Digital Content Store

DCS Course Content URL API

Technical specification

Version 1.1

© The Copyright Licensing Agency 2017

Document Control

Change Control

<table>
<thead>
<tr>
<th>Version</th>
<th>Name</th>
<th>Position</th>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>Kieran Burke</td>
<td>Business Analyst</td>
<td>04 Oct 2016</td>
<td>First draft</td>
</tr>
<tr>
<td>0.2</td>
<td>Alex Cole</td>
<td>Senior Business Analyst</td>
<td>11 Oct 2016</td>
<td>Review and changes</td>
</tr>
<tr>
<td>1.0</td>
<td>Alex Cole</td>
<td>Senior Business Analyst</td>
<td>19 Oct 2016</td>
<td>Final version</td>
</tr>
<tr>
<td>1.1</td>
<td>Kieran Burke</td>
<td>Business Analyst</td>
<td>07 Apr 2017</td>
<td>Removed description from '/GetCourseContent' call</td>
</tr>
</tbody>
</table>

Table of Contents

- Digital Content Store
  - Document Control
    - Change Control
  - Table of Contents
  - Introduction
    - Purpose
    - Scope
    - Glossary
  - Service Description
    - Overview
    - Getting Started
    - Architecture
    - Subscription
    - Get Institutions
    - Get Courses
    - Get Course Content
    - State Changes in Courses and Content
    - Starting a new academic year
- Technical Specification
  - Introduction
  - Authentication
  - Workflow and sequence diagrams
Introduction

Purpose

This document describes the Digital Content Store (DCS) Course Content URL API which allows a client to present course data and the content links and metadata from subscribed institutions. It is intended that this service will be used by Third Party Vendors of technology to CLA Higher Education Institutions (HEIs) or by CLA Licensed HEIs to serve content links in their systems.

Scope

This document includes the three methods currently implemented in the API, along with specifications and examples. Changes that may occur in the API will provoke a new document being provided. This document will not include any information about how the data should be used or manipulated in a system, though sequence diagrams are provided to show call order scenarios.

The DCS API uses HTTP GET methods for the requests. Responses are given as JSON data.

Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCS</td>
<td>The Digital Content Store, a web application provided to CLA licensees in the Higher Education Sector</td>
</tr>
<tr>
<td>API</td>
<td>Application Programming Interface is a source code based specification intended to be used as an interface by software components to communicate with each other</td>
</tr>
<tr>
<td>HTTP</td>
<td>Hypertext Transfer Protocol. An application protocol for exchanging files (text, graphic images, sound, video, and other multimedia files) on the Web</td>
</tr>
<tr>
<td>GUID</td>
<td>Globally Unique Identifier</td>
</tr>
<tr>
<td>JSON</td>
<td>JavaScript Object Notation</td>
</tr>
<tr>
<td>URL</td>
<td>Uniform Resource Locator. The address of a resource accessible on the Web. The URL includes the name of the protocol required to access the resource, identifies the address of a specific server on the Web, and contains a hierarchical description of a file location on the server</td>
</tr>
<tr>
<td>DOI</td>
<td>Digital Object Identifier</td>
</tr>
<tr>
<td>ISBN</td>
<td>International Standard Book Number. Uniquely identifies a single manifestation of a work. ISBNs are used for monographic publications</td>
</tr>
<tr>
<td>ISN</td>
<td>International Standard Number. A generic term referring to both ISSN and ISBN identifiers</td>
</tr>
</tbody>
</table>
**Service Description**

This section gives an overview of the Digital Content Store (DCS) Course Content URL API and gives a high level description of the functionality. The subsections will include details on the methods in place, some information on the interaction of the DCS with the API, and important knowledge about system activity and periodic events.

**Overview**

The DCS is a workflow tool for librarians in HEIs to help with the process of sourcing and clearing digitisation requests required for course packs and reading lists. Documents are stored in the system as PDF files. When a request is completed a URL is published that allows authorised users at the institution to access the content. Each content URL is specified by course and requires the end user (typically students) to authenticate to access the content. The DCS API allows external systems to publish the content links and associated metadata on a course by course basis. CLA will create the client system credentials within the DCS so that administrators of the HEI accounts can let the clients subscribe to their courses.

The API is comprised of three methods: `GetInstitutions`, `GetCourses` and `GetCourseContent`. Content retrieval is only possible using the HEI and course codes provided by the relevant methods.

**Getting Started**

The following steps need to be completed

1. Terms and Conditions agreed
2. CLA sets up the API user account the on DCS system
3. Testing can commence on the demo (aka sandbox) environment
4. Once testing is complete, HEI administration users can create the subscription of the API system for their institution on the production environment

**Architecture**

The DCS API uses the HTTP protocol and returns data in the JSON format.

**Subscription**

Once the API account has been added to the DCS system, HEI users (HEI administrator level only) will have the option to subscribe the API to their institution. Once they have allowed the subscription their details and courses and content will be available to the calling API system.

As HEIs only appear in the `GetInstitutions` response when they have allowed the subscription to the API system, you will have to notify the HEI that at your system is now available, as CLA do not have the right to subscribe on behalf of an HEI. HEIs can be referred to our helpdesk and knowledgebase available at [https://cla.zendesk.com/hc/en-us](https://cla.zendesk.com/hc/en-us).

**Get Institutions**

This service can be used to check which HEIs have enabled access to your platform/service and to retrieve the HEI IDs needed on the subsequent calls.
Get Courses

Using the retrieved HEI ID, this method will provide you with a list of courses in the current academic year for the nominated HEI and data relating to the course. Each course has a course code that is used for identifying the course in the GetCourseContent calls.

Get Course Content

The links for a nominated course are provided with associated metadata, including bibliographic details and some details relating to the file that is accessed through the link. The HEI ID and the course code must both be provided.

State Changes in Courses and Content

See Appendix 2 for content status definitions.

Courses may be in one of the following states:

- Active
- Archived
- Deleted

Calling the GetCourses method will only return courses in the Active state. As courses transition to Archived or Deleted from being Active they will be removed from subsequent updates whilst remaining in any non-active states.

Courses in the DCS can be Active, Archived or Deleted, and can be done at the HEIs’ discretion. Only active courses are shown in the GetCourses method, and therefore an archive action or deletion will remove it from the API output. In addition to this, if the GetCourseContent method is called with an archived course as the parameter, then the call will return a success status code, but the response will not contain any content data, as the change in course state causes the content states to change.

Content items in the DCS may be in one of the following states:

- Active
- Archived
- Deleted
- Pending

HEIs are responsible for managing the changes of state. All states except Deleted will be returned in the GetCourseContent call, as long as they have been made public by the institution at some point. This is to say, if a piece of content is moved from active to archived or pending and vice versa, then this change will be seen in the content status field. However, a change from active, archived or pending to deleted will mean that the content is not shown in the response. Links to content are only provided for content items in the active state.

Starting a new academic year

See Appendix 2 for content status definitions.

See Appendix 3 for a chart explaining the timeline of events.

When an institution wants to move into the next academic year, their courses will also all be carried over, and this will be shown by the academic-year field in the GetCourses call. This is typically performed between 1st June and 31st July. Since the choice of timing is in the HEIs’ control, it is their responsibility to inform any external technology providers, if required. Since the GetCourses call can only be used to get the courses in the current academic year, once an institution has started the next academic year, courses in the previous academic year will no longer be able to be queried via the API.

In the first week of August, CLA will recheck all the items used under the CLA licence to verify that course content is covered in the following academic year under the CLA licence or permissions need to be sourced elsewhere. This may cause course content to go from active to pending if the permissions change. Librarians and academic support staff will be able to source alternative permissions or content at this time, triggering changes in content status and availability of content.

Technical Specification

Introduction

Authentication
The DCS API uses a standard “basic” authentication scheme. All calls to the API need to include an HTTP header with appropriate credentials. The authorization field consists of a username and a password combined with a single colon encoded in base64 prepended by the word “Basic”. The HTTP request’s Authorization header should then be set with the resulting string.

The process to create a proper authentication header is as follows:

1. Define a string in the following format: “username:password” (ignore quotes).
2. Convert this string into a base64-encoded string.
3. Prefix the word “Basic” (again, ignore quotes) to the base64-encoded string, separated by a space.
4. Set the HTTP request’s authorization header with the final string.

For example, if the caller has the username “Foo” and the password “Bar” the resulting header would look like:

```
authorization: Basic Rm9vOkJhcg==
```

In order to ensure the security of this scheme, all calls will be over SSL using a recognized security certificate, thus resulting in only encrypted data sent to and from the DCS Course Content URL API. More information on basic authentication can be found here (RFC2617): https://tools.ietf.org/html/rfc2617#section-2

**Workflow and sequence diagrams**

### Get all institution IDs

This sequence will retrieve the IDs that are associated with the HEIs that have allowed subscriptions. This can be used to confirm that an institution has granted access.

![Workflow for getting all institution IDs](image)

### Get all course and content updates for a particular institution

For a specified institution obtain the list of courses and the course content. Also used to check for new courses, updates to courses and updates to content.

This sequence can be used to update a specific institution courses and content items.
Get all course and content updates for all institutions

This sequence can be used to provide the full set of data so the client system can be completely updated to match the current status of the Digital Content Store. This is to be used at times when all institutions need to be updated, such as after CLA annual permissions recheck.
### Get Institutions

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>status</td>
<td>string</td>
<td>The status of the API response (&quot;ok&quot;, &quot;error&quot;)</td>
</tr>
<tr>
<td>status-code</td>
<td>int</td>
<td>Status-code of the API response</td>
</tr>
<tr>
<td>status-message</td>
<td>string</td>
<td>Message related to the API status code (if applicable)</td>
</tr>
<tr>
<td>total-results</td>
<td>int</td>
<td>Number of Institutions returned</td>
</tr>
<tr>
<td>Institutions</td>
<td>object[]</td>
<td>List of institutions associated to the API user</td>
</tr>
<tr>
<td>id</td>
<td>int</td>
<td>Institution ID</td>
</tr>
<tr>
<td>name</td>
<td>string</td>
<td>Institution Name</td>
</tr>
</tbody>
</table>

```json
{
    "status": "string",
    "status-code": 0,
    "status-message": "string",
    "total-results": 0,
    "institutions": {
        "id": 0,
        "name": "string"
    }
}
```

### Error Response

```json
{
    "status": "error",
    "status-code": 3,
    "status-message": "Could not authenticate user"
}
```
Sample Call
GET https://{host}/GetInstitutions HTTP/1.1
authorization: Basic YourBase64encodedusername:password

Sample Response

GetInstitutions Response

{
    'institutions':
    [
        {'id': 209, 'name': 'API_TEST'},
        {'id': 210, 'name': 'API_TEST_OTHER'}
    ],
    'status': 'ok',
    'status-code': 100,
    'status-message': 'Success',
    'total-results': 2
}

Get Courses

<table>
<thead>
<tr>
<th>Get Courses</th>
</tr>
</thead>
<tbody>
<tr>
<td>URL</td>
</tr>
<tr>
<td>Method</td>
</tr>
<tr>
<td>URL Parameters</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Authentication</td>
</tr>
</tbody>
</table>
Success Response

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>status</td>
<td>string</td>
<td>The status of the API response (&quot;ok&quot;, &quot;error&quot;)</td>
</tr>
<tr>
<td>status-code</td>
<td>int</td>
<td>Status-code of the API response</td>
</tr>
<tr>
<td>status-message</td>
<td>string</td>
<td>Message related to the API status code (if applicable)</td>
</tr>
<tr>
<td>total-results</td>
<td>int</td>
<td>Number of Institutions returned</td>
</tr>
<tr>
<td>courses</td>
<td>object[]</td>
<td>List of institutions associated to the API user</td>
</tr>
<tr>
<td>academic year</td>
<td>string</td>
<td>Academic year the course is currently in</td>
</tr>
<tr>
<td>id</td>
<td>int</td>
<td>Id of the current course</td>
</tr>
<tr>
<td>course-code</td>
<td>string</td>
<td>The course code of the current course</td>
</tr>
<tr>
<td>name</td>
<td>string</td>
<td>Name of the course</td>
</tr>
<tr>
<td>duration</td>
<td>int</td>
<td>Duration of the course in weeks</td>
</tr>
<tr>
<td>lecturer</td>
<td>string</td>
<td>The lead lecturer assigned to the course</td>
</tr>
</tbody>
</table>

```json
{
    "status": "string",
    "status-code": 0,
    "status-message": "string",
    "total-results": 0,
    "courses": {
        "academic-year": 0,
        "id": 0,
        "course-code": "string",
        "name": "string",
        "duration": 0,
        "lecturer": "string"
    }
}
```
Error Response

```
{
    "status": "error",
    "status-code": 1,
    "status-message": "Institution not found"
}
```

```
{
    "status": "error",
    "status-code": 5,
    "status-message": "User not subscribed to HEI"
}
```

```
{
    "status": "error",
    "status-code": 6,
    "status-message": "Invalid Parameter"
}
```

Sample Call

```
GET https://{host}/GetCourses?hei=206 HTTP/1.1
authorization: Basic YourBase64encodedusername:password
```
GetCourses Response

{
    "courses": [
        {
            "academic-year": '2016-2017',
            "course-code": 'HIST101',
            "duration": 52,
            "id": 70668,
            "lecturer": 'Jane Bunt',
            "name": 'Introduction to World History'
        },
        {
            "academic-year": '2016-2017',
            "course-code": 'ENG101',
            "duration": 26,
            "id": 70918,
            "lecturer": 'Steve McGill',
            "name": 'Introduction to English Language'
        },
        {
            "academic-year": '2016-2017',
            "course-code": 'LIT500',
            "duration": 52,
            "id": 70919,
            "lecturer": '',
            "name": 'English literature and the works of Shakespear'
        }
    ],
    "status": 'ok',
    "status-code": 100,
    "status-message": 'Success',
    "total-results": 3
}
### URL
/GetCourseContent? hei={hei}&code={code}

### Method
GET

### URL Parameters
**Required:**
- hei=[integer]
- code = [string]

### Authentication
Required

### Success Response

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>status</td>
<td>string</td>
<td>The status of the API response (&quot;ok&quot;, &quot;error&quot;)</td>
</tr>
<tr>
<td>status-code</td>
<td>int</td>
<td>Status-code of the API response</td>
</tr>
<tr>
<td>status-message</td>
<td>string</td>
<td>Message related to the API status code (if applicable)</td>
</tr>
<tr>
<td>total-results</td>
<td>int</td>
<td>Number of Institutions returned</td>
</tr>
<tr>
<td>HEI</td>
<td>string</td>
<td>Name of the HEI</td>
</tr>
<tr>
<td>course-ID</td>
<td>int</td>
<td>ID of the course the data comes from</td>
</tr>
<tr>
<td>content-items</td>
<td>object[]</td>
<td>List of content items associated to the given course</td>
</tr>
<tr>
<td>content-GUID</td>
<td>guid</td>
<td>Unique Id for the content link</td>
</tr>
<tr>
<td>content-status</td>
<td>string</td>
<td>Content's status, possible outputs:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Pending</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Active</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Archived</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Deleted</td>
</tr>
<tr>
<td>last-modified</td>
<td>DateTime</td>
<td>Date and time the content item was last modified.</td>
</tr>
<tr>
<td>content-URL</td>
<td>string</td>
<td>Content item’s current DCS link</td>
</tr>
<tr>
<td>bibliographic-details</td>
<td>object[]</td>
<td>Content item’s bibliographic data</td>
</tr>
<tr>
<td>type</td>
<td>string</td>
<td>Possible outputs:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Book</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Journal</td>
</tr>
<tr>
<td>identifier</td>
<td>string</td>
<td>Content Item's ISN</td>
</tr>
<tr>
<td>DOI</td>
<td>string</td>
<td>Journal's DOI if the content is a journal</td>
</tr>
<tr>
<td>title</td>
<td>string</td>
<td>Content's title</td>
</tr>
<tr>
<td>extract-title</td>
<td>string</td>
<td>Content's article or chapter title</td>
</tr>
<tr>
<td>publication-form</td>
<td>string</td>
<td>Possible outputs:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Print</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Digital</td>
</tr>
<tr>
<td>year</td>
<td>string</td>
<td>The year the content was published</td>
</tr>
<tr>
<td>volume</td>
<td>string</td>
<td>The content's volume number</td>
</tr>
<tr>
<td>issue</td>
<td>string</td>
<td>The content's issue number</td>
</tr>
<tr>
<td>page-range</td>
<td>string</td>
<td>The page range extracted from the content</td>
</tr>
<tr>
<td>Field</td>
<td>Type</td>
<td>Description</td>
</tr>
<tr>
<td>---------------</td>
<td>---------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>author</td>
<td>string</td>
<td>The author of the content</td>
</tr>
<tr>
<td>colour-scale</td>
<td>string</td>
<td>The content’s colour scale. Possible outputs:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• BlackAndWhite</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Greyscale</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Colour</td>
</tr>
<tr>
<td>publisher</td>
<td>string</td>
<td>The content’s publisher</td>
</tr>
<tr>
<td>extract-author</td>
<td>string</td>
<td>The content’s extract author</td>
</tr>
<tr>
<td>chapter-number</td>
<td>string</td>
<td>The content’s chapter number</td>
</tr>
<tr>
<td>edition</td>
<td>string</td>
<td>The content’s edition</td>
</tr>
<tr>
<td>book-pages</td>
<td>int</td>
<td>The content’s book page count</td>
</tr>
<tr>
<td>publication-place</td>
<td>string</td>
<td>The content’s publication place</td>
</tr>
<tr>
<td>OCR</td>
<td>bool</td>
<td>Does the content have optical character recognition?</td>
</tr>
<tr>
<td>file-size</td>
<td>double</td>
<td>The content’s file size in KB</td>
</tr>
</tbody>
</table>
Error Response

{
  "status": "error",
  "status-code": 1,
  "status-message": "Institution not found"
}

{
  "status": "error",
  "status-code": 2,
  "status-message": "Course not found"
}

{
  "status": "error",
  "status-code": 6,
  "status-message": "Invalid Parameter"
}

Sample Call

GET https://{host}/GetCourseContent?hei=209&code=ENG101 HTTP/1.1
authorization: Basic YourBase64encodedusername:password

Sample Response

GetCourseContent Response

{
  'content-items':
  [
    
    'bibliographic-details':
    {
      'DOI': None,
      'OCR': False,
      'author': 'ed. by Michael Lewis; Jeannette M. Haviland.
      'book-pages': None,
      'chapter-number': None,
      'colour-scale': "BlackAndWhite",
      'edition': None,
    }
  ]
}
Examples

Demo Environment

The DCS demo environment can be accessed here can be signed up to on the CLA API portal, and using the API here.

The following institutions and courses have been set up on the demo environment in order to demonstrate the possible return values of calls

Institutions

Example Response: Institutions

```
'institutions':

[
  {'id': 209, 'name': 'API_TEST'},
  {'id': 210, 'name': 'API_TEST_OTHER'}
]
```

Courses

The API_TEST University will have the following courses
Course Content

The course with code ENG101 under API_TEST University will retrieve the following items:

```
Example Response: Course Content

'content-items':
  [{
    'bibliographic-details':
      {
        'DOI': None,
      },
    'lecturer': '',
    'name': 'English literature and the works of Shakespear'
  }]
```
Appendix

Appendix 1 - Error Message

Response fields

<table>
<thead>
<tr>
<th>Field</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>status</td>
<td>string</td>
<td>The status of the API response (&quot;ok&quot;, &quot;error&quot;)</td>
</tr>
<tr>
<td>status-code</td>
<td>int</td>
<td>Status-code of the API response</td>
</tr>
<tr>
<td>status-message</td>
<td>string</td>
<td>Message related to the API status code</td>
</tr>
</tbody>
</table>

Status Codes

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Institution not found</td>
</tr>
<tr>
<td>2</td>
<td>Course not found</td>
</tr>
<tr>
<td>3</td>
<td>Could not authenticate user</td>
</tr>
<tr>
<td>4</td>
<td>Internal server error</td>
</tr>
<tr>
<td>5</td>
<td>User not subscribed to HEI</td>
</tr>
<tr>
<td>6</td>
<td>Invalid Parameter</td>
</tr>
<tr>
<td>100</td>
<td>Success</td>
</tr>
</tbody>
</table>

Appendix 2 - Content Statuses

<table>
<thead>
<tr>
<th>Status</th>
<th>Description</th>
<th>Available in GetCourseContent call?</th>
<th>Content can be accessed through links by students?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>This means the content has been published and is usable by students</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Pending</td>
<td>Pending content is being prepared by librarians in order to become active</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Archived</td>
<td>Archived content is temporarily put into this state so that it can be used at another time without being setup again</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Deleted</td>
<td>Deletion occurs when content is not going to be used in the foreseeable future</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Appendix 3 - Academic Year Timeline
Key dates

- 1st August - new licence year starts and the academic year starts, if not already initiated by the HEI
- May to July - course content rolled over into next academic year. This process will be completed by the end of July in most cases
- Beginning of August - CLA Permissions re-check