# Introducing

# ✱DIAD

✱DIAD or StarDIAD (Domain-name Identifier and Directory) is an innovative approach to electronic identifiers using the Domain Name System (DNS) to produce identifiers that are globally unique but locally generated, efficient but personalised, a name not a number. (✱DIAD patent application has been lodged).

## *FUNCTIONAL DESCRIPTION OF ✱DIAD*

✱DIAD is a method of providing electronic identifiers for large target populations, be they persons, organisations, places or objects in any combination.

The population to be named has the following characteristics:

- The population can readily be arranged into a hierarchy of entities. In one aspect, each entity falls into one of two categories; it is either a *simple entity* or a *naming authority entity*. Naming authority entities will hereinafter be referred to as *naming authorities*. Any nameable entity can be a simple entity, but in general only organisations will be naming authorities.
- Each entity except one, the *root naming authority*, is assigned its identifier by a responsible naming authority. Each naming authority also has an identifier. There is no difference in kind between the identifiers assigned to naming authorities and to simple entities.
- Each naming authority may issue identifiers to simple entities, and it may issue identifiers to other naming authorities. There may, consequently, be a multi-level hierarchy of naming authorities and simple entities in the form of a tree, with branches emanating from each naming authority, and with the end points of all branches being simple entities. These end points are known as the leaf nodes of the tree. The identifier of the naming authority that issues any particular identifier is known as the *parent* of the particular identifier. Correspondingly, the identifiers which are issued by a naming authority are known as the *children* of the authority's identifier. Note that, strictly speaking, the terms *parent* and *children* refer in this discussion to *identifiers*, not *entities* as such. However, the entities may loosely be referred to as parents and children, while keeping the strict meaning in mind.
- Only one naming authority is responsible for any identifier. The root naming authority issues an identifier to itself. Note that while a naming authority is responsible for all of the identifiers it issues, it is not responsible for identifiers issued by any naming authorities it has identified. Such children are in turn responsible for identifiers they issue.

As a consequence of this structure, each entity can be uniquely named by combining the identifier of the issuing naming authority with a name component which is unique amongst the children of the issuing authority's identifier. That is, the problem of unique naming is, by this naming structure, reduced to the problem of uniquely naming only the direct children of each naming authority.

Naming, and associated responsibility, is thereby localised to each naming authority.

While the naming responsibility is localised, in order to be functional as a population-wide identifier, the identifier must be able to be checked by anyone with an interest in the population.

✱DIAD addresses these concerns by a novel application of an existing technology – the Domain Name System (DNS). The DNS is a distributed database that associates small sets of data with a key. The key is a hierarchically-organised name which is readily comprehensible and communicable by both human beings and computers. DNS names are ubiquitous on the Internet and on the World Wide

Web. They form part of all modern email addresses, and part of every textual Universal Resource Locator (URL) used on the Web. Without DNS, the Web could not exist, and email could not function. DNS is consulted across the world many millions of times every day.

The key to the success of DNS is its hierarchical and distributed nature. There is no central repository of DNS names. Nor could there be. No centralised system could keep track of the constantly evolving space of DNS names. Technical and commercial experience of DNS systems is correspondingly widely distributed.

DNS was designed for a specific purpose: to associate human-comprehensible and human-communicable names with the numerical Internet addresses that underlie all Internet communication. These associations are described by *A* (address) and *AAAA* (IPv6 address) records within the DNS database. Many other types of records have been defined, and continue to be defined for DNS. To discover the Internet address associated with a given domain name, the DNS database is queried using the name as a key, and, if the name is present in the database, the address is returned.

The ***novelty*** of ✱DIAD is that the database for an entire sub-tree, or *zone*, of the DNS name space associates no Internet addresses with its domain names. A query using the domain name as a key will return a possibly empty set of *directory* data, not including A or AAAA records. The most important information returned from the query of a name within the zone is the fact of its existence or non-existence. The ***domain-names*** are just names; ***identifiers*** associated with human and corporate agents or other nameable entities. In a name of the form *A.B.C*, the structure of the identifier implicitly asserts that the entity named *A.B.C* has been *named by* the entity *B.C*. This implication provides a foundation for the hierarchical assertion of identity, because only the controller of the domain *B.C* can issue subdomains such as *A.B.C*.

The ***directory*** associated with each name is optional. If used, it is defined primarily in *TXT* (text) and *NAPTR* (naming authority pointer) resource records in the DNS database. Other resource records, for example *SRV* (server) and *LOC* (location) records may also be used. All information in the directory is public. However, such information may be represented directly and publicly, as in the form of TXT records, or may be represented indirectly, as in the form of NAPTR records, which re-direct enquiries to systems entirely external to DNS, where appropriate levels of access security may be enforced. That is, whilst the re-direction details are public, the information available from the re-directed query may be public or private.

The hierarchy of names would be based on functional, legal and practical categories determined by the nature of the application domain. Individual entities (be they persons, organisations, places, objects or any other nameable entity) would obtain or be assigned identifiers at their point or points of contact with the application domain. The characteristics of individual names would, again, be determined by the requirements of individual components of the application domain. For example, in the e-health system, individual patients could have "opaque" identifiers congruent with their need for privacy. On the other hand, registered professionals may be required to use an identifier which could be readily confirmed as belonging to a particular person.

DNS offers security at a number of levels. All of the names and associated information for a zone is regularly transferred between the master and slave nameservers of each zone. It is common practice to validate requests for this transfer by signing such requests, and to validate the transferred zone data by signing the return message. The signing is done with a key shared by the master and slaves. This process ensures that the complete zone data cannot be readily obtained by parties other than zone administrators. As a result, whilst the zone can be queried for individual names, the complete zone contents could only be discovered by exhaustive queries for all possible names within the zone.

In standard DNS, guaranteeing the validity of all of the data returned in response to queries is more difficult. It can be achieved by applying DNSSEC.bis (DNSSEC hereafter) security to the zones. Whether DNSSEC is used would depend on the security requirements of each application domain. However, in the case of the e-health application domain, it would be essential. When used, DNSSEC would have to be applied to *all* zones within the application domain sub-tree.

## *Examples*

**N.B. All of the following examples are notional. The particulars of domain names, and the syntax and semantics of directory information are beyond the scope of the ✱DIAD. They would be determined as part of the specification of the application domain.**

### Example: GP Clinic and Patients

Identifiers are not restricted to persons; organisations may also possess them. For example, a medical general practice within the Sunshine Coast Division of General Practice, may have the identifier:

*dh-clinic.scdgp.gpq.ehealth.id.au*

A patient called James M. Brown, attending a general practice clinic might have the identifier:

*james-brown002.dh-clinic.scdgp.gpq.ehealth.id.au*

Another patient of the same clinic might choose the the identifier:

*thegecko.dh-clinic.scdgp.gpq.ehealth.id.au*

In the first case, the patient is not concerned that a casual observer of the identifier would deduce that a person called James Brown, one of at least two living in the area served by the clinic, was the one referred to by this identifier.

In the second case, the patient does not want such conclusions to be drawn, and chooses a *handle*, i.e. pseudonym, as the individual component of the identifier. Such pseudonyms are now familiar from social networking sites on the Web. Note that in both cases, the usual medical records will be maintained, with the addition of this identifier. At the clinic, the patients will be known to the staff as usual. However, no casual observer would be able to make that association.

The identifying entity for both identifiers is *dh-clinic.scdgp.gpq.ehealth.id.au*.

### Example: Professional Study and Registration

Professional health care workers will generally have different levels of authorisation at different stages of their careers. For example, student, intern, registrar, specialist. If an identifier were to be associated with each stage, the identifying entity would be different in each case.

Assume health-care professionals are registered by a national body, known as National Registration and Accreditation Scheme (NRAS), and that accrediting educational institutions and training hospitals are affiliated with the NRAS.

If a student named John Stephen Smith were to enrol in medicine at Queensland University, for example, he might be assigned an identifier of the form:

*smith-john-s-19731106-01.uq.nras.ehealth.id.au*

On graduation, John Smith obtains an internship at Princess Alexandra Hospital, a recognised tertiary teaching hospital. He might be assigned a new identifier of the form:

*smith-john-s-19731106-01.pah.nras.ehealth.id.au*

Once his internship has been successfully completed, John Smith is eligible for registration with NRAS, when he might receive the identifier:

*smith-js-19731106-01.mbbs.nras.ehealth.id.au*

The person is identified as *smith-js-19731106-01*, whose identifier was provided by the entity *mbbs.nras* within the Australian health-care application domain. In the case of professional identifiers, it would probably be considered advantageous to have an identifier which is readily associated with the person identified, as in this example.

### Example: Directory Linking Multiple Identifiers

The professional registration example shows that multiple identifiers may both be required, and required to be linked. It would be necessary to be able to trace the path to accreditation of professionals within the system.

Linking could be achieved in a number of ways through DNS records. One possible method is to use NAPTR records to establish a double-linked list. The Application Unique String for this application would the identifier for which links were required. The First Well Known Rule would use the identifier unchanged to request NAPTR records. The expected service in the NAPTR records would be *alias* with either *next* or *prev* as secondary service protocols. The possible values of the service field would then be *alias+next* or *alias+prev*. An application program can then derive the complete set of links starting from any point in the chain.

If there were no aliases, no appropriate NAPTR records would be returned. If there were one alias only, both the *next* and *prev* entries would transform to the same alias. In other cases, the chain could be followed in either direction.

The same process allows the optional linking of multiple identifiers that individual patients may choose to have.

### Example: Location capability for service or care facilities location finding

Where a specific location is required, for example the exact location of a hospital, DNS provides a Location (LOC) Record function (known as LOC RR) that might be assigned or attributed to the e-identifier of the service. The LOC RR permits the addition of latitude, longitude and altitude to e-identifier – should this be required.

### Example: Directory Publicly Available Health Information for Patient

If a patient chose to have some demographic or medical information associated with his or her identifier, the TXT records associated with the identifier might contain the following strings:

birthyear=1960
eyes=blue
height=170cm
blood=Oneg
allergies=penicillin,bee-sting
medications=Warfarin:2mg

Such information is *public*; it would be available to anyone with Internet access.

### Example: Directory Private Health Information for Patient

If a patient chose to participate in a service supplying certain information from his or her medical records as held by the clinic, in an Internet-accessible service whose interface were well-known, a combination the patient-specific label (the first label of the domain-name) and the clinic identifier (the remainder of the domain-name) could be used in conjunction with NAPTR and SRV records associated with the clinic domain-name to access the service with a query about the patient. Note that the NAPTR and SRV records are publicly readable, but the service to which they point can be public or private. If private, this access would be subject to the authentication and access policies governing such sensitive data.

**Example: Directory Information for Professionals**

For a professional identifier, some publicly accessible information may be associated with the identifier by means of TXT records. It might, for example, be considered that a professional registration number should be published with the identifier. In addition, because the professional may also use his or her professional identifier as a patient identifier, the same basic personal medical facts may also be recorded.

Professional registration bodies may define interfaces to web services with confidential information concerning the professional's registration and history. Such services might be accessed through NAPTR and SRV records on the registration body's identifier, using that identifier and the professional's identifier to locate and query the service. As with patient data, this access would be subject to the authentication and access policies governing such sensitive data.

✱DIAD – a name, not a number.

For further information please contact either:

**Phil Johnson**

+61 (0) 412 561 453

phil@ehealth.id.au

**Peter West**

+61 (0) 431 665 100

pbw@ehealth.id.au