Insert STM Newsletter 95, November 1994

Papers available of presentations made at the Informal Part of the 26th STM
General Assembly, held on 4 October 1994 in Frankfurt

by
Mr. Douglas Armati, UDID
Mr. Bernard Naylor, University Librarian, University of Southampton
Mr. Richard Patterson, Cascade Systems Ltd

with an introduction by Mr. Stephen White, Sweet & Maxwell, who chaired the
informal part of the meeting.

Introduction by Stephen White

The intention of anyone organising this session must be to offer knowledge, interest and that frisson of excitement which comes from being told about a development fundamentally damaging to the continued well being of the publishing industry as we know it. In the last few years we have had knowledge galore, interest and enough fear and anxiety to make us all retire to a bar somewhere warm.

Well, I've never seen it is anything but my duty to follow where the great and the good lead and I am delighted to be able to present three speakers who will enlighten, interest and frighten you during the next hour and a half.

The topic is On-line text: delivery, cost and usability. Whatever is technically feasible, the key question for us in this room is how to make money today, if we're lucky, and certainly tomorrow. To make money we have to be at least as good as the best, appeal to customers and charge them. On all three counts I suspect mos of us fail and are floundering.

Douglas Armati knows that all the concentration, thought and angst are pointless unless we can be paid.

Bernard Naylor is an eminent librarian who has long sounded alarm bells for academic publishers about the value for money element of paper journals and will today focus our attention on whatthe modern university wants.

Richard Patterson will talk about a cousin industry, newspapers, which has taken more steps to develop on-line text delivery than the companies represented in this room because, rightly or wrongly, it sees a greater to its traditional product base.

A Uniform Approach to Identification of
Digitized Copyright Content(?)
by Douglas Armati

The problem: How to identify with certainty the origin and ownership of
copies of valuable data in the perfect reproduction environment afforded by
digital technologies? How to control copying?
How to prosper commercially, keep shareholders and customers happy, as we
move from the relative security of the existing physical production and
distribution system to the apparent digital anarchy of the Net, Inter or
otherwise?
These pressing questions demand an answer. When Stephen White invited me to
speak with you today he described it as the search for the holy grail.
Well, I have some good news. The end of the search may be in sight.
There is still some work to do. Commercial reality suggests you may be just the
group to do it!
The holy grail could turn out to be remarkably simple.

A solution: It takes the form of a technique for indelibly, yet unobtrusively,
encoding all forms of digitized data with a unique structured identifier.
Using this type of approach it becomes feasible to create a uniform system for
identifying all valuable data, including rights based information assets.

A Universal Data Identification (UDID) system.
Why is this a good idea?
This session is about the practicalities of making money from copyright based
assets in the digital world.

The Rationale: One simple argument underpins the rationale for using a common
approach to the identification of digitized copyright works, no matter what the
eventual shape of the digital delivery system. A uniform, effective, well-used
digital data ID system will add value to the portfolios of all participants. It will
do this by facilitating open trade and enabling automatic global management of
rights based assets.
There are some good reasons for adopting this approach.

Changing threats in opportunities:
1. Electrons do not respect borders.
A digitized work is vulnerable to quick, virtually perfect copying. It is also
subject to transmission locally, nationally or internationally over broadband
networks, without trace. In the absence of reliable identification of copies and
fragments of copies, this is a serious threat to portfolio value.
In the presence of easily "read" unique copy and fragment identifiers, coupled
with automatic transaction management systems, those attributes change from
threats into opportunities.
Removing the threat posed by mobile data encourages trade, opens up physical
and electronic barriers and should add value.

It becomes realistic for rights owners to encourage copying and cross border
retransmissions, secure in the knowledge that they will be paid.
An added benefit from using network resident rights management is the
principle of national treatment can be upheld no matter where the reproduction
takes place.

The present commercial answer to the digital challenge is to offer access to valuable data only within the framework of tight proprietary "closed circuit" systems.

Simplify access to core value This works only moderately well for the seller. For the users it is decidedly uncomfortable. They need to navigate through many different, often quite complex, hardware and software boundaries before getting to the core value. The more suppliers used, the more complexities experienced.

Taking a uniform approach to identification of all copyright works is one important way in which to facilitate the emergence of open systems based on automatic network-resident rights management.

Wherever possible the goal should be to simplify the transaction between supplier and user, by using rather than inhibiting the ready mobility of electrons.

Common medium: common identifiers
2. **Electrons do not know the difference between a book and a film.** In a digital environment, there is a common medium carrying copyright works withich differ in content but not in form.
Convergent communications and computing technologies are similarly blurring the content boundaries between different types of copyright works.

Historically, different industrial groupings have grown up around the special tools required to capture creative works, the different media used to store those works and the different equipment used to communicate the stored contents to end users. Film and paper based industries, for example, have quite different needs.
In that context it was sensible to use identification systems that reflected the industrial structure -- each specialised industrial cluster developing its own international codes.

New industrial structure: However, when digital tools are used to capture the whole range of creative works, digital media are used to store those works and digital equipment used to communicate the stored contents, then the differences between media vanish -- there is just one medium. What were once thought of as images on a film or pages in a book are now created and expressed in a common electronic language and communicated through the same networks and devices.

In that convergent industrial context it makes sense to use just one identification system, especially when more and more works are created in digital form and expressed using digital devices.

An ongoing role for the existing systems: This does not mean discarding old systems. Rather these existing systems (such as ISBN/ISSN, ISMN and UPC/EAN) will continue to identify books, periodicals, cassettes, sheet music videos, CD's, videodisks and so on in their physically stored form. The current estimate is that only 1% of the world's existing knowledge base is in digital form. There will be plenty of scope for using these physical coding systems when coping with the necessary links between digitized and physical objects. No doubt, too, the existing code management organisations will have a great

4

deal to contribute to the cross industry norm setting processes and subsequent administration of the standards.

The new rules: Any successful uniform system must, however, now be able to identify all forms of digitized data, no matter the chosen format or distribution platform.

3. Solutions are available

The proprietary present: At present only fairly crude proprietary systems are available to facilitate identification of a limited range of propietary content within very tight "closed circuit" parameters.
This is the antithesis of what is promised by the technology of open, fully distributed broadband communications.

Tools for an open future: Earlier this year, in an effort to provide a catalyst to unlock the full power of the technology, to open up the huge value potential for rights owners, I defined a functional specification for a Universal Data Identification (UDID) system.

The UDID experiment: I took this specification to Professor Laurie Turner, head of the Digital Communications section at the Imperial College in London. Making modifications to earlier inventions of his, now in the public domain, his team experimented with identifying various forms of digitized copyright works.

To meet the specification I asked him to explore such thorny issues as the effects of digital to analogue conversion and reconversion to digital, whether the identification would survive compression, encryption, signal distortions and so on.

No code / No value One of the most significant features of the resulting system is that if the UDID code is altered or removed then the data structure is corrupted, rendering the data economically valueless.

As an example, in the tests done with musical material, if the UDID code is removed, the CD quality digital recording suddenly painfully resembles the sound of a very scratched, totally unacceptable old 78. Yet, even in the crispest silences the coding is inaudible even to the most finely tuned ears.

The Network Connection: Although the physical work on the UDID system is still to be completed and some technical questions remain, the design was well enough advanced to introduce it to BT. This is the next crucial step. For a UDID system to provide the foundation for automatic network-resident rights management systems, the network designers need to be involved.

I understand discussions between the parties are proceeding.
Technically, Professor Turner believes the problems of permanently identifying and controlling the uses of digitized copyright works (and all other digitized information) in open communications networks can be solved.

Network-resident rights management

Design of network-resident individual transaction rights management systems then becomes possible and economically feasible.

The UDID specification provides a foundation on which to build a comprehensive technological response to the dematerialisation of copyright works and other valuable information.

I want to stress I have no commercial interest in the Imperial College/BT project. The goal has been to prove a practical, uniform coding system is technically feasible.

Proceed with caution

Indeed, I would counsel you not to commit to using any system until it can be clearly established there is an economic case for doing so and all possible solutions have been thoroughly reviewed.

4. To realise its full potential economically, the chosen international copyright work identification code will require extensive technological support.

The will be considerable additional investment required in digital communications infrastructure and network management systems, software support systems, computer and reprographic hardware systems and in increasingly automated international rights management system.

History indicates these investments are most successful when preceded by standardisation.

5. Standard identification systems evolve to assist in driving transaction costs lower.

The standardisation and coordination of international telephone line numbering, for example, enabled the provision of International Direct Dialing. This easier access made international trade simpler to negotiate and transact.

Growing volumes of business based on internationally standarized network protocols, coupled with growing competition, have fueled a virtuous circle of cost and price reductions. Increased facility for the user has led to dramatic reductions in transaction costs that have brought us to the dawn of the information century.

The Uniform Product Code

The Uniform Product Code and European Article Number (bar code) systems, developed in the early 1970s reflected a transition phase from largely human management systems to an increasing reliance on process automation. Its global implementation was founded on development of machines capable of "reading" and processing standardised information.

Its evolution required the creation of a uniform coding structure, new printing techniques, standards for bar-code scanners and bodies to administer standards. It succeeded in reducing industry costs by between 1.5% and 4%.

A revolution in logistics management The successful internationalisation of this technology has revolutionised the production, warehousing and distribution of physical products (including most mass market copyright content packages).

Digital Distribution logistics A similar change is underway. The approach to a uniform code must reflect the technological era into which we are moving. The digital exchange of economically valuable copyright content is becoming more prevalent. As this trend accelerates, automatic, on line recognition and processing of these standard codes will become essential. The communication of copyright works is one of the foundations of the shift into an increasingly "virtual" global economic era.

Making the new distribution structure work. In our global communications networks we are evolving a large, complex technological structure. This structure will be more than capable of coping with the management of transactions involving copyright based economic value, but only if there are core identifiers to facilitate this. With these in place, digital technologies can then be employed from conception to collection.

6. A uniform coding protocol and its basic supporting technologies are part and parcel of the same system.
The practical approach is to harmonise their design, testing and implementation.

STM: Taking the lead?
7. How? Someone has to take the lead.
In the case of Uniform Product Code standards, it was the US grocery industry, seeing the chance to lower transaction costs, which led the, way.

Because it was clear the industry would buy the equipment needed to implement the system, it received tremendous cooperation from hardware and software suppliers in developing the core technologies.

Reasons for action In the case of the rights based industries, you were the first group exposed to mass reprography. Now you are the first exposed to the growing threat of uncontrolled network mediated exchange and reproduction of your key assets - valuable information.
It seems logical then, to expect you to take the lead in order to protect your asset portfolios.

Some side benefits By sponsoring the development of a uniform approach to the coding of valuable copyright works you will do a great deal to protect the value of your own assets. Significantly, you will simultaneously establish the path for others to follow as they too face the challenges of digital distribution and open exchange. This knowledge may offer commercial opportunities in and of itself. Recently I discussed this problem with Tom Wilson, who for 20 years was the McKinsey & Co consultant who guided the development of the UPC system. Drawing on his insights, if I was advising you on how to do this, I would recommend the following implementation programme:

Recommended programme
First of all, form an ad hoc committee composed exclusively of Chief Executive Officers.
This committee should be drawn from those with the willingness and ability to commit themselves and their corporations to the eventual group decision. It should be small, no more than a dozen people. It should be led by someone with universal respect and influence. Together, the group should control companies that collectively have sufficient market share to influence the entire rights based

industry.

From the earliest days of the project it is vital to have a professional support team willing to perform beyond the call of duty. They will be needed to steer the project throughout the process.

A decision should be taken to keep governments at arms' length, informed but not involved. The process should proceed on a strictly commercial basis.

Next, I would suggest commissioning an economic analysis of the costs and benefits of uniform identification.
Once the economic case is established, the main committee should set the broad goals of the system and then form a technical sub-committee to oversee the process of achieving them.

In deciding the best solution, this committee should establish a technical quest specification covering a structure for the code and the necessary encoding, decoding and monitoring wares. It should then request submissions globally from technically competent organisations. All submissions should be on the basis relevant intellectual property be put into the public domain.

Next, the sub-committee should receive and review the submissions before making its report and recommendations to the main committee. The main group would subsequently review this information, making its own decision and announcement of the chosen solution.

Rapid deployment Immediate implementation should follow.

Being in the public domain, the key technologies on which the system is built will be quickly internationalised. The keys to the kingdom should be available to all at low cost. The code structure should be usable by anyone who chooses to participate in the system.

There will be plenty of experience to draw on from within the various industry sectors and from cross industry bodies such as the EAN. There is positive interest from these quarters in assisting with the numbering structure and advice on the administration of global standards.

Formal standards later In time it will no doubt become desirable to proceed to formalise the standards at the SEN and the ISO, but that process will long follow the commercial implementation.

Focus on strategic importance It is important not to lose sight of the strategic importance of the project. It is not just an exercise in standards, but rather the core technology of a radically new way of doing business.
Time is of the essence. This is a 30 to 36 month project. The sooner it is begun the sooner you can breathe easily about the long term future of your industry.

Beware VAN dependence In the interim I would counsel caution on being an early adopter of dazzling proprietary solutions. There will be impressive Value Added Network (VAN) distribution solutions on offer from the new network alliances. Locking into these may seriously limit your later options. It may even deliver effective control of the exploitation of your assets into their hands.

<u>Strength in numbers</u> Choose rather to join with your natural allies, rights owners with a common problem to develop a distribution structure that suits your mutual needs. This will put you in a position to negotiate the terms with all those controllers of digital distribution networks who so desperately need your valuable traffic to justify their colossal investments.

They need you much more than you need them.

<u>The new game is a monomedium</u> In the quest for a universal data identification system, form alliances of common interest as broadly as possible. Cross the old industrial boundaries. The new game is a monomedium. It requires a different strategic vision at its centre, a new set of rules to allow publishers and others to do what they have always done best -- finding and delivering value to real customers.

As John Dill said in his inaugural speech as Chairman of your Association : Take inspiration from your founders and pursue "unity well in advance of your most challenging need."

On-Line Text: Delivery, Usability and Cost
An academic library viewpoint
by Bernard Naylor

Introduction
The scholarly journal has been an important feature in the communication of research for more than three hundred years. Its origins are traced to the Philosophical Proceedings of the Royal Society and the French Journal des Scavants, both founded in 1665. A slow-growing phenomenon for many years, its more recent spectacular expansion has not surprisingly mirrored the world-wide growth in research activity. There is quite a number of ways in which the ability to access text - and indeed other forms of non-textual information-online is likely to impact on university libraries.
Journals, however, take over twenty five per cent of my library's total spend and I therefore have every reason to speculate vigorously about the possible effects of on-line access to the information now carried by print-on-paper journals.

This paper refers to the measure of consensus now felt by UK university librarians about the future of the print-on-paper journal, and the inroads on it which are likely to be made by alternative electronic forms of access. It then describes some of the features of the possible electronic forms of access. It goes on to mention some of the opportunities and constraints which at present surround this potentially new way of working for scholars. Finally, it says something about the possible financial context in which this new way of disseminating scholarly information will operate. It is a very short paper for such a large subject and is therefore inevitably highly selective.