

Specification of an architecture and user interface for the Directory Schema Registry

TERENA project Directory Schema Registry, Deliverable E

Peter Gietz, DAASI International Ltd

peter.gietz@daasi.de

Version 0.9, May 2003

1. Status of this document

This is deliverable E of the TERENA project Directory Schema Registry, which is co-funded by TERENA, JISC (Joint Information Systems Committee, UK), REDIRIS (Spanish National Research Network), CESNET (Czech National Research Network), POZMAN SUPERCOMPUTING (Poznan Supercomputing and Networking Center, Poland) and DAASI International and performed by DAASI International. Together with four other deliverables [RegIntro], [RegPolicy], [RegSchema], [RegBusiness] and the bibliography [RegBib] it forms the documentation of this project.

In this document the software of the Directory Schema Registry (DSR) is described. This version 0.9 reflects the status in May 2003 and is the first public version. A final version 1 will be published after all other project documents have been completed.

In the last chapter of this document, enhancements of the DSR not relevant for the currently planned registry are discussed which points to a possible future broadening the functionality of the registry after the end of this project. A version 2 of this document is planned for the time after the end of the project. For this future version comments and additions to the current document are welcomed. Please send them to the email address of the author.

Table of Contents

1. Status of this document	1
2. Introduction.....	1
3. General architecture of the DSR	2
4. The server infrastructure	3
5. The MIME interface for including new schema.....	4
6. The administration interface	5
7. The search and browse interface	5
8. Possible usage of the LDAP interface provided by the server.....	6
9. Possible enhancements of the system	6
10. Pointer to References	7

2. Introduction

[SchemaProc] gives a good introduction into the ldap schema registration:

"There is a growing number of places where schema for Internet Directory Services are being defined, with varying degrees of documentation. This plethora of schemas is unavoidable in the light of the needs of different service communities, but it makes it difficult for directory service builders to find and make use of an existing schema that will serve their needs and increase interoperability with other systems. A listing service providing a single point of discovery for

directory service schema will promote schema reuse, reduce duplication of effort, and thus promote directory service interoperability."

This document specifies a software architecture for such a schema listing service called the Directory Schema Registry (DSR). The software specification is based on the schema specification [RegSchema] and the policy specification [RegPolicy] and thus the subsequent chapters only describes the general architecture, the server infrastructure, the search and browsing interface, the possible usage of the LDAP interface provided by the server, the administration interface and the MIME interface for inclusion of new schema. In a last chapter possible future enhancements of the system are discussed.

3. General architecture of the DSR

The requirements for the schema of the DSR that are listed in chapter 3 of [RegSchema] also apply to the software design. In addition, other requirements that are to be considered in the software design are listed in the beginning of the chapters on the single components.

The DSR in its current implementation can be divided into the following components:

- an internal LDAP server for not yet published data
- an external LDAP server (farm) for the published data, providing an LDAP interface
- a Web-based user interface
- a MIME interface for adding new data
- an administration interface to ease the administrative tasks of the DSR operator
- a mailing list for the discussion about the single schema submissions

Diagram 1 shows how these components interact:

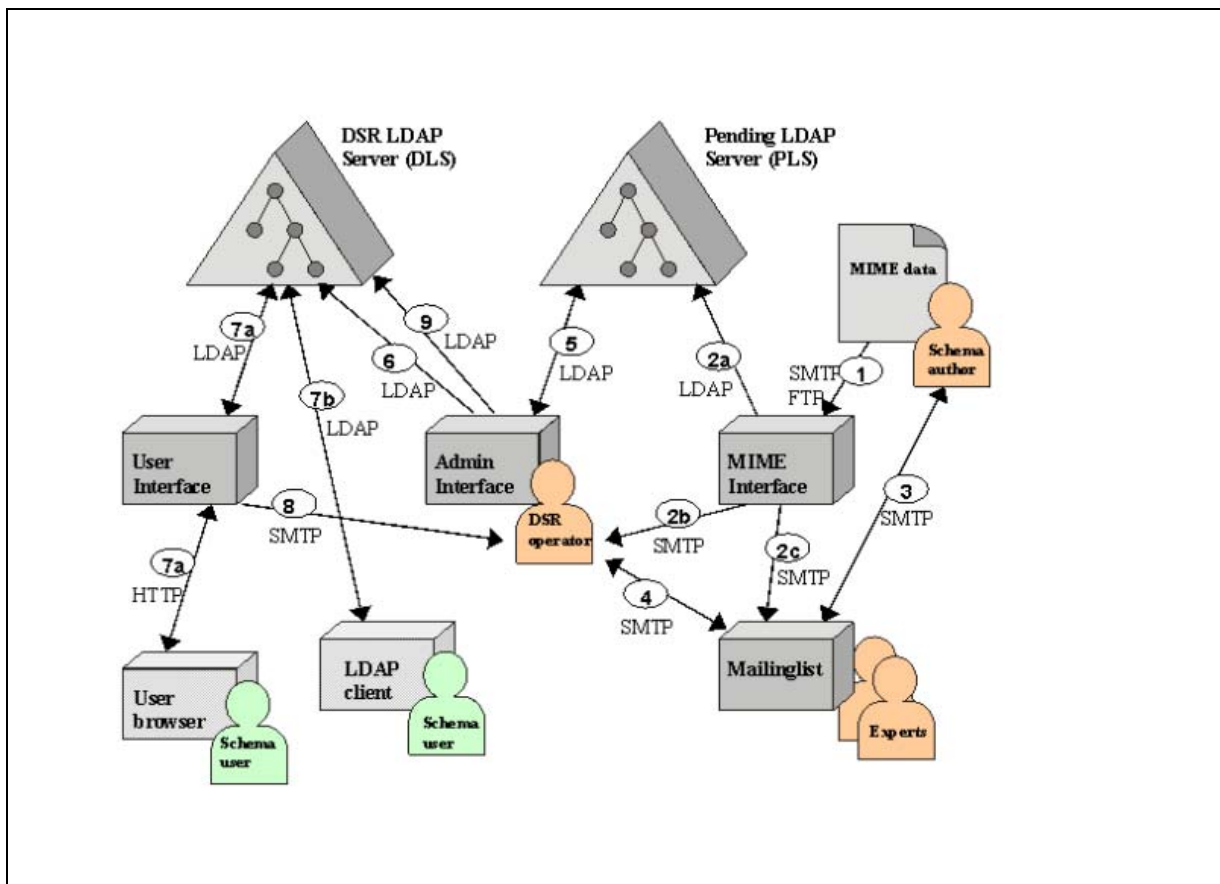


Diagram 1 Workflows in the DSR

Following interactions (marked in the diagram with a number) occur in the current DSR:

- 1 A schema author or her representative sends a MIME formatted schema registry request either via email or FTP
- 2a The MIME interface stores the data in the pending LDAP server
- 2b The MIME interface sends a notification email to the DSR operator
- 2c The MIME interface sends a notification email to the DSR mailing list
- 3 The submitted schema is being evaluated and discussed on the DSR mailing list in which experts, the schema author and the DSR operator interact
4. The DSR operator checks consensus for the decision for acceptance
5. The DSR operator checks the correctness of the data and adds additional metainformation
6. The DSR operator publishes the schema in the public service
- 7a. A schema user retrieves schema information via the web search and browse interface
- 7b. A schema user retrieves schema information via LDAP client
8. A schema user posted an informational comment on a schema and the user interface sends a respective notification email to the DSR operator
9. The DSR operator checks and publishes the informal comment in the public service

The following chapters discuss the single elements of this architecture.

4. The server infrastructure

Following requirements have to be fulfilled in the implementation of the server infrastructure:

- clear separation of non-published and published data
- security for the non published data
- scalability and fail-safe of the service

All LDAP servers of the DSR are implemented with OpenLDAP 2.1.x.

Two master server provide the data base for storing all information of the DSR both using the same LDAP schema as defined in [RegSchema].

One server is an internal server called Pending LDAP Server (PLS) that contains all data to be published. This server provides only write access to the MIME interface as well as read and write access to the DSR operator. It should be located in an internal network, not accessible from the outside.

A second server called DSR LDAP Server (DLS) contains the published data. This server provides public anonymous read access but write access only for the DSR operator. For scalability and fail-safe of service access, the DLS can and should be replicated onto several slave servers. In such an architecture, the master DLS could be located in an internal network and only the slave servers made available to the public. These then only need write access for the replication mechanism, in the case of OpenLDAP this is the stand-alone update and replication daemon called SlurpD. The public available LDAP interface of the DSR is available at `ldap://ldap.schemaraeg.net`. Diagram 2 shows such a server architecture. Only the LDAP communication between the SlurpD and the slave servers has to go through a firewall.

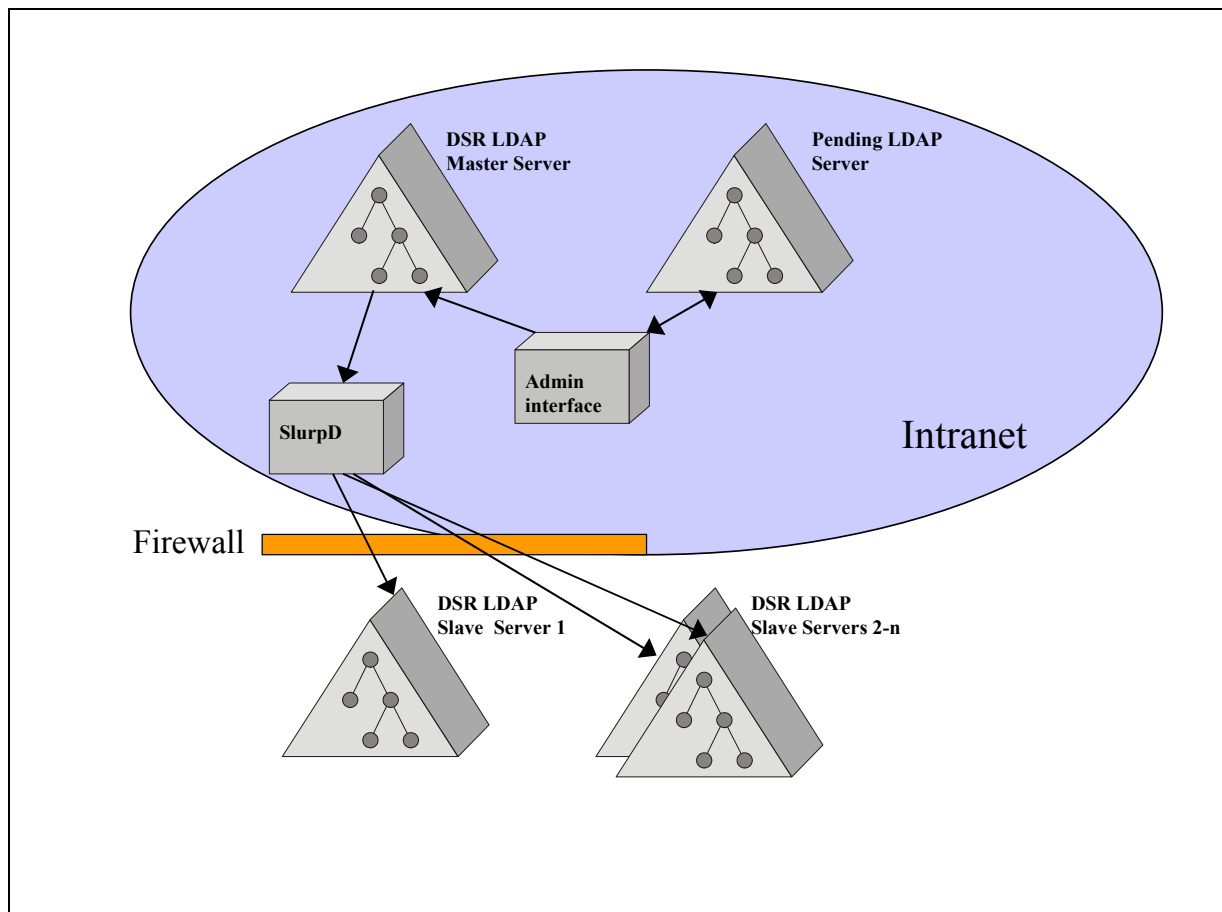


Diagram 2 DSR Server architecture

5. The MIME interface for including new schema.

Following requirements have to be fulfilled in the implementation of the MIME interface:

- The MIME-interface has to include workflow processes, namely sending emails to the DSR operator and to the DSR mailing list.
- The MIME analysis as well as the LDAP add operation have to be logged so that errors can be noticed by the DSR operator.

The MIME interface is constructed as a number of modules integrated into one daemon program.

Following modules are included in the MIME interface program:

1. The daemon and dispatcher module listens to the email inbox of the address mime_input@schemareg.net and dispatches a new process for each email arriving there.
2. The MIME module is able to process either single MIME parts for document metadata, schema metadata and the actual schema data, or multipart MIME messages that include all three parts. It analyses the MIME parts and checks whether the syntax is ok. This module writes a dedicated log file for the MIME errors.
3. The data enhancer takes over the data sent in the MIME format and creates LDIF files which already separate the information into several LDAP entries. It also adds additional data that is not provided by the schema author like the unique OID for the schema and internal data needed for the maintenance of the schema.
4. The LDAPadder takes this LDIF file and adds the data into the PLS. Any errors given back by the server are logged in a dedicated log file.

5. The workflowmanager writes three emails:
 - a. one to the DSR operator including the two mentioned error files
 - b. one to the mailing list notifying the acceptance of a new schema submission
 - c. one to the schema author to acknowledge the receipt of the schema submission

6. The administration interface

The administration interface helps the DSR operator in her day to day work. Thus it consists of a number of tools for different tasks of the DSR operator. Following administration tools are used:

1. For doing manual changes to the data in either the PLS or the master DLS a standard LDAP client will be used.
2. For proving the correctness of the syntax of the entries of one schema before moving it to the DLS, checks have to be performed. This can be done semi automatically by a dedicated script tool.
3. For the workflow of moving a schema from the PLS to the DLS, which represents a subtree search in the PLS, a tree add in the DLS and a tree delete in the PLS, a separate script tool will be used.
4. Another tool is used to do referential integrity checks. This is needed, since there are a lot of internal pointers in the data. This script can be started regularly and automatically by a crontab job.
5. For adding informal comments to the DLS another dedicated script will be used.
6. To operate the mailing list a respective Open Source tool will be used by the DSR operator. Due to its very good web interface, Mailman has been chosen. The name of the mailing list is schema_discussion@schemareg.net.
7. Following other email addresses are maintained by the DSR operator:
 - info@schemareg.net for general information about the project and the DSR
 - operator@schemareg.net the email adress of the operator which is used in the mailinglist, in the communication with the schema authors and for the requests to publish informal comments.
 - mime_input@schemareg.net an automatic account to which the MIME formatted schema are to be sent

7. The search and browse interface

Following requirements have to be fulfilled in the implementation of the search and browsing interface:

- It has to include an intuitive search and browse facility that provides all information about the schema as well as links to external information.
- It should also provide the possibility to retrieve schema information in the format of an OpenLDAP schema file, so it can easily be imported into the user's own LDAP server.
- It has to provide a means for the user to add additional informal comments as described in [RegPolicy] and [RegSchema]

The Search and browse interface is based on W2L, a generic LDAP/Web gateway developed by DAASI International. This Programm written in Perl had to be amended to fulfill the needs for the schema registry.

Following features will be implemented during the first projekt phase:

1. The overall weblayout has to be simple but intuitive.
2. Facility to browse the DIT.
3. Facility to search in names, descriptions, keywords and in numeric OIDs at every point in the DIT. The results of the search will be displayed in a list with links to the respective entries. Other search possibilities may be added in future versions.
4. Below each node the list of child entries is devided in, e.g., schema elements, persons and comments.
5. Facility at each node that gives the user the oportunity to post informal comments to the DSR-operator.
6. Facility to retrieve a whole schema in a text file in the format of an OpenLDAP schema file.
7. Facility to retrieve all information about a schema in an ascii text file.
8. A list of named elements in a left frame as known from the LDAP Schema Viewer as an additional navigation possibility.

The Web interface of the DSR is available at <http://www.schemareg.net>.

8. Possible usage of the LDAP interface provided by the server

Due to the fact that the DLS provides a standard LDAP interface, additional retrieval possibilities are given for LDAP clients. Following list gives some usage examples. Within this project no such implementations are planned.

- A client could be written that takes over the function as search and browse interface and that could have additional features that a Web-Gateway cannot provide.
- A client could supply itself with schema information to understand a schema element unknown to the client in the first place.
- Some clients only operate with the descriptors instead of the OID. If communicating with an LDAP program that needs the OID, the latter could be retrieved from the schema registry.
- a client program for schema design could also make use of the LDAP interface of the DSR.

9. Possible enhancements of the system

This document describes the first software implementation of the DSR created within the project. After the project enhancements are planned. This section describes some possible future additions and enhancements.

- The possible usages of the LDAP interface of the DSR by new dedicated LDAP clients as described in chapter 8 can create added value in the use of the DSR.
- Value added services provided by the DSR like newsletters and notification services are another field of enhancements.
- In future versions of the DSR additional input formats may be supported. Thus it could be possible to use the XML format for RFCs as specified in [RFC2629] instead of providing this data via MIME. This means that the whole XML formated RFC can be sent to the DSR and the input interface extracts from the XML elements all metadata needed. That this is possible was

shown in the sections 4.3 and 4.4. of [RegSchema]. The modularity of the input component of the DSR described in chapter 5 makes it easy to integrate the support of other input formats.

- On the organisational side more specialised mailing lists may be introduced, e.g. a separate discussion list for person related schema.
- Other possible enhancements will become obvious from the expectations of the schema writers and the DSR users. The informal comment mechanism that also works on the top nodes of the registry may be used to point out such additional needs.
- The not yet specified business model might also influence the further development of the DSR

10. Pointer to References

[RegBib] Gietz, Peter, "Bibliography for the Directory Schema Registry Project", Version 1, Deliverable B of the TERENA project Directory Schema Registry, January 2003

[RegBib] contains all references used in the project documents.