

Eprints Archive Software System Documentation

Robert Tansley

November 29, 2000

1 Introduction

This document describes the workings of the Eprints archive software, and how it interacts with the other components to provide services.

What this document does not contain is an API specification. Nor does it contain detailed information about what each file in the distribution does. The code is richly commented; this information is by and large not duplicated in this document. Producing such a document would have delayed the *Eprints* release by weeks. The best way to learn about the system, and how to achieve something, is to look at the code, read the comments therein, and follow the code's example. The comprehensiveness of this document should improve over time.

Before I delve in, a short note about programming style. The code is designed to be viewed with a tab width of 3. Tabs are *always* used for indentation. Indentation of eight spaces is far too much, and using a combination of spaces and tabs just causes problems.

In your site specific code, and especially if you're developing the core code itself, please follow the style of the existing code, and comment richly, to enable the *Eprints* software to be easily maintainable in the future.

1.1 Further Help and Bug Reports

Technical queries (*not*

HTML files are filled in" with appropriate values in place of the placeholders, and given the site "skin" (look and feel) if appropriate.

2.2 The html Directory

The contents of the html directory are controlled entirely by the software; you should not edit or add things directly. If you want to add an HTML page, graphic or other miscellaneous file to your site, add it to the static directory and run `generate_static` (or `update_laf`.) The file will be added to the archive Web server.

boolean	true or false
date	a date (day, month and year)
email	an e-mail address
enum	one of a specified set of values
eprinttype	a type of eprint (generally only used by system)
int	integer value (unsigned)
multitext	(potentially) several lines of text
multiurl	several WWW site addresses
name	a person's name, or list of people's names
pagerange	a range of pages (or a single page)
set	. integer value (unsigned) —

Argument	Types it applies to	Description
type	ALL	the type of field, from table 1
required	ALL	is the field required? (YES or NO)
displayname	ALL	the name of the field to display to the user.
editable	ALL	is the field editable by users (as opposed to purely staff members?) (YES or NO)
visible	ALL	is the field globally visible?
help		

required/editable/visible Correspond to the arguments for the verbose specification. A `1' means yes, a `0' means no.

indexed This means the field will be indexed by the MySQL database. For this to be possible, the field must ALWAYS be non-null, and unique. It is up to the code to ensure that this is the case. For this reason, only

users Hold user records.

i nbox

Field Type	Stored as MySQL Type	Internal Representation (MySQL & <i>Eprints</i> code)
boolean date	SET('TRUE','FALSE')	`TRUE' or `FALSE'

before Open Archives compliant services become commonplace. In the very near future, more powerful search and linking engines will be able to provide far more

5.1.3 email, multiurl and url elds

HTML Form: Rendered as a text input eld, accepting the internal representation directly.

5.1.6 set and subject fields

HTML Form: Rendered as a list of possible values, with a popup menu on the right hand side from which the user can select "Any of these" or "All of these." If the user selects "Any of these," the search will retrieve any record which has any value the user has selected in the relevant field. If the user selects "All of these", a retrieved record must have all of the values the user has selected in the list. The name of this extra input field is fieldname

Whether records should satisfy *all* search fields to be retrieved, or just one. Basically, this just means are the search fields AND'd or OR'd?

How the results should be ordered.

The default ordering of results.

The ordering is specified in terms of SQL, for example:

`year DESC, authors, title`

This sorts results by descending year, then authors, then title. The default order if ASCending or DESCending are not specified is ascending.

SearchExpressi on.pm can also produce a text string which, at a later time, can be fed back into SearchExpressi on.pm to recreate the same search. This technique is used by the subscriptions mechanism to store and retrieve use3 Tf 0 -11.955 Td[(subscription

code will remain the same, so sites running the *Eprints* software are protected against changes in the protocol.

Using the protocol, an *Eprints* archive can export several metadata sets. By default, Dublin Core metadata and OAMS (Open Archives Metadata Set) metadata are exported.

Two methods in the `SiteRoutines.pm` module are used to configure the metadata your instance of *Eprints* will export. The `oai`

If the email isn't recognised, a return is sent with instructions on how to use the *Eprints* automatic mail processing core. You can change the wording of this text, including the "introduction to the archive" and the "email successfully changed" by changing the template files in the `config` directory.

9 Scripted Importing of Data

It's likely that you have a large amount of data in another database or format that you want to populate your *Eprints* database with. The *Eprints* Perl modules make this importing trivial.

The best place to find more information is in the `59(y)00(nd)-297(information)-298(ab)-27(out)-299(metho)-28(ds)-299(is)-299(in)-299(5h`

