Draft

Standardization of LocalityName andStateOrProvinceName

Prepared by: Eva Van Steenberge
October 2019 V0.7

GlobalSign®
GMO INTERNET GROUP

Table of Contents

Overview

Recently, multiple Certificate Authorities have reported problems with unverified information in the subject:localityName and the subject:stateOrProvinceName. In its report, GlobalSign stated its support for the effort to standardize practices, with the ultimate goal being a common set of whitelist values for each country, including supported synonymous values for those countries with language variations.

In order to support this effort, GlobalSign is in the process of conducting a categorization and quantification exercise of what it considers "synonymous values (or functional equivalents)" on a relevant sample of requests, and sharing some of the methods used in in establishing these "synonymous values (or functional equivalents)". By doing this, GlobalSign intends to lead the discussion which will should lead to further standardization.

This investigation has not been completed and this document is not final. This draft has been made available to request feedback to the wider PKI community and to explore what additional information could be provided that may be of use in this discussion.

Introduction/background

The Baseline Requirements do not provide a lot of information about the acceptable values in subject:localityName or in subject:stateOrProvinceName (unlike the reference to ISO 3166-1 for subject:countryName). In both cases, the Baseline Requirements merely say that at least one of these fields is required if the subject:organizationName field, subject:givenName field, or subject:surname field are present and the other field is present.

The Baseline Requirements provide a limited set of rules on how the validation of the contained value must be performed. This set of requirements covers a wide range of validation options with different levels of assurance – assurance levels that will partially depend on the interpretation of the Certificate Authority of certain key terms like "Reliable Data Source". The requirements even allow "any other option" in the last line of section 3.2.2.1 which ends in "*or other form of identification that the CA determines to be reliable*."

Additionally, unlike the EV Guidelines, that include an appendix which "provides alternative interpretations of the EV Guidelines for countries that have a language, cultural, technical, or legal reason for deviating from a strict interpretation of the EV Guidelines.", no such guidance is provided for the Baseline Requirements. It is worthy of note that even in the EV Guidelines, this alternative interpretation is only limited to the Organization name, and some Japan specific procedures – and therefore not of help when it comes to the subject:localityName or the subject:stateOrProvinceName fields. It may be argued that this allowance is not needed for the Baseline Requirements, as sufficient options are provided.

While on the subject of the EV Guidelines, another related but at the same time distinctly different problem arises there with the following fields: subject:jurisdictionLocalityName, subject:jurisdictionStateOrProvinceName and subject:jurisdictionCountryName. For these fields, the EV Guidelines prohibit "information that is not relevant to the level of the Incorporating Agency or Registration Agency". This Juridiction of Incorporation is defined as follows: "*In the context of a Private Organization, the country and (where applicable) the state or province or locality where the organization's legal existence was established by a filing with (or an act of) an appropriate government agency or entity (e.g., where it was incorporated). In the context of a Government Entity, the country and (where applicable) the state or province where the Entity's legal existence was created by law*.". It is important to note that this cannot necessarily be established by the territorial cover of the Incorporation source used (although it may be a good indicator). GlobalSign is of the opinion that standardization in this area is also important, and will contribute to this effort as well. However, this topic is out of scope for the current version of this document.

These two challenges are currently barriers to meet the Applicant's needs in a standardized way in certain countries, let alone introduce a common set of whitelist values for each country.

We recognize that in some countries, the definition of a Locality or State/Province is more clear than others – as is the way to validate this information. For these countries it would be more straightforward to introduce a common set of whitelisted values. In future versions of this document we would like to propose certain candidate countries for this common set of whitelisted values, as well as ways to identify other candidate countries to introduce this set of whitelisted values in a phased approach.

BR: Certificate Field: subject:localityName (OID: 2.5.4.7)

Required if the subject:organizationName field, subject:givenName field, or subject:surname field are present and the subject:stateOrProvinceName field is absent.

Optional if the subject:stateOrProvinceName field and the subject:organizationName field, subject:givenName field, or subject:surname field are present.

Prohibited if the subject:organizationName field, subject:givenName, and subject:surname field are absent.

Contents: If present, the subject:localityName field MUST contain the Subject's locality information as verified under Section 3.2.2.1. If the subject:countryName field specifies the ISO 3166-1 user-assigned code of XX in accordance with Section 7.1.4.2.2(g), the localityName field MAY contain the Subject's locality and/or state or province information as verified under Section 3.2.2.1.

BR: Certificate Field: subject:stateOrProvinceName (OID: 2.5.4.8)

Required if the subject:organizationName field, subject:givenName field, or subject:surname field are present and subject:localityName field is absent.

Optional if the subject:localityName field and the subject:organizationName field, and subject:givenName field , or subject:surname field are present.

Prohibited if the subject:organizationName field, subject:givenName field , or subject:surname field are absent.

Contents: If present, the subject:stateOrProvinceName field MUST contain the Subject's state or province information as verified under Section 3.2.2.1. If the subject:countryName field specifies the ISO 3166-1 user-assigned code of XX in accordance with Section 7.1.4.2.2(g), the subject:stateOrProvinceName field MAY contain the full name of the Subject's country information as verified under Section 3.2.2.1.

BR: 3.2.2.1 Identity

If the Subject Identity Information is to include the name or address of an organization, the CA SHALL verify the identity and address of the organization and that the address is the Applicant's address of existence or operation. The CA SHALL verify the identity and address of the Applicant using documentation provided by, or through communication with, at least one of the following:

1. A government agency in the jurisdiction of the Applicant's legal creation, existence, or recognition;
2. A third party database that is periodically updated and considered a Reliable Data Source;
3. A site visit by the CA or a third party who is acting as an agent for the CA; or
4. An Attestation Letter.

The CA MAY use the same documentation or communication described in 1 through 4 above to verify both the Applicant's identity and address.

Alternatively, the CA MAY verify the address of the Applicant (but not the identity of the Applicant) using a utility bill, bank statement, credit card statement, government-issued tax document, or other form of identification that the CA determines to be reliable.

EVG: Subject Jurisdiction of Incorporation or Registration Field

Certificate fields:

Locality (if required):

subject:jurisdictionLocalityName (OID: 1.3.6.1.4.1.311.60.2.1.1)

State or province (if required):

subject:jurisdictionStateOrProvinceName (OID: 1.3.6.1.4.1.311.60.2.1.2)

Country:

subject:jurisdictionCountryName (OID: 1.3.6.1.4.1.311.60.2.1.3)

Required/Optional: Required

Contents: These fields MUST NOT contain information that is not relevant to the level of the Incorporating Agency or Registration Agency. For example, the Jurisdiction of Incorporation for an Incorporating Agency or Jurisdiction of Registration for a Registration Agency that operates at the <u>country</u> level MUST include the country information but MUST NOT include the state or province or locality information. Similarly, the jurisdiction for the applicable Incorporating Agency or Registration Agency at the <u>state or province</u> level MUST include both country and state or province information, but MUST NOT include locality information. And, the jurisdiction for the applicable Incorporating Agency or Registration Agency at the <u>locality</u> level MUST include the country and state or province information, where the state or province regulates the registration of the entities at the locality level, as well as the locality information. Country information MUST be specified using the applicable ISO country code. State or province or locality information (where applicable) for the Subject's Jurisdiction of Incorporation or Registration MUST be specified using the full name of the applicable jurisdiction.

Synonymous values (or functional equivalents)

While we have established the difficulties any CA may encounter in the validation of the subject:localityName or the subject:stateOrProvinceName fields, we have not yet defined what we mean with "synonymous values (or functional equivalents)" for locality or state/province. In our reports, we have stated: "*Customers may have a perfectly good reason not to want the exact value that our Validation Specialists can find in a QGIS, but a synonymous value, or a functional equivalent".*"

When we say a certain Locality and State/Province is a synonymous value (or a functional equivalent) of another Locality and State/Province value, we say that it has the same meaning as this other Locality or State/Province. It fulfils the same function: when you look at the map, both values will point to the same place. Our Customers request these synonymous values (or functional equivalents) for a variety of reasons: commercial, legal, and sometimes political.

A synonymous value (or a functional equivalent) requires validation and obviously cannot be misleading. In many of the circumstances below, we point out that the synonymous value (or a functional equivalent) may be more clear to relying parties where to situate the Subscriber.

Below, we set out to describe both broad categories as well as describe some available methods used to establish this "synonymous value (or functional equivalent)" for these different categories. Please note: all methods described are permissible for the validation of these fields for OV certificates, but not all methods are permissible for these fields in EV certificates.

Transliteration
Transliteration in this context means the conversion of a text from one script to another. This should be treated as a pure translation.

In the reviewed sample, transliterations were always applied from a non-Latin script to Latin script (for example, Cyrillic to Latin in Москва to Moskva), although it would also (at least in principle) be possible in a different direction. Our hypothesis for the lack of transliteration from Latin script to non-Latin script in our reviewed sample is that there is little practical need for customers to do this. We would allow it if this was required by our Subscribers.

Examples:
- Москва to Moskva
- Αθήνα to Athína
- 東京 to Tōkyō
- 广州市 to Guangzhou

The following methods of validation of the synonymous values (or functional equivalents) can be used for this category:
- Confirm the information via another Qualified Information Source (mostly QIIS).
- Confirm the transliteration in a Government source (not a QGIS, but a Government source) or a source of a supranational organization (e.g. EU, UN).
- Applying a transliteration system recognized by the International Organization for Standardization.
    - ISO 9 - Cyrillic (Russian, Bulgarian, Belarusian, Ukrainian, Serbian and Macedonian)
    - ISO 233 - Arabic
    - ISO 259 - Hebrew
    - ISO 843 - Greek
    - ISO 3602 - Japanese
    - ISO 7098 - Chinese
    - ISO 9984 - Georgian

- o ISO 9985 - Armenian
  - o ISO 11940 - Thai
  - o ISO 11940-2 - Thai
  - o ISO 15919 - Indic scripts
- Applying a transliteration system recognized by the United Nations (http://www.eki.ee/wgrs/)
- For Japan, Audited Financial Statements (Financial Services Agency).
- Documentation supporting the information: utility bill, bank statement, credit card statement, government-issued tax document, Articles of Incorporation (the Articles of Incorporation must be verified to be authentic – either because they were received by GlobalSign directly from the incorporating or registering agency, or because they are accompanied either: by a document, signed with the original Japanese Corporate Stamp of which GlobalSign must verify the authenticity, that proves that the Articles of Incorporation are authentic and current, or by a Lawyer's Opinion or Accountant's Letter (which is itself verified in line with the EV Guidelines).
- A Lawyer's Opinion or Accountant's Letter (which is itself verified in line with the EV Guidelines) confirming the proper transliteration.

Endonyms and exonyms

For geographical places, whether it be Localities or States/Provinces, there is no such thing as a translation. Instead, there are two types of synonymous values (or functional equivalents):
- Endonyms, which are names of a geographical place in an official or well-established language occurring in that area where the feature is located. This usually happens in countries with multiple (un)official languages.
- Exonyms, which are external names for a geographical place. Exonyms are usually in English.

Endonyms and exonyms of the locality or the state/province are more prevalent for jurisdictions where non-Latin character sets are used, however, it is not exclusive to those jurisdictions. While English exonyms are the most prevalent It would also in principle be possible to use an exonym in a different language than English if so required (for example, for a localised version of the website). As with the transliteration, we found no examples of this in the sample size we reviewed.

Examples:
- Москва to Moscow
- Αθήνα to Athens
- 東京 to Tokyo
- 广州市 to Canton

The following methods of validation of the synonymous values (or functional equivalents) can be used for this category:
- Confirm the information via another Qualified Information Source (mostly QIIS).
- Confirm the information in a Government source (not a QGIS, but a Government source) or a source of a supranational organization (e.g. EU, UN). For countries that are multi-lingual there may be official resources available for translations of geographical place.
- Geonames has some exonyms/endonyms, but the information is not exhaustive, and the values are marked as alternative names only.
- Other reliable data sources.
- For Japan, Audited Financial Statements (Financial Services Agency).

- Documentation supporting the information: utility bill, bank statement, credit card statement, government-issued tax document, Articles of Incorporation (the Articles of Incorporation must be verified to be authentic – either because they were received by GlobalSign directly from the incorporating or registering agency, or because they are accompanied either: by a document, signed with the original Japanese Corporate Stamp of which GlobalSign must verify the authenticity, that proves that the Articles of Incorporation are authentic and current, or by a Lawyer's Opinion or Accountant's Letter (which is itself verified in line with the EV Guidelines).
- A Lawyer's Opinion or Accountant's Letter confirming the translation.

Abbreviations

Sometimes the customer will request a shortened form of a locality or state. This happens more for Province/States than for localities, although we sometimes find (part of) the locality name abbreviated as well.

Examples:
- BW for Baden-Württemberg
- MA for Massachusetts
- Saint Louis - St. Louis

The following methods of validation of the synonymous values (or functional equivalents) are used for this category:

- Confirm the information via another Qualified Information Source (mostly QIIS).
- Confirm the abbreviation in a Government source (not a QGIS, but a Government source – at GlobalSign, we distinguish between QGIS and Government sources) or a source of a supranational organization (e.g. EU, UN).
- Confirm that the abbreviation is an ISO 3166-2 code (or the second element of that code) for the verified State. Please note that this is not always desirable, as some of the ISO 3166-2 codes are not easily recognized as such, nor is their meaning obvious.
- Geonames has some abbreviation, but the information is not exhaustive.
- Other reliable data sources.
- For Japan, Audited Financial Statements (Financial Services Agency).
- Documentation supporting the information: utility bill, bank statement, credit card statement, government-issued tax document, Articles of Incorporation (the Articles of Incorporation must be verified to be authentic – either because they were received by GlobalSign directly from the incorporating or registering agency, or because they are accompanied either: by a document, signed with the original Japanese Corporate Stamp of which GlobalSign must verify the authenticity, that proves that the Articles of Incorporation are authentic and current, or by a Lawyer's Opinion or Accountant's Letter (which is itself verified in line with the EV Guidelines).

Different subdivision

Locality, State and even Country may have different meanings in different jurisdictions.

Even something as uncontested as country may lead to confusion – Wales, Scotland, England and Northern Ireland are countries of the United Kingdom. This doesn't usually pose a problem because the acceptable value for Country is well-defined – however, in the cases a country doesn't have an ISO 31622 2 character code, and then the lack of definition of what it means to be a "country"

becomes a problem. This mirrors a difficulty in international law, where the constitutional elements of what makes a "Sovereign State" are an ongoing discussion (e.g. whether recognition is required, and if so, what constitutes enough recognition. Or what is required for an acceptable declaration of Independence. For example Kosovo vs. Transnistria).

In a lot of Incorporation or Registration Sources, the "State or Province" will simply not be mentioned. This is usually the case when incorporation happens at national level, in countries where there is no concept of "State or Province" or where the terms "State or Province" will have multiple meanings. It is important to note that not all of the official subdivisions are covered by ISO in their ISO 3166-2 standard (e.g. Belgium doesn't just have 3 regions and 10 provinces, but also 3 communities).

Examples:
- Boulonne-su-Mér in Hauts-de-France (region) and Pas-de-Calais (department)
- Bournemouth in Dorset (Ceremonial county), South West England (Region) and England (Country)
- Seville in Seville (Province) and Andalusia (Autonomous Community)
- Baguio in Cordillera Administrative Region (Region), Benguet (Province) and Lone district of Baguio City (Districts).

Locality suffers from the same problem, although usually something that could be considered a locality is mentioned in the QGIS. Sometimes the QGIS does have multiple values that could be considered Locality – we find that this mostly happens in jurisdictions where the QGIS performs limited checks on self-reported information.

Examples:
- Kessel-Lo – Leuven. The previous address in the QGIS for GlobalSign NV in Belgium showed Leuven as the locality, while the physical address in the Federal Gazette showed Kessel-Lo. Kessel-Lo is a subdivision of Leuven.

The following methods of validation of the synonymous values (or functional equivalents) can be used for this category:
- Confirm the information via another Qualified Information Source (mostly QIIS).
- Confirm the information in a Government source (not a QGIS, but a Government source – at GlobalSign, we distinguish between QGIS and Government sources) or a source of a supranational organization (e.g. EU, UN).
- Geonames has some subdivisions of Localities for certain jurisdictions, but even when this information is provided, it is not complete or consistently applied, not in a consistent format, and are just marked as an alternative name along with other alternative names
- Confirm Locality or State or Province via a check on postal code or street address in reliable data sources. For example
  - via national mail providers;
  - sources that determine the jurisdiction of courts over a specific territory.
- For Japan, Audited Financial Statements (Financial Services Agency).
- Documentation supporting the information: utility bill, bank statement, credit card statement, government-issued tax document, Articles of Incorporation (the Articles of Incorporation must be verified to be authentic – either because they were received by GlobalSign directly from the incorporating or registering agency, or because they are accompanied either: by a document, signed with the original Japanese Corporate Stamp of which GlobalSign must verify the authenticity, that proves that the Articles of Incorporation are authentic and current, or by a Lawyer's Opinion or Accountant's Letter (which is itself verified in line with the EV Guidelines).
- A Lawyer's Opinion or Accountant's Letter confirming the translation.

Other alternatives (to be completed)

This category is mostly applicable to Locality, although we have seen examples in State or Province as well. We already touched on the example of 's-Gravenhage vs Den Haag (and similarly "'s-Hertogenbosch" and "Den Bosch", although "Den Bosch" is only colloquial), which is an official alternative.

Some subcategories are as follows:
- Official Locality or State/Province name may include an element which says what type of locality or State/Province it is. This element would sometimes be excluded. For example:
  - In Scotland: Glasgow City - Glasgow
  - In Ireland: County Laois - Laois
  - In the USA: New York City – New York
  - In Japan:
    - 渋谷区 - Shibuya-ku – Shibuya
    - 東京都 - *Tōkyō-to* - Tokyo Metropolis - Tokyo
- Elements of the official locality name may be abbreviated, even not the whole locality name is abbreviated "Frankfurt am Main", which is often referred to as "Frankfurt (Main)" or "Frankfurt a. M.".
- Politically sensitive: Derry and Londonderry.

Draft conclusion

We have highlighted a challenge which is a barrier to meet Applicants' needs in a standardized way in certain countries, let alone introduce a common set of whitelist values sets of Locality and State/Province for each country. This is caused in part by a lack of defined alternative interpretations for the validation of the Locality and State/Province for any countries that have a language, cultural, technical, or legal reason for deviating from a strict interpretation of the Baseline Requirements or the EV Guidelines.

In order to overcome this barrier, we have shared a list of categories of "synonymous values (or functional equivalents)" for locality or state/province that we have encountered in our reviewed sample. We believe the Applicant should be entitled to request these in their certificates. We have also provided an overview of our validation methods of these synonymous values (or functional equivalents).

We recognize that in some countries, the definition of a Locality or State/Province is more clear than others – as is the way to validate this information. These countries would be better suited to introduce a common set of whitelisted values.

We would like to propose a phased approach where we tackle different problems in parallel:

- Standardized approach in validation of "synonymous values (or functional equivalents)".

- Whitelist values: In future versions of this document we would like to propose certain candidate countries for this common set of whitelisted values, as well as ways to identify other candidate countries to introduce this set of whitelisted values in a phased approach.

  A whitelist would need to include the permitted values for both fields. It may look something like this.

Country	Locality	Locality alternative	State	State alternative
US	New York City	New York NYC	New York	NY
RU	Благовещенск	Blagoveshchensk	Amurskaya oblast	Amur oblast RU-AMU

We have not discussed in depth the related but at the same time distinctly different problem for the following fields: subject:jurisdictionLocalityName, subject:jurisdictionStateOrProvinceName and subject:jurisdictionCountryName. GlobalSign is of the opinion that standardization in this area is also important, and will contribute to this effort as well. However, this topic is out of scope for the current version of this document.