2003

CHAOS -- Standards Column

Priscilla Caplan

Member of the NISO Standards Development Committee, Co-chair of the NISO/EDItEUR Joint Working Party, pcaplan@ufl.edu

Follow this and additional works at: http://docs.lib.purdue.edu/atg

Part of the Library and Information Science Commons

Recommended Citation
Caplan, Priscilla (2003) "CHAOS -- Standards Column," Against the Grain: Vol. 15: Iss. 6, Article 34.
DOI: http://dx.doi.org/10.7771/2380-176X.4225

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Stretching ONIX for Serials: The Joint Working Party on the Exchange of Serials Subscription Information

by Priscilla Caplan (Member of the NISO Standards Development Committee, Co-chair of the NISO/EDiteEUR Joint Working Party) <caplan@ufl.edu>

Column Editor: Pat Harris (NISO; Phone: 301-654-2512) <pharris@niso.org>

In the movie Willy Wonka and the Chocolate Factory, little Mike TeeVee is sent overhead “in a million pieces” like a television image. When Mike is reassembled on the other end, he’s only a few inches tall, because as everyone knows when you transmit something by television, it’s always smaller on the other end!

Serials subscription and holdings information always makes me think of little Mike TeeVee. I envision lots of little bits of information flying in all directions. But, when I manage to grab some and pull it down, it never is quite complete.

That’s also the picture the National Information Standards Organization (NISO) had in mind in early 2002 when it commissioned Ed Jones to do a study on the exchange of serials subscription data. There was a sense that a lot of serials data was being exchanged in a lot of formats between a lot of partners for a lot of purposes. There was also speculation that a common format for exchanging this information might save the community as a whole both time and trouble. But it was necessary first to know exactly which parties were involved and what their needs were in order to determine if a standard format would be useful or even feasible.

Defining the Problem

Ed Jones’s study was partially funded by the Digital Library Federation and released in September 2002 as a NISO white paper, “The Exchange of Serials Subscription Information” (http://www.niso.org/standards/resources/SerialsWP.html). Jones’s paper focused on information pertaining to what ejournals a library licensed or subscribed to, individually or in aggregations. Ed found that among the parties exchanging such data were libraries, e-journal aggregation services, publication hosting services, publishers, subscription agents, library systems vendors, and publication access management services or PAMS, a term Ed coined to refer to services like TDNet and Serials Solutions. Data was communicated in order to verify subscription access, to create A-Z title lists, to link to libraries’ catalog holdings, to control online display options (e.g., to suppress document delivery options for titles held by a library), and to populate link resolution servers, among other things. Jones’s study also confirmed that there were nearly as many transmission formats in use, as there were transmissions.

The study noted that most of the bits of information being circulated already existed as data elements in ONIX for Serials. ONIX for Serials is the younger sibling of ONIX for Books, developed by EDiteEUR. At the time of the study three specifications were published in draft form: the serial title record, the serial item record, and the subscription package record. Unlike ONIX for Books, which was developed primarily for communications from publisher to retailer, ONIX for Serials was intended from the start to be useful to libraries as recipients of information about serial issues and articles. Ed recommended ONIX for Serials be evaluated for use in some of the transactions he described.

NISO/EDiteEUR JWP Formed

NISO and EDiteEUR immediately established the Joint Working Party on the Exchange of Serials Subscription Information, fondly known as the JWP. Co-chaired by Richard Gedye of the Oxford University Press and myself, the JWP was charged with recommending changes to ONIX for Serials to enable its use in this context, and to demonstrate the utility (or lack thereof) of the format by piloting its use.

Having worked in the standards arena for nearly ten years, I can testify that normally you have to bribe, threaten or trick people into joining standards committees. With the JWP, it was more a case of beating people away with broomsticks. The group has 35 official members, the majority of whom participate actively, and a number of unofficial participants as well. Membership is divided fairly evenly between the US and the UK, and represents libraries, library systems vendors, publishers, PAMS, subscription agents, and online content providers. EDiteEUR has provided the invaluable services of its ONIX consultants, while NISO has provided communications and staff support.

The JWP was initially charged to complete its work within a year, which would have it self-destruct before you read this issue of ATG. However, the pilots are just now getting off the ground, and the group will probably remain active through the spring of 2004.

Serials Formats Defined

By late-2003 the JWP had defined two ONIX for Serials formats (XML schema): Serial Products and Subscriptions format (SPS) and the Serial Online Holdings format (SOH). These two formats support six different types of transactions (transmission of data from one party to another).

First, some working definitions. A product is an entity representing one or more electronic and/or print resources to which a library may subscribe. Or, more simply, anything sold with a price. So, for example, if a publisher sells separately the print version of a journal, the electronic version of a journal, and a bundled print-and-electronic package, these are three products. WorldCat is a product, as is Academic Search Premier. A work is an entity representing the intellectual content of a publication. The Journal of Mathematical Logic is a work with two versions, one electronic and one print. A version, or title-version, is a specific published format of a journal, such as the electronic version.

The SPS supports at least four different transactions:

a) an unpriced list of products available from the sender, showing works and title-versions in each product;
b) a priced list of products available from the sender, showing works and title-versions in each product;
c) a list of products with prices for one

continued on page 89

<http://www.against-the-grain.com>
particular subscriber, including current
prices and prices last paid;
d) an unpriced list of products to which
a library currently subscribes, through
an agent or from the publisher.
Transactions (a) and (b) could be used by
a publisher to send to a library its catalog and
price catalog respectively. They both answer
the questions, what works and versions of
works does this publisher publish, and in
which products are they available?
Transaction (c), the subscriber’s priced
product list, would also be sent from the pub-
lisher to a library. It identifies which prod-
ucts the publisher has on record as being sup-
pied to the library either directly or through
an agent, and what title-versions are included
in those products. It could be used by a li-
brary to compare a publisher’s price list to
what the library is actually paying, or to en-
sure that the library is not paying more for a
bundled package than it would for individual
subscriptions. It could also be used by a li-
brary to ascertain which of the print publica-
tions it receives have electronic versions, in
order to expand the number of electronic re-
sources the library has.
Transaction (d) could be sent by a library
to a publisher, and would identify which prod-
ucts the library believes it subscribes to
from that publisher. Along with transaction (c), it
could be used to reconcile the library’s order
file and catalog with the publisher’s records.
The SOH is designed for use by a PAMS
or subscription agent to send information to
a library. It supports two transactions, one
allowing the library to create an A-Z list of
ejournal titles with links, and one allowing
the library to populate its link resolution
server. The SOH contains a construct for
“online service,” a named host through
which e-journals are available, and a “hold-
gings record” which identifies for each title-
version what holdings are included.

Testing Phase Begins
The ONIX formats are a bit more com-
licated than the ad hoc, local formats being
used to communicate similar data now, many
of which are just spreadsheets, HTML, or
comma-delimited files. Since the benefits
the SPS and SOH offer depends upon their wide-
spread adoption, they won’t materialize if the
formats are too difficult to produce or pro-
tected. Therefore, it is essential to prove
the transactions are workable in the real world.
JWP participants are testing the SPS and
SOH transactions now. In one pilot, Serials
Solutions has produced an ONIX transaction
that the Library of Congress used to create
an A-Z list. Other pilots involve one or more
publishers creating SPS transactions for li-
braries, and at least one PAMS producing an
SOH transaction for importation to a link
resolution system.

One problem in the pilot period is that li-
braries have not automated way of processing
the ONIX for Serials
transactions. Even if pro-
viders can produce them, li-
brary systems cannot im-
port them. The good news
is that several library sys-
tems vendors have been partici-
pat ing in the JWP,
and one of them, Innova-
tive Interfaces, is involved
in the piloting of the SOH.
The ILS vendors recognize
that standardization of these transactions is in
their best interest also, and if li-
braries spread the word and
insist on support from their
vendors, it seems reason-
able to expect that this sup-
port will be developed.

Identifier Issues
This work has exposed
some of the limitations in
our current system of serials identifiers. Ideally, serial
titles (works), title versions,
online services, libraries and publishers
would have unique iden-
tifiers making it easy to col-
locate information, to link
different applications,
and to implement up-
date and replace function-
ality. In practice, there are
problems with each of these.

For example, the ISSN
should function as a title
version identifier, because
of the rule that print and electronic versions
should receive different ISSNIs. However, not
all serials have ISSNs, not all publishers assign
multiple ISSNIs, and not all systems can
capture and display multiple ISSNIs correctly.
Therefore, many serial records have missing
or incorrect ISSNIs.

At the same time, there is no serial work
identifier that can bring together all versions
of a title. There are some possible solutions
on the horizon: using the new International
Standard Text Code (ISTC), which is cur-
rently not being assigned to serials; using the
print ISSN; or using a DOI that publishers
can assign at the work level. Additional work
is needed to determine if any one of these
options is realistic.

The best solution to both problems, in my
view, might be to revise the ISSN system to
reform it as a base number plus suf-
fix. The base number would be assigned
to all versions of a title, functioning as a work
identifier, while the suffix could identify the
title version uniquely. While this change is
not likely to happen, it is within the realm of
the possible, as the ISSN standard is currently
being revised by ISO.

While the pilots are in process, an Identifi-
iers Subgroup of the JWP is looking at these
and other identifier issues.

Against the Grain / December 2003 - January 2004 <http://www.against-the-grain.com> 89