



newspaper licensing agency

My Archive “How To”

Contents

OVERVIEW	3
DATA SET COVERED	3
CONDITIONS	3
SEARCHING	4
DIRECT SALES.....	4
FEES.....	5
SECURITY.....	5
STANDARD AUTHENTICATION.....	6
IP AUTHENTICATION.....	6
MYARCHIVE API CALLS.....	6
MANAGE ARCHIVE	6
<i>Supported Parameters</i>	6
<i>Error Handling</i>	7
LIST ARCHIVE	7
<i>Supported Parameters</i>	7
<i>Error Handling</i>	7
SEARCH ARCHIVE.....	7
<i>Supported Parameters</i>	8
<i>Examples</i>	8
OUTPUT	9
<i>XML Example:</i>	9
<i>HTML Example:</i>	10
<i>ASCII Example:</i>	10
<i>Error Handling</i>	11
RESEARCH.....	12
DIGITAL CUTTINGS (84 RESPONDENTS)	12
THE CENTRAL DATABASE (164 RESPONDENTS)	12

Overview

This document is a “how to” for PCAs wishing to use the My Archive functionality of the new NLA eClips service.

Before undertaking the eClips project, the NLA did extensive research with our licensed organisations to determine what they need and want from a digital service. A significant number (~60%) of users identified that in a fully digital environment – unlike with paper based offerings – they would lose their archive unless it is transferred to paper. Therefore, the NLA is offering a digital archive to all end-clients of the eClips digital service.

The archive will serve as a digital ‘binder’ where users can store copies of their digital cuttings for later review. Current plans include saving cuttings for a 12-month period during which users will have unlimited access to view their stored cuttings. Users will be able to search their own archive to locate old cuttings, but will not be able to search the entire database.

Data set covered

All titles available through both the XML feed and the Access model will be covered by the ‘My Archive’ service. This includes all national titles and all additional key regional and non-London national titles.

All clippings may be archived immediately from the day any title goes into production.

As the archive product matures, the eClips team will look into the possibility of modifying the maximum archive time, but there are no current plans to extend beyond the 12 month limit¹.

Conditions

No more than 25 users from a single organisation are allowed – this service is designed for PR and communications professionals only.

No more than 25 cuttings from any one paper on any day can be included in the service – it does not facilitate creation of a local research archive.

¹ Please remember that the archive can only be created going forward.

The NLA feels that these limits should be more than adequate for 99.9% of all clients.

'My Archive' works only through licensed Press Cuttings Agencies with access to the NLA database.

PCAs using the eClips service can flag (automatically) all cuttings sent to a user with links back to the NLA database. Users with 'My Archive' rights enabled on their ID and password will be able to access these cuttings and search them by date, headline, newspaper, and through a free text search.

Access to My Archive clippings can be via a PCA created user interface or via a generic NLA My Archive user interface.

Searching

My Archive is not intended to be used as a research tool, so users will only be allowed to search all of *their own* archived cuttings for up to 12 months. Archive searches are intended to locate only specific cuttings already added to the archive.

An NLA user interface will be provided to facilitate searching, the details of which will be provided as soon as they are available.

Direct Sales

It became obvious during the consultation phase that most of the UK PCAs would like to take responsibility for marketing 'My Archive' directly to their clients and collecting all fees.

'My Archive' can be sold by the NLA as part of a digital licence or by authorised PCAs. It can only be offered if a PCA is using the NLA digital database because the service requires that article copies or links sent by a PCA are flagged on the NLA database.

We are happy to support direct sales by authorised PCAs, and have drafted a supplement to the PCA Licence Agreement that will outline in detail the terms and conditions of such sales.

However, the eClips team thought it would be good to clarify the basic rules here.

- PCAs will only be allowed to sell "My Archive" to their client base.
- The archive will be owned by the client, not the PCA.
- The PCA will be responsible for reporting all new My Archive clients to the NLA on a monthly basis.

- Upon termination of a client contract, the PCA must agree to facilitate the transfer of My Archive information for the client

PCAs should report all client activity through existing NLA contacts in Tunbridge Wells. The details of the reporting process are still being finalised by the NLA. Once accurate information is available, it will be added to this document.

Fees

'My Archive' will build to allow access to up to one year's data. The archive can only be created going forward, so will tend to be even more valuable during renewal years. For peace of mind on pricing we offer a two year fixed price, with an option to cancel after one year.

The 'My Archive' fees are as follows.²

1 user	£500
5 users	£1,000
10 users	£1,500
25 users	£2,000

All fees will be payable in full to the NLA within 30 days of enrolment.

Security

Access to My Archive will be allowed only on an individual user basis. Each user must have a unique ID and password and may not share their ID with other users.

The eClips team will regularly examine the user logs to look for suspicious activity. If any is found, an Account Executive will contact the relevant PCA to investigate further. We do not plan on contacting users directly unless it is absolutely necessary. All client interaction should be handled by the PCA.

There will be several different approved security models in place to accommodate for differing levels of technical sophistication.

² Fees are payable annually in advance.

Standard Authentication

The standard method of authentication is direct user authentication.

Users who have already viewed an eClips clipping will not need to re-enter their username and password to view archived clippings if they are clicking an existing link. The only time users will be required to re-authenticate will be when they wish to use the NLA's My Archive site. All users must authenticate on the NLA's User Management Interface (UMI).

IP Authentication

IP Authentication can be used for access to clippings as set out in the Technical Specification, but is not a valid authentication method for My Archive, since all access must be via either an individual cookie or manual authentication on the NLA UMI.

MyArchive API Calls

To allow PCAs to easily manage users' archives, the NLA has provided three API calls:

- managearchive – allows the addition and deletion of articles to an archive
- listarchive – allows a complete list of all included articles to be shown
- searcharchive – allows searching within a single archive

Manage Archive

Users may need to be able to add / delete items from their MyArchive, and display the contents of a MyArchive. The ability to use the Archive must be validated against the user's licence.

Supported Parameters

Parameter	Mandatory?	Default	Example	Notes
UserID	Yes	None	123456	The unique user ID
ObjectID	Yes	None	987654	The NLA ID for the object (i.e. the article, picture, page etc) – mandatory unless Action=List
Action	Yes	None	Add/Delete	Can be Add or Delete

Example:

<http://www.nla-clips.com/nlaapi.dll/ManageArchive?UserID=123&ObjectID=456&Action=Add>

Adds article 456 to the MyArchive of user 123

Error Handling

Error messages are returned as plain text strings. The following error results are returned:

- **“Invalid User ID”** : no valid userid has been supplied
- **“No object specified”**: no objectID was supplied
- **“Invalid object specified”** : no valid objectID has been supplied
- **“User does not have archive rights”**: the user does not have my archive rights

List Archive

The 'list archive' API call will return a full list of all articles in a user's archive.

Supported Parameters

Parameter	Mandatory?	Default	Example	Notes
UserID	Yes	None	123456	The unique user ID
Format	None	XML	ASCII	Can be XML, ASCII or HTML

Examples

<http://www.nla-eclips.com/nlaapi.dll/ListArchive?UserID=123&Format=HTML>

Displays the MyArchive for user 123 in HTML format

Error Handling

Error messages are returned as plain text strings. The following error results are returned:

- **“Invalid User ID”** : no valid userid has been supplied
- **“User does not have archive rights available”**: this user does not have an archive right
- **“No articles found in the archive”**: this user's archive is empty

Search Archive

This call searches the entire article text (including headline, sub-heading and byline) for a match on a supplied text string. It returns a table of contents listing all the matching articles. Matching is done on a case-insensitive exact pattern match, eg a search for “Jack” would find a match on “hijack”.

Search terms must be two or more characters in length (eg 3i or O2 is supported, but X is not). A “Stop Word” list also exists which prevents certain words from being searchable. These include restricted words (AND, OR, NOT) which are used by the application and also a limited number of

common words (THE, OF, IN, HE etc). The NLA client services team is able to discuss these should they present an issue.

Supported Parameters

Parameter	Required	Default	Example	Notes
SearchTerm	Yes	None	Iraq	URL encoded search string. Includes support for parentheses and Boolean operators AND, OR and NOT
ObjectType	No	Article	Article	Object types that should be searched (provides future support for photos, adverts etc This initially will only support "Article")
Source	No	None	GDN	The 3 letter code for a specific newspaper (see Appendix for lists)
SourceSection	No	None		A URL encoded free text string referencing a specific newspaper section
SearchArea	No	Article		May be Article, Body, Headline, SubHeading, Byline
StartDate	No	7 Days ago	25122003	Format DDMMYYYY
EndDate	No	Tomorrow	01042004	Format DDMMYYYY
Edition	No	-1	0	-1 is the last edition, 0 is all editions, 1 is the 1st, 2 is the 2nd and so on
Format	No	XML	ASCII	Defines the return type as XML, HTML or ASCII
UserID	Yes	None	4590	The userID dictates whose archive to search

Note: if a search term features an ampersand (&), it must be enclosed in quotes, eg
[http://www.nla-eclips.com/nlaapi.dll/SearchArchive?SearchTerm=Tony Blair&SourceSection="Leaders & Letters"](http://www.nla-eclips.com/nlaapi.dll/SearchArchive?SearchTerm=Tony Blair&SourceSection=Leaders & Letters)

Examples

Example 1: [http://www.nla-eclips.com/NLA-API.dll/SearchArchive?SearchTerm=Tony Blair or Gordon BrownWar or \(Saddam Hussain\)&StartDate=10052005&userid=4590](http://www.nla-eclips.com/NLA-API.dll/SearchArchive?SearchTerm=Tony Blair or Gordon BrownWar or (Saddam Hussain)&StartDate=10052005&userid=4590)

Returns all stories in the archive published after 1-Feb-2004 featuring the term "Iraq War" (as an exact, case insensitive match) or the words "Saddam Hussain".

Example 2: <http://www.nla-eclips.com/NLA-API.dll/SearchArchive?SearchTerm=Microsoft&SearchArea=Headline&Source=TIM&format=ASCII&userid=4592>

Returns all stories in the archive from The Times featuring Microsoft in the headline, returning the content in ASCII format

Output

The results from this call include the following fields for each found article:

- ArticleID
- Headline
- Source (eg Guardian, The Sun etc)
- Newspaper Section (eg Lex, Letters Page, War in Iraq etc)
- Newspaper SubSource (i.e. the Supplement) if appropriate
- Date
- Byline
- Edition (eg 1st, Last Edition)
- Word count

XML Example:

```
<NLAML>
  <SourceMetaData>
    <Property FormalName="Publication_Name" Value="The Guardian" />
    <Property FormalName="Publication_SubSource" Value="G2" />
    <Property FormalName="Edition" Value="1" />
    <Property FormalName="Publication_Date" Value="20041225" />
    <Property FormalName="Day" Value="Monday" />
    <Property FormalName="Page_Section" Value="" />
  <Property FormalName="NLAArticleID" Value="450" />
</SourceMetadata>
  <DescriptiveMetadata>
    <HeadLine>
      <strong>Doubt cast on PM's 'nuclear threat' claim</strong>
    </HeadLine>
```

```
: <ByLine>
    <strong>Richard Norton-Taylor</strong> in London
</ByLine>
<Property FormalName="Wordcount" Value="635" />
</DescriptiveMetaData>
</ARTICLE>
</NLAML>
```

HTML Example:

```
<p>
<span class="toc_articleid">Article ID 1233: </span>
<a href="NLA-API.DLL/GetObject?userid=123&objectid=1233" class="toc_headline">Another
night to forget for troubled Benitez</a>
<br>
<span class="toc_source">Daily Express</span>,
<span class="toc_edition">1<sup>st</sup> edition</span>,
<span class="toc_section">Sports Section</span>,
<span class="toc_page">page 84</span>,
<span class="toc_date">>Thursday 20-Jan-2005</span>
<br>
<span class="toc_author">Paul Joyce</span>,
<span class="toc_words">1,412 words</span>
</p>
```

Which might be rendered in a browser as:

Article ID 1233: [Another night to forget for troubled Benitez](#)

Daily Express, 1st edition, *Sports Section*, page 84, Thursday 20-Jan-2005

Paul Joyce, *1,412 words*

ASCII Example:

ArticleID: 1233

Headline: Another night to forget for troubled Benitez

Source: Daily Express

Edition: 1st

Section: Sports Section

Page: 84

Date: Thursday 20-Jan-2005

Author: Paul Joyce

Wordcount: 1412

Error Handling

Error messages are returned as plain text strings. The following results are returned:

- **“No results found”**: no matching articles have been found
- **“No search key specified”**: no valid search term has been supplied
- **“Invalid (start or end) date specified. Format is DDMMYYYY”**: invalid start or end dates supplied
- **“Invalid source specified”**: an unrecognised 3 letter newspaper source has been entered

Note that there is no error checking on the searcharea, edition or format fields (valid values are defined above) nor on the subsource field (which may have many possible values)

Research

Here is some of the detail from research conducted in Oct 2004 in relation to barriers to entry for digital clippings, and again in May 2005 in relation to the central database.

Digital Cuttings (84 respondents)

The table below lists the factors considered important for not using digital copies. Whilst user preferences and internal culture were important reasons why organisations went digital, the barriers restricting growth of use of digital were all within the control of the NLA shareholders.

	Very Important	Very or Fairly Important
Only possible to hold articles for one week	62%	84%
Some publications not available	56%	71%
Costs of licence too prohibitive	47%	86%
Users happier with paper format	18%	83%
Better circulating summaries or using alternative means	13%	42%
Do not have infrastructure to support it	14%	31%
Does not fit with culture of organisation	4%	16%

When asked which of the above factors would need to change to make them more likely to provide digital copies, costs and the 7 day limit were considered most important:

- 19 mentioned cost;
- 10 mentioned the 7 day file storage limit;
- 3 mentioned inclusion of News International titles;
- 3 infrastructure;
- 2 'unavailable publications';
- 2 a change in internal culture;
- 7 said there was nothing they could think of; and
- 2 did not answer.

The Central Database (164 respondents)

25% of those with a digital licence were aware of the development of the new database.

When the database was explained to them and they were asked about their interest in using it as an electronic archive in which they could search all articles they had been sent by their PCA,

- 65% said they would be interested;
- 15% said they would not be interested; and
- 19% said maybe.

The highest proportion of those interested in this service would have:

- 44% expecting fewer than 5 users;
- 20% expecting 5-9 users;
- 16% expecting 10 to 25 users; and
- 11% 50 or more users.
- 6% were unsure or gave other answers.

Views on the fee structure were as follows:

- 47% thought fees should be higher, but based on the size of archive
- 33% thought fees should be the same for all organisations;
- 9% preferred higher fees for larger organisations;
- 2% were unsure; and
- 9% - made other comments.

The table below shows the proportion of organisations who said they would use the service at each price point for an annual single user licence:

- 56% said "yes" at £500 (7% maybe);
- 24% at £1,000 (13% maybe);
- 11% at £1,500 (2% maybe); and
- 2% at £2,000 (2% maybe).

Typically, they would expect to pay £50 to £200 per additional user (11% less than £50, 49% £50-£199, 27% £200+). 40% would have less than 5 users, 7% 5-9, 29% 10+ (24% not sure, other response or not answered).