Date: January 1, 2002

To: ALCTS Committee on Cataloging: Description and Access

From: Michael Chopey

Subject: Comments on ISO Draft ISO/CD 21047
            International Standard Textual Work Code

These are my comments on the ISO Committee Draft ISO/CD 21047 (International Standard Textual Work Code).   I intend for these to be possible starting discussion points for us as we decide whether to recommend a yes vote or a no vote on the draft standard to Paul Weiss (ALA's NISO voting member).   None of these comments has to be included in our final comments to him -- after we have had a week to discuss this, I will plan on including comments that the Committee seems to agree upon.

Would you please send your own comments, and responses to these comments, no later than next Tuesday, January 8?   After your comments are received, Brad Eden and I will compile them into a report (or a letter, perhaps) that Kristin can send to Mr. Weiss after the Committeevotes on it and votes whether to recommend a yes vote or no vote.

I assume that you have all received and read the draft standard, so I don't think I need to summarize or quote extensively from the document. First some general comments, and then some section-by-section comments.

General Comments

My first comment is that I like the idea behind this standard. If ISTCs are assigned accurately and according to the specifications in this draft, and are printed on bibliographic items or otherwise provided with them (e.g., in the ONIX metadata accompanying them), we will have one more piece of important bibliographic data to aid catalogers around the world in identifying and collocating related manifestations of the same text-based expression. [This raises an important point about the use of the word and the concept "work" in the draft standard — see below for further comments on that).

One instance where having an ISTC included on/with a bibliographic item might be useful to a cataloger is in the case where the cataloger is working in an unfamiliar foreign language. Another instance would be when a lot of time would otherwise have to be spent by the cataloger in determining that the item in hand is indeed a manifestation of a given expression. If adopted widely enough, I think that eventually our cataloging code could prescribe transcription of the ISTC (in Area 8), and MARBI could assign it a MARC field. If ISTCs were assigned widely enough (especially if retrospectively), and consistently included in bibliographic records, this could provide a very simple and powerful collocation point in our bibliographic utilities and union catalogs, and one that might prove more efficient in some ways than uniform title headings. I also envision that ISTCs could eventually possibly lay

the groundwork for an international authority file of works and their expressions and manifestations.

Having said all of that in favor of adoption of this standard, I also see some fundamental problems in the standard as written.

One of the first things you will have noticed in this draft standard is that the all-important term and concept "work" here actually corresponds to the FRBR definition of "expression," not the FRBR definition of "work." This disturbs me because it seems like a real disservice to the monumental work of the IFLA Study Group on the Functional Requirements for Bibliographic Records (and to all of the subsequent work that has built upon the definitions in FRBR) to introduce such a fundamental discrepancy in terminology. The authors of the draft standard are aware of the the discrepancy (and they cite FRBR in the draft's bibliography). This is the rationale they offer for rejecting the use of the term "expression":

> "In generic bibliographic terminology, the term expression is defined in the International Federation of Library Associations and Institutions' publication Functional Requirements for Bibliographic Records (section 3.2) as an abstract entity between the concept of a work and the concept of its physical manifestations. The following is a brief excerpt from the FRBR description of these concepts:
>
> '… The intellectual or artistic realization of a work in the form of alpha-numeric, musical, or choreographic notation, sound, image, object, movement, etc., or any combination of such forms …
>
> An expression is the specific intellectual or artistic form that a work takes each time it is "realized." …
>
> Inasmuch as the form of expression is an inherent characteristic of the expression, any change in form … results in a new expression. Similarly, changes in the intellectual conventions or instruments that are employed to express a work (e.g. translation from one language to another) result in the production of a new expression. …'
>
> For the purposes of assigning ISTC, this relation is covered by the term derivation. For that reason and to prevent any confusion, the term expression has been avoided in drafting the ISTC standard.
>
> Examples of specific kinds of derivations which result in new works (for the purposes of assigning ISTC) and/or new 'expressions' (in FRBR terms) are provided in an annex to the ISTC standard and will also be covered in greater detail in future user manuals for the ISTC system."

[This is not included in the draft standard. You can read this, and other facts about ISTC on an FAQ on the Working Group's Web site at http://www.nlc-bnc.ca/iso/tc46sc9/istc.htm]

Another reaction that I think many of us who are steeped in the FRBR Work-Expression-Manifestation-Item model will have is that the idea of a "textual work" as defined here is too simplistic and does not sufficiently anticipate all of the

complexities that can arise in delineating bibliographic relationships. I certainly believe that this is true, but after some thought on this, I'm thinking that this could end up being a plus, rather than a minus, in that the idea of the ISTC is to assign a numeric identifier to every new textual expression, and then link those that are bibliographically related in the ISTC metadata that will express relationships between the expressions that will have ISTCs assigned to them. The ISTC itself (I mean the 16-digit number) is not designed to be expressive of any bibliographic relationships between a given expression (what the draft standard refers to as a "work") and any other related expression. The idea is that a textual derivation of an existing expression would be assigned a new ISTC, and these two expressions would be linked by the ISTC metadata that is "declared and recorded as part of the ISTC registration process." Any shortcomings in this system will be in the metadata, not in the ISTC number. Perhaps we librarians might someday make use of the number and the supplied metadata telling what kind of relationship this entity bears to another, to overcome these shortcomings in our own systems? I think that is possible, obviously.

The categories of "origination type" that the standard prescribes for ISTC metadata are: original (the default), abridgement, annotated, critical, expurgated, additional non textual content, translation, revision, compilation, excerpt, unknown, and unspecified modification. (Definitions of these categories are on pages 10-11 of the draft standard.) More than one "origination type" value can be recorded.

Another general concern I have is that not enough attention has been paid to the possibility of anonymous works and works of unknown or uncertain authorship. This is understandable I guess, in that this standard is probably designed for rights management more than anything else, and the idea behind it is that the party interested in registering a "work" will most often be an author or someone else who has something to gain by getting credit for the expression. Still, for our purposes, if this code ends up being widely assigned and adopted by the publishing industry, I think we will want to leave room for the (retrospective) assignment of ISTCs to existing anonymous works and works of unknown or uncertain authorship down the road. (In my section-by-section comments, I point out where this could be addressed).

A final general concern I have is whether registering parties and/or the ISTC registration agency will really be capable of identifying (and motivated to identify) related bibliographic expressions.

Section-by-Section Comments

Here are my section-by-section comments on the draft. Again, this is probably more than we want to tell Paul Weiss, and more than he wants to hear from us. These comments are meant mostly as discussion points for us as we decided whether to recommend a yes or a no vote.

Note: I find the use of the word "work" and the phrase "textual work" troublesome throught the text of this standard. I would prefer that the entire document were rewritten (and the standard code renamed) to replace "work" with "expression" throughout. I won't comment on this anywhere else here.

1. Scope

The first sentence of the second paragraph says "The ISTC may be applied to any textual work, when there is an intention to disseminate such a work, at any point between creator and user."   This language seems to preclude the possibility that an ISTC will be assigned to a textual work (textual expression, actually) which has no named creator.

3. Terms and definitions

3.2 Author. A creator wholly or partly responsible for the intellectual content of a textual work.

Comment: We would probably prefer something like "A creator to whom the the intellectual content of a textual work is wholly or partly attributed."

3.4 Contributor. A person contributing to the making of a textual work in whole or in part.

Comment: "Person" is defined below in this list as "an individual or organization." Maybe there is some good reason for using the term "person" in this sense, but I find it confusing every time it occurs in this document. I would prefer "party." I.e., Contributor. A *party* contributing to the making of a textual work in whole or in part.

Also, there is a contradiction to this definition later in the document (Annex D.3) when a publisher is listed as one type of contributor. A publisher does not contribute the making of a textual work (expression) — a publisher contributes only to the making of a manifestation of that work (expression).

3.5 Creator. A contributor originating content in a work.

Comment: Inserting the word "intellectual" between "originating" and "content" here might make this more clear. (Later in the standard, in Annex D.3, the different contributor roles besides creator are enumerated. It becomes more clear at that point that "creator" does not include translator or compiler or excerptor, etc.)

3.11 Person. An individual or organization.

Comment: C.f. 3.4, 3.13, etc., I would prefer "party" to express this meaning, unless there is some good reason why "person" is preferable.

Annex D (normative) ISTC metadata for registration of textual works

D.1 Before allocating an ISTC to a textual work, an ISTC registration agency shall capture the essential ISTC metadata for that work, as described in D.2 — D.6 below.
These tables of metadata elements are subject to change at the discretion of the ISTC Registration Authority in consultation with the ISTC registration agencies. In particular, the ISTC Registration Authority may establish rules for the registration of different genres of textual work and the specific data

elements that must be registered before allocating an ISTC to a textual work in that genre.

Comment: This is a good thing, I think. As I mentioned in my general comments, if there are shortcomings in this whole ISTC scenario, they will be in the metadata, not in the number itself (unless the number itself could be made more expressive and meaningful — I don't argue that it can't, I just don't have any suggestions for making it so).

D.2 Title information

Publication date. The date of publication of the first published manifestation of the work.

Comment: Insert "known" between "first" and "published" here.