Donantonio, an architecture for automated electronic publishing (draft)

Ismael Olea
<mailto:olea@hispalinux.es>
v0.01 1999-09-02

The very basic ideas for an architecture for publishing and maintaining all kind e-libraries using known standards as XML, RDF, Dublin Core, etc.

1 Introduction

Please read the introducction about a global document project and the technologies involved see the document «Electronic libraries: documentation technology».

tecnologia-doc-es
: the Spanish version.

tecnologia-doc-en
: the English version.

2 Donantonio design

2.1 basic planning

- based on true free software, using the GPL license or compatible if it's possible;
- open especifications clearly documented;
- based on true standards as RFC's, ISO and 3WC standards.
- simple, modular and flexible to become the layer for distributed applications

2.2 The design goals are:

- automatic publishing;
- distributed publishing along the Net;
- distributed publishing ¿responsabilidad?;
- automated and programmable view organization —system and user—;
- mirroring facilities;
- automatic updating facilities;
- flexible relations between **Donantonio** agents.

2.3 Necessary elements (some of them):

- meta-data declaration conforming a normalized schema (description);
- the file as the atomic transfer data unit;
- use of internet file transfer protocols;
- simple and flexible query system (¿SQL?).

3 Architecture versus application.

This paper is focused in architecture only. The idea is to design a set of conventions and rules to build a working environment to develop applied publishing tools.

For this cause I'll try not to say too much about the Metalab/LDP publishing system. Indeed, the designing work for these appliances will help to refine Donantonio to get a multi true purpose and generalized architecture.

4 Donantonio definitions

4.1 Elements

Resource

file or a document located via an URL.

Description

a RDF file meta-data, located via an URL.

Repository

collection of resources or collection of resources' instances.

Library

collection of descriptions or collection of descriptions' instances.

4.2 Agent tasks

These tasks aren't completely defined yet.

Server

software acting as a server

Client

software acting as a client

Proxy

Proxy between two servers or between a server and a client

4.3 Actions

These are the "atomic actions":

Resource download

the act of downloading a resource referred by an URL via an Internet transfer protocol.

Resource mirroring

the act of mirroring (and sincronicing) a resource referred by an URL via an Internet transfer protocol.

Description download

the act of downloading a description referred by an URL via an Internet transfer protocol.

Description mirroring

the act of mirroring (and sincronicing) a description referred by an URL via an Internet transfer protocol.

Petition

a query on a transport layer (mime+http??).

Answer

of a query on a transport layer (mime+http??). It can contains a view.

And these are the "composed actions"

Library mirroring

TACA

Repository mirroring

TACA

Distributed query

TACA

4.4 Relations

Projection

relational representation in SQL language of a RDF schema.

Query

SQL query to a projection.

View

a set of one or more description Urls.

5 The Donantonio metaphor

5.1 The astral metaphor

An incomplete metaphor of the things Donantonio could support is the ªstral metaphor".

If we believe to some astrologists, the physical body has an astral body (the soul). In some circumstances the soul can travel without its physical body. But soul and body are linked with a silver bow. The astral body (the soul) can travel to any place of the world and ever is linked to his body, which is still sleeping somewhere on the Earth. But if the silver bow get broken, the soul is not linked to its body anymore. It's the death for its body.

5.2 Equivalence between Donantonio metaphor and the astral metaphor

For Donantonio thinking, documents are the body, descriptions are the astral bodies of its documents and are linked to it by a URL. If the link get broken the description lost all is usefulness and the documents die forgotten in somebody hard disk.

The principal idea is that while resources are sleeping in some place, their souls (meta-data descriptions) are traveling along the world, being processed and shared between Donantonio applications. Donantonio add the ubiquity faculty letting several duplicated souls for a only resource.

The soul's movememt capabilities are configured by the relations between agents. Movement makes happen in the form of actions, performed by agents.

5.3 Practical consequences

- Resources had fixed publishing points (one or many) referenced by a URL.
- A resource never travel across the net.
- A resource can be downloaded. The downloaded resource is an instance of the resource and it hasn't bibliographic relevance.
- Descriptions had a one and only fixed publishing point referenced by a URL.
- Descriptions instances can travel across the net to by archived or processed. A description instance has bibliographic relevance while they had a valid link with their master description.
- A repository can be a Donantonio repository while all its resources had its description in a Donantonio application somewhere in the net.
- A Donantonio library is a true library while all its descriptions or descriptions instances are valid (they refer to their own resources).
- It's too easy to replicate libraries only replicating its descriptions files. A way to distribute load distributing server accesses.
- The way for building indexing clusters or indexing servers trees is distributing the descriptions their self between the servers. An information petition on one of the servers (the cluster case) or in the main server (the tree case) is distributed to the others with a Donantonio query. Them, the systems acts as one only server.
- Descriptions can be grouped into views with SQL queries and explicit Urls.

- A distributed/monolithic library can be replicated with only a query (a view) into a monolithic mirror or with several well designed queries (several views) into a net of servers.
- The resources referenced by a library (or many libraries) can be instanced locally (cached) with views and URL downloads.

6 Who's Don Antonio?

Don Antonio is the librarian of the high school were I had my bachelor degree.

I like this name because it's different and sounds funny :-)

7 Bibliography

- XML
- DC
- RFC 2413 (DC)
- RDF
- Los de la clasificacion automatica.
- URL de LuCAS
- de insflug
- de Ofelia
- del borrador del libro de estilo de Insflug
- al glosario de LuCAS
- ldp-core.dtd: http://casbah.org/cgi-bin/cvsweb.cgi/XML/DTDs/ldpCoreMetaML.dtd Creada por Kendall Clark en XML. Se supone un paso previo a la versión RDF.

8 Change-log

v0.0, first release in English