A Common Schema for the Internet White Pages Service

1.0 Introduction

The Internet community has stated that there is a need for the development and deployment of a White Page service. This service would be used to locate information about people in the Internet. To facilitate interoperability and a common user expectation the Internet White Pages Service (IWPS) needs to have a common set of information about each person.

This Document will focus only on common information modeling issues to which all IWPS providers must conform. To insure a consistent user experience of this service we need to define a common user object. This will allow a user to go between different implementations of the service and have a consistent expectation as to what information can be found about people on the Internet. Developers of this service need to have an
unambiguous method of representing the Information objects managed by
the service. This will help facilitate interoperability and data management.

2.0 Scope

This document will establish the set of attributes that specify the common
user information object for the IWPS. It will not attempt to be an
exhaustive specification of all objects that will be stored in the IWPS.
The process used by this document to define the user object will be used to
define all other information objects used in the IWPS.

All conforming implementations must support at the minimum, the core
attributes listed in Appendix A. Implementations may include additional
local attributes and be considered in conformance as long as they support
the core set of attributes.

This document will not specify rules with respect to information privacy.
Each country has its own set of laws and practices. Work covering
this area was done by North American Directory Forum (NADF) [NADF92]. In
this area recommendations for registrants rights for both the USA and Canada.

3.0 IWPS Schema Considerations

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

1. Syntax for definition/representation of Information
   Object Templates.
2. Registration procedures for information object
   Templates, etc.
3. Database structure or schema.

Items 1 and 2 will be covered in this Document. Database structure
can potentially restrict implementations (i.e. X.500 schema based verses
DNS schema based) and will not be defined in this document. This area is
a separate Research topic and will be covered in its own document.

3.1 Syntax for Definition/Representation of Information objects

A clear, precise and consistent method must be used when information
object Templates and their associated attributes are discussed within
the context of IWPS. There are two possible methods to do this. i.e.

1. BNF
2. ASN.1

The Working Group has recommended the use of BNF. BNF is widely used by
the Internet community and is well understood. It is used by the
LDAP work on attribute definitions. This document makes use of the
previously defined syntaxes use by LDAP. They are included in Appendix
B for convenience.

The use of Object inheritance is not used or specified by this document.
The IWP person object specifies a set of recommended attributes that a
White Page Service should include. This draft suggests storage sizes,
but does not recommend storage types. Storage of user attributes is a
local issue. The Syntax listed with the attributes are provided so the
developers of user interfaces (UIs) may have a consistent expectation.
This document does not specify a Directory access protocol (i.e. whoi++,
LDAP, DAP, etc.) or how the UI is to display these attributes.

Attributes that contain textual information that must be split over multiple
lines (i.e. Postal address) will use the procedure defined in RFC 822 in section 3.1.1 on "folding" long header lines [RFC-822].

For International localization it is recommended that attributes (except email addresses) used to identify people must follow the DirectoryString syntax defined by LDAP [LDAP-A].

3.2 Publishing of IWPS Information object Templates.

The Working Group recommends that all information object Templates used for the IWPS be published as an RFC at the minimum. To facilitate distribution of IWPS information object Templates they should be made available on the Internet information server (i.e. InterNIC).

Individual organizations may define information object Templates that are only local in scope. This may be needed to meet local organizational needs. All information that the organization wishes to be part of the IWPS must use an IWPS published information object Template.

4.0 Data Privacy

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

As suggested in RFC 1355 (4) each Directory provider should provide a clear statement of the purpose of the directory, the information that will be contained in it, and a privacy policy associated with the information that is stored within the directory. This policy should include restrictions for data dissemination.

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

5.0 Data Integrity

Data Integrity was first addressed in RFC1107 [KS89]. Which states, that if the information is out of date it is useless and the service will not be used. Therefore, a clear requirement is that any production IWPS provider must insure that all data is reasonably correct and current.

Ancillary attributes have been included which state the source of origin and the current party responsible for the data in addition to date; such that the owner and the freshness of the data can be easily determined.

To facilitate the user in determining the quality of the data that has been retrieved it is recommended that the optional Ancillary attributes of the IWPS person Template be supported. This would require that the IWPS User Agent be able to retrieve and display this information. This may be done as a separate operation from the fetch of the information object. The Ancillary Attributes are defined in Appendix A. It is further recommended that any new information object Template include as a minimum the Ancillary attributes as an optional set of attributes. It would then be left to the IWPS servers to optionally support the storage and retrieval of this data.

The Ancillary attributes have been designed to provide the following information about the information object with which it is associated:

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.

6.0 References


7.0 Authors Address

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

Appendix A Information Object Template Definitions

This appendix contains the IWPS Person Information Object Template and its associated attributes. The Person Object is a simple list of attributes, no structure or object inheritance is implied. All size recommendations are in bytes.
The following size recommendations should be used as an indication of the largest size of a particular attribute that an IWPS client application would see in practice. In particular instances, actual user attributes may be larger or smaller than these recommendations, and applications should be written to accept any size attribute returned from a server.

-- 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 sky page #123456
+1 882 8080 ext 30852

Email address: Is multivalued and uses the otherMailbox syntax to identify the different email addresses.

Certificate: Is multivalued.

Common Name: Is multivalued.

-- THE INFORMATION OBJECT TEMPLATE FOR THE IWPS PERSON --

--General Attributes--

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Size</th>
<th>Syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Email</td>
<td>360</td>
<td>otherMailbox</td>
</tr>
<tr>
<td>Cert</td>
<td>4000</td>
<td>Certificate</td>
</tr>
<tr>
<td>Home Page</td>
<td>128</td>
<td>URI</td>
</tr>
<tr>
<td>Common Name</td>
<td>64</td>
<td>DirectoryString</td>
</tr>
<tr>
<td>Given Name</td>
<td>48</td>
<td>DirectoryString</td>
</tr>
<tr>
<td>Surname</td>
<td>48</td>
<td>DirectoryString</td>
</tr>
<tr>
<td>Organization</td>
<td>64</td>
<td>DirectoryString</td>
</tr>
<tr>
<td>Locality</td>
<td>20</td>
<td>DirectoryString</td>
</tr>
<tr>
<td>Country</td>
<td>02</td>
<td>DirectoryString (ISO3166)</td>
</tr>
<tr>
<td>Language Spoken</td>
<td>02</td>
<td>DirectoryString (ISO 639)</td>
</tr>
</tbody>
</table>

--Personal Attributes--

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Size</th>
<th>Syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telephone Number</td>
<td>30</td>
<td>PrintableString</td>
</tr>
<tr>
<td>Fax</td>
<td>30</td>
<td>PrintableString</td>
</tr>
<tr>
<td>Mobile Phone</td>
<td>30</td>
<td>PrintableString</td>
</tr>
<tr>
<td>Pager Number</td>
<td>30</td>
<td>PrintableString</td>
</tr>
<tr>
<td>Postal Address</td>
<td>255</td>
<td>PostalAddress</td>
</tr>
<tr>
<td>Description</td>
<td>255</td>
<td>DirectoryString</td>
</tr>
</tbody>
</table>

--Organizational Attributes--

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Size</th>
<th>Syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>64</td>
<td>DirectoryString</td>
</tr>
<tr>
<td>Office Phone</td>
<td>30</td>
<td>PrintableString</td>
</tr>
<tr>
<td>Office Fax</td>
<td>30</td>
<td>PrintableString</td>
</tr>
<tr>
<td>Office Mobile Ph</td>
<td>30</td>
<td>PrintableString</td>
</tr>
<tr>
<td>Office Pager</td>
<td>30</td>
<td>PrintableString</td>
</tr>
<tr>
<td>Postal Address</td>
<td>255</td>
<td>PostalAddress</td>
</tr>
</tbody>
</table>

--Security--
Appendix B INWS Person Information Object Template Syntaxes

This Appendix contains the definitions of the syntaxes used by the INWS Person Information Object Template. They are copied in whole from the LDAP attribute working document. Some modification to the LDAP attribute text was done for completeness.

Certificate:

Do to differences from version X.509(1988) and X.509(1993) and additional changes to the ASN.1 definition to support certificate extensions, no string representation is defined, and values with Certificate syntax must only be transferred using the binary encoding, by requesting or returning the attributes with descriptions "userCertificate;binary" or "caCertificate;binary". The BNF notation in RFC 1778 for "User Certificate" is not recommended to be used.

DirectoryString:

A string with DirectoryString syntax is encoded in the UTF-8 form of ISO 10646 (a superset of Unicode). Servers and clients must be prepared to receive arbitrary Unicode characters in values.

For characters in the PrintableString form, the value is encoded as the string value itself.

If it is of the TeletexString form, then the characters are transliterated to their equivalents in UniversalString, and encoded in UTF-8 [Davis].

If it is of the UniversalString or BMPString forms [UCS], UTF-8 is used to encode them.

Note: the form of DirectoryString is not indicated in protocol unless the attribute value is carried in binary. Servers which convert to DAP must choose an appropriate form. Servers must not reject values merely because they contain legal Unicode characters outside of the range of printable ASCII.

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

OtherMailbox:

Values of the OtherMailbox syntax are encoded according to the following BNF:

<otherMailbox> ::= <mailbox-type> ‘$’ <mailbox>

<mailbox-type> ::= an encoded Printable String
<mailbox> ::= an encoded IA5 String

In the above, <mailbox-type> represents the type of mail system in which the mailbox resides, for example "MCIMail"; and <mailbox> is the actual mailbox in the mail system defined by <mailbox-type>.

Password:

Values with Password syntax are encoded as octet strings.

PostalAddress:

Values with the PostalAddress syntax are encoded according to the following BNF:

<postal-address> ::= <dstring> | <dstring> $ <postal-address>

In the above, each <dstring> component of a postal address value is encoded as a value of type DirectoryString syntax. Backslashes and dollar characters, if they occur in the component, are quoted as follows:
A backslash quoting mechanism is used to encode symbol character such as ''' or '#'. The backslash is followed by a pair of hexadecimal digits representing the next character. A backslash itself in the string which forms part of a larger syntax is always transmitted as '\5c' or '\5C'.

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 discribed by the following BNF:

<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> | ' ' | '(' | ')' | '+' | '-' | ':' | '?' | '.'