

A Common Schema for the Internet White Pages Service

Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

Abstract

This work is the result of the IETF Integrated Directory Services (IDS) Working Group. The IDS Working Group proposes a standard specification for a simple Internet White Pages service by defining a common schema for use by the various White Pages servers. This schema is independent of specific implementations of the White Pages service.

This document specifies the minimum set of core attributes of a White Pages entry for an individual and describes how new objects with those attributes can be defined and published. It does not describe how to represent other objects in the White Pages service. Further, it does not address the search sort expectations within a particular service.

1.0 Introduction to IWPS

The Internet community has stated a need for the development and deployment of a White Pages service for use in locating information about people in the Internet [PA94]. To facilitate interoperability and to provide a common user experience, the Internet White Pages Service (IWPS) must have a common set of information about each person.

A common user object would allow a user to go between implementations of the service and to expect consistency in the types of information provided. A common user object would also provide developers with an unambiguous method of representing the information managed by the service.

This document will focus only on common information modeling issues to which all IWPS providers must conform.

2.0 Scope

This document establishes the set of attributes that specify the Common User Information Object for the IWPS. It does not attempt to be an exhaustive specification of all objects that may be stored in the IWPS. The process used by this document to define the user object is recommended to be used to define other information objects used in the IWPS.

All conforming implementations must support at the minimum, the core attributes listed in Section 5.0. Implementations may include local attributes in addition to the core set and still be considered "in conformance".

This document will not specify rules with respect to information privacy. Each country has its own set of laws and practices. Previous work covering this area has been done by the North American Directory Forum (NADF), whose publication [NADF92] contain recommendations for registrants' rights in both the USA and Canada.

This document does not specify a Directory access protocol (i.e. whois++, LDAP, DAP, etc.).

3.0 IWPS Schema Considerations

The description of the IWPS information object consists of the following requirements:

1. Syntax for definition/representation of information object templates.
2. Publication of information object templates, etc.
3. Database structure or schema.

Items 1 and 2 will be covered in this document. Because database structure can potentially restrict implementations (i.e. X.500 schema based versus DNS schema based) it will be treated as a separate research topic and will not be defined in this paper.

4.0 Syntax for Definition/Representation of Information Object Templates

A clear, precise, and consistent method must be used when discussing information object templates and their associated attributes. Therefore, this document makes use of the previously defined syntax used by LDAP. To avoid restrictions on implementations of the IWPS,

some syntax are listed as requirements vs specific encodings. The general IWPS syntax is included in section 6.0 for reference.

The IWPS Person Object specifies a limited set of recommended attributes that a White Pages Service must include. Storage of user attributes are a local issue, therefore, this memo suggests storage sizes but not storage types.

This document lists the syntax with the attributes for developers of user interface (UIs) to use as a reference, but it does not specify how the UI should display these attributes.

Attributes that contain multiple-line text (i.e. Address) must use the procedure defined in RFC 822 in section 3.1.1 on "folding" long header lines [RFC-822].

5.0 Information Object Template Definitions

This section describes the IWPS Person Information Object Template and its associated attributes. The Person Object is a simple list of attributes, no structure nor object inheritance is implied.

IWPS client applications should use the following size recommendations as the maximum sizes of the attributes. However, applications should be able to handle attributes of arbitrary size, returned by a server which may not comply with these recommendation. All size recommendations are in characters.

Note: Because many characters in many encodings require more than one byte, the size recommendations cannot be interpreted as sizes in bytes.

This set of attributes describes information types, and are not defined attributes in a particular schema. Any technology deploying a White Page service (WHOIS ++, LDAP, vCard, etc.) will need to publish as a companion document, their specific schema detailing how the general attributes of the White Pages schema are expressed.

SPECIAL CONSIDERATIONS

Phone number: The full international form is recommended; i.e. +1 206 703 0852. The field may contain additional information following the phone number. For example:

+1 800 759 7243 #123456
+1 505 882 8080 ext. 30852

Email address: Is multivalued.

Certificate: Is multivalued.

Common Name: Is multivalued.

Language Spoken: Is multivalued.

THE INFORMATION OBJECT TEMPLATE FOR THE IWPS PERSON

--General Attributes --

Field Name	Size	Syntax
Email	360	Mailbox
Cert	4000	Certificate
Home Page	128	URI
Common Name	64	WhitepageString
Given Name	48	WhitepageString
Surname	48	WhitepageString
Organization	64	WhitepageString
Locality	20	WhitepageString
Country	2	WhitepageString (ISO 3166)
Language Spoken	128	WhitepageString (RFC 1766)

--Personal Attributes

Personal Phone	30	PrintableString
Personal Fax	30	PrintableString
Personal Mobile Phone	30	PrintableString
Personal Pager Number	30	PrintableString
Personal Postal Address	255	Address
Description	255	WhitepageString

--Organizational Attributes

Title	64	WhitepageString
Office Phone	30	PrintableString
Office Fax	30	PrintableString
Office Mobile Phone	30	PrintableString
Office Pager	30	PrintableString
Office Postal Address	255	Address

--Ancillary

Creation Date	24	GeneralizedTime
Creator Name	255	URI
Modified Date	24	GeneralizedTime

Modifier Name	255	URI
---------------	-----	-----

6.0 IWPS Person Information Object Template Syntax

This section defines the syntax used by the IWPS person information object template. It is copied in whole from the LDAP attribute working document with some modification for completeness.

Certificate:

The certificate field is intended to hold any kind of certificate; X.509 certificates are one example. A specific implementation will specify how to indicate the type of certificate when describing the mapping of the IWPS schema onto the implementation schema.

WhitepageString:

This syntax must be able to encode arbitrary ISO 10646 characters. One such encoding is the UTF-8 encoding [UTF-8].

GeneralizedTime:

Values of this syntax are encoded as printable strings, represented as specified in X.208. Note that the time zone must be specified. It is strongly recommended that Zulu time zone be used. For example:

199412161032Z

Mailbox:

here are many kinds of mailbox addresses, including X.400 and Internet mailbox addresses. The implementation must clearly distinguish between different types of mailbox address, for instance by using a textual refix or a set of attribute types. There must be a way to represent any mailbox type.

Address:

According to Universal Postal Union standards, this field must be able to represent at least 6 lines of 40 characters.

PrintableString:

The encoding of a value with PrintableString syntax is the string value itself. PrintableString is limited to the characters in production <p>. Where production <p> is described by the following:

```

<a> ::= 'a' | 'b' | 'c' | 'd' | 'e' | 'f' | 'g' | 'h' | 'i' |
        'j' | 'k' | 'l' | 'm' | 'n' | 'o' | 'p' | 'q' | 'r' |
        's' | 't' | 'u' | 'v' | 'w' | 'x' | 'y' | 'z' | 'A' |
        'B' | 'C' | 'D' | 'E' | 'F' | 'G' | 'H' | 'I' | 'J' |
        'K' | 'L' | 'M' | 'N' | 'O' | 'P' | 'Q' | 'R' | 'S' |
        'T' | 'U' | 'V' | 'W' | 'X' | 'Y' | 'Z'

```

```

<d> ::= '0' | '1' | '2' | '3' | '4' | '5' | '6' | '7' | '8' | '9'

```

```

<p> ::= <a> | <d> | ''' | '(' | ')' | '+' | ',' | '-' | '.' |
        '/' | ':' | '?' | ' '

```

7.0 Publication of IWPS Information Object Templates.

The Working Group recommends that all information object templates used for the IWPS be published.

Individual organizations may define information object templates that are local in scope as required to meet local organizational needs. All information that the organization wishes to be part of the IWPS must use a published IWPS information object template.

8.0 Data Privacy

Each country, and each state within the US, has legislation defining information privacy. The suggested attributes in Section 5.0 may be considered private and the directory administrator is strongly advised to verify the privacy legislation for his domain.

As suggested in "Privacy and Accuracy in NIC Databases" [RFC-1355], each directory provider should provide a clear statement of the purpose of the directory, the information that should be contained in it, and a privacy policy associated with that information. This policy should include restrictions for data dissemination.

This policy is strongly recommended for the US and Canada and required by many countries in the European Community for data sharing.

9.0 Data Integrity

Data Integrity was first addressed in RFC 1107 [KS89], which states "a White Pages service will not be used, if the information it provides is out of date or incorrect." Therefore, any production IWPS provider must insure that all data is reasonably correct and up-to-date.

The Ancillary Attributes of the IWPS person template denote the information's source and date of origin, and the source and date of its latest modification. They provide the user with some measurement of the quality of data making it easy to determine the owner and freshness of the data retrieved.

The IWPS User Agent must be able to retrieve and display Ancillary Attributes. Retrieval and display may be done as separate operations.

The Ancillary Attributes are recommended as the minimum set of attributes for any new information object template. Each IWPS server may individually decide whether to support the storage and retrieval of this data.

The Ancillary Attributes (also defined in Section 5.0) provide the following information about its associated information object:

1. The date and time the entry was created; Creation Date.
2. Owner or individual responsible for the data creation; Creator Name.
3. The date and time of the last modification; Modified Date.
4. Individual responsible for the last modification; Modifier Name.

10.0 Security Considerations

Security is implementation and deployment specific and as such is not addressed in this memo. Security must ensure that the constraints mentioned in the Data Privacy Section 8.0 are complied with.

11.0 References

[KS89] Sollins, K., "A Plan for Internet Directory Services", RFC 1107, Laboratory for Computer Science, MIT, July 1989.

[NADF92] North American Directory Forum, "User Bill of Rights for entries and listings in the Public Directory", RFC 1295, North American Directory Forum, January 1992.

[PA94] Postel, J., and C. Anderson, "WHITE PAGES MEETING REPORT", RFC 1588, University of Southern California, February 1994.

[RFC-822] Crocker, D., "Standard for the Format of ARPA Internet Text Messages", STD 11, RFC 822, August 1982.

[RFC-1355] Curran, J., and A. Marine, "Privacy and Accuracy Issues in Network Information Center Databases", FYI 15, RFC 1355, August 1992.

[UCS] Universal Multiple-Octet Coded Character Set (UCS) - Architecture and Basic Multilingual Plane, ISO/IEC 10646-1, 1993.

[RFC-1766] Alvestrand, H., "Tags for the Identification of Languages", RFC 1766, March 1995.

[UTF-8] Yergeau, F., "UTF-8, a transformation format of ISO 10646", Work in Progress.

11.0 Authors' Addresses

Tony Genovese
The Microsoft Corporation
One Microsoft Way
Redmond, Washington 98007
USA

Phone: (206) 703-0852
Email: TonyG@Microsoft.com

Barbara Jennings
Sandia National Laboratories
Albuquerque, New Mexico 87106
USA

Phone: (505) 845-8554
Email: jennings@sandia.gov

