Key issue

Automated Content Access Protocol (ACAP)

MARK BIDE
Project Co-ordinator
ACAP

An issue which has been quietly on the agenda for well over a decade has risen to prominence over the last couple of years: improving the communication of access and use permissions.

A number of different projects have arisen, each dealing with a different facet of this challenge. Creative Commons (http://creativecommons.org/) will be familiar to all readers of serials; the work that EDItEUR (http://www.editeur.org/) has been undertaking in the development of ONIX for Publication Licences (ONIX-PL) probably only marginally less so. The work of the PLUS Coalition (http://www.useplus.com/) in standards for the licensing of images is also attracting growing attention.

ACAP, a 12-month pilot project which launched in January of this year, is something of the new kid on the block, but is nevertheless raising some significant waves. This is perhaps a reflection of the considerable and very diverse support that it has managed to bring together in a relatively short period. Not only does ACAP bring together the enormous diversity of what we normally think of as 'the publishing industry' – journals, books, magazines, newspapers, data – other organizations as diverse as the International Federation of the Phonographic Industries, the Motion Picture Association, and the UK's Office for Public Sector Information have also become members.

What has united these eclectic interests? The challenge the publishers face in communicating permissions to online aggregators – starting with the search engines.

If there seems to be a strong initial focus on search engines in ACAP, it is no coincidence. They are currently the focus of much attention for publishers. All sectors of publishing have a 'search engine dilemma'. The search engines are an unavoidable and valued port of call for anyone seeking an audience on the Internet. Search engines sit between Internet users and the content they are seeking out and have found brilliantly simple and effective ways to make money from the audience that they attract. As a result, they have become so dominant that even the largest website owners are not large enough to have any serious impact on their commercial fortunes.

The benefits of powerful search technology to both users and providers of content are well recognized by publishers. Few, if any, publishers want simply to block them or to keep their content off the web. At the same time, publishers are aware that search engines are, in following their business logic, inevitably and gradually moving into a publisher-like role, initially merely pointing, then caching and, finally, aggregating and 'publishing' and perhaps even creating content themselves, while using other people's content as the basis for many of their services.

With the current technology, there can be none of the differentiation of terms of access and use that characterizes content-trading relationships in publishing environments, whether electronic or physical. The search engines can and do reasonably argue that, since their systems are completely automated, and they cannot possibly enter into

and manage individual and different agreements with every website they encounter, there is no practical alternative to their current modus operandi.

But publishers do not always agree with the way in which the search engines use their content, and would like to be able to communicate rather more sophisticated permissions than are allowed by the 'Robots Exclusion Protocol' (REP), a positively ancient specification for communicating a relatively small menu of permissions to search engines (ancient in Internet terms: it was released as a draft in 1994 and never finalized or revised). See http://www.robotstxt.org/wc/exclusion.html#robotstxt .

REP is very widely implemented; however, there are a number of problems. Not only is it nothing like sufficiently expressive for today's very much more sophisticated and diverse world wide web publishing and search models – both entirely unimaginable in 1994 – it is also very inconsistently applied by the search engines.

Under the leadership of the World Association of Newspapers (http://www.wan-press.org/), the International Publishers Association (http://www.internationalpublishers.org/) and the European Publishers Council (http://www.epceurope.org/), the ACAP project has been established with the remit of seeking a solution to the challenge. A group of major publishers from Europe, the US and South Africa (including Holtzbrinck Macmillan, Reed Elsevier and John Wiley & Sons) are working with representatives of the major search engines and other intermediaries to find better ways of communicating permissions online.

Among the technical partners of the pilot project is the British Library, which is actively seeking a way of providing publishers with the capability of communicating differential permissions for web harvesting – an application which points beyond the search engines towards the other areas of application which ACAP sees as falling within its scope.

ACAP is committed to the development of an open standard, as useable by the smallest self-publisher as by the largest media conglomerate. It is also committed to avoiding reinventing wheels – and to this end, has sub-contracted some of its technical work to EDItEUR to ensure inter-operability with ONIX-PL.

For more information about the project, visit the website (http://www.the-acap.org/), or contact the Project Co-ordinator Mark Bide (mark.bide@rightscom.com).

Article © Mark Bide

To view the original copy of this key issue, published in *Serials*, the journal of the UKSG, click here:

   http://serials.uksg.org/openurl.asp?genre=article&issn=0953-0460&volume=20&issue=2&spage=148

For a link to the table of contents for the issue of *Serials* in which this article first appeared, click here:

   http://serials.uksg.org/openurl.asp?genre=issue&issn=0953-0460&volume=20&issue=2