

Community Scoring in ICANN's Draft Applicant Guidebook: How to Make Sure that New Top-Level Domains are Meaningful and Useful

May 22, 2009

Summary: The new top-level domains (TLDs) from ICANN have the potential to usher in a much more useable Web, but ICANN needs to define a "community" TLD better so that existing communities of interest can create and manage their own TLDs. As it stands, these communities will be shut out and many new TLDs will become meaningless replications of .COM. We recommend tweaking the scoring in ICANN's Draft Applicant Guidebook to make it easier to qualify as a community.

Table of Contents

Overview	2
Background	3
Background: Domain Name Scarcity and Pricing in .COM	4
Background: The Impact of Domain Speculation	5
Background: The Impossibility of Policing the .COM Namespace	8
Background Summary: The Chaos of .COM	10
How to Assure Meaningful TLDs	10
Community TLDs and Search Engines	11
Demand for new TLDs	11
The Definition of Community	12
Community Scoring in Detail	14
Recommendation	15
Conclusion	15

Overview

The simultaneous introduction of hundreds of new Top Level Domains (TLDs), as proposed in ICANN's second revision of the Draft Applicant Guidebook, offers the very real possibility of an evolution of the Internet namespace into semantically meaningful zones, where the TLD extension is a reliable guide to the content found on the Web sites contained within it. For instance, .EDU is a meaningful extension: every web site that ends with the .EDU extension belongs to an accredited institution of higher education, and the content there has a meaningful connection to academic pursuits. As a result, .EDU is prized by search engines as an authoritative source, and users rarely run into nasty surprises.

In contrast, ".COM" is an extension that conveys no meaning at all; its openness has allowed for much experimentation, but its lack of rules, standards, and enforcement mechanisms, combined with the lack of meaningful alternatives, have made it a perfect testing-ground for a host of familiar domain name ills: phishing, malware, spam, trademark abuse, and parking pages and other auto-generated content. It would be a grave mistake to let new TLDs replicate these ills, especially when there is a much better alternative.

Under the current Draft Applicant Guidebook rules for what constitutes a "community" application, many new TLDs would be denied to the communities they clearly represent, and instead be turned over to the highest bidder at an auction. ICANN should make sure that the Draft Application Guidebook definition of "community" allows real-world communities to get the TLDs that clearly apply to them.

Applications that would be denied community status under current Draft Application Guidebook rules include: .ACTOR, .AIRLINE, .AUTHOR, .BIKER, .BOATING, .CHRISTIAN, .CLIMB, .FISHING, .GAY, .HIKING, .ISLAM, .OCEANIA, .REALTOR, .SKI, .SURFER, .WOMEN. There are hundreds more. These are communities in every ordinary sense of the word, representing passionate interest groups comprising millions of people. They should be given the opportunity to create TLDs that represent their interests, to use the TLD name by which they know themselves, and to manage the TLD with policies that reflect their values. By doing so, they will create meaningful namespaces that will be useful to everyone.

By adopting just a slightly more expansive definition of "community" in its Draft Applicant Guidebook, and by giving greater weight to endorsements from leaders and members of these communities, ICANN can encourage the creation of new TLDs, managed and regulated by the people the TLD represents (many of them with vast number of adherents), which will do away with most if not all of these ills by creating policies and policing mechanisms that are proper to the aims of that community.

A definition of "community" that more closely resembles the common meaning of the term will reduce brand confusion, expand the availability of reasonable domain names, reduce the incentives for cyber- and typo-squatting, allow the relevant community to come up with its own rules and enforcement mechanisms, improve search engine results, and improve upon the overall quality of the navigation experience for Internet users. If the introduction of new TLDs is largely reserved to speculators, however, as the current Draft Applicant Guidebook would encourage, all of the problems that plague the Web under .COM will be replicated in each new TLD, and the Internet will be littered with innumerable clones of .COM.

Up to a point, the process described by the Draft Application Guidebook favors an application filed by a "community" over an "open" application in case of contending applications. While this is the right approach, Draft Application Guidebook criteria for deciding which application qualify for community status is too restrictive and fails to take account for the many communities which do not meet its narrow definition.

Under the Draft Application Guidebook, scores if not hundreds of community-focused TLDs with broad support from their communities will likely be classified as "open" and thereby thrown into an auction, where community wishes are completely subordinated to financial considerations. This will result in a large number of legitimate community TLDs going to auction, where deep-pocketed speculators will force out legitimate community contenders. If that happens, well-regulated namespaces, managed by the legitimate representatives of the relevant communities, will be replaced by the wild-west lawlessness that characterizes .COM.

ICANN can make the Web navigable, comprehensible, and litter-free, simply by changing its scoring system to define a community application as one that scores 12 or better, instead of the current 14, and by giving more weight to the most important criteria, which is endorsement and support by leaders and members of a community.

Background

The effect of the seemingly small change from a score of 14 to a score of 12 can hardly be exaggerated. The difference will be between a series of .COM clones on the one hand, and a responsive, responsible, and comprehensible set of TLDs on the other. To understand why this is so, we need to examine the problems of the current "open" generic top-level domain (gTLD) namespace. Most readers are so used to considering the state of affairs in .COM as "normal" that a re-examination of the effect of the

hands-off management of .COM, .NET and the other "open" TLDs is essential for understanding how creating more of them could be disastrous, and how a better administration of the top-level domain spaces is crucial.

Background: Domain Name Scarcity and Pricing in .COM

Every day, thousands of ideas for new web sites are conceived. They may be ecommerce sites, personal blogs, photo albums, web applications, directories, review sites and so on. While very different in nature, these businesses face a common challenge in the existing .COM world – the inability to get a reasonable domain name at a reasonable price. The commonly advertised price of \$10 for a .COM name is simply not true – *the real price for a usable .com domain is several hundred or thousand dollars*. Anyone who has tried to register a decent name for a new web site knows this well.

With over 80.5 million .COM names registered, every single word in each major language dictionary has already been taken. Furthermore, common combinations of several words, or domain names made by adding popular prefixes or suffixes, are also generally not available.

To illustratrate the scarcity in the .COM domain name space, suppose a new business specializing in Plasma TVs wants a reasonable .COM address. A search for "plasmaTV.com" on GoDaddy.com reveals that it is not available, but that a wide array of subpar alternatives are available – at exhorbitant prices. Someone not familiar with the unavailability of names in .COM might be forgiven for thinking that GoDaddy was engaging in bait-and-switch marketing:



Even the completely uninspiring domain name "PlasmaToday.com" fetches a price of \$1,988 – a high starting price to create a new web presence, and a far cry from the advertised \$7.49 .COM retail price advertised by GoDaddy. To say that you can buy a .COM name for \$10 is to be very economical with the truth.

Take another example. Consider the average resale price of three-letter domain names in the five top generic TLDs. All combinations of three-character names in .COM, and

most if not all in .NET, have been registered. In 1999, the last three-character .COM domain name containing only letters or numbers was taken. By 2006 there were no longer any three-character .COM domains available. Today, three-character .COM and .NET domain names are only available on the aftermarket, as shown in the table below:

>TLD						
YEAR	.COM	.NET	.MOBI	.ORG	.INFO	Avg
2007	\$1,110.42	\$265.50	\$212.50	\$114.17	\$93.75	\$510.56
2008	\$1,431.67	\$284.38	\$206.25	\$117.50	\$93.75	\$628.24
2009	\$1,414.58	\$315.00	\$220.00	\$106.67	\$180.00	\$743.89
Avg	\$1,318.89	\$285.86	\$210.56	\$112.78	\$111.00	\$618.99

Average price of three-character domains

Background: The Impact of Domain Speculation

At present, about 30% of all domains in .COM are "parked"; they consist of nothing more than pages filled with ads and links to ads, and very few of the many people who end up there intend to do so. Every day, people are frustrated in trying to find the information they are looking for on the web, and instead end up on a parked page. This is due to a confluence of five factors:

- The low carrying cost of .COM domains, \$10/year or less
- The ability to monetize domains through advertisements from the two major search engines, Google and Yahoo!
- The development of an intermediate market of "parked page" providers who facilitate the monetization of domains.
- In the U.S. market especially, the dominant position of .COM as the *de facto* TLD, due primarily to early stifling of the .US country-code domain and the early lack of viable gTLD alternatives.
- The role of registrars, who have in the last few years kept expired names to sell at auction instead of releasing the names back into the public pool.

These parked pages are owned by companies with large domain portfolios. For each and every domain name they own, there is a corresponding ad-filled parked page.

Many of these domain names are valuable generic terms that describe entire industries.

Some portfolio holders have invested money and time in developing automated, highspeed registration systems that register domain names in high quantities, largely to monetize them en masse via parked pages, but sometimes to "flip" the domains for quick resale at higher prices, or to "warehouse" them for resale at a much later date. These high-volume registrations severely decrease the availability and choice of domain names for new registrants.

For a new business, getting the right name is crucial, and today it is not uncommon for startups to pay hundreds of thousands of dollars for names that originally cost under \$10. Registrars now regularly feature aftermarket names for resale, invariably at prices many times higher than a new registration.

The frustration and cost of .COM domain names is not limited to new businesses. In addition to owning many valuable generic names, speculators have seized on many of the typographical variants of popular websites and brands. As an example, the research at <u>http://www.domaintools.com/domain-typo/</u> uncovers the following registered typos of the extremely popular destination "myspace.com."

⊖ Domain	O Registrant
Kyspace.net	1&1 Internet, Inc http://1and1.com/contact
Myspace.us	Aaron Kluge
Mjyspace.com	Abadaba S.A.
Myspac3.com	Abadaba S.A.
Nyspace.us	Alan Lin
Mhspace.com	Alan Lin
Nyspace.biz	Alan Lin
Myepace.com	Alan Lin
Mygspace.net	Allison Lopez
Myspasce.net	Ana Wiatkima
Mysoace.net	Andrew Miller
Myspaec.net	Andrew Miller

Note that in this example, the domain "myspac3.com" is a typo of myspace.com on a US-English keyboard, due to the position of the "3" key directly above the "e" key.

In fact, the site at myspac3.com is a parked page:



This page is generated algorithmically by a parking service, and makes money each and every time a user clicks on one of the "sponsored listings" (note also that the "Other Searches" list on the right sidebar leads to other parked pages filled with ads. As is often the case, the first sponsored listing is in fact paid for by News Corp (the owner of Myspace), who is paying twice: once for the lost traffic to their real site, and again because they are buying ads that are placed contextually on "relevant" sites.

Consider the estimated traffic to this site, as provided by the service compete.com:



As we see, between 100 and 600 people a month go to this typo-variant of the real site. Assuming 20% of these people click on a link (about the average for such pages), this parked page will generate 10 to 60 clicks per month, or about 2 to 12 dollars in revenue. The revenues for this site, while small, are entirely sufficient to cover the low annual renewal fee (about \$10) for a .COM domain name, and to generate a small profit. From here, it's a numbers game. To make a lot of money, simply register every possible typo of every high-traffic site – the aggregate profit has made some speculators very rich.

With instant access to traffic data, high-volume domain speculation has evolved into a specialized industry with a handful of extremely artful and successful players making vast amounts of money, ultimately from users who never meant to visit their pages in the first place.

Background: The Impossibility of Policing the .COM Namespace

Although such clear-cut, unambiguous domain typo-squatting is rampant in .COM, there are no rules preventing it, and only burdensome mechanisms in place to police or combat it. More important, the .COM registry has no interest in helping clean up the namespace, because their single financial interest is to sell more names – they make the same amount regardless of the quality or legality of the registration. In the case of .COM, there is no community to insist on a well-regulated namespace; the entire burden of policing is placed on the aggrieved party. As an analogy, imagine a neighborhood where people drive by and throw litter out their window, but the entire

burden of stopping them rests with owner of the particular lawn where the garbage landed – there is no help from the neighbors, and there is no police force. Even when the trademark holder acts, it is often thwarted by obfuscation and deceit. Because .COM is, by historical artifact, a "thin" registry (meaning that data on who owns a domain is held by many different parties, with different access rules), it is difficult to verify ownership.

For example, the owner of the previously-mentioned typo site "myspac3.com" has an address at a post-office box in Panama:

Whois Record

```
Domain myspac3.com:

Abadaba S.A.

P.O. Box 0819-09689

Panama, Panama 6 PA

Administrative contact:

Technical contact:

Billing contact:

Abadaba S.A.

Domain Admin

<u>abadaba@eml.cc</u>

P.O. Box 0819-09689

Panama, Panama 6 PA

Phone: +1.2063090626

Fax: +1.2063090626
```

The listed phone number, which is actually a Seattle phone number, unsurprisingly goes directly to voice mail.

Reasonable people can disagree on the extent to which trademark owners should have power to compel others to recognize their asserted rights, although we find it hard to imagine in the clear-cut case presented above.

But what about malware, phishing, or spam, which are also rampant in lightly-managed gTLDs such as .COM, .NET, .INFO and .BIZ. The managers of these TLDs recuse themselves from responsibility and refer enforcement to ICANN, whose procedures are lengthy and ineffective. In contrast, many country-code top-level domains (ccTLD), administrators combat these activities effectively by suspending the offending domain names upon a verified report of wrongdoing. These managers care about the TLDs they administer, and they are empowered to fix the problems. This is what a

community does, and it's an important difference between a community TLD and an open one.

Background Summary: The Chaos of .COM

As we have seen, the current system of a single dominant extension, .COM, has severe drawbacks:

- Useful .COM names are phenomenally expensive
- It is next to impossible to find a useful .COM domain name for under \$100 the "real price" of a .COM domain name
- Approximately 30% of all .COM websites are ad-filled parked pages
- Domain speculators rarely produce useful web sites, preferring easy-to-install parked pages
- Speculators warehouse names for years, demanding high prices from new registrants, desperate for a name
- Manager indifference, ICANN rules and the thin WHOIS model of .COM makes
 enforcement difficult

Because .COM is by far the largest single extension for Internet addresses, and because it has evolved over the years in symbiosis with the eco-system of domain speculators and parked pages, its understandable that there is no practical way, at this point, of "undoing" the vast proliferation of meaningless and confusing websites in the .COM namespace. For various reasons of history, policy, and fairness to those who operated legitimately under the rules then in force, it isn't possible or even desirable to shut down what amounts to about 30% of all existing .COM web addresses, in the name of cleaning up the Internet.

So what is to be done?

How to Assure Meaningful TLDs

The solution is to create new top-level domains with meaningful extensions, so that consumers and registrants can use the top-level extension to identify content type. To re-iterate: *A meaningful TLD is one where the extension is a reliable guide to the content found on the web sites contained within it.* Creating meaningful TLDs is best accomplished by encouraging applications from communities who will police their TLDs according to the values of that community. This will lower the real price of a good domain name, provide real choice among TLDs, make policing of abuse possible, make parked pages rare and unprofitable, and get rid of most warehousing. If ICANN rules prevent this, many new TLDs will quickly resemble the polluted .COM namespace.

Community TLDs and Search Engines

Search engines will give greater weight to meaningful TLDs. We know that because they already do. All else being equal, Google will rank .GOV sites higher than .COM on a search engine results page than those in .COM. A German-language search query is more likely to return a .DE (the German country-code TLD) site than any other TLD.

If someone searches for "opticians in New York," and .NYC has regulated its namespace so that .NYC sites have a meaningful connection to the city of New York, then, other factors being equal, Google will absolutely rank optician.nyc higher than optician.com for that search. On the other side, searchers will be more likely to click on a domain name ending in .NYC for searches related to New York – and as that happens, Google will rank .NYC names even higher. This feedback cycle, already shown to exist for the regulated namespaces of .EDU and .GOV, would pertain for new TLDs as well.

Furthermore, links from regulated namespaces are worth more. A link from a .EDU or .GOV is worth more, and has more "Google juice," than links from other TLDs – quite a bit more. The reason is that Google knows that these namespaces are not filled with link farms and other attempts to outsmart the Google algorithm – a direct consequence of enforcement of rules by the .EDU and .GOV registries.

The impact of new TLDs – if they are meaningful – is therefore quite significant for marketing and search, and this is worth many millions of dollars.

Demand for new TLDs

One criticism of the new TLD process says that there is no demand for the new TLDs, citing the low level of adoption of the limited number of new TLDs introduced in 2000 and 2005. This is a spurious assertion – you cannot expect high demand from TLDs that are either as meaningless as .COM (e.g., .INFO, .BIZ) or appeal to restricted groups with very few members (.MUSEUM, .AERO).

Even so, while not comparable to .COM, the numbers for some recently-introduced TLDs are quite large. With over 1 million .MOBI names, 5 million .INFO names, and 2 million .BIZ extensions, even new TLDs with a very attenuated semantic meaning have proven to have demand.

Previous TLD rounds have been too limited in scope and therefore have not had a chance to develop the critical mass necessary to effect an overall change in how users semantically decipher Internet addresses. Furthermore, they have been hampered by the expectation that they would be either (a) just like .COM, or (b) extremely restrictive. Community TLDs, endowed with policies that reflect the degree of openness that their

community desires, will change this. In 2010, as hundreds of new TLDs are released, many of them with policies as to their registrants and content, users will be able to use the TLD extension as a guide to the content of web sites, and communities will encourage their members to use them.

This leads us to the heart of our argument: the definition of community.

The Definition of Community

ICANN's Draft Applicant Guidebook defines only two types of top-level domain applications: "community" applications, where the applicant represents a community signified by the TLD string, and "open" ones, where it does not. Community applications trump non-community applications. If there are no valid community applications for a string, the tiebreaker "trial by fire" selection method where there is more than one applicant for the same TLD is an auction.

As we noted earlier, ICANN has devised a point system to determine what constitutes a community. The more defined, restricted, and narrow the community is, the higher the score.

This is a good approach, as far as it goes. The point test will allow extremely welldefined and organized communities such as the AARP to get .AARP without much risk of auction. The NFL would also qualify, as a brand, as would .NYC, as a geographical TLD. Numerous TLDs which represent passionate and easily identifiable communities, however, will not qualify. Under existing rules, these communities will have to reach deep in to their pockets to outbid speculators.

For example, suppose .BIKER is applied for by a consortium of some of the large number of motorcycle groups, put together for the express purpose of obtaining the TLD. Suppose further that they decide that domain names under .BIKER should be limited to those who own a motorcycle, or repair motorcycles, or sell motorcycles and motorcycle parts. Finally, suppose they decided that .BIKER sites should contain only motorcycle-related content. If that happened, .BIKER would be promoted by motorcycle organizations and would be a natural home for bikers everywhere. Current Draft Application Guidebook rules would deny this consortium community status, and instead award this TLD to the group with the deepest pockets, who would have no obligation to respond to the concerns of motorcycle enthusiasts, motorcycle manufacturers, or motorcycle repair shops. In fact, they would have no obligation at all, except to pay the auction fees. They could simply open up .BIKER to all comers, including speculators, and make .BIKER a useless copy of .COM. How does that advance the interests of anyone at all except the auction winner?

In another example, .GOLF, even if backed by the leading golf associations, would under current rules be auctioned off to the highest bidder. .WINE, even if supported by the major vineyards and wine associations, could be outbid by a complete outsider. In these cases, the high bidder could then fashion policies that could fill the .GOLF or .WINE namespaces with parked pages that bring little or no value to these communities. We contend that this is not the way to make new TLDs meaningful and useful, and that it disenfranchises legitimate communities simply because they are not a brand (NFL), or a geographical grouping (NYC), or tightly organized membership organization (AARP).

Affinity and group allegiance will be key factors in the adoption and natural use of a new TLD. Therefore we question the results that will come from ICANN's current rules. If ICANN's rules mean that an application for .GOLF, backed by the leading world golf associations, has the same worth as an application for .GOLF backed by a speculator, then there's something wrong with the rules.

In the example above, there are several reasons for preferring the .GOLF application backed by golfers and golf associations over the speculator's application:

- Golfers, who are passionate about their sport, will want a .GOLF domain if they know it is legitimate and backed by the respected organizations in the golf world.
- Golfers, knowing that the .GOLF TLD will be used primarily by other golfers, will strive to produce useful web sites.
- A leading golf association, or a coalition of golf associations, will market and brand the new top-level domain better than the domain speculator.
- Golfing associations will protect the brands of golf-related companies far better than the domain speculator.
- The industry-backed .GOLF TLD is far less likely to fail because it has backing from financially stable organizations.

This is not to say that the PGA, or the USGA, or the LPGA, are the only valid golf association. If there were two applications, both backed by significant golfing organizations, an auction or an evaluative panel might be a reasonable way to settle the dispute. But in the case of the PGA (for example) versus an unknown company, the PGA has a stronger claim to .GOLF.

Another example is .SCOUT. Although there are many scouting organizations, and the word "scout" is not an identical character-for-character match with "Boy Scouts of

America," if the BSA applied for ".SCOUT" – a name by which they are known to many – should trump the application of an applicant with no connection to the name "scout."

It is worth noting that neither the Boy Scouts or the PGA have large war chests to win auctions. Domain speculators, who follow the ICANN process, do. Therefore, while auctions may be a good method for distributing scarce bandwidth among cell phone operators, it is not a good way to allocate resources when not-for-profits, industry consortiums, trade groups, or other community groups are involved.

Community Scoring in Detail

There are 4 criteria for scoring in the Draft Applicant Guidebook. Each of the following criteria is graded on a score of 0-4, from "isn't even close" to "perfect match." A perfect score is 16, or 4 points on each question. Under current rules, to qualify as a community, a score of 14 points is necessary.

- Criterion 1 The nexus between the string and the name of the community (e.g., .SCOUT versus .BoyScoutsOfAmerica)
- Criterion 2 Dedicated registration policies (how restrictive is it)
- Criterion 3 Community establishment (how big the community is, and how long it has been around)
- Criterion 4 Community endorsement (e.g., letters of support from community leaders)

Let's work through the scoring for .GOLF, backed by a consortium of the PGA, the LPGA, and the U.S. Golf Association, and the Japan Golf Association, several golfing manufacturers, along with letters of support from hundreds of ordinary golfers (in other words, a very strong application from a significant portion of the golfing community).

Criterion 1 – Nexus (name similarity). Because the name .GOLF isn't an exact match to any of the backers' names, under current ICANN rules it most likely scores a 3 or possibly a 2.

Criterion 2 – restrictive policies. The score could vary here. If .GOLF were restricted to professional golfers who were members of the associations in the consortium, it might get a 4. If the consortium wanted to make it available to any member of a golf club and to golfing equipment manufacturers, as they might be expected to do, then it would score at most a 3.

Criterion 3 – community establishment. There are a lot of golfers and golf equipment manufacturers, but the newly-formed consortium is brand new. They might score only a 2, possibly a 3.

Criterion 4 – community endorsement. The weight of an endorsement from important members of the community should count significantly more than the other three criteria. The fact that the major golfing associations and many ordinary golfers support an application for .GOLF, should in and of itself qualify the .GOLF application as a "community application." Industry-backed applications, under the current rules, would score a 4 for this criterion.

Overall score for .GOLF, backed by leading organizations, manufacturers, and ordinary golfers: from 12 to 14 points. Under the rules in the current Draft Application Guidebook, this .GOLF application would probably fail to achieve community status. In our opinion, this is a travesty: a solid community application, backed by the community's leading organizations, run a very real risk of being outbid by a speculator with no ties at all to the community.

Recommendation

There is far more danger, for the Internet and for the new TLD process, in being too restrictive on the community definition than on being too lenient. *We recommend lowering the threshold for community standard from 14 to 12, and giving more weight to community endorsements.* This will allow applications from well-established communities and interest groups the chance they deserve.

Conclusion

The issue of the community scoring system has profound implications for Internet naming and the future of the Internet generally. If we keep the bar high, new TLDs will simply be auctioned to the highest bidder, who may well copy the .COM model and increase confusion. We contend that this is not the way to make new TLDs meaningful and useful – it disenfranchises legitimate communities which do not qualify as a brand (NFL), a geographical grouping (NYC), or a tightly organized membership organization (AARP).

If instead we lower the bar, valid community applications will emerge, and semantic meaning will be restored to the top-level domain space. This will lower the cost for registrants, reduce squatting and trademark abuse, improve the comprehensibility of domain names, make searching easier and more relevant, and improve the experience of using the Internet for everyone.