

# Executive Proposal for IDN Policy

## Contents

1. Introduction .....	1
1.1 Background.....	2
1.1.1 International experience .....	2
1.1.2 Explanation of terms .....	2
1.1.3 The new phishing threat.....	2
2. Policy Principles.....	2
3. Policy Details.....	3
3.1 Supported languages and scripts .....	3
3.1.1 Maintain a list of allowed scripts/characters that will be as wide as possible. ....	3
3.1.2 Punctuation, mathematical symbols, box drawing characters, etc are not allowed .....	3
3.1.3 No restrictions on domains that differ only in diacriticals .....	3
3.1.4 Separation of different scripts.....	4
3.2 Registration .....	4
3.2.1 Phased introduction of allowed characters.....	4
3.2.2 No sunrise period.....	4
3.2.3 Decreasing price introduction.....	4
3.2.4 Requirement for providing the internationalised and registry forms upon registration .....	4
3.2.5 No requirement for providing a Latin equivalent .....	5
3.3 Support services.....	5
3.3.1 DRS will be extended to cover the internationalised form, not just the registry form .....	5
3.3.2 Public register search and lookup services to include the internationalised form as well as the registry form, where appropriate .....	5
3.3.3 Communications sent to the registrant about their domain will, wherever possible, include the internationalised form. ....	5
3.3.4 All other support interactions will require the registrant to use the registry form of the domain name... ..	5
3.3.5 Public tools will be provided to allow transformation between internationalised and registry forms.....	5
4. Conclusion.....	5
5. Recommendations .....	6
6. References.....	6

## 1. Introduction

This document is the Executive response to the Internationalised Domain Names (IDN) Policy proposal from the Policy Advisory Board (PAB), which itself followed from the public consultation on IDN usage within .uk. This document does not follow the structure of the previous PAB paper, as it endeavours to separate policy principles from practice and implementation details.

This document includes:

- a clear split into policy principles and associated policy details (2);
- a different proposal for domains that differ only in the diacriticals (3.1.3);
- a different proposal for domains with characters taken from different scripts (3.1.4);
- a new proposal regarding Latin equivalent (3.2.5);
- a new proposal to modify the Dispute Resolution Service (DRS) to handle some potential IDN disputes (3.3.1);
- a new proposal for search and lookup services (3.3.2);
- a new proposal regarding communications to registrants (3.3.3).

The recommendation is for the PAB to convene a special sub-committee, including Executive members to discuss the two documents and develop joint proposals for a way forward.

Following this the Executive will draw up a technical and operational implementation document. This will incorporate, where appropriate, many of the technical details given in the first PAB proposal.

## 1.1 Background

### 1.1.1 International experience

In addition to the local consultation, the policy details proposed here have also been discussed with other registries that have introduced IDNs to ensure that they meet with the broad agreement of the international community.

The International Corporation for Assigned Names and Numbers (ICANN) IDN Guidelines v.2.1 have also been consulted for this document, but they contain so many problem areas, namely the confusion of languages and scripts, that very little of those guidelines are incorporated.

### 1.1.2 Explanation of terms

Just to be clear, this document refers to the 'internationalised form' of the domain, which is the form that uses the full international character set, and refers to the 'registry form' for the form that begins xn---, which is what is actually registered and then appears as the published Domain Name System (DNS) entry.

### 1.1.3 The new phishing threat

One of the most talked about threats, that comes from the introduction of IDNs is phishing using domain names constructed from different scripts. The main example often cited is the registration of a domain that looked like paypal.com, but used a combination of Latin and Cyrillic characters. Any IDN policy must address this new threat.

## 2. Policy Principles

The following principles are proposed, from which the details are extrapolated. These principles can be used, if needed, to answer further issues or questions that have not arisen yet. They are

1. Domain names are just domain names, not words.  
It should be remembered that the purpose of domains is merely to assign labels. Some of the restrictions that are proposed will reduce the number of words, in various languages, that can be expressed as domain names. This is inevitable and not a sign that this proposal is flawed.
2. We support characters (or scripts), not languages.  
Even with our limited character set currently in use, people are able to register domains that have a meaning in a different language other than English. Our position on this is entirely neutral, we neither explicitly support, nor restrict the support for certain languages, but focus instead on the individual characters.
3. We aim to be as liberal as practically possible in our support for the use of additional characters.  
Of course there will be restrictions and procedures to be followed so complete support of every available character will not be possible. There will also undoubtedly be some restrictions on combinations. But this principle does mean that we work on the basis of restricting for a sound reason, rather than only allowing things when we can understand how they might reasonably be used.
4. Mirror the use of letters, digits and hyphen (LDH) in pre-IDN domain names.  
All non-IDN domain names are restricted to LDH, which means that many other characters such as the question mark, colon or equals sign are not allowed in domains names. This proposal follows the same principle wherever possible.
5. Minimise the potential for phishing using mixed characters.  
Our policies must be developed to minimise, where practical, the potential for mixed character phishing. This is unlikely ever to be eliminated, but we need to be clear how our policies will or will not affect it. It is quite likely that new means of exploiting this vulnerability will be developed as well as increased functionality of browsers to combat these attacks and so the policy details will need to adapt over time.

6. Cost of IDN support to be kept distinct, where appropriate, otherwise to be absorbed.  
For some areas of support of IDN it may be the case that there are distinct costs that are significantly in excess of those for non-IDN domains. In which case it should be possible to identify those distinct costs and pass them on to the customer. However, in most cases the additional cost of supporting IDNs will be absorbed.
7. Registrant to see both forms, wherever possible.  
The existence of two very different forms of the same domain is going to be confusing to the registrant. We need to do as much as possible to reduce that confusion and ensure that all interactions that we have with registrants are as confusion-free as possible.

## 3. Policy Details

### 3.1 Supported languages and scripts

#### 3.1.1 Maintain a list of allowed scripts/characters that will be as wide as possible

Whilst the principles mean that there is a general presumption that all scripts are allowed, in practice this will mean a list of allowed scripts as each new script will need to be checked for any characters that need to be excluded for the reasons given below, before it can be included.

#### 3.1.2 Punctuation, mathematical symbols, box drawing characters, etc are not allowed

In each script that is to be included a set of characters will be excluded because they are not generally recognised as letters or numbers or special characters that are used in for the same purposes (such as marks and ideograms). There may also be some characters excluded because they can be mistaken for punctuation.

Please note, this is a very simplified description of a complex technical area. The full technical interpretation of this policy detail is likely to be the same that proposed by the PAB.

#### 3.1.3 No restrictions on domains that differ only in diacriticals

One proposal that circulated in the Internet community some time ago, was the idea of bundles, which can be used as a way of controlling the registration of names that differ only in diacriticals. A similar proposal was made by the PAB, which recommended restrictions on names that differed only in diacriticals.

However the practical issues with bundles mean that this is no longer considered a viable mechanism by the international registry community. The point was eloquently made during the consultation that words that differ only in diacriticals may have completely different meanings. For example, in Welsh, “fèis” means a summer school/festival and “feis” means sexual intercourse.

The view expressed in the PAB proposal, namely that this is too confusing, is not felt to properly recognise the role of registrants who are comfortable with other scripts and so would be disadvantaged by this proposal, because they might not be able to register the domains that they want that have not been registered by anyone else.

The other view expressed in the PAB proposal, that these restrictions would remove the need for defensive registrations, is felt not to fit with the general principle that registrants decide for themselves what defensive registrations they should make. There is no precedent within the open Second Level Domains (SLDs) for restricting available names to help the holders of intellectual property, instead the onus is on the use of the DRS, after the fact.

There is also the possibility that imposing restrictions of this nature could lead to some very awkward corner cases. For example someone may register a name without diacriticals in which they have clear Intellectual Property (IP) rights but by doing so prevent the registration of a diacritical version, in which another party has clear IP rights. The dispute resolution process in this example could be intractable.

So this proposal is to follow the lead taken by almost all other registries and to have no restrictions at all on domains that differ only in diacriticals.

### 3.1.4 Separation of different scripts

To ban names from containing characters from different scripts would cause problems for those languages that have adopted words in different scripts. The alternative, which is to allow a free for all, would introduce too much scope for mixed character phishing. The solution chosen by most registries so far has been to go for the former and ban mixed script domains. Some of those that started by allowing mixed scripts have since moved to the complete restriction, with some small exceptions for scripts that are generally considered interchangeable.

We have considered a position between the two, which is to allow different scripts to be used within a single label of the domain name, but to ensure that different scripts are separated by the hyphen character. Discussions with other registries and IANA have leant support to this proposal but it would mark a more liberal approach than is generally being taken by other registries.

To be clear, this proposal would not eliminate all possible avenues for phishing using mixed characters sets, nor would it allow words in languages that use mixed characters in a single word to be expressed as domain names. It is likely that this would allow greater scope for phishing than complete separation of scripts.

If the PAB is uncomfortable with this proposal then the position of complete separation (with the small exceptions proposed by the PAB) could be adopted.

## 3.2 Registration

### 3.2.1 Phased introduction of allowed characters

As the PAB has noted, it is far harder to take names away once they have been released than to prevent them being registered in the first place. If the introduction of IDNs generates unforeseen problems then it is better that these are resolved with a limited set of names.

This proposal is therefore for a phased introduction that start by only allowing the registration of names using the full Latin script for a period of time. Once any problems that appear have been resolved then this should be opened up to all scripts. If any substantive problems emerge then it may be necessary to consult with the PAB further on how they should be tackled.

### 3.2.2 No sunrise period

IDNs have not been available for some time and so registrants have been forced to register names that are similar to those they want, but differ only in the absence of diacriticals. Some registrants think that non-diacritical versions are easier to use since most UK keyboards do not easily give access to characters with diacriticals. However there are almost certainly a large number of names that have been registered without a thought to a version with diacriticals.

Taking this together with the fact that some words that differ only in diacriticals have very different meanings, it is clear that a sunrise period could be fraught with difficulties. It is felt that it is better to follow the principle of First-Come-First-Served and allow the market to decide as we do with other domains, with no sunrise period.

### 3.2.3 Decreasing price introduction

This proposal is that IDN registrations should initially start at a higher than normal price, that drops over time to the normal price. The effect of this, as seen in the launch of me.uk, would be to limit the rate of new applications, both overall and by individual people and companies.

### 3.2.4 Requirement for providing the internationalised and registry forms upon registration

The registration of an IDN must include both the internationalised and registry forms and they must match when correctly processed through the transformation algorithm. The reason for this requirement is to ensure that the registry form being registered is that which was intended, eliminating any differences in algorithm implementations. This will have the side effect of preventing any domains beginning xn-- from being registered unless they are a registry form of an internationalised domain.

It should be recognised that this will mean that systems used by any tag holder that wishes to register IDNs will need to support international characters

An operational note is that both of these forms would be stored in the registry to make operations on them easier.

### **3.2.5 No requirement for providing a Latin equivalent**

It is the practice in some registries to ask registrants to provide a Latin equivalent for any name they register that does not use the Latin alphabet. This is generally intended to make it easier for them to work with those domains. However there are a number of difficulties with this. First, Latin equivalents are different for different people, for example we now use Beijing for Peking even though the Chinese name of the city has not changed. Second, we have no way to verify whether the Latin equivalent provided is correct or not.

The proposal therefore is that this will not be a requirement. In those rare circumstances where a Latin equivalent is required, such as in certain DRS cases, then independent opinion will be sought.

## **3.3 Support services**

### **3.3.1 DRS will be extended to cover the internationalised form, not just the registry form**

The proposal is to make two modifications to the DRS.

This first is to allow disputes based on an examination of the internationalised form not just the registry form. So, for example, if someone were to register a name using the Arabic script, which someone else felt infringed their trademark, then they might launch a dispute and it would not be a defence for the registrant to claim that all they had done was register a name like xn--???????.co.uk.

The second is that we will add provisions to the DRS to deal with translation, transliteration or phonetic translation of the internationalised form, if required. This will involve passing the cost of any translation required on to the parties at cost, in accordance with the current policy position that we provide administration for free but outside expertise is paid for by the parties. The exact details of who pays for what and in what circumstances this will be used is for the Executive to determine at a later date.

### **3.3.2 Public register search and lookup services to include the internationalised form as well as the registry form, where appropriate**

We currently provide three register search and lookup services. The proposal is to extend these, where appropriate to include the internationalised form of the name. This may mean changing the character encoding that these systems use.

### **3.3.3 Communications sent to the registrant about their domain will, wherever possible, include the internationalised form.**

To avoid confusing the registrant, communications sent directly to them by Nominet will include the internationalised as well as the registry form.

### **3.3.4 All other support interactions will require the registrant to use the registry form of the domain name**

It will not be possible for us to employ people who understand all of the different scripts in the world. As a result all day-to-day support interactions, such as transferring a name, will require the use of the registry form of the name. This may be made easier by the use and integration of transformation tools.

### **3.3.5 Public tools will be provided to allow transformation between internationalised and registry forms.**

The proposal is to provide tools for registrants, probably a web form, that allows transformation between the internationalised and registry forms. However this is not the same as tag holder libraries for inclusion in registration systems, which we will not provide.

## **4. Conclusion**

This document sets out a proposed policy framework for the implementation of IDNs within .uk. If agreed, then the next steps are to

- Produce a technical and operational implementation document
- List the allowed scripts for the second phase of introduction

- List the disallowed characters
- Begin the development of the supporting systems

It is likely that this will take some time to complete. In particular, there is some concern that an anti-phishing policy will need to be in place before IDNs are introduced. Though it is unlikely that the introduction of IDNs with controls on mixed scripts, will add significantly to the problems of phishing that people are already experiencing.

## 5. Recommendations

The PAB is asked to convene a sub-committee, including Executive members, for a single meeting to discuss this document and prepare an amended policy proposal for presentation to the September PAB meeting.

## 6. References

- The main IETF standard for IDNA (Internationalised Domain Names in Applications)  
<http://www.ietf.org/rfc/rfc3490.txt>
- ICANN IDN Guidelines v2.1  
<http://www.icann.org/general/idn-guidelines-22feb06.htm>
- Nameprep  
<http://www.ietf.org/rfc/rfc3491.txt>
- .de FAQ about their introduction of IDNs  
[http://www.denic.de/en/faqs/idn\\_faqs/index.html](http://www.denic.de/en/faqs/idn_faqs/index.html)