Network Working Group
Request for Comments: 2148
BCP: 15
Category: Best Current Practice

H. Alvestrand
UNINETT
P. Jurg
SURFnet

September 1997

Deployment of the Internet White Pages Service

Status of this Memo

This document specifies an Internet Best Current Practices for the Internet Community, and requests discussion and suggestions for improvements. Distribution of this memo is unlimited.

1. Summary and recommendations

This document makes the following recommendations for organizations on the Internet:

(1) An organization SHOULD publish public E-mail addresses and other public address information about Internet users within their site.

(2) Most countries have laws concerning publication of information about persons. Above and beyond these, the organization SHOULD follow the recommendations of [1].

(3) The currently preferable way for publishing the information is by using X.500 as its data structure and naming scheme (defined in [4] and discussed in [3], but some countries use a refinement nationally, like [15] for the US). The organization MAY additionally publish it using additional data structures such as whois++.

(4) The organization SHOULD make the published information available to LDAP clients, by allowing LDAP servers access to their data".

(5) The organization SHOULD NOT attempt to charge for simple access to the data.

In addition, it makes the following recommendations for various and sundry other parties:

(1) E-mail vendors SHOULD include LDAP lookup functionality into their products, either as built-in functionality or by providing translation facilities.

(2) Internet Service providers SHOULD help smaller organizations follow this recommendation, either by providing services for hosting their data, by helping them find other parties to do so, or by helping them bring their own service on-line.

(3) All interested parties SHOULD make sure there exists a core X.500 name space in the world, and that all names in this name space are resolvable. (National name spaces may elobarate on the core name space).

The rest of this document is justification and details for this recommendation.

The words "SHOULD", "MUST" and "MAY", when written in UPPER CASE, have the meaning defined in RFC 2119 [17]

2. Introduction

The Internet is used for information exchange and communication between its users. It can only be effective as such if users are able to find each other's addresses. Therefore the Internet benefits from an adequate White Pages Service, i.e., a directory service offering (Internet) address information related to people and organizations.

This document describes the way in which the Internet White Pages Service (from now on abbreviated as IWPS) is best exploited using today's experience, today's protocols, today's products and today's procedures.

Experience [2] has shown that a White Pages Service based on self-registration of users or on centralized servers tends to gather data in a haphazard fashion, and, moreover, collects data that ages rapidly and is not kept up to date.

The most vital attempts to establish the IWPS are based on models with distributed (local) databases each holding a manageable part of the IWPS information. Such a part mostly consists of all relevant IWPS information from within a particular organization or from within an Internet service provider and its users. On top of the databases there is a directory services protocol that connects them and provides user access. Today X.500 is the most popular directory services protocol on the Internet, connecting the address information of about 1,5 million individuals and 3,000 organizations. Whois++ is the second popular protocol. X.500 and Whois++ may also be used to interconnect other information than only IWPS information, but here we only discuss the IWPS features.

Note: there are other, not interconnected, address databases on the Internet that are also very popular for storing address information about people. "Ph" is a popular protocol for use with a stand-alone database. There are over 300 registered Ph databases on the Internet. Interconnection of databases however, is highly recommended for an IWPS, since it ensures that data can be found. Hence Ph as it is now is not considered to be a good candidate for an IWPS, but future developments may change this situation (see section 12).

Currently X.500 must be recommended as the directory services protocol to be used for the IWPS. However, future technology may make it possible to use other protocols as well or instead.

Since many people think that X.500 on the Internet will be replaced by other protocols in the near future, it should be mentioned here that currently LDAP is seen as the surviving component of today's implementations and the main access protocol for tomorrow's directory services. As soon as new technology (that will probably use LDAP) becomes available and experiments show that they work, this document will be updated.

A summary of X.500 products can be found in [14] (a document that will be updated regularly).

The sections 3-7 below contain recommendations related to the publication of information in the IWPS that are independent of a directory services protocol. The sections 8-11 discuss X.500 specific issues. In section 12 some future developments are discussed as they can be foreseen at the time of writing this document.

3. Who should publish IWPS information and how?

IWPS information is public address information regarding individuals and organizations. The IWPS information concerning an individual should be published and maintained by an organization that has a direct, durable link with this individual, like in the following cases:

- The individual is employed by the maintainer's organization
- The individual is enrolled in the university/school that maintains the data
- The individual is a (personal) subscriber of the maintainer's Internet service

The organization that maintains the data does not have to store the data in a local database of its own. Though running a local database in the X.500 or Whois++ service is not a too difficult job, it is recommended that Internet service providers provide database facilities for those organizations among its customers that only maintain a small part of the IWPS information or don't have enough system management resources. This will encourage such organizations to join the IWPS. Collection of IWPS information and keeping it up-to-date should always be in the hands of the organization the information relates to.

Within the current (national) naming schemes for X.500, entries of individuals reside under an organization. In the case of Internet service providers that hold the entries of their subscribers this would mean that individuals can only be found if one knows the name of the service provider. The problem of this restriction could be solved by using a more topographical approach in the X.500 naming scheme, but will more likely be solved by a future index service for directory services, which will allow searches for individuals without organization names (see section 12).

4. What kind of information should be published?

The information to be published about an individual should at least include:

- The individual's name
- The individual's e-mail address, in RFC-822 format; if not present, some other contact information is to be included
- Some indication of the individual's relationship with the maintainer

When X.500 is used as directory services protocol the last requirement may be fulfilled by using the "organizationalStatus" attribute (see [3]) or by adding a special organizational unit to the local X.500 name space that reflects the relation (like ou=students or ou=employees).

Additionally some other public address information about individuals may be included in the IWPS:

- The individual's phone number
- The individual's fax number
- The individual's postal address

- The URL of the individual's home page on the Web

In the near future it will be a good idea to also store public key information.

More information about a recommended Internet White Pages Schema is found in The Internet White Pages Schema [16]

Organizations should publish the following information about themselves in the IWPS:

- The URL of the organizations home page on the Web
- Postal address
- Fax numbers
- Internet domain
- Various names and abbreviations for the organization that people can be expected to search for, such as the English name, and often the domain name of an organization.

Organizations may also publish phone numbers and a presentation of themselves.

5. Data management

Data management, i.e. collecting the IWPS information and keeping it up-to-date, is a task that must not be underestimated for larger organizations. The following recommendations can be made with respect to these issues:

- An organization should achieve an executive level commitment to start a local database with IWPS information. This will make it much easier to get cooperation from people within the organization that are to be involved in setting up a Directory Service.
- An organization should decide on the kind of information the database should contain and how it should be structured. It should follow the Internet recommendations for structuring the information. Besides the criteria in the previous section, [3] and [4] should be followed if X.500 is used as directory services protocol.

- An organization should define criteria for the quality of the data in the Directory, like timeliness, update frequency, correctness, etc. These criteria should be communicated throughout the organization and contributing entities should commit to the defined quality levels.

- Existing databases within an organization should be used to retrieve IWPS and local information, to the greatest extent possible. An organization should involve the people who maintain those databases and make sure to get a formal written commitment from them to use their data source. The organization should rely on these people, since they have the experience in management and control of local, available data.

- The best motivation for an organization to join the IWPS is that they will have a local database for local purposes at the same time. A local database may contain more, not necessarily public, information and serve more purposes than is requested for in the IWPS. In connecting to the IWPS an organization must "filter out" the extra local information and services that is not meant for the public IWPS using the directory services protocol.

6. Legal issues

Most countries have privacy laws regarding the publication of information about people. They range from the relaxed US laws to the UK requirement that information should be accurate to the Norwegian law that says that you can't publish unless you get specific permission from the individual. Every maintainer of IWPS information should publish data according to the national law of the country in which the local database which holds the information resides.

Some of these are documented in [5] and [1].

A maintainer of IWPS information should also follow some common rules, even when they are not legally imposed:

- Publish only correct information.

- Give people the possibility to view the information stored about themselves and the right to withhold information or have information altered.

- Don't publish information "just because it's there". Publish what is needed and what is thought useful, and no more.

Given the number of data management and legal issues that are involved in publishing IWPS information, good consulting services are vital to have smaller companies quickly and efficiently join the IWPS. Internet service providers are encouraged to provide such services.

7. Do not charge for lookups

In the current IWPS it believed that due to today's technological constraints, charging users is harmful to the viability of the service. There are several arguments for this belief:

- Micropayment technology is not available at the moment.
- Subscription services require either that the customer sign up to multiple search services or that the services are linked "behind the scene" with all kinds of bilateral agreements; both structures have unacceptably high overhead costs and increase the entry cost to the service.
- The current directory services protocols do not support authentication to a level that would seem appropriate for a service that charges.

Therefore it is strongly recommended that all lookups by users in the IWPS are for free. This, of course, does not limit in any way the ability to use the same IWPS dataset to support other services where charging may be appropriate.

8. Use X.500

The IWPS based on the X.500 protocol has a relatively wide deployment. The current service contains about 1,5 million entries of individuals and 3,000 of organizations. It is coordinated by Dante, an Internet service provider in the UK, and known as "NameFLOW-Paradise".

Though X.500 is sometimes criticized by the fact that its functionality is restricted by the hierarchical naming structure it imposes, it provides a reasonably good functionality as has been shown in several pilots by organizations [5], [2], [6], [7] that are now running a production X.500 IWPS. User interfaces also determine the functionality the X.500 IWPS offers. Usually they offer lookups in the IWPS based on the following user input:

- The name of a person
- The name of an organization this person can be related to

- The name of a country

As a result they will provide the publicly available information about the person in question. Most user interfaces offer the possibility to list organizations in a country and users in an organization to help users to make their choice for the input. It may also be possible to use part of the names as input or approximate names.

Specific user interfaces can provide lookups based on other input, like e-mail addresses of people or postal addresses of organizations. Such possibilities may however violate privacy laws. Providers of directory services services may then be held responsible.

The X.500 naming scheme imposes the requirement on an interconnected IWPS that all entries stored in it must have unique names (the "naming scheme"). This is most easily fulfilled by registering all entries in a "naming tree" with a single root; this is the reason why the totality of information in an X.500 IWPS is sometimes referred to as the "Directory Information Tree"
or DIT.

Organizations are strongly encouraged to use the X.500 protocol for joining the IWPS. The current service is based on the X.500 1988 standard [8] and some Internet-specific additions to the protocol that connects the local databases [10] and to the access protocol [9]. Organizations should use X.500 software based on these specifications and additionally supports [11] for the transportation of OSI protocols over the Internet.

Organisations may connect to the NameFLOW-Paradise infrastructure with 1988 DSAs that don’t implement [10], but they will lack automatic replication of knowledge references. This will be inconvenient, but not a big problem. The 1993 standard of X.500 includes the functionality from [10], but uses a different potocol. Hence organisations that connect to the infrastructure with a 1993 DSA will also encounter this shortcoming. Section 12 "Future developments" explains why the infrastructure doesn’t use the 1993 standard for the moment.

For recommendations on which attributes to use in X.500 and how to use them (either for public IWPS information or additional local information the reader is referred to [3] and [4]. For specific non-public local purposes also new attributes (and object classes) may be defined. Generally it should be recommended to use as much as possible the multi-valuedness of attributes in X.500 as this will improve the searching functionality of the service considerably. For example, the organizationalName attribute which holds the name of an

organization or the commonName attribute which holds the name of a person should contain all known aliases for the organization or person. In particular it is important to add "readable" variants of all attributes that people are expected to search for, if they contain national characters.

Another recommendation that can be made is that replication of data [10] between local databases is used in order to improve the performance of the service. Since replicating all entries of a part of the IWPS from one local database in another may violate local privacy laws, it is recommended to restrict replication to country and organizational entries and knowledge references (which tell where to go for which part of the IWPS). Of course privacy laws are not violated when the replicating database is managed by the same organization as the one that masters the information. So local replication between two databases within the same organization is highly recommended.

In general replication within one country will usually be less a legal problem than across country borders.

Recommendations for the operation of a database in the X.500 infrastructure can be found in [12].

X.500 is not recommended to be used for:

- A Yellow Pages service with a large scope. See [5].

- Searching outside the limited patterns listed here, in particular searching for a person without knowing which organization he might be affiliated to.

- Publishing information in other character sets than ASCII, some of the Latin-based European scripts and Japanese (the T.61 character sets). While support for these character sets is available in revised versions of X.500, products that support the revision aren’t commonly available yet.

9. Use the global name space

Some people, for instance when using Novell 4 servers, have decided that they will use X.500 or X.500-like services as an internal naming mechanism, without coordinating with an outside source.

This suffers from many of the same problems as private IP addresses, only more so: your data may need significant restructuring once you decide to expose them to the outer world.

A globally accessible X.500 service requires a globally connected X.500 name space. See [3] and [4] for recommendations on how create a local part of the global name space.

Though the standard is not very clear about this and the most recent version (93) appears not to support it, in practice the X.500 name space is only manageable if there is a single root context operated under a cooperative agreement. However, one can be sure that there will be turf battles over it's control.

If those turf battles aren't decided outside the actual running service, the effect on the service quality will be ruinous.

This document appeals to all players in the field to let existing practice alone until a better system is agreed and is ready to go into place; at the moment, the root context of the day is operated by the Dante NameFLOW-Paradise service.

More information on the Dante NameFLOW-Paradise service is found at the URL

http://www.dante.net/nameflow.html

10. Use LDAP

At the moment, LDAP as documented in [9] is the protocol that offers the most X.500 functionality in places where it is not feasible to implement the full OSI stack.

It is implemented on a lot of platforms, including several PC-type platforms, and is popular in a multitude of commercial offerings.

A concerted effort to make LDAP available is the publication method that gives the widest access to the data.

In addition, X.500 DSAs must implement the necessary linkages to make sure they are properly integrated into the naming/referral tree; in most cases, this will mean that they should implement the X.500 DSP protocol at least.

(The question of whether one gateways LDAP to DAP or DAP to LDAP is irrelevant in this context; it may be quite appropriate to store data on an LDAP-only server and make it available to the DAP/DSP-running world through a gateway if the major users all use LDAP)

11. Make services available

The technical investment in running an X.500 service is not enormous, see for example [5].

12. Future developments

Today [October 1996] there are several enhancements to be expected with respect to IWPS technology.

The most important one to be mentioned here is the creation of a "Common Indexing Protocol" that must enable the integration of X.500, Whois++ and protocols that use stand-alone databases. Such a protocol would not only enable integration but would offer at the same time the possibility to explore yellow pages services and enhanced searches, even if used for X.500 only.

In the context of the Common Indexing Protocol the stand-alone LDAP servers should be mentioned that are announced by several software developers. These are stand-alone address databases that can be accessed by LDAP. Currently also a public domain version is available from the University of Michigan. Also announced is an LDAP-to-DAP gateway that can integrate a stand-alone LDAP server in an X.500 infrastructure.

Other improvements include defining a common core schema for multiple White Pages services, leading to the possibility of accessing data in multiple services through a single access protocol.

The 1993 version of the X.500 standard has already been implemented in several products. It is an enhancement over the 1988 standard in several ways, but has not been implemented in the NameFLOW-Paradise infrastructure yet. The main reason is that the standard doesn't recognize the existence of a single root DSA, but assumes that the managers of first-level DSAs (the country DSA's) make bilateral contracts for interconnection. In the case of NameFLOW-Paradise such a situation would be unmanageable. In [13] an enhancement of the 1993 standard is proposed that makes a single root possible. As soon as implementations of [13] are available, NameFLOW-Paradise will experiment with 1993 DSAs. This is expected in 1997.

Once these developments reach stability, they may be referenced by later versions of this BCP document.

13. Security considerations

The security implications of having a directory are many.

- People will have a standard way to access the information published.
- People will be able to gather parts of the information for purposes you never intended (like publishing directories, building search engines, headhunting or making harassing phone calls).
- People will attempt to access more of the information than you intended to publish, by trying to break security functions or eavesdropping on conversations other users have with the Directory.
- If modification over the Net is possible, people will attempt to change your information in unintended ways. Sometimes users will change data by mistake, too; not all undesired change is malicious.

The first defense for directory security is to limit your publication to stuff you can live with having publicly available, whatever happens.

The second defense involves trying to impose access control. LDAP supports a few access control methods, including the use of cleartext passwords. Cleartext passwords are not a secure mechanism in the presence of eavesdroppers; this document encourages use of stronger mechanisms if modification is made available over the open Internet. Otherwise, modification rights should be restricted to the local intranet.

The third defense involves trying to prevent "inappropriate" access to the directory such as limiting the number of returned search items or refuse list operations where they are not useful to prevent "trolling". Such defenses are rarely completely successful, because it is very hard to set limits that differentiate between an innocent user doing wasteful searching and a malicous data troller doing carefully limited searches.

Future enhancements may include using encrypted sessions, public key logins and signed requests; such mechanisms are not generally available today.

14. Acknowlegdements

The authors wish to thank the following people for their constructive contributions to the text in this document:

Peter Bachman <peterb@suport.psi.com>

David Chadwick <D.W.Chadwick@iti.salford.ac.uk>

William Curtin <curtinw@ncr.disa.mil>

Patrik Faltstrom <paf@swip.net>

Rick Huber <rvh@att.com>

Thomas Lenggenhager <lenggenhager@switch.ch>

Sri Saluteri <sri@qsun.ho.att.com>

Mark Wahl <M.Wahl@critical-angle.com>

15. Glossary

DAP Directory Access Protocol; protocol used between a DUA and a DSA to access the Directory Information. Part of X.500.

DSP Directory System Protocol: the protocol used between two DSAs

DSA Directory System Agent - entity that provides DUAs and other DSAs access to the information stored in the Directory

LDAP Lightweight Directory Access Protocol - defined in RFC 1777

Further terms may be found in RFC 1983.

16. References

[1] Jeunik, E. and E. Huizer. Directory Services and Privacy Issues. Proceedings of Joint European Networking Conference 1993, Trondheim, http://www.surfnet.nl/surfnet/diensten/x500/privacy.html

[2] Jennings, B. Building an X.500 Directory Service in the US, RFC1943, May 1996.

[3] Barker, P., S. Kille, T. Lenggenhager, Building Naming and Structuring Guidelines for X.500 Directory Pilots, P. Barker, S. Kille, T. Lenggenhager, RFC1617

[4] The COSINE and Internet X.500 Schema. P. Barker & S. Kille, RFC1274

[5] Introducing a Directory Service, SURFnet report 1995 (see URL: http://info.nic.surfnet.nl/surfnet/projects/x500/introducing/)

[6] Paradise International Reports, University College London, April 1991 - April 1994

[7] Naming Guidelines for the AARNet X.500 Directory Service, Michaelson and Prior, RFC 1562

[8] CCITT Blue Book, Volume VIII - Fascicle VIII.8, November 1988

[9] Lightweight Directory Access Protocol, W. Yeong, T. Howes, S. Kille, RFC1777

[10] Replication and Distributed Operations extensions to provide an Internet Directory using X.500, S. Kille, RFC1276

[11] ISO transport services on top of the TCP: Version: 3, M. Rose, D. Cass, RFC1006

[12] Recommendations for an X.500 Production Directory Service, R. Wright et al., RFC1803

[13] Managing the X.500 Root Naming Context, D. Chadwick, RFCxxxx

[14] A Revised Catalog of Available X.500 Implementations, A. Getchell, S. Sataluri, RFC1632

[15] A Naming Scheme for c=US, The North American Directory Forum, RFC1255

[16] A Common Schema for the Internet White Pages Service, T. Genovese, B. Jennings, Work In Progress.

[17] Key words for use in RFCs to Indicate Requirement Level, S. Bradner, RFC 2119,

17. Authors address

Harald Tveit Alvestrand
UNINETT
P.O.Box 6883 Elgeseter
N-7002 TRONDHEIM
NORWAY

+47 73 59 70 94
Harald.T.Alvestrand@uninett.no

Peter Jurg
SURFnet
P.O.Box 19035
NL-3501 DA UTRECHT
THE NETHERLANDS

+31 30 2305305
Peter.Jurg@surfnet.nl