

# irisEduPerson: Recommendations for Modeling a University into a Directory Schema

Carmen López<sup>1</sup>, Ignacio Diaz<sup>2</sup>, Domingo Ortega<sup>3</sup>, Liborio Revilla<sup>4</sup>, Victoriano Giralt<sup>5</sup>

<sup>1</sup>ICT Services, University of Seville, Avda. Reina Mercedes, ES41012 Sevilla, Spain, [carmen@us.es](mailto:carmen@us.es). <sup>2</sup>IT Services, Carlos III University, ES28071 Madrid, Spain, [nacho@di.uc3m.es](mailto:nacho@di.uc3m.es). <sup>3</sup>ICT Services, University of Seville, Avda. Reina Mercedes, ES41012 Sevilla, Spain, [domingo@us.es](mailto:domingo@us.es). <sup>4</sup>ICT Services, University of the Basque Country, ES20018 Gipuzkoa, Spain, [libo.revilla@ehu.es](mailto:libo.revilla@ehu.es). <sup>5</sup>Central IT Services, University of Malaga, ES29071 Malaga, Spain, [victoriano@uma.es](mailto:victoriano@uma.es).

## Keywords

Directory, schema, identity, interoperability.

## 1. EXECUTIVE SUMMARY

The identity federations, and the inter-federations built between them, that are spreading all over Europe and the world, together with the coming European Higher Education Area (a.k.a. Bologna Process) constitute an interesting technological challenge for the interchange of personal information in an interoperable way. This requires setting up common syntaxes and semantics for describing persons and other relevant educational information in the context of their institutional relationships.

This information interchange is also useful inside the institutions, as there are many applications that access personal information through diverse mechanisms with different formats, but the person described by that data is always the same. So, it is of capital importance to keep all those identity bits connected and synchronised.

### 1.1. Background

RedIRIS (the Spanish National Research and Education Network) has hosted and supported a working group for defining a common way of describing a higher education institution, such as an university, using a directory schema. This paper will present the works and results of this task force.

The working group started by locating relevant international schemas that described persons, institutions and educational information. As a result of this search, the group decided to use the attributes available from several international schema recommendations: person, organizationalPerson, interOrgPerson, eduPerson and SCHAC, as well as the Spanish national iris schema. A document was produced - irisEduPerson - with recommendations on which attributes to use for which information and how the information should be stored in them.

### 1.2. Conclusions

The most prominent and successful result of the group works has been a thorough collection of use cases describing how to solve problems common in the constituency, using the schema proposed in the irisEduPerson document. And, of course, made available through an open wiki (<http://wiki.rediris.es/gtschema/>) with the rest of the materials, like basic glossaries or deployment guides.

## 2. INTRODUCTION

A corporate directory has been in fashion or a cool thing for some time now, and many people like to brag about having one, even if they do not know what it is useful for. The aim of the work group was two pronged: the shed some light in this space for this unknowingly ones and for newbies and seasoned experts alike; as well as producing a common schema for the Spanish higher education institutions.

A corporate directory can be as useless, or useful, as a traditional relational database, if the user has no schema for the data to be put into and no use for that collected data, so our group started up collecting a gamut of uses where a corporate directory has demonstrated to be useful.

Once we have common schema, we can even leave the directory space and risk into federated space, using the same attributes for information exchange between organizations. This approach even allows for mapping data stored in diverse sources, including classical relational corporate databases, into a common set of syntaxes and semantics.

It is important to point out that our group meas by directory, for the purpose of our work and this paper, a protocol to access the data, not the electronic repository where this data is stored, that could be anything form a flat file through a relational database, even what many people regard as a directory: a repository for unstructured data that follows the X.500 recommendations.

## 3. PROCESS

The started off studying the schemata available for defining persons and their academic data and the broad categories for attributes defined in “Higher-Education Person: A Comparative Analysis of Collaborative Public LDAP Person Object Classes in Higher-Education”.

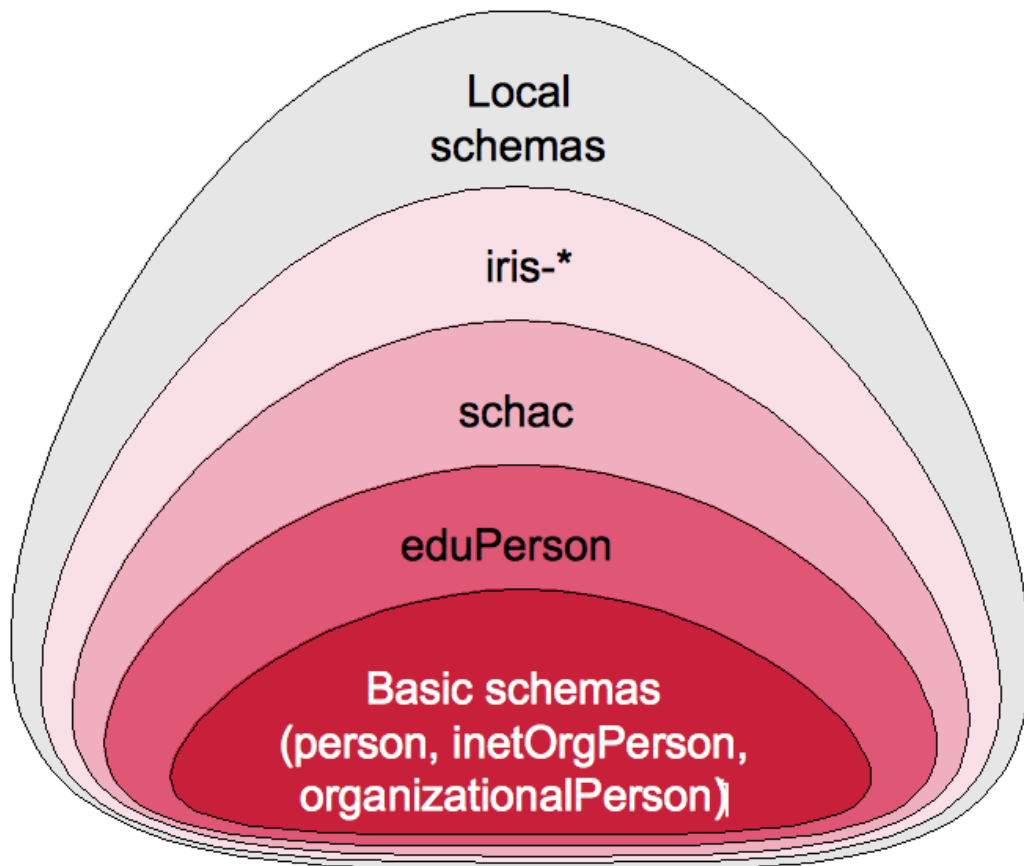
Studied person schemata included:

- person
- organizationalPerson
- inetOrgPerson
- eduPerson
- SCHAC
- several countries person definitions

The categories for person attributes are:

- Personal Characteristics
- Contact / Local Information
- Student Information
- Employee Information
- Linkage Identifiers / Foreign Keys
- Entry Metadata / Administration Information
- Security Attributes and Keys
- Confidentiality / Attribute Release (Visibility)
- Authorization, Entitlements
- Group-related Attributes
- Attributes outside the previous categories

The relationship among the several attribute categories could be seen like the layer of an onion, with the most general schema as the core, as seen in figure 1.



**Figure 1: The schema onion**

Then, the group found a list of common use cases in the Spanish higher education space, and defined the set of applicable attributes from the available schemata. At the present state, there has been no need to define new attributes, but in case they are needed, the group will define if the use case has an applicability outside the Spanish space, for pursuing definition at a layer closer to the core.

The use cases belonged in a list of applications that frequently use corporate directories:

- white pages (in a very broad sense)
- e-mail services
- workstation access authentication
- network access
- User services

Use case have been fundamental for the group's tasks, as we decided that no work will be started without a clear use case. The group's motto has been "no use case, the problem does not exist".

The work has worked via teleconference, face to face meetings, mailing list and a wiki for doing collaborative editing of documents.

#### **4. DELIVERABLES**

As the group work proceeded on the wiki and we felt that some document was ready for public consumption, it was moved into the public area. This has allowed us to follow the successful opensource *release early, release often* method.

## 4.1. Public wiki

This is the main deliverable, as other deliverables are published in there. It has been provided by RedIRIS the Spanish NREN. It is a read only wiki. Anyone from the Spanish higher education community can join the group, and, then, get write access to the wiki.

The wiki, at present, is available only in Spanish, but we have a firm intention for translating the most relevant contents into English.

## 4.2. irisEduPerson document

We are not going to reproduce the whole document here, more so, as it is available in the wiki. We have described every attribute with a metadata set, similar to the one used in the SCHAC. Table 1 presents the attribute metadata definition.

*Table 1: Attribute metadata*

Name	Attribute name
Description	Short description
Format	Attribute format, including allowed values if a controlled vocabulary exists for the attribute.
Relevance	Base, recommended, optional, obsoleted
Source	Standard which the attribute is taken from
OID	Object identifier
URN	Identifier in URN format
LDAP Syntax	Attribute LDAP syntax according to RFC4517
Indexing	Type of indexes, though searching modes
No of vales	Single valued or multivalued
Examples	Examples in LDIF format
Notes	Additional notes or information about attribute usage

The document provides an exhaustive list of recommended attributes for categories listed in the process section. Although academic and grouping attributes are still a work in progress.

## 4.3. Use Cases

The group has produced up to now an extensive list of use case that are available in the public wiki. The list is not at all final. As it is our intention to continue working, but this list follows the approach that we have adopted for the whole process: making the community aware of our work as soon and broad as possible in order to get as much input as possible.

The list is classified in the following main categories, and most of them include several use case descriptions:

- Entry naming. Recommendations on how to name entries in a directory.
- Relationships. Use cases for describing the relationships of a person to the institution.
- Privacy. Use cases describing how to manage and protect users privacy.
- Backups and replication. How to preserve the integrity of the data.
- E-mail.
- Mailing lists
- Group management

- Non institutional user accounts

## 5. RESULTS

The most relevant result, apart from the deliverables presented above, has been the keen take up of the recommendations by the community. The use of the proposed schema is spreading fast and many institutions are adopting in some extent.

So, we should consider <http://wiki.rediris.es/gtschema/index.php> to be the most important result of the group's work.

The excellent level of adoption of the schema and its derivative works, has prompted the formation of another group with closer ties to the corporate database arena, but connected to the schema work group through common members. This systems constitute very important sources of identity and person data in our institutions.

## 6. FUTURE DEVELOPMENT

We do not consider our work a finished one, actually, we consider it a never ending task. The main tasks for the close future are:

- Increasing the visibility of the results
- Translating the materials into English
- Attacking the task of complex institutional relationships
- Dealing with groups and group management
- Describing academic information with common shareable syntax and semantics

## 7. REFERENCES

eduPerson & eduOrg Object Classes, retrieved May 25<sup>th</sup> 2009 from: <http://middleware.internet2.edu/eduperson/>

GT-SCHEMA Public wiki, retrieved May 25<sup>th</sup> 2009 from: <http://wiki.rediris.es/gtschema/index.php>

“Higher-Education Person: A Comparative Analysis of Collaborative Public LDAP Person Object Classes in Higher-Education”, retrieved May 25<sup>th</sup> 2009 from: <http://middleware.internet2.edu/dir/docs/draft-internet2-mace-dir-higher-ed-person-analysis-latest.htm>

Legg, S, ed. (2006) RFC4517 “Syntaxes and Matching Rules”, retrieved May 25<sup>th</sup> 2009 from: <http://www.rfc-editor.org/rfc/rfc4517.txt>

Sciberras, A, ed. (2006) RFC4519 “Schema for User Applications”, retrieved May 25<sup>th</sup> 2009 from: <http://www.rfc-editor.org/rfc/rfc4519.txt>

RedIRIS Directory Schemas, retrieved May 25<sup>th</sup> 2009 from: <http://www.rediris.es/ldap/schema/>

SCHAC, SCHEMA for Academia, retrieved May 25<sup>th</sup> 2009 from: <http://www.terena.org/activities/tf-emc2/schac.html>

Smith, M (2000), RFC2798 “Definition of the inetOrgPerson LDAP Object Class”, retrieved May 25<sup>th</sup> from <http://www.rfc-editor.org/rfc/rfc2798.txt>