

Summary of ALM alerts

Scott Chamberlain 10 November, 2014

Introduction

The [Lagotto application](#) collects and provides article-level metrics data for scholarly articles. As part of a data integrity process, various alerts are given from Lagotto that help determine what may be going wrong with the application, data sources used in Lagotto, and any problems with users requesting data from the Lagotto application. Analyzing these alerts is helping to determine what errors are the most common, and what may lie behind errors.

Scott Chamberlain has been working on an R client to work with Lagotto application data, called `alm`. This R client can also interact with alerts data from Lagotto.

As other publishers are starting to use Lagotto, the below is a discussion mostly of PLOS data, but includes some discussion of other publishers.

How to interpret alerts

Alert class name	Description
Net::HTTPUnauthorized	401 - authorization likely missing
Net::HTTPForbidden	xxxxxx
Net::HTTPRequestTimeout	408 - request timeout
Net::HTTPGatewayTimeout	xxxxxx
Net::HTTPConflict	409 - Document update conflict
Net::HTTPServiceUnavailable	503 - serve is down
Faraday::ResourceNotFound	404 - resource not found
ActiveRecord::RecordInvalid	title is usually blank, and can't be
Delayed::WorkerTimeout	xxxxxx
DelayedJobError	xxxxxx
TooManyErrorsBySourceError	xxxxxx
SourceInactiveError	xxxxxx
TooManyWorkersError	xxxxxx
EventCountDecreasingError	Event count decrease too fast, check on it
EventCountIncreasingTooFastError	Event count increasing too fast, check on it
ApiResponseTooSlowError	Alert if successful API responses took too long
HtmlRatioTooHighError	HTML/PDF ratio higher than 50
ArticleNotUpdatedError	Alert if articles have not been updated within X days
SourceNotUpdatedError	xxxxxx
CitationMilestoneAlert	Alert if an article has been cited the specified number of times

PLOS

PLOS has 137753 articles available in their Lagotto instance as of 2014-11-10.

```
## Error in alm_alerts(page = x): server error: (504) Gateway Timeout
```

```
## Error in lapply(res, "[", "data"): object 'res' not found
```

```
## Error in eval(expr, envir, enclos): object 'resdf' not found
```

Crossref

Crossref has the biggest collection of articles in any Lagotto application, with 11678341 as of 2014-11-10.

class_name	n
ActiveRecord::RecordInvalid	6067
Net::HTTPServiceUnavailable	814
Net::HTTPRequestTimeout	412
Faraday::ClientError	285
Net::HTTPInternalServerError	155
Net::HTTPUnauthorized	79
Net::HTTPConflict	58
StandardError	28
FaradayMiddleware::RedirectLimitReached	21
Faraday::ResourceNotFound	10
ArgumentError	1
Net::HTTPBadRequest	1
TooManyErrorsBySourceError	1

PKP

PKP has a growing collection of articles, with 158095 as of 2014-11-10.

class_name	n
Net::HTTPForbidden	7533
Net::HTTPClientError	6928
DelayedJobError	2621
Net::HTTPServiceUnavailable	1623
Faraday::ResourceNotFound	697
Net::HTTPInternalServerError	226
Net::HTTPRequestTimeout	125
TooManyErrorsBySourceError	34
StandardError	19
Net::HTTPNotAcceptable	17
Net::HTTPBadRequest	13
ActiveRecord::RecordInvalid	12
ActiveRecord::StatementInvalid	11
NoMethodError	11
Net::HTTPBadGateway	5
ActionView::MissingTemplate	2