Response to Keith Drazek’s note re the KIND data element

Steve Crocker

18 August 2021

Keith wrote:

1. There appears to be momentum to support a standardized data element that MAY be used by Registrars if they choose to differentiate between legal and natural persons, and/or whether a registration data set contains personal data. That said, based on recent input and discussion, moving this forward to a consensus recommendation appears likely only if the full group agrees that any disclosures or use of the data element(s) would occur within a restricted system such as SSAD.

1. It has also been suggested that the KIND data element within RDAP, which already exists, could be modified to become fit for purpose. Whether it is or not, additional specificity is required on how the data element(s) will achieve the purpose of our policy objective. Without that specificity, the benefit of Recommendation #3 may be difficult to define or implement.

1. Doing this extra work now depends in part on whether we can reach consensus around the development/use of a standardized data element within a restricted system such as SSAD, so that may be a gating question to be answered in short order.

**Assignment:**

If there is general agreement among the full team that a standardized data element would be used within a restricted system, the small team will develop a proposal that can better inform Recommendation #3 and, if adopted, better inform the parties who would implement it. As such, for the purpose of this very focused work, the topics of transfer of the data element from Registrar to Registry and the publication of data element(s) to a public directory are out of scope. If there’s general agreement to proceed, the proposed specific tasks of the small group are:

• Is the KIND RDAP data element fit for purpose of differentiation or the indication of whether the registration data contains personal data?

• If not, what data element(s) need to be created?

• What are the value types for these data element(s) and its respective definition?

• Revision of Recommendation #3 text.

• Indication of what ICANN Org must do vs. IETF or other standards bodies.

Crocker’s response

There are two possible approaches. They’re more or less equivalent each with mild pros and cons.

One approach is to extend the allowed values for KIND. The other is to create a new data element

In brief, it appears we need a way to express the legal status of the registrant, i.e. whether the registrant is a Natural Person or a Legal Person. It also appears there needs to be a way to express whether the registrant data contains personal information. It’s probably safe to assume the data about a Natural Person does contain personal information, although I can also imagine arguments that it’s possible for a Natural Person to be a registrant and nonetheless declare the data does not contain personal information. The situation re Legal Persons is more straightforward. The data about a Legal Person may or may not contain personal information, and there’s no simple way to guess based on the name of the registrant.

In a fashion similar to the treatment of the rules governing collection, in addition to the binary choice of Natural vs Legal and Personal vs Non Personal, there are two forms of incomplete information that have to be included in the set of values. One form of incompleteness arises if the registrant insists on saying they don’t want to specify which category they belong in. A different form of incompleteness arises if the question is not asked.

The definition of these values is separate from the policy questions of whether this data must, may, or must not be collected, and is also separate from how this data is to be used if it is collected.

Here are suggestions for the two approaches. The bit patterns are concrete suggestions for those who like to see this level of detail, but they are just suggestions. The details should be left to the implementation team. That said, there is a bit of complexity in the use of these values. A policy statement may need to refer to a set of possible values and not just to a single value. The details are similar to what I showed in my memo on collection policies.

**Combined Approach -- Extend the set of values defined for Person**

000000 Blank, indicating the question either wasn't asked or wasn't answered.  This is distinct from the registrant declaring they won't specify.

000001 Unspecified.  The registrant says they won't say.

000010 Natural person.

000100 Legal person, but the registrant either has not been asked or has not responded as to whether there is personal data in their contact data

001100 Legal person who has said they won't say whether there is personal data

010100 Legal person with personal data

100100 Legal person with no personal data

(The order of the bits doesn't matter, but the structure does.)

**Two Data Element Approach**

A similar but different approach is to use two data elements, one for the legal status of the registrant, and one for whether the registrant data contains personal information

Legal Status (KIND)

000 Blank, indicating the question either wasn't asked or wasn't answered.  This is distinct from the registrant declaring they won't specify.

001 Unspecified.  The registrant says they won't say.

010 Natural person

100 Legal person

Personal Data

000 Blank, indicating the question either wasn't asked or wasn't answered.  This is distinct from the registrant declaring they won't specify.

001 Unspecified.  The registrant says they won't say.

010 There is personal data

100 There is no personal data