OIDs: how can I express you? Let me count the ways

Steven J Steindel

ABSTRACT
An object identifier (OID) has a central utility in providing a traceable source for the meaning of an identifier appearing in a cross-system communication. The views in this paper illustrate the problems with using the present OID registration system as a reliable source for the identifier, the confusion that the use of an OID introduces in messages, and the redundancy that the OID introduces at the expense of increased message size and no new content. In promoting clearly defined cross-system communication identifiers, Health Level 7 developed a standard that required use of OIDs outside of network addressing. This standard and its propagation by others may have paradoxically added more confusion than clarity.

An object identifier (OID) has a central utility in providing a traceable source for the meaning of an identifier appearing in a cross-system communication. From an International Standards perspective, OIDs derive from recommendations of the International Telecommunication Union Technical Sector (ITU-T Rec X.660) on interacting between networks and specific International Standards Organization (ISO 9834 Series) standards.2

ITU-T X.667 defines universally unique identifiers (UUID) to include the subset of globally unique identifiers (GUID) that are used mostly as short-lived instance identifiers, for example, for real-time internet transactions.5 GUIDs are not the focus of this discussion. The focus is on OIDs that are meaningful permanent hierarchically assigned UUIDs, registered and based on the ASN.1/ISO 9834 standards under a node jointly maintained by ITU-T and ISO. Under this system, US bodies registered to assign OIDs all start with the US identifier: 2.16.840 indicating a base under the commonly maintained ITU-T/ISO node of the registry system. Prior to 1991 the US identifier was maintained under the ISO node (1.2.840). The nodes that had been established in that branch still exist, but the organizations involved were given new identifiers under 2.16.840 and requested to make all new additions there. US OID registry responsibility lies with the American National Standards Institute (ANSI, the US ISO representative) and the Department of State (US ITU-T representative).

The primary use of identifiers, either UUID or OID, is in computer networks. OIDs name object types in X.509 certificates5 issued for secure identification in connection with X.500 networks5 and the similar structure found in lightweight directory access protocol (LDAP) schemas,7 and as identifiers in simple network management protocol (SNMP) based systems.6 Health Level 7 (HL7) was one of the first healthcare standards bodies to recognize and embrace the need for OIDs. They have a node designation of 2.16.840.1.113883 (joint-iso-itu-t-country (16) us (840) company7 Health Level 7 (113883))6 with in excess of 5000 OIDs assigned for use in US healthcare.9 The Centers for Disease Control and Prevention (CDC) similarly maintains an OID tree under 2.16.840.1.11422.4 (joint-iso-itu-t-country (16) us (840) company5 Centers for Disease Control and Prevention (11422) PHIN6) for US Public Health and their Public Health Information Network (PHIN).10

USES OF OIDS IN HEALTHCARE
Intra-institution exchange of healthcare information generally proceeds without the need to electronically exchange OIDs. The owner of an identifier is usually well known or, if needed, is found through direct communication. Extension to point-to-point communication between trading partners, as is common with reporting of laboratory results from an independent laboratory to a provider, also does not generally need OIDs. When communication extends beyond these limited but common situations, the need to know who owns the identifier is critical in case questions arise concerning the meaning of the identifier.

In establishing the Public Health Information Network, CDC embraced the use of OIDs to unambiguously identify owners of objects. When formulating its OID registry, CDC decided to distinguish among many different types of entities relevant to electronic interchange in the public health environment, as noted in table 1.11 An organization identifier, assigned by a known authority, is critical, first for validating the uniqueness of an organization, and second for providing a source for metadata if contact is required. Note that this construct implies that it is the couplet of base root plus the specific identifier that is unique and not just the identifier. CDC found that a requirement exists to uniquely identify applications as they have permitted functions, known attributes, and defined behaviors within a network. CDC chose to separate those we have direct control over from those we did not. Unique application identification extends beyond the need for potential identification for troubleshooting errors. In a networked environment, applications will be independently communicating under...
Table 1 CDC healthcare OID usage areas

<table>
<thead>
<tr>
<th>CDC OID root</th>
<th>Symbolic use</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.16.840.1.114222.4.1</td>
<td>Partner organizations</td>
<td>Root OID for PHIN organizations</td>
</tr>
<tr>
<td>2.16.840.1.114222.4.3</td>
<td>Assigning authorities</td>
<td>Root OID for identifier assigning authorities</td>
</tr>
<tr>
<td>2.16.840.1.114222.4.3.1</td>
<td>General assigning authorities</td>
<td>Root OID for registering general assigning authorities for identifiers</td>
</tr>
<tr>
<td>2.16.840.1.114222.4.3.2</td>
<td>CDC software applications</td>
<td>Root OID for registering CDC software application instances</td>
</tr>
<tr>
<td>2.16.840.1.114222.4.3.3</td>
<td>External software applications</td>
<td>Root OID for registering external software application instances</td>
</tr>
<tr>
<td>2.16.840.1.114222.4.5</td>
<td>CDC coding systems</td>
<td>Root OID for CDC-authored and maintained coding systems</td>
</tr>
<tr>
<td>2.16.840.1.114222.4.8</td>
<td>Surveillance program areas</td>
<td>Root OID for surveillance program areas</td>
</tr>
<tr>
<td>2.16.840.1.114222.4.11</td>
<td>CDC value sets</td>
<td>Root OID for CDC defined value sets</td>
</tr>
<tr>
<td>2.16.840.1.114222.4.20.2</td>
<td>Persons</td>
<td>Root OID for person IDs, assigned through PHIN</td>
</tr>
</tbody>
</table>

Application OIDs identify the actual application requesting access. Generally, network protocols will assign other identifiers of limited lifetime for session use.

A second use of people and places OIDs involves who gave care and where. One prominent physician in the electronic communication arena anecdotal noted he now had at least 15 different identifiers for his practice needs. In HL7 identifiers for people and places are easily accommodated through the use of the universal service identifier data type, the first option of the list of allowed data types for people-place identifiers. Will OIDs add another? It has taken over two years to enumerate the provider community for the purpose of another identifier, the national provider identifier (NPI), and hopefully we do not want to repeat that process. If we do not use the NPI, what would we use that allows specific unique identification of the person and would be electronically available, if needed, while meeting privacy concerns? Similar observations exist regarding places, many of which have an NPI.

To suggest a resolution, we should look at the now decade-long transformation of the US health payment system from fragmented local standards to national ones as mandated by the Health Insurance Portability and Accountability Act of 1996 (HIPAA). Electronic transaction standards under HIPAA derive mostly from two groups: the American Standards Committee (ASC) X12 and the National Council for Prescription Drug Programs (NCPPD). Neither uses OIDs and each has strict guidance for what appears in fields. Extensive, expensive testing, however, is required when change occurs to validate that all adhere to the guidance. Fortunately all are undergoing this change with the introduction of the NPI for people and places; successful testing was completed in 2008 and NPI use is now required.

While it seems logical to use an OID representing the NPI as the standard for people and places, none exists at this time nor does one exist for the Centers for Medicare and Medicaid Services (CMS), the actual enumeration organization. If we agree on using the NPI in this role, why is an OID needed? If we do not use a common, readily accessible identifier here, how will we enumerate and provide access to the information? Finally, using a known identifier for this OID may not resolve the network address use case.

The next most common need in healthcare for OIDs is for code sets. HIPAA standards call for use of only a limited defined set of codes for each field and validation of correct use through extensive testing. Unlike the confined use case of HIPAA, the broader healthcare industry cannot limit use to a small subset of terminologies. Hence the need exists for OIDs to identify the specific code set down to the version level and in HL7 the use of an OID is suggested at essentially the same level as people and places. The need for an OID is exacerbated when a portion of one or more codesets, sometimes called a value set, best describes the data. Fortunately OIDs do not change and, once identified, can be cached for reuse at a site. Users needing more information on received specific terms can use the cached OID to find the exact definition from the source site or contact information for those maintaining the terminology. We do not know if the persistent and repetitive use of OIDs in messages will yield increased semantic knowledge of the term without extensive testing, as required by HIPAA, but the possibility exists.

The internet based services as commonly associated with Service Oriented Architecture can but do not use OIDs for addressing. Addressing on the internet is based on the universal reference indicator composed of two basic forms: a universal reference locator or a universal reference name. Physical locations of the service are found using a domain name service and can change. If either of these forms are assigned a permanent address by the service they may be expressed as an OID.
The last area of healthcare OIDs, as noted by the CDC list, can be collectively called “things” including other OID enumeration groups and network attributes like system or program identifiers. Hopefully this is a relatively small, bounded list, easy to enumerate, provide, and maintain. Most users would only use a subset of this small list that they could cache locally.

**THE OID REGISTRY**

Permanent, findable OIDs are essential for confidence building and any needed communication with the holder of the identifier. A semi-formal OID registry is maintained by France Telecom (http://www.oid-info.com/). On visiting this website one has the immediate impression that it is not meant to serve robust real-time needs of providing OIDs or their associated metadata. Review of their policy indicates that rapid posting of new OIDs or correction of existing ones is not facilitated, based in part on issues associated with validation of OID ownership.

In healthcare we see two general states of OIDs at a use site: cached and needed. Cached OIDs are those widely repeated in a site’s messages and, outside of possibly the first look-up, a registry is not needed. It is suspected that the most common OIDs in this class would pass from site to site in readily available formal or informal lists. Needed OIDs are those generally for people or places not routinely communicated to or reused by a site. As OIDs do not change, it is suspected that these would also be cached in case of later reuse, making a registry necessary only for the initial look-up. As essentially all sites will cache the OIDs they use, the need for a registry is not rapid return of an OID or its metadata, but completeness.

We can conclude from the above that the essential features of an OID registry are timeliness of generating new OIDs and completeness of the contained list. Neither are attributes of current registries or their associated registration authorities. The standards are quite clear that there are essentially none to minimal standards for registration authorities. While most countries have one, it is not required. Indeed, formal methods exist to have any UUID registered directly as an OID at http://www.oid-info.com, bypassing any registration authority. Investigation of the country nodes found at http://www.oid-info.com shows that each has a different approach. Looking at the CDC within that registry shows that none of the OIDs in table 1 are registered. Discussion with those maintaining the various nodes of the CDC registry and the corresponding HL7 one reveals many stories of time delays and errors in posting the OIDs. Evidence shows that none of the current OID registries or registration authorities directly involved in healthcare are timely or complete.

**THE DEFAULT OID**

The knowledge of an OID across essentially all sites, especially as time evolves and they are cached, and the associated prescribed code sets along with implied meanings of objects that appear in standards leads to the suggestion of a concept I call the default OID. The implication of the default OID is that an OID is known and does not need to be stated explicitly.

An example of a default OID is the HL7 registered OID for SNOMED CT (2.16.840.1.113883.6.96). In figure 6 in the article describing clinical document architecture (CDA), release 2, this OID appears four times in the space of 37 XML tags. Does that redundancy add any clarity or new information? While this specifically points to a publication being used for another purpose, let us assume that the CDA document referred to had a specific implementation guide that called for the use of SNOMED CT for these tags. Indicating the OID for SNOMED CT just indicates we are conforming to the standard and provides even less new information.

In instances like this, usefulness and readability of the document is increased if we indicate the OID not by the actual value but by some default indicator. For those who feel that the explicit OID needs stating in the transmitted message or document, I suggest we introduce an OID macro-like syntax similar to define in many programming languages such as:

```
#OID SNOMED_CT code system="2.16.840.1.113883.6.96"
codeSystemName="SNOMED CT"
```

to appear before any statements in the message or document. The repetitive

```
<code code="27597001" code system = "2.16.840.1.113883.6.96"codeSystemName = "SNOMED CT"
displayName="Family history of cancer"/>
```

could be replaced by the more readable and simpler

```
<code code="27597001" SNOMED_CT displayName="Family history of cancer"/>
```

without loss of information content, and would have added human readability and less message bandwidth, an increasing issue with XML formats. While this macro-construct is not currently available in underlying standards, I believe the utility is evident and wide usage in programming languages would make addition easier.

Another syntax that is feasible for the default OID is one that makes use of the ubiquitous nature of the OID. In this syntax the explicit macro is not added to the message, but those using healthcare OIDs agree that SNOMED_CT represents the registered OID for that code set when it appears in a message location that requires an OID.

The above example is a bit trivial and, while it has utility, is really replacing something that is self-evident with another syntax for the same self-evident item. The real utility of the default OID is not for code or value sets but in people and places. To illustrate this example I will use the recently passed HL7 Implementation Guide for Version 2.51 Laboratory Results Reporting; those readers interested in the actual HL7 content should refer to the Guide.

A clinical laboratory result is normally reported on a person and that person’s reference is through an identification number. A person having testing done by provider A may have a number of 1254 and the same person may have a number of 4567 with provider B. Hence the unique person ID is really the couplet of the provider ID plus that provider’s assigned person identifier. As called for in the implementation guide, the provider ID is represented by an OID. As noted above, our present assignment of provider OIDs is at best unclear and in reality non-existent. One can argue that in essentially all laboratory result reporting, whether point-to-point or otherwise, at least two default identifiers exist for provider’s ID. The first is the required receiver’s address in the message envelope, defined in HL7 by the MSG segment as it would be the exception that the receiver was not the originator of the test and the holder of the patient identifier. I propose a default OID of MSG_HEADER for this example. The second example has more certainty of being the actual party who would be a source for the patient ID, but reporting is not required by the reporting laboratory: the actual Ordering Provider. For this example I propose a default OID of REQUESTOR. One can express these default OIDs either using

---

13Until 2003 Harold Alvestrand maintained an OID registry that was subsumed with

15Does that redundancy add any clarity or new information? While this
the macro syntax for explicit identification in each message or by inference that assumes the use stated in this paragraph.

A circular OID issue exists with either of the above examples for assigners of patient identification numbers. The default OIDs presume an assignment of underlying OID. In the MSG_HEADER example this has to be a minor issue as some registration authority, which may not be associated with healthcare, has to have responsibility for the receiver’s address or the transport system will not know how to deliver it. The REQUESTOR assumes that an OID exists for the registration authority for that ID. As discussed above with OIDs for people and places, what constitutes a valid OID in this area is unclear. Under recent regulations associated with HIPAA, CMS has essentially made the NPI the authoritative identifier for people and places that provide healthcare. If CMS had an OID or one existed for the actual NPI enumeration system, the National Provider System, we would solve this issue, but the NPI exists mostly for X12 HIPAA transactions that do not require OIDs. Even if we resolved the OID issue for an NPI, it still exists for the many other identifiers people and places have, including the common taxpayer ID. It seems simplest to resolve this circular OID issue by not requiring that the underlying OID exist for a set of defined, common people and place identifiers. Doing so would make the use of a default OID easier.

Reasonable boundaries need to exist around OIDs for people and places when a common identifier such as the NPI is used. If an ambulatory care electronic health record, or the provider’s office itself, is assigning patient numbers, does it need a unique OID? Can we set a reasonable default OID such as SITE_EHR? The issue is not just confined to office practices. Multi-site facilities with a master patient index have a similar need. Without boundaries it is easy to envision each office and maybe each examination room having an OID, an absurd but possibly technically correct case.

SUMMARY
While it seems axiomatic that the owner of an identifier should be unambiguously identified when using that identifier, I have demonstrated here the difficulties associated with the seemingly simple task when using an OID for that task. A major, simple problem is that we have used many different identifiers for the same or similar roles for too many years, none having OIDs but all with known assigning authorities. Creating OIDs just to have OIDs, as HL7 sometimes does, for these legacy identifiers or their assigning authority serves no useful purpose. Even if we did create these OIDs, there is no timely, complete central registry that one could consult to retrieve them. While the expense of mechanically creating such a registry could be low, the cost of populating and maintaining it would not be.

I have also questioned the basic premise that OIDs are needed as the unambiguous identifiers in many situations. In many instances where an OID is required, the knowledge of that OID is known by default. Using or repeating it multiple times lends little clarity or completeness to the information. For those that still feel OIDs are required, introduction of a default OID that self-defines the obvious is a useful course to consider but could require changes in the underlying standards.

One instance where we cannot avoid the use of true OID was discussed above and needs repeating. OIDs are the basis of network address and, in messaging outside of web-based services, they must continue to be used for that purpose. These OIDs generally come from nodes of the OID registries not associated with healthcare, like the ITU-T node or the PHIN LDAP (partner organizations in table 1).

Finally, I have suggested that in the real world OIDs would be implemented by rote copying of lists from others for those commonly used and caching them to eliminate look-up. That system is feasible because OIDs will not change, by definition. If we are dealing with an unchanging quantity that represents a known, why use it when the information is already present in another form?

Implementation of interoperable systems across networks is a complex undertaking. Unless an understandable and reproducible way of identifying the owner of an identifier becomes standard, the patient safety issues of having incorrect information or having a false sense of security that the OID is correct and meaningful are vast. Those that use OIDs today, including Public Health, are finding that one of the most difficult steps is to determine the “correct” OID. Instead of introducing an operational barrier to interoperable systems, we need to re-explore the necessity of the OID and provide a more workable, reasonable, and simpler solution.

Acknowledgments The views in this paper are those of the author but represent years of communication with peers on the difficult and poorly understood question of OIDs. In particular he wishes to note the contributions of Austin Kreisler and Dan Russler for their careful reading and comments on the manuscript during development.

Competing interest None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES