs can encourage or change identity and community more than the legacy set or the intervening additions? Identity work and community building are not processes that most humans embark on with great intentionality (Baym, 2010, pp. 51, 84). As such, technological affordances are an essential part of these social shaping processes. Affordances are essentially elements of technology that make them intuitive (Baym, 2010, p. 51). Humans build identity and community inside technological systems by metaphor. The COM metaphor was simple, you could locate any real world entity's website by

adding COM to the end of their common name. Expanding the TLD space requires that the new gTLDs be as respectful to the human need for metaphor. In adding new strings that have intrinsic meanings the intuitive processes of identity and community are re-established, but it does require dismantling the COM metaphor to overcome initial resistance to this change in the Internet's naming system. As Baym points out from her research, the most successful evolutions in Internet technology are those that users lead themselves (Baym, 2010, p. 52). While the new gTLDs were not lead by the users, many were created by organizations that represented groups of users and their styles of interacting.

Internationalized Top-Level Domains

One feature of the new generic top-level domains process that will further stratify the named Internet is the addition of non-English domain names. Non-English domain names have been supported by and allowed in DNS since November 2009, but there was no immediate action to add top-level domains utilizing non-English characters. When the new generic top-level domain process was opened in 2011 it was open to any technically valid string, including the extended international character sets. All recognized sovereign nations were previously assigned their own nation-based two-character (English characters) based on their International Standards Organization country-codes (Postel J. B., 1994). For example, China has been assigned CN, Japan is assigned JP, United Kingdom is assigned UK and Germany is assigned DE. These county code top-level domains (ccTLDs) were simple allocations based on a perceived need for each country to have its own Internet naming space. This is a purely logistical and political consideration and does not allow for any cultural expression (Baasanjav, 2014). English has

become somewhat standard for academic, scientific and some commercial communication, but does not replace the local languages for daily use. The assumption of English as a communications standard has limited access to the Internet for some groups. Extending the available character sets to include non-Latin characters has increased global inclusion of the Internet and perhaps will allow additional groups to utilize it as a communication and information medium. Non-English Internet users will be able to access Internet resources using addresses in their native language (Baasanjav, 2014). Much of identity is established through a common sense of meaning. Whereas historically DNS has excluded everyone who did not use the accepted standard character set, those who were previously excluded now have the capability to establish Internet communities in their own language.

Applicant Definition of Identity and Community

ICANN's expansion of the generic top-level domains has been greatly critiqued on numerous aspects, most notably in its respect or lack of respect for intellectual property rights holders and in its choice to handle all new gTLD applications fairly, regardless of the string that that applicant has requested. Both of these factors can affect the identity and communication value of the new generic top-level domains, but do not pose a direct threat to the changing of the cultural conception of domain names. Domain names function as identities and with additional options for top-level domains, those identities have the capability to be more clearly defined. Also with the further refinement of identity it may also be possible for a common identity between domains under a single top-level domain to be established. Each top-level domain application contains a section that describes the purpose and value of the new top-level domain

as perceived by the applicant (Internet Corporation for Assigned Names and Numbers, 2015). There is no certainty that the usage by the registrants will match what is described by the applicants, or would it be expected in the evolution of an organic self-selecting social organization.

Methods

A central overarching question to guide my research into identity and community in the new gTLDs is: Are the new gTLDs making a difference in the way that individuals and organizations function in their online representations? The specific question will not be answered by the research, but it will guide its structure. This research proceeds with a fundamental assumption that the top-level domains are most beneficial if they function as categorical labels that contain meaning. Community formation is one mechanism by which the meaning, beyond the linguistic interpretation of the string, can be imparted upon new gTLDs.

The immediate questions of interest to this research are: 1. Are informal communities forming under top-level domains that are related to the top-level domain string and its registered statement of purpose? And 2. Do the presented identities of entities who register in the new gTLDs resemble have common themes?

Scope

The overall frame for this research is all new gTLDs that have been delegated and are operational on the Internet and are registerable by entities other than the TLD applicant. For

logistical reasons only a small subset of gTLDs will be utilized. Specifically, I have restricted my sampling frame to TLDs that were delegated in the first 6 months of active delegation in the 2011 set, from October 2014 until April 2014. Inside that frame I reviewed the statement of purpose for each TLD meeting the previously stated criteria and performed an intentional sample. Statement of purpose was obtained from Question 18A on the new gTLD application submitted to ICANN. I selected to review the following TLDs: CLUB, PHOTO and XYZ. My selections were directed by specific objectives I wished to accomplish through evaluating them.

CLUB was selected because it represents a TLD with potential wide appeal and has various lexical uses, but has at least one identifiable use-case. The mission for CLUB is to establish a TLD namespace for

> "an association of persons for some common object, usually jointly supported and meeting periodically; also : a group identified by some common characteristic b: the meeting place of a club c : an association of persons participating in a plan by which they agree to make regular payments or purchases in order to secure some advantage d : a nightclub e : an athletic association or team." (LaMantia, 2012).

The application lays out four distinct uses of the term 'club' and as such specifies four distinct styles of CLUB registrants.

PHOTO was selected because it is a TLD with a clear niche. The mission for PHOTO is described as

> "An identifiable style of work is key to a professional or even amateur photographer. It defines them and draws attention and support from individuals who admire or relate to their work. The Internet has made it possible for photographers to instantly display their images to a larger crowd of potential clients and followers. As Internet use continues to rise and even more Internet users share photos, we believe .PHOTO will serve as an optimal domain space with which professionals can identify their work and individuals can upload and share photographs amongst family, friends, and a wider Internet audience." (Schilling, 2012)

In the mission for PHOTO there are no variations in the anticipated registrants. It is expected to be a unified TLD representing only photographers. The expected registrants would share at least the commonality for an interest in a particular creative or documenting medium which implies a shared value of the visual image.

XYZ was selected because it does not align to a text string with a particular lexical meaning. The XYZ TLD application specifies the mission as:

> "To provide business owners across multiple verticals a cost-effective means of establishing a viable online presence, allowing these entities to register names that are appropriate to their business name or service offering. The TLD will provide simple, industry all-inclusive, memorable domain names for every business type. It will serve as a designated name space where these entities can register domain names spanning across virtually all industries, and in a format that is easily recalled, recognizable and located by consumers." (Negari, 2012).

The domain's mission states specifically that it wishes to cross all industries. The TLD assigns its mission as being a commercial TLD similar to the COM legacy TLD or the BIZ TLD from the 2000 TLD expansion. This TLD intentionally distances itself from any concept of a unifying identity in favor of being an open alternative.

Cases

Due to the abbreviated nature of this research a simplified selection process for individual domain names was utilized. The 'site:' control in Google's Advanced Search functionality was utilized to select top-scoring webpages[5] in each of the candidate gTLDs. The searches were

[5] Webpages (user-interpretable content delivered over the HTTP or HTTPS protocols) were used as a proxy for domain names as a convenience.

performed on three distinct Internet connections[6] utilizing a clean private browsing session[7] for each search. Google does not publish the criteria or algorithms used for their search system, so the exact nature of the sample can only be estimated from the results displayed consistently. For the purpose of this simplistic study it will be assumed that Google applies some weight to search results based on the number of clicks received, thus results that were beneficial or interesting to other Google users in the past.

Search results had to appear in the top 10 eligible results on at least 2 of the 3 searches to be included. Non-English search results were excluded for logistical reasons and results that had duplicates in the COM TLD were excluded to avoid contaminating data with results not related to the mission of this research.

Data Form

Websites that were selected for analysis were viewed and open-coded with relation to the type of entity that they represent. *See Appendix I.*

Preliminary Findings

The CLUB TLD reflected almost exactly the types of entities that the application had predicted. In the sample I performed I coded three of the ten results as 'gaming' clubs based on their self-description, but these sample items could potentially be split into the sports club

[6] 1. Personal Time Warner Cable broadband subscription, 2. Fiber connection over NCREN provided by UNC Charlotte, 3. EDUROAM access provided by UNC Charlotte.
[7] A private browsing session is a browser session that has no prior cookie or cache data

category for the online gaming website, whereas the gaming clubs with casino characteristics may be categorized as nightclubs for the purpose of the study. The ten sample websites did not reveal a coherent identity, but did have characteristics similar to the application mission. An interesting finding in the sample was that two of the ten websites were redirected to elaborate Facebook pages. It is possible these entities are utilizing the new gTLD as a marketing tool to establish their own brand of web presence that is not directly reliant upon the Facebook brand as well. There were no individuals represented in the CLUB sample. Many of the websites in the sample listed information for gaining membership and benefits of that membership, showing a clear objective of growing membership and embracing a group identity or expanding an existing community.

The PHOTO TLD is more extremely niche, as expected. Six of the ten sample websites were for individual photographers who have developed a web presence exclusively on the PHOTO TLD. Five of the six are professionals, with the remainder being an amateur. The websites not operated by photographers are operated by stock photo services, with the exception of one that is a gallery curated by a space photograph enthusiast who maintains an archive of images specific to his interest. The photographer websites are mostly established as portfolio and 'business card' websites with a similar arrangement, but there does not seem to be an underlying template that is common between them. This may indicate a type of isomorphism in the photography community which is being embodied in the PHOTO TLD.

The XYZ TLD has an unexpected theme across many of the sample domains. Many of the domains sampled references 'new', 'different' or 'unique' in their descriptions, perhaps an indication of an emerging meaning of the XYZ TLD. Three of the ten sampled websites were

coded as ‘magazine’ websites, containing web-based magazine-style content. These three had an emphasis on fashion and other cultural items.

ABC.XYZ, owned by Alphabet Inc, the parent company of Google is perhaps one of the most prominent owners of a second-level domain in XYZ. Alphabet utilized the new gTLD to express their identity in a creative way. XYZ was the only TLD sampled that had examples of online tools/web apps residing on them rather than simple content pages.

Of all of the samples to be created, XYZ had the most overlapping websites with COM that had to be eliminated from the sample. The relative popularity of XYZ, holding 16.3% of registrations in the new gTLDs (greenSec Solutions, 2016), to some of the other new gTLDs may have sparked an urgency among intellectual property rights holders to claim their mark in the TLD.

The samples indicate that there may be a common identity forming in the PHOTO TLD. XYZ is more concentrated by organizations and individuals with a common interest in innovation than was originally expected when beginning this research. CLUB was more heterogeneous than expected, but is forming various distinct communities.

Formal Research Procedure

The research procedure detailed above is adequate for a preliminary exploration of the themes of community and identity in the new gTLDs. However, lack the scale and distribution of coverage that would be needed to make any definitive claims related to the development of

community in the new gTLDs or the success of the social success of the new gTLD program itself.

In ideal circumstances, I would perform a random selection of 20-50 of the eligible TLDs as specified in the initial scope previously. As Halvorson, et. al. created the only readily available established research standard for evaluating differences between TLDs, replicating their study for each TLD would be ideal. Halvorson, et. al. utilized specific DNS queries, WHOIS data and differential algorithmic page content analysis to determine uniqueness factors for domain name registrations (Halvorson, et al., 2012). A limitation of implementing the Halvorson, et. al. method is that their comparison was designed to be a like-with-like comparison, rather than differentiating between different scales and intentions of TLDs. This limitation primarily affects assumptions that can be made in deciphering the meaning of the quantitative results. While identical content and redirects are still not favorable for a TLD, the overlap in WHOIS data should be expected in some cases and would allow for a two-leg study to evolve from the initial differential analysis.

A key factor of interest from the Halvorson, et. al. study that is of particular value to this research is the percentage of domains that are NOT present in a new gTLD that are present in COM. This is because one of the missions of the new gTLD program is to provide additional ‘space’ for new registrations. This mission relates to the research questions because open domain names that are consistent with existing identities are beneficial in translating the identity across the technological metaphor. The other items of interest are the lists of domains that do overlap in WHOIS (but not in content or as a redirect) and a list of domains that are unique. These lists provide the sampling frame to explore the utilization of new gTLDs for new missions of identity and community formation. The list of overlapping WHOIS data allows for examining how

existing COM registrants are using the new gTLDs to expand their identity into other spaces. The unique list offers an opportunity to discover the types of uses for domain names that may have not been previously engaged in under COM.

From the two lists it will be possible to randomly select an adequate sample (perhaps 100-250 domains) to investigate. Qualitative content analysis would be an effective way to develop the themes, however using such a large data set would potentially create problems for data aggregation and filtering. Utilizing an open coding technique with a qualitative analysis software, such as QRS's NVIVO application may be a more practical implementation of this procedure.

Conclusion

The new gTLDs open up many possibilities for Internet users and content producers. The preliminary research conducted in this paper provides only a small sample of the variation in uses that have been undertaken by Internet users. Community and identity are not easily defined and looking for markers of the existence, especially between many entities presents a challenge.

The results of this preliminary study are inconclusive. The literature highlighted at the beginning of this paper provides a foundation for further research into this area.

Appendix I.

CLUB

Site	Type of Entity	Type of Presence	Notes
ELECTRICMONKEY.CLUB	Nightclub	Redirect to Facebook	Developing web presence, may be migrating to a web identity from a social media only presence
ASHLEYBRIDGET.CLUB	Retailer	Redirect to Facebook	
NITROGENSPORTS.CLUB	Gaming Club	Website	
WNATV.CLUB	Sports Organization	Website	Specific club for a geographic location, mission statement on homepage
KENKODO.CLUB	Health/Fitness Club	Website	Membership information, basic facility details
MEFIGHT.CLUB	Gaming Club	Website	Fully developed website, description mentions moving from a forum
WAC.CLUB	Sports Organization	Website	Community sports organization
THEDUDES.CLUB	Gaming Club	Website	
YOUNGEQUESTRIANS.CLUB	Sports Organization	Website	
COFFEE.CLUB	Subscription Commerce	Website	

PHOTO

Site	Type of Entity	Type of Presence
HEALTHERSAUNDERS.PHOTO	Professional Photographer	Website
AEROSPACE.PHOTO	Gallery -Not Photographer Specific	Tumblr
GNAM.PHOTO	Professional Photographer	Website
MYCOMPANY.PHOTO	Stock	Website
JBM.PHOTO	Professional Photographer	Website
KTMILLER.PHOTO	Professional Photographer	Website
DJC.PHOTO	Stock	Website
F18.PHOTO	Ametur Photographer	Website
AROMERO.PHOTO	Professional Photographer	Website
ALIVE.PHOTO	Photographer Community	Website

XYZ

Site	Type of Entity	Type of Presence	Notes
ZAYO.XYZ	Internet service provider	Website	
MILK.XYZ	Magazine	Website	different' emphasis
ABC.XYZ	Cloud service provider	Website	Google
GADGETDAILY.XYZ	Magazine	Website	
MEDIAOPS.XYZ	Author	YouTube	
RISE.XYZ	Software	Website	new' emphasis
ESCAPISTMAGAZINE.COM	Magazine	Website	
UMAKE.XYZ	Software	Website	new' technology
READER.XYZ	Tool	Website	tool to estimate reading time
WAY.XYZ	Tool	Website	guidebook for various cities

References

Alramahi, M. (2008). New gTLDs – Pandora's Box is open. *International Review of Law, Computers & Technology*, 183-192.

Attaran, A., Kohler, J., Liang, B., & Eysenbach, G. (2014). A call for a moratorium on the .health generic top-level domain: preventing the commercialization and exclusive control of online health information. *Globalization and Health*, 1-11.

Baasanjav, U. B. (2014). Linguistic diversity on the internet: Arabic, Chinese and Cyrillic script top-level domain names. *Telecommunications Policy*, 961-969.

Baym, N. K. (2010). *Personal connections in the digital age.* Cambridge, UK.

boyd, d. (2012). Privacy and Security: The Politics of Real Names. *Viewpoints*, 29-31.

Brown, N. A. (2012). Six Degrees of Disputation: New DRPs for new gTLDs on the Internet. *Dispute Resolution Journal*.

Cerf, V. (1998, October). *I Remember IANA.* Retrieved from Internet Request for Comments Series: https://tools.ietf.org/rfc/rfc2468

EDUCAUSE. (2009, August 27). *Dot-Edu Policy.* Retrieved from Dot-Edu Administration: http://net.educause.edu/edudomain/policy.asp

Goffman, E. (1959). *The Presentation of Self in Everyday Life.* Anchor Books.

Goodnight, G. T., & Green, S. (2010). Rhetoric, Risk, and Markets: The Dot-Com Bubble. *Quarterly Journal of Speech*, 1479-5779.

greenSec Solutions. (2016, May 10). *nTLD Statistics*. Retrieved from https://ntldstats.com/

Halvorson, T., Szurdi, J., Maier, G., Felegyhazi, M., Kreibich, C., Weaver, N., et al. (2012). The BIZ Top-Level Domain: 10 Years Later. In N. Taft, & F. Riciatto, *Passive and Active Measurement* (pp. 221-231). London: Springer.

Harrenstien, K., White, V., & Feinler, E. (1982, March 1). *Request for Comments Series.* Retrieved from Internet Engineering Task Force: https://tools.ietf.org/rfc/rfc811.txt

Hering, I. (2004). The sky's the limit ICANN is set to roll out new generic top-level domains next year, sparking debate in the internet community over their merits. *Managing Intellectual Property*, 22-25.

Internet Corporation for Assigned Names and Numbers. (2010, January). *Public Comments on New gTLD Program.* Retrieved from Internet Corporation for Assigned Names and Numbers: https://www.icann.org/resources/pages/draft-eoi-2009-12-18-en

Internet Corporation for Assigned Names and Numbers. (2015). *About the Program*. Retrieved from New Generic Top-Level Domains: https://newgtlds.icann.org/en/about/program

Internet Corporation for Assigned Names and Numbers. (2016). *Delegated Status.* Retrieved from New Generic Top-Level Domains: https://newgtlds.icann.org/en/program-status/delegated-strings

Jarassriwilai, T., Dauber, T., Brownlee, N., & Mahanti, A. (2015). Understanding evolution and adoption of Top-level Domain names. *Proceedings - Conference on Local Computer Networks*, 687-694.

Jones, S. G. (1998). The Internet and its Social Landscape. In S. G. Jones, *Virtual Culture: Identity & Communication in Cyberspace* (pp. 7-35). London: Sage Publications.

King, R. (2012, June 13). *New gTLD Application: INK.* Retrieved from New Generic Top-Level Domains: https://gtldresult.icann.org/application-result/applicationstatus/applicationdetails:downloadapplication/1454?t:ac=1454

LaMantia, P. A. (2012, June 12). *New gTLD Application: CLUB.* Retrieved from New Generic Top-Level Domains: https://gtldresult.icann.org/application-result/applicationstatus/applicationdetails:downloadapplication/1856?t:ac=1856

Negari, D. M. (2012, June 13). *New gTLD Application: XYZ.* Retrieved from New Generic Top-Level Domains: https://gtldresult.icann.org/application-result/applicationstatus/applicationdetails:downloadapplication/126?t:ac=126

Postel, J. B. (1994, March 1). *Domain Name System Structure and Delegation.* Retrieved from Request for Comments Series: https://tools.ietf.org/html/rfc1591

Postel, J. B., & Reynolds, J. K. (1984, October). *Domain Requirements.* Retrieved from Internet Request for Comments Series: https://tools.ietf.org/rfc/rfc1920.txt

Prahl, D., & Null, E. (2011). The New Generic Top-Level Domain Program: A New Era of Risk for Trademark Owners and the Internet. *Law Journal of the International Trademark Association*, 1757-1800.

Schilling, F. T. (2012, June 13). *New gTLD Application: PHOTO.* Retrieved from New Generic Top-Level Domains: https://gtldresult.icann.org/application-result/applicationstatus/applicationdetails:downloadapplication/1916?t:ac=1916

Turkle, S. (2011). *Alone Together : Why We Expect More from Technology and Less from Each Other.* New York: Basic Books.

Watson, N. (1998). Why We Argue About Virtual Community. In S. G. Jones, *Virtual Culture: Identity & Communication in Cyberspace* (pp. 102-132). London: Sage Publications.

Weinberg, J. (2002). ICANN, "Internet Stability," and New Top Level Domains. In L. F. Cranor, & S. Greenstein, *Communications Policy and Information Technology: Promises, Problems, Prospects* (pp. 3-24). MIT Press.