



# **Managing Copyright Infringement**

February 2019

## Contents

**3- Introduction to OATS**

**4- OATS in action**

**7- How does it work?**

**8- Successes and case studies**

**11- FAQs**

**12- About NLA**



## OATS – Introduction

A large volume of publisher's web content is currently being republished online without publisher permission or credit. The NLA's Online Article Tracking Service (OATS) identifies and manages removal of content from freely accessible infringing websites on the publisher's behalf, saving time and money. OATS can analyse every article that's produced by an online newspaper and then find cases of where these articles have been copied by a particular domain. To do this manually would take days of work. As well as this, using OATS also makes reporting and removing pirated content a faster process.

OATS tracks content republished on commercial and non-commercial sites, including blogs and forums, who in many instances, may not be aware that they are infringing copyright.

The OATS service aims to educate infringing domain owners on what can or cannot be copied and offer alternatives, such as direct publisher permission. By the end of 2018, NLA media access has contacted over 3,400 domains and had an 89% success rate of removal of infringed content for subscribing publishers.

### **What is copied?**

The scale and variety of abuse can make categorisation difficult but there are clear hotspots.

As one would expect, news content is the largest category as there are a high volume of news aggregators / businesses that republish articles. Sport-related content features highly especially for news, fixtures, league tables etc. A high volume of sport content is also used by blog/forum sites and in particular football and other sports clubs. Entertainment/showbiz and celebrity content are also popular topics for infringement.

Specialist B2B magazine, travel and tourism, and academic content was also seen to be infringed although at a lower volume - which is consistent with the nature of the specialist content being infringed. It was also noted that there were also peaks and troughs of infringement when specific events occurred that created a large amount of interest, for example, the football world cup saw a spike in the volume of football content that was infringed.

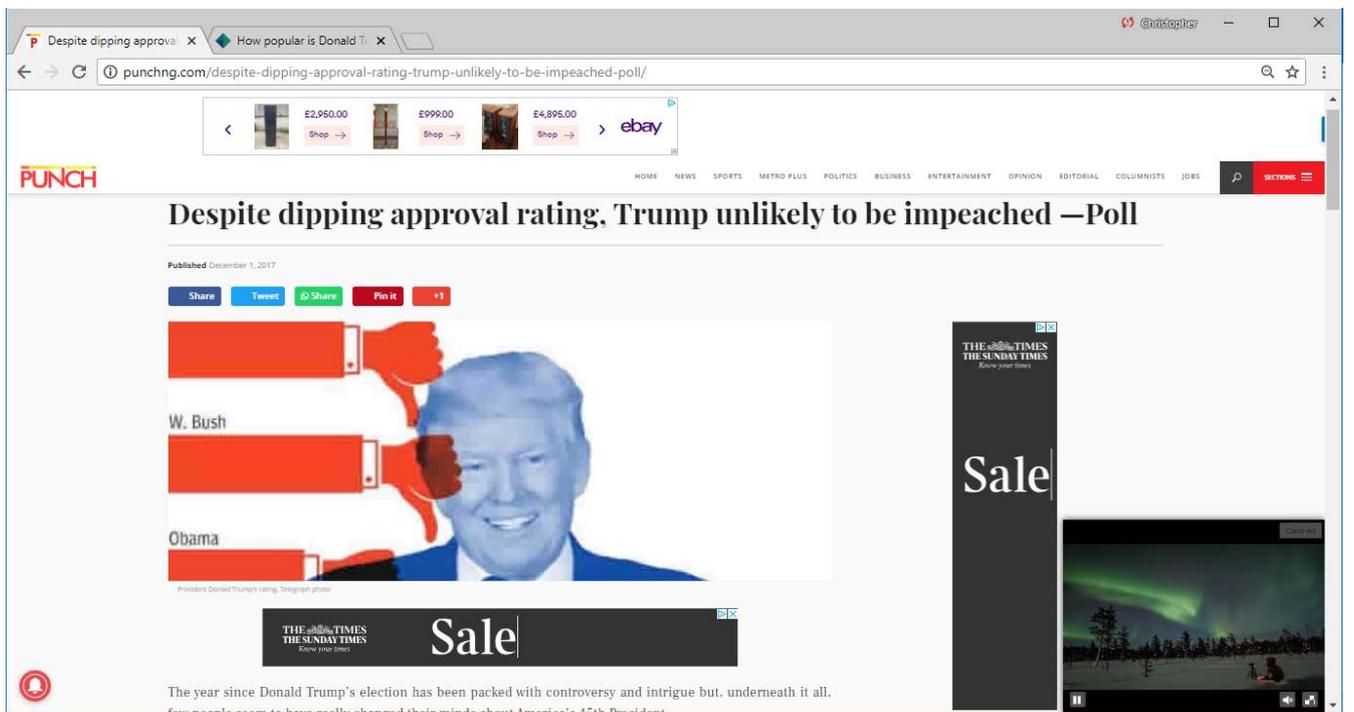
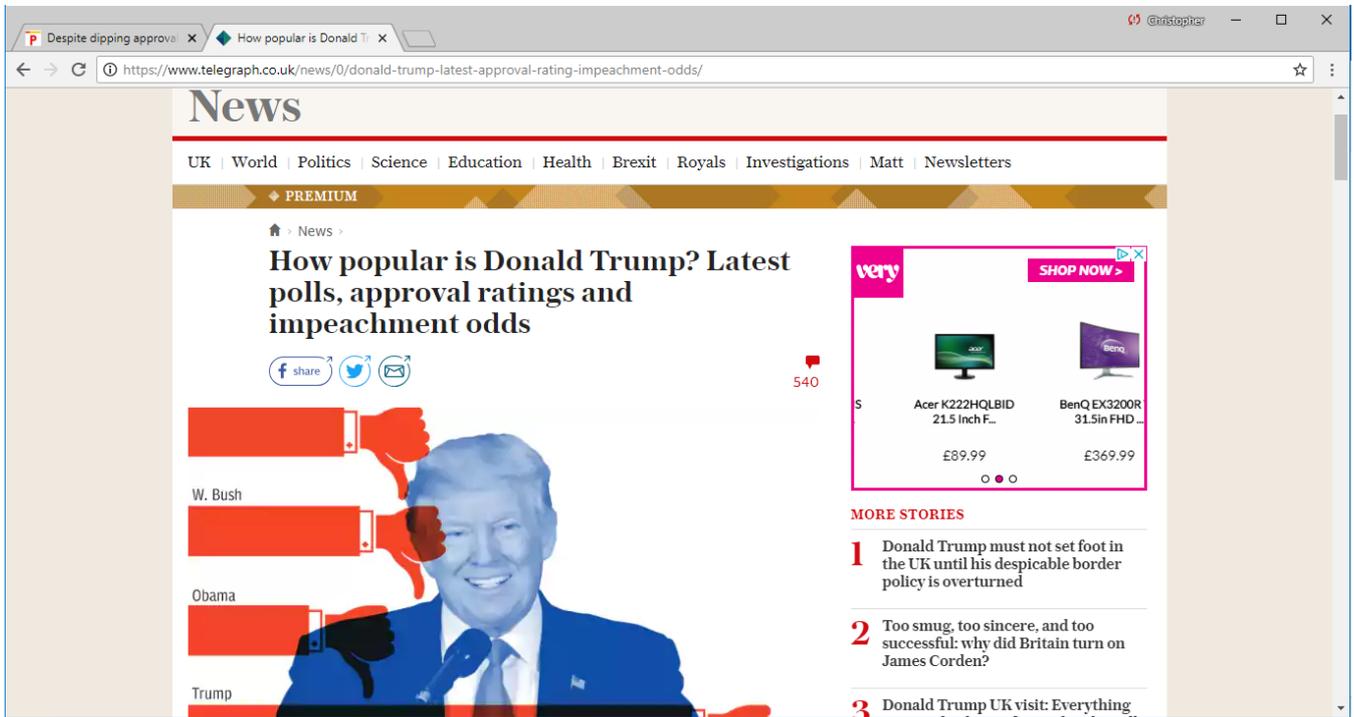
### **What type of publisher content can OATS detect and remove?**

OATS detects and deals with a variety of copyright abuses which range from parts-of, to full articles, as well as photographs or captions that have been copied and posted without permission. The most common types of copyright abuse are:

- **Copying / lifting text from an online article and adding a credit to the name of the original publisher.** This usually occurs because there is an ignorance around copyright law and terms and conditions. In our experience, once you notify the infringing domain, they remove the content.
- **Copying / lifting parts of (or all of) the text from an online article and not adding credit / linking to the original publisher.** These sites tend to design the layout to look as though it's their own original work by their own reporters.
- **Occasionally, websites that replicate a major news site in terms of design and URL.** Sometimes the content is kept the same however content can be changed, often altering it slightly which gives the article a new tone or message. These instances of slightly-altered content contribute to the issue of fake news.

## OATS in action

Below is a typical example of a website, which has copied an original publisher article, including the images. As you can see, the website in question is also hosting advertising on this website in an attempt to monetise this infringement.

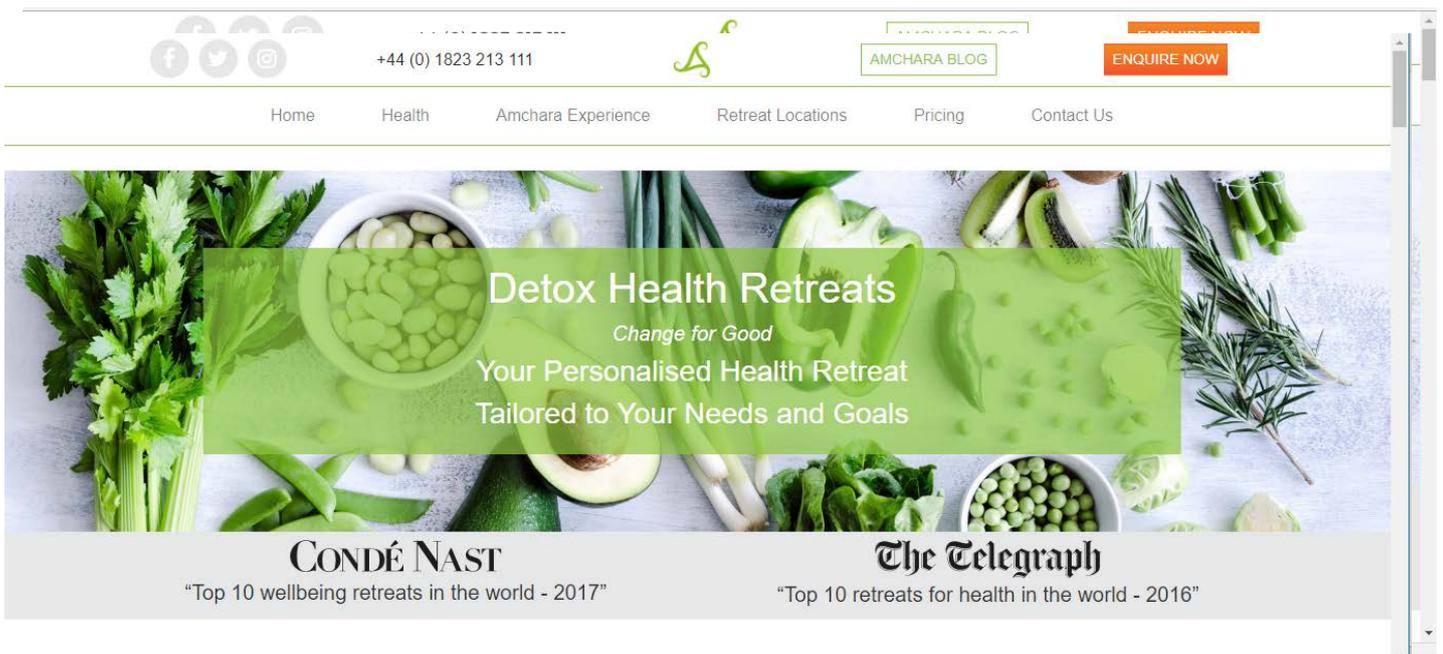


Below is an example of a fake news website, cloning the Mirror.co.uk webpage. The original Mirror logo and banner is used at the top of this article. If it wasn't for the different URL, it could easily be confused for the original Mirror.co.uk website.

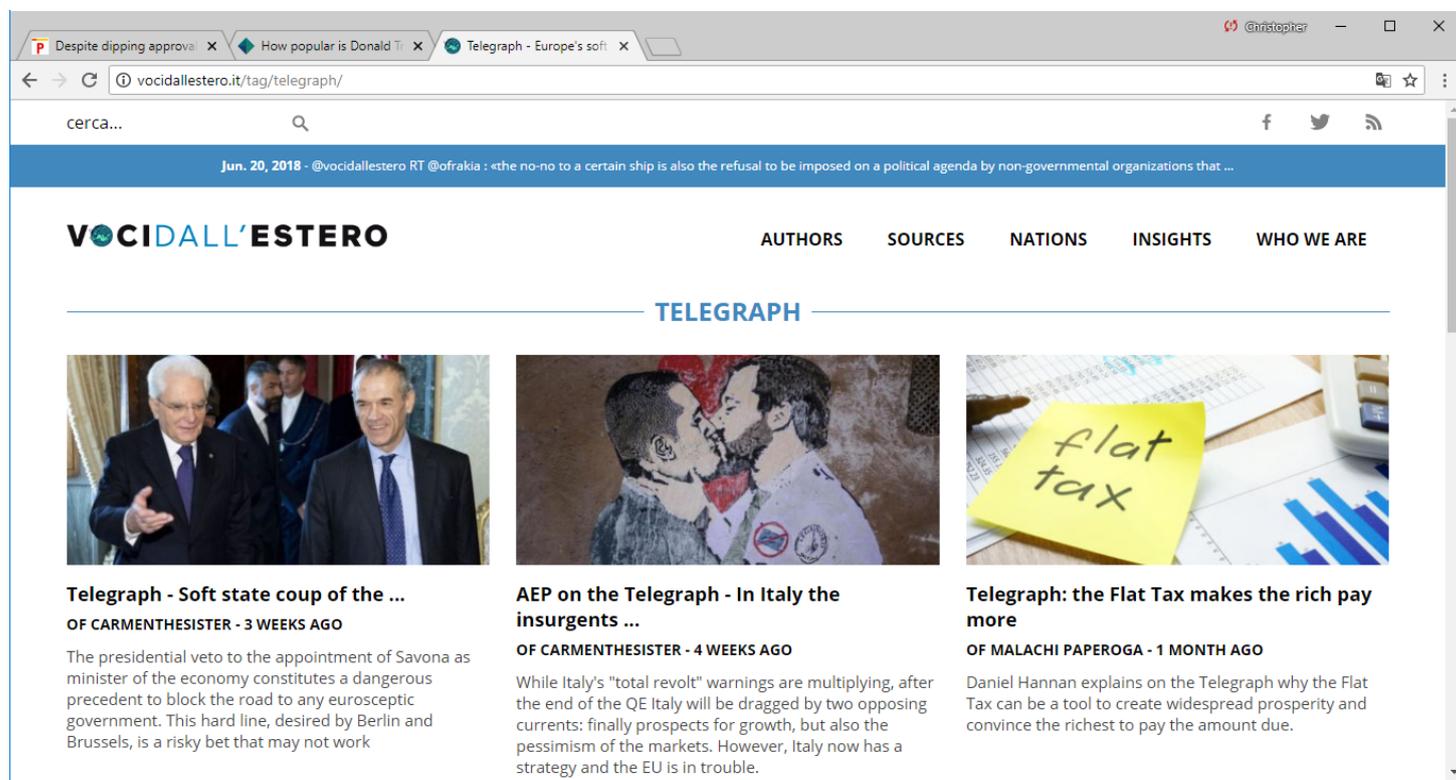


This below example demonstrates how the OATS team were able to remove Logo infringement. In this case the Daily Telegraph logo had been used without permission by this particular website. Many publishers are rightly protective of their intellectual property (in this case a logo) being used to endorse a website without permission.

The original page before OATS intervened: The amended page, after OATS intervention:



In this below example, a website had dedicated an entire section to posting copyrighted publisher articles. This entire section was removed by the OATS team.



## How Does OATS Work?

The OATS system is able to utilise NLA's unique article database combined with our bespoke web crawling software to search the internet for websites that have copied publisher articles.

OATS is currently set to a copy threshold of 75% - which means that it's highly accurate in finding content where over 75% of an article has been copied. This threshold can be changed.

OATS can detect an article infringement within 24 hours of an article being published and search for several months' worth of data at a time.

Once the OATS system has identified copyrighted material, NLA will work to directly target the domain who are hosting the copyrighted content directly. NLA then inform them that they are in breach of copyright and ask them to remove the content. This process reduces the chance of the same person and/or domain repeating the offence. However, if after multiple attempts NLA are unable to contact the owner of the domain, NLA contact the domain hosting provider and issue a Digital Millennium Copyright Act (DMCA) notice to remove the content.

## The OATS process

A simplified workflow below includes constant publisher monitoring.

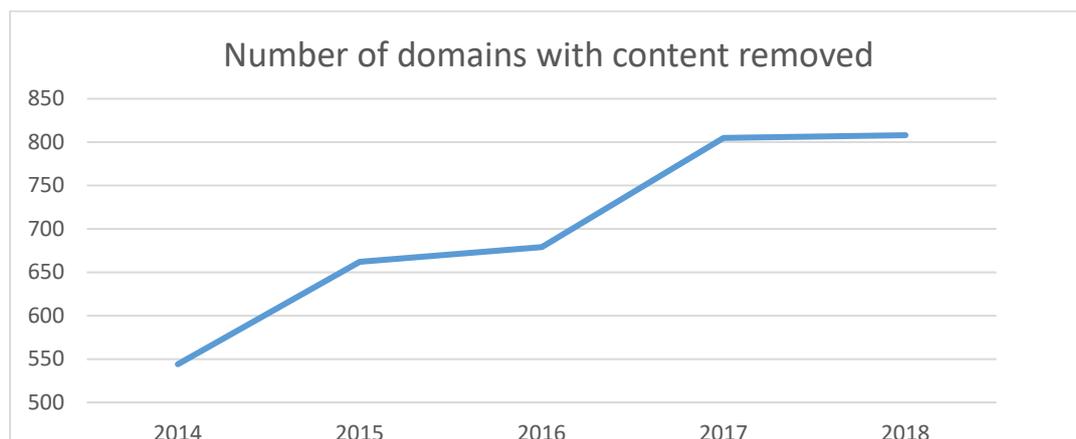
<b>1. Base data</b>	<ul style="list-style-type: none"> <li>•Content from participating publishers is collected daily into a common pool</li> </ul>
<b>2. Scanning</b>	<ul style="list-style-type: none"> <li>•Data is matched against all web data using search engines, resulting in suspect sites being ranked and sorted by the extent and frequency of identified copying</li> </ul>
<b>3. White list</b>	<ul style="list-style-type: none"> <li>•Legitimate sites are removed from suspect lists</li> </ul>
<b>4. Contact 1</b>	<ul style="list-style-type: none"> <li>•Initial educational enquiry email is sent, stressing fair linking rules and requesting proof of permission to re-use, and offering syndication contacts</li> </ul>
<b>5. Contact 2</b>	<ul style="list-style-type: none"> <li>•(If use continues without permission) take down request sent</li> </ul>
<b>6. Contact 3</b>	<ul style="list-style-type: none"> <li>•(If use continues without permission) cease and desist</li> </ul>
<b>7. Contact 4</b>	<ul style="list-style-type: none"> <li>•(If no response and continued use) via Registrant or Host Provider, submitting DMCA notices where required</li> </ul>
<b>8. Referral to publisher</b>	<ul style="list-style-type: none"> <li>•For legal action or further process.</li> </ul>

## Successes

In **2018** OATS removed over **89,400 articles** from **802** sources with a **90%** success rate of removing content.

Since its inception in 2014 OATS has removed over **400,000** articles from over from over **3,400** domains.

OATS has steadily been building its success rate since it started in 2014. A lot of OATS success is the result of our takedown processes and growing understanding of copyright infringement.



In **2014** our article takedown success rate was **73%**. This is the percentage of articles that we were able to remove after we had started our takedown process.

In **2016**, this number grew to **79%**. There are always a small percentage of articles, which can't be removed, due to factors such as foreign law and host providers that do not recognise our copyright laws. However, as we continue to learn and improve our processes we hope to close this gap even further.

## **Case Studies**

*(Names removed as infringement has now ceased)*

### **Case study 1**

The OATS database highlighted that a news aggregation website was routinely republishing articles from many national newspapers that use the OATS service. This site was taking thousands of articles a month and reposting the full text without permission. In addition to this, the website was attempting to make profit from advertising revenue by driving traffic away from the original publisher/copyright holder to their own website.

Using our OATS takedown process, NLA contacted the website in question, requesting that all of the copyrighted publisher content is removed ASAP. This was followed up by 2<sup>nd</sup> and 3<sup>rd</sup> emails noting that the website was in breach of our publisher's copyright and asking for immediate removal of the content. Usually in this scenario, a website will remove the content in question and the case is resolved. However, our requests were ignored.

The OATS team then contacted the websites host provider to notify them of the breach of publisher copyright and asked them to remove the content in question. After back and forth communication with the host provider, highlighting the scale of the copyright abuse - the entire website was removed. This website was breaching many international publisher's copyright by republishing news content without a licence.

### **Case study 2**

A national newspaper publisher discovered that a website was taking scans of physical newspapers from many National, regional and international sources around the world. This website had built up a huge database of thousands of newspaper PDF's, going back nearly 2 years.

The OATS team contacted both the website directly and their host provider to ask for this website to remove the copied content. The website ignored our requests and the host provider, which existed outside of EU/US copyright laws refused to act.

The NLA OATS team then called a meeting of all OATS publishers and we orchestrated a group approach to getting the content removed. Our aim was to highlight the gravity of this gross abuse of publisher copyright and note how the infringement affects a large group of major publishers. All the OATS publishers agreed to individually send this notice directly. NLA then shared this information with a few key international publishers who were also affected by the website as asked them to send their own direct notices.

As a result of this joint publisher act, the host provider & website both received a large volume of individual copyright complaints, from multiple national and international publishers. A few weeks later the website and all of its data was removed.

### **Case study 3**

A US based news aggregation site was found to be posting a large volume of articles from a UK newspaper publisher. Usually, news aggregation websites will republish other publisher's articles, credit the original publisher and assume that it's acceptable practice to do so. The OATS team spend a lot of time removing this content and informing news aggregators that this is a breach of publisher copyright, unless they have a specific licence or permission to do so.

However, not only was this news website using publisher's content, but they were crediting the article with their own journalist as the original author.

The OATS team contacted this website and supplied a full list of the URLs's that had been copied. The website in question, responded, with a full apology and removed the content right away. It turned out that that one of their staff journalists had been routinely plagiarising publisher content and claiming it as their own work. Undetected until the OATS system matched the text with the original articles.

## **Our publishers agree:**

**Sophie Hanbury** – *Content Partnership Director, Telegraph Media Group*

“Working with OATS reiterates The Telegraph's commitment to protecting the copyright of our valuable news content whilst reinforcing a consistent strategic approach to our metered digital products and paid services. OATS has been highly effective in the drive to police large-scale systematic infringements and works in tandem with our Legal and Syndication teams to reassure our commercial partners that the contracted products and services they pay for are valuable, of a high quality and worth protecting. Endorsed by The Telegraph; NLA Media Access is actively educating the market place on the necessity to seek a legitimate license for content; it also clearly demonstrates that the terms and conditions we display on our website actually mean something. Policing unauthorised usage through OATS positively supports our journalists, contributors, advertisers and commercial partners.”

**Mike Pocock**– *Director of Content Partnerships, at ESI Media*

“NLA discovered and removed over 10,000 individual articles that had been lifted from The Independent and Evening Standard websites across 2018 and republished on over 100 infringing websites without our permission. OATS has been a great resource for identifying these illegal usages and provides a solution that allows multiple publishers to efficiently and proactively address such copyright infringement without the extensive resource that would be necessary to cater for that individually.”

## **FAQs**

### **How often does OATS check for a copy of each article?**

OATS carries out a single check for infringing content, 24 hours after the article is added to OATS

### **What format does OATS search for?**

OATS searches for HTML content only. It cannot search across PDF's or Word documents.

### **What is regarded as an infringement in OATS?**

OATS only picks up infringements that have over 75% copied text. The reason we set this threshold is because OATS provides a huge volume of content. There are thousands of websites sites that might take 50% of an article, however, we prioritise the big offenders who taking larger volumes of articles where over 75% has been copied. These websites tend to be the ones where large volumes of articles have been routinely scraped/copied.

### **Can the OATS team reactively help us remove content from a website that's been reported to us externally?**

We regularly get requests from clients to help them remove content that's been reported externally (often by readers or journalists). Often, articles that have received a libellous or legal complaint get copied and the OATS team will help remove them from 3rd party websites.

### **How long does it take to remove content?**

Usually an article(s) takes no more than 3 weeks to be removed from an infringing site. Sometimes it can be a lot faster, sometimes it can take much longer. It really varies depending on the website in question. For example, websites that exist outside of the EU or US, that don't immediately respect UK copyright can be unresponsive. Some host providers (such as Google) can take longer to remove content as they receive such a high volume of complaints. Websites are generally keen to be compliant once NLA draw their attention to copyright rules.

### **Can you remove content that another NLA publisher has taken?**

We can't get involved, due to conflict of interest if another newspaper/magazine publisher has taken content from another publisher. Most of the content we remove is hosted by news websites who take our publishers content. If an NLA publisher owns the site, we wouldn't be able to pursue it.

### **Can I use OATS to find opportunities to licence our content to 3<sup>rd</sup> parties?**

When OATS first started back in 2015, we believed that it would be a great way to source licensing opportunities for our publishers. We assumed that many 3rd parties who take our publisher content would be willing to pay for a licence and continue using it, once we contacted them. The idea changed when we discovered that most websites would rather remove content and stop sourcing publisher content altogether than have to pay for it. This is not true of all websites – there are many exceptions, and we do get requests for a licensing quote (which we always pass to the original publisher) however, most of the time, infringing websites don't want to pay for a licence.

## **About NLA**

NLA Media Access (“NLA”) is a collective licensing and publisher services business. Its mission is to support journalism. Under the mandate of several hundred newspapers, magazines and digital-only brands – NLA licenses global media monitoring organisations (e.g. Cision, Kantar, Meltwater) and their clients (UK businesses) to use publishers’ copyrighted content. NLA collect royalty payments through this licensing activity and return royalties to copyright owners (publishers).

As well as consistently growing royalties for publishers through our collective licensing business, NLA have developed a portfolio of cost saving and revenue enhancing publisher services – OATS included.

For more information, please contact [publisherservices@nla.co.uk](mailto:publisherservices@nla.co.uk)