RIOXX Plugin for EPrints

User Guide

# Introduction

## What is RIOXX?

"The RIOXX Guidelines and Application Profile provide a mechanism for institutional repositories in use in the UK Higher Education sector to comply with the RCUK policy on open access[, focusing on] applying consistency to the metadata fields used to record research funder and project/grant identifiers."

http://rioxx.net/

For the rest of us, this is how I tried to explain RIOXX to a colleague (who has an Arabic background):

*The word "Amma" means "mother" in Urdu, but "aunt" in Arabic.*

*So an Urdu-speaking person might find it hard to talk to an Arabic-speaking person about the concept of "mother" - unless they can come to a mutual understanding about the meaning of "Amma".*

*Shared understanding is the key to standards like RIOXX. Just as in human conversation, unless computer systems agree about the of the meaning of the data they exchange all sorts of problems can arise - misinterpretation, mistakes, and ultimately malfunction.*

*However whereas human understanding is often implicit and tacit, a mechanism is required to formally and explicitly define the meaning of concepts and data shared between computer systems. This definition allows each system to expose it's own data in a way that can unambiguously understood by other systems, and to likewise be able to interpret the data exposed by other systems.*

*Enter RIOXX, an application profile - or formal definition of the concepts, policies and guidelines underpinning a computer application - for communicating and exchanging information about research publications: how and where they are published, how they are funded and ultimately how they are made available to the tax-paying general public.*

*Because when it comes to research grants, you don't want to be overlooked because the funding council confused your mother with your aunt.*

## What does the RIOXX plugin for EPrints do?

The RIOXX plugin gives your EPrints repository the capability to expose its publication data using the RIOXX application profile.

## How does it work?

The plugin works by adding 21 new metadata fields to your EPrints repository, one for each of the elements in the RIOXX application profile (you can count them all here: http://rioxx.net/v2-0-beta-1/). These fields let you capture the necessary data about each publication in order to create a RIOXX representation of that publication. The plugin will check the data that you enter in these fields against the application profile and warn you if it finds any problems.

## But I'm already capturing some of that data, does that mean I have to enter it twice?

No - we've created a special kind of metadata field that can work out (derive) a suitable value for RIOXX by looking at the data in your existing metadata fields. This field doesn't even have to appear in your workflow, so you would just enter data about each publication as normal.

## That sounds easy, is that all there is to it?

Unfortunately it's slightly more complicated than that, for two reasons:

1. There is some data that RIOXX requires that is simply not captured by the (default) EPrints workflow

The dcterms:dateAccepted element is a good example. The RIOXX application profile tells us that this data is mandatory, however there is no field in the (default) EPrints workflow specifically designed to capture this information (by using a combination of the Date and Date Type fields you can enter either the date of publication, submission or completion; but not the date of acceptance). One solution could have been to ask every institution in the UK to update their Date Type field to include an "Accepted" option - but what if institutions wanted to capture both acceptance AND publication dates?

Therefore we decided that we would create a new dedicated field for capturing the date of acceptance and automatically insert this field into the workflow when the RIOXX plugin is installed.

1. Sometimes derived values can only be a "best guess"

In most cases we can be confident that the value we derive for RIOXX is accurate. Take for example the RIOXX dc:title and dc:description elements - if we copy the publication title into the dc:title element and the abstract into the dc:description element it follows that - as long as the title and abstract are entered correctly in the first place - RIOXX will get the right data.

In other cases, we cannot be as confident about the value we derive. The RIOXX rioxxterms:project element for example, is mandatory and "designed to collect the project ID(s), issued by the funder(s), that relate to the resource, together with the name and/or global identifier for the funder(s)". The default EPrints workflow has 2 fields that we can use to derive the value of this element - Projects and Funders. Both fields accept multiple entries, but there is no explicit link between them ie. users are not forced to enter a funder for each project or vice versa.

Given the following entries for Projects and Funders (and assuming that the user has entered project IDs into the Projects field and funder names into the Funders field - again this is not enforced):

|  |  |
| --- | --- |
| Projects | Funders |
| 1. JA/139K/02 2. GLX\_456\_P 3. JP/894Y/77 | 1. EPSRC 2. NERC |

The RIOXX plugin will derive the following value:

|  |
| --- |
| rioxxterms:project |
| 1. JA/139K/02 funded by EPSRC 2. GLX\_456\_P funded by NERC 3. JP/894Y/77 funded by NERC |

..which may or not be accurate!

Our solution in this case was to give the user the option to override the value derived from Projects and Funders by entering a new value of their own. This required another new field to be added to the EPrints workflow (automatically added when the RIOXX plugin is installed) which is designed specifically for RIOXX. The user can enter a list of project IDs and for each entry a funder name or funder ID is explicitly required.

It is up to the Institution to decide how best to adopt this approach, for example:

* Continue using the Projects and Funders fields and use the RIOXX override to enter a separate value for RIOXX
* Adjust the use of the Projects and Funders fields - for example so that the first Funders entry is the funder for the first Projects entry, the second Funders entry is the funder for the second Projects entry, and so on - and let the plugin derive the RIOXX value
* Discontinue use of the default Projects and Funders fields (remove them from the workflow) and use the field added by the RIOXX plugin to enter project and funder data (no derived value necessary)

## Which fields will be automatically added to our workflow?

Installing the RIOXX plugin automatically adds a new "RIOXX" stage to the EPrints workflow with the following fields:

|  |  |  |
| --- | --- | --- |
| **Field** | **How is value derived?** | **How is (override) value entered?** |
| Type | Predefined mapping between default EPrints deposit types and RIOXX types | Drop down list of RIOXX types |
| Coverage | Not derived | Free text |
| Language | Language of selected document[[1]](#footnote-2) or default repository lanaguage | Language selection list |
| Date accepted | Not derived | Year, month, day entry |
| Free to read | Selected document OA or restricted with embargo date | Year, month, day entry for start and end dates |
| License ref | TODO | License URL and year, month, day for start date |
| APC | Not derived | RIOXX APC selection list |
| Author | TODO |  |
| Project | Combined Projects and Funders fields |  |
| Publication Date | TODO |  |
| Version | TODO |  |

## But I've already customised my EPrints workflow to capture some/all of that data, what can I do?

EPrints is a flexible platform and it is straightforward to add additional metadata fields and/or refine the default workflow - therefore it is quite reasonable to assume that most of the UK institutions that use EPrints will have customised their workflow to some extent.

We cannot anticipate all the different changes individual institutions will have made so we have focused instead on providing mechanisms to allow institutions to align the RIOXX plugin with the data that they are already collecting.

As an example, let's imagine that an institution had already recognised EPrints' shortfall in capturing acceptance dates and had decided to create a new "Date of Acceptance by Publisher" field and add it to their workflow. In this, having installed the RIOXX plugin, the institution would simply need to adjust the plugin configuration so that the plugin's own "Date accepted" field (a) no longer appeared in the workflow and (b) derived its value directly from the new "Date of Acceptance by Publisher" field added by the institution.

The steps required to achieve this, as well as other configuration changes, are described in the Configuration Guide.

## Shouldn't the default EPrints workflow just be updated to reflect the RIOXX requirements?

Probably, but the effort involved in migrating all UK EPrints repositories to a common workflow would be outside the scope of this project. Our hope is that we've designed the RIOXX plugin for EPrints in such a way that it will be possible to take advantage of any workflow refinements in future EPrints releases, or even new features added by other plugins, by making straightforward adjustments to the RIOXX plugin configuration.

# Using the RIOXX Plugin

## Installation

## Changes to Workflow

## RIOXX Tab

## Compliance Reporting

1. See "What about publication records with more than one document?" [↑](#footnote-ref-2)