ISO/TC 46/SC 9 N 285
2000-05-03

THE DIGITAL WORLD AND THE ONGOING DEVELOPMENT OF ISBN

A Position Paper Developed by the US ISBN Agency
RR Bowker, New Providence, New Jersey, USA

At the most recent meeting of the International Advisory Panel of ISBN Agencies (October 1999) consideration was given to the issues facing ISBN in the digital world. Since that time the US ISBN Agency has sought the advice of industry experts in identifying and proposing ways of dealing with these issues. The following represents a discussion of our work to date.

We welcome your input on the further development of these ideas. We would also like to be advised if we have neglected to identify and address any other relevant points.

This document has been provided to

- The International ISBN Agency Berlin <www.isbn.spk-berlin.de> for consideration by the International Advisory Panel, as well as to the international network of ISBN registration agencies
- ISO Technical Committee 46, Subcommittee 9 (ISO/TC46/SC9): <www.nlc-bnc.ca/iso/tc46sc9>
- The National Information Standards Organization (NISO): <www.niso.org)>
- Book Industry Study Group (BISG): <www.bisg.org>
- Book Industry Communication (BIC): <www.bic.org.uk>

We will keep our constituency in the United States informed of these developments.

(I) ISBN PRODUCTS DISTRIBUTED & SOLD AS DIGITAL FILES

The basic guidelines for assigning ISBNs to digital files have been established:

1. Format/means of delivery are irrelevant in determining whether or not a product is assigned an ISBN. If a product is a monograph, it should be identified with an ISBN.
2. Each format of a digital publication should get a new ISBN.

The definition of new edition is normally based on one of two criteria:

1. A change in the kind of packaging involved: the hard cover edition, the paper cover edition and the library-binding edition would each get a separate ISBN. The same applies to different formats of digital files.
2. A change in the text, not including packaging art or minor changes such as correcting a spelling error. Again, this criterion applies whether the publication is print on paper or digital.

To date the US ISBN Agency has identified the following information regarding digital packages that should receive ISBNs when they contain text-based monograph publications:

FORMAT

EVB = EveryBook
LBS = Librius
MLL = Millennium
PPT = PalmPilot
RKT = Rocket eBook
SBK = SoftBook
GLS = Glassbook
TXT = ASCII Text
HTM = HTML
WRD = MS Word
PDF = Portable Document Format
RTF = Rich Text Format
OEB = Open EBook
MSR = Microsoft Reader
MP3 = MP3 spoken word files
WAV = Wave (audio) files

FILE SIZE

in real numbers

WORD COUNT

In KB

(II) METADATA

The forms that Bowker uses for collecting information about ISBN products include profiles of the metadata needed by the customers for these products. Certain categories of metadata are always relevant: the title of the product is needed whether the book is printed or on audiotape. But the number of pages is relevant only to the printed version, while running time is relevant to the audio version.

The International ISBN Agency is currently establishing the core metadata needed for both print on paper and digital ISBN products in conjunction with EDItEUR International <www.editeur.org>

(III) AUTHORITY TO ASSIGN ISBNs

The publisher, not the distributor of the files or any other entity in the supply chain, is responsible for the management of the ISBN assignment process. It should be noted that when a distributor acquires the appropriate rights it becomes a publisher.

(IV) THE EVOLVING CONCEPT OF PUBLISHER

The growth of publishing on the Internet has resulted in an explosion of small publishing enterprises that publish a limited number of monographs online. The US ISBN Agency would like to encourage this development, and a policy is evolving that will provide ISBNs to small publishers (of both print on paper and digital monographs) on a cost effective basis to facilitate their entry into the ISBN system.

(V) PRINT ON DEMAND PUBLICATIONS

A special instance of digital files includes those used for On Demand publishing.

The following are categories of On Demand publications that have been identified jointly by the Managing Agent of Book And Serial Industry Communications (BASIC): <www.bisg.org> and the US ISBN Agency <www.bowker.com/standards/> or <www.isbn.org/>:

1. A copy of some edition of a title reproduced on demand via rudimentary photocopy, the text of which is exactly the same as the edition copied, while physical aspects of the book are different.

2. A copy of some edition of a title, reproduced on demand via the new reproduction technology, by a publisher that has bought On Demand rights to another publisher's title.

3. A copy produced via the new technology in one of the following three categories:

    a) A copy produced by an organization other than the publisher on behalf of the publisher (e.g., no rights have changed hands);

    b) A copy produced by the publisher;

    c) A copy produced by a retailer.

Of the above, categories "1" and "2" should be given a new ISBN. This ISBN should be applied by the rights holder of the On Demand edition, who may or may not be the original publisher.

It is recommended that EDItEUR devise a subset of On Demand metadata as part of the core ISBN activity currently underway.

(VI) THE USE OF ISBNs TO IDENTIFY PARTS OF MONOGRAPHS AND ONGOING PUBLICATIONS

Given the proliferation of the sale and distribution of discrete parts of monographic and ongoing publications (chapters, articles etc.), the US ISBN Agency recommends the amendment of the ISBN standard to allow for the assignment of ISBNs to these subsets when they are put into the supply chain by their publishers.

Note: We do not see any conflict between the use of the ISBN (ISO 2108) as the assigned identifier for monographic publications that have entered the supply chain independent of their original appearance as parts of larger works and the NISO standards:

> Book Item and Content Identifier (BICI--which identifies parts of monographs that are not independent of their original package;)
>
> Serial Item and Contribution Identifier (SICI—which identifies parts of ongoing publications that are not independent of their original package.)

(VII) THE INTERNATIONAL STANDARD TEXTUAL WORK CODE

Over the past few years the music and movie industries have been engaged in developing identifiers for works. These codes have potential uses both in rights management processes and in e-commerce. It has been deemed necessary to establish a similar kind of

work code for text creations, as well. This code is being developed by the International Organization for Standardization Technical Committee 46/Subcommittee 9 (ISO tc46/sc9.) The US Agency will keep its constituency informed of the progress of this activity.

(VIII) EXPANSION OF ISBN TO THIRTEEN DIGITS: PROPOSAL FROM BASIC

For many years publishers have created Bookland EAN bar codes for their products. These codes are created by adding the prefix "978" to the beginning of the ISBN and updating the check digit. The 978 prefix is one of two prefixes allowed for use in the international book-publishing sector by contract with EAN International Brussels <www.ean.be> and the Uniform Code Council, Lawrenceville, New Jersey <www.uc-council.org>. (The other prefix, 979, has been used to date only for published sheet music.)

For the past two meetings of the International Advisory Panels of ISBN Agencies (1998 and 1999), there has been a discussion of utilizing this bar code scheme as the basis for expanding the ISBN to thirteen digits. The Managing Agent of Book And Serial Industry Communications (BASIC), the New York based publishing industry standards organization has led this discussion, at the invitation of the Director of the International ISBN Agency. The essence of the proposed scheme is that all ISBNs currently in circulation, and all those issued until the supply of ten digit ISBNs runs out, will be considered "978" ISBNs. Then we will begin to reissue the same ISBNs, but with the prefix "979", thus insuring their uniqueness.

There are several reasons that argue in favor of this development:

1. It will incorporate the numbering of sheet music (via the ISMN) into the overall ISBN scheme.
2. It will make the ISBN compatible with both the EAN and U.P.C. codes.
3. It will greatly increase the pool of ISBNs available to publishers.

The Director of the International Agency (Berlin) has also proposed that we do away with both the country/regional identifier and the publisher identifier, turning the ISBN into a completely "dumb" number. This change would also greatly increase the number of ISBNs available for product numbering. This change will lead to problems in managing ranges of ISBNs given to publishers unless a management scheme based on logical divisions is set up. The US Agency (see proposal below) is in the process of setting up such a scheme.

The following timetable has been proposed for the development of the thirteen digit ISBN:

October 1999-April 2001: Consideration of the proposal by the International Advisory Panel and by ISO

May 2001: Approval of the modified standard by ISBN

January 1, 2005: All organizations in the ISBN chain to be thirteen digit compliant

It should be noted that the Uniform Code Council is following this same timetable for requiring that everyone in the North American supply chain be able to accommodate at least a thirteen-digit item identifier in his or her computer systems.

(IX) EXPANSION OF ISBN TO THIRTEEN DIGITS: PROPOSAL FROM THE US ISBN AGENCY

Introduction

At the most recent meeting of the International Advisory Panel of ISBN Agencies (October 1999) consideration was given to the issues facing the ISBN in the digital world. There are other issues concerning the long-term health of the ISBN that present themselves. Following is a decidedly select list:

- Proliferation of authors/publishers given the ease and ubiquity of web publishing (evolving concept of publishers)
- Proposal for identification of discrete parts of monographs (chapters) and ongoing publications (articles)
- Multitude of digital formats available for publication (ebooks)

Others, no doubt, will be identified, as the discussion of these proposals grows wider.

Shortcomings

Under Utilization of Number Possibilities

One of the arguments in favor of the BASIC proposal is, "… it will greatly increase the pool of ISBNs available to publishers." This is debatable. Theoretically, the increase in amount of available numbers at best doubles, from 1 billion to 2 billion. (Check digit not shown; spaces included for readability.)

- 978 000 000 000 to 978 999 999 999 (most of these already assigned)
- 979 000 000 000 to 979 999 999 999

In reality, the theoretical pick up would amount to a mere 90%. Because of an error in ISMN assignments, 979 0 (or 979 3 – research pending) is reserved for the 13 digit

ISMN. Thus, 100 million possible numbers are off limits – reserved as space for tens (possibly hundreds?) of *thousands* of printed music titles.

One could argue that an increase in the length of the number by three digits should yield a benefit of 3 orders of magnitude (to 1 trillion). As proposed, the publishing industry would realize 1/500, or .2% (2 tenths of 1%) of the potential benefit.

We do not feel this is an acceptable proposal. The ISBN – as presently constituted – has served this industry, and has been the envy of other industries, for over 30 years. It is precisely because of the 3 initial bullet points of this paper, namely: 1) the evolving concept of publisher; 2) identification of discrete parts of monographs; 3) digital formats of books, that it is felt that the current proposal is not sufficient. Little hope is held that the proposed scheme could last 35 years. Given the increased stress on the number, 5 to 10 years would be hopeful.

Unintelligent-Number Proposal

If the physical scheme as we now know it (insertion of spaces or hyphens to improve human readability and transcription) goes away, the need for a "virtual" hyphenation scheme incorporating variable size number blocks will remain. Two reasons for this are:

- Manageability
- Efficiency of Number Allocation

Manageability: The US Agency recommends the continued use of large and small blocks of numbers (via virtual division of numbers) so the possibility of double assignment is far less – both at ISBN headquarters and at the agency level. If the ranges are not "virtually implied" the possibility for hidden or oblique double assignment is far greater than under today's system.

Efficiency of Number Allocation
It is envisioned that certain varying "blocks" of numbers will need to be set aside for small, medium, and large entities.

Additional Proposal
Many possibilities exist that would extend the ISBN into the future, and retain backward compatibility with the existing ISBN/EAN number that has served this industry well, and continues to do so. Below are some that were considered. (One is proposed; three rejected.)

Proposed Addition to BASIC Proposal

- The US Agency recommends that the International ISBN Agency petitions EAN/UCC for at minimum one 2 digit prefix (preferably more), thus allowing for a gain of at least one order of magnitude over the BASIC proposal.

The proposed addition ensures backward compatibility, as well as enough headroom to provide a growth path well into the future. It incorporates aspects of the present proposal and extends them where possible.

For instance: the suggestion that ISBNs currently in circulation, as well as those to be issued until the 10 digit number runs out, be considered "978" ISBNs is a good one and is to be maintained. The use of the 9790 (or 9793, as the case may be) to be reserved for ISMNs should suffice.

Conclusion
It is the US Agency's hope that this paper leads to a discussion that – in the end – creates a more durable ISBN for the future. Please feel free to comment.

Also, please feel free to take the ideas under "Rejected Alternatives" (see below) and rework them if possible!

Thank You.

The US ISBN Agency
27 APR 2000

Rejected Alternatives

- Simply declare a 13 digit numeric and use it. Not quite so simple. This number would have little support from international bodies, such as EAN. Upside potential: quick, yields many more numbers. The downside: its usefulness as a scanning number would seriously be in question.
- Take the proposal as floated, and use the 3rd digit (97x) in a hexadecimal representation. (Examples: 97A 000 000 000, 97B 000 000 000, etc.) Excepted, of course, would be 97* prefixes previously assigned by EAN. The upside: a minimum increase of 600% of the present possible ISBNs. The downside: bar coding symbology for books does not support alpha characters.

- Take the proposal as floated, use the 3rd digit to encode the modulus and/or weighting algorithm used to ascertain the check digit for a particular ISBN. The upside: virtually unlimited. The downside: complexity.