

## **New Directions in Information Organisation: submission from Alan Poulter**

### **Filling in the blanks in RDA or remaining blank? The strange case of FRSAD**

#### **Introduction**

RDA (Resource Description and Access) was released in July 2010, and made available for use, either in an online form, the RDA Toolkit (<http://beta.rdatoolkit.gvpi.net/>) or in printed form, in a large loose-leaf binder. In July 2011, the Library of Congress, the National Library of Medicine, and the National Agricultural Library announced the decision to adopt RDA after conducting trials. The decision to adopt RDA though carried riders on certain perceived issues to be resolved, related to rules readability, online delivery issues of the RDA Toolkit and a business case outlining costs and benefits of adoption. It appears though that, allowing for these issues to be dealt with, RDA will begin adoption in 2013 and will gradually replace the aged AACR2 (Anglo-American Cataloguing Rules, Second Edition).

Unlike AACR2, RDA was intended to also provide subject access. As RDA currently stands, Chapters 12-16, 23, 33-37 are intended to establish guidelines for providing subject access, but only Chapter 16, 'Identifying Places' is complete..

This paper will outline possible strategies for moving forward in completing the remaining blank chapters, based on the model given in the recent Functional Requirements for Subject Authority Data (IFLA Working Group, 2010), here-after referred to as FRSAD.

#### **Paper overview**

This paper begins by outlining significant developments prior to the appearance of FRSAD which was formerly known as FRSAR. This involves coverage of the two preceding reports, the Functional Requirements of Bibliographic Records (FRBR, IFLA 2008) and the Functional Requirements for Authority Data (FRAD, IFLA 2009), which was formerly known as FRANAR.

The final version of FRSAD, released in 2009, will be contrasted to earlier efforts to extend the FRBR/FRAD models to fully cover subject access.

Finally, a prospective proposal to take FRSAD forward to implementation using PRECIS (Preserved Context Indexing System) will be examined, as well as the general reception of FRSAD.

#### **Before FRSAD**

The roots of FRSAD go back to a critical juncture in the revision of AACR2. In April 2004 two bodies managing the development of a revision of AACR2, the funder, the CoP (Committee of Principals) and the developers, the JSC (Joint Steering Committee) decided that the level of change was no longer at the amendment level and was instead a comprehensive revision of AACR2. In April 2005 it was decided that AACR2's structure should be abandoned and that a new alignment with two abstract models of publication, based on ER (entity-relationship) models, FRBR and FRAD was to be used as the basis for the new rules to replace AACR2: their name was changed to RDA to indicate this fundamental shift.

## **New Directions in Information Organisation: submission from Alan Poulter**

An 'entity' is a thing which is capable of an independent existence and which can be uniquely identified. Every entity must have a minimal set of uniquely identifying attributes, which is called the entity's primary key. A 'relationship' expresses how entities are related to one another. Entities and relationships can both have 'attributes', named features. The intention in using ER modelling was to make explicit what was being described and how the elements of the model related.

The entities in FRBR were split into three Groups. Group 1 was for "intellectual products" and there were four entities for these: 'works', 'expressions', 'manifestations' and 'items' (WEMI). The 'work' entity was a distinct intellectual creation, for example Daniel Defoe has the idea of a story about a man stranded on an island. The 'expression' entity is the realization of a work in some form (a language, music, etc). Defoe thinks of the story in English but it can be realised in other languages and media. The 'manifestation' entity is the embodiment of an expression of a work, e.g. the first edition in English, a later English version in the Penguin Classics etc. The 'item' entity represented a single physical copy of a manifestation, e.g. an owned copy of the Penguin Classic. Using ER relationships, a work can have many expressions, each expression can have many manifestations and each item can only come from one manifestation. Generally, most works will have one expression and one manifestation of that expression. Manifestations of the same expression may have identical content but will vary in some other detail e.g. publication date. Manifestations of different expressions equate roughly to editions.

Group 2 entities were those responsible for intellectual/artistic content, that is 'persons', 'corporate bodies' and 'families', while Group 3 entities were proposed to represent subjects: 'concepts', 'objects', 'places' and 'events' as well as all entities in Groups 1 and 2. Thus, a place can be the subject of a travel guide, a person can be the subject of a biography and a poem can be the subject of a critical text. However, the Group 3 entities were only intended as place holders, to indicate a future desire to represent subjects.

FRBR was explicitly designed to support user tasks. It does this by defining a set of user tasks:

Find: find entities that match a need

Identify: confirm that entities match a need and be able to distinguish them

Select: find the entity most appropriate

Obtain: get access to the required entity

and then explicitly highlighting particular attributes of WEMI entities as being required for one or more of the above tasks. Again, as far as subject access was concerned, these tasks were insufficient.

### **Precursors to FRSAD**

Prior to the appearance of FRSAD there were two significant attempts to extend the FRBR/FRAD models to subject access. Pino Buizza and Mauro Gerrini had been involved in creating and testing an Italian version of PRECIS (Preserved Context Indexing System) for Italian libraries and in their paper (Buizza and Guerrini 2002) they outlined a potential subject access model for FRBR. Tom Delsey, the author of RDA also examined the problem of adding subject access.

## **New Directions in Information Organisation: submission from Alan Poulter**

Buizza and Gerrini note that, uniquely, FRBR tried to bring cataloguing and subject access together, rather than consider them as distinct, as in the past. There was also an international aspect, which tried to make subject access a feature not restricted by language:

“While certain aspects of semantic indexing have necessarily national characteristics...It is indispensable for the theoretic development to take place within international debate, and that the new working instrument be conceived as part the logic of international cataloguing co-operation and integration” (Buizza and Guerrini 2002, p33)

Buizza and Guerrini note that subject is not an entity present in an item nor does it exist in its own right, it is a mediator between the topic of a work and the universe of enquiries which seek answer. Rather, subject persists independently and allows us to recognise common themes and distinguish competing claims of relevance.

They point out that because of the relationship between work and expression, manifestation and item, there was no need to investigate entities other than work as they would inherit their subject from the source work. In FRBR they recognise that the expression of Group 3 subjects is not meant to be exhaustive. For example, there is no category for living organism. The entities in the subject group, even when supplemented by the Group 1 and 2 entities correspond to a very simple categorisation, which is there as a placeholder, and which is intended to be built upon and expanded. While FRBR does not perform an analysis of publication models but rather defines a practical generic structure, it makes no claim to be a semantic model. Unlike the other entities, subjects are presented as individual instances of atomic units, with no attributes.

They attempt to extend the ER model to indexing by proposing two new entities: ‘subject’, the basic theme of a work, and ‘concept’, each of the single elements which make up the subject. The entity types making up subject are suggested as: ‘object’, ‘abstraction’, ‘living organism’, ‘material’, ‘property’, ‘action’, ‘process’, ‘event’, ‘place’ and ‘time’. ‘Person’, ‘corporate body’ and ‘work’ are also included from FRBR. This is a much more extensive model and appears to cover the full range of potential classes of entities.

Having two distinct entities (‘subject’ and ‘concept’) allowed statements of the subjects of works, as well as allowing for recurring elements of subjects and the generic set of relationships (broader/narrow, related, use for etc) between them. The main attribute of ‘subject’ is defined as ‘verbal description’, the statement of the subject. Further attributes would include ‘identifier’ and ‘language’. Both these attributes would be required for managing multi-lingual systems. For ‘concept’ the main attributes are given as ‘term for the concept’ and ‘qualifier’, for example for a limited date range. An example ‘subject’ might be ‘training dogs’ in which there are two ‘concepts’, ‘dogs’ as an entity type of ‘living organism’ and ‘training’ as an ‘action’ type entity.

They proposed three types of relationship to exist. There is the primary relationship of the ‘subject’ to its constituent ‘concept’ elements. The second relationship was between the potentially different constituent ‘concepts’ in ‘subjects’ which are identical. Finally, there would be relationships between the concepts themselves. These would be hierarchical, associative and synonymous/antonymous. They also proposed to expand the set of user tasks given in FRBR to add some appropriate for subject access, for example ‘search for a known topic’.

## **New Directions in Information Organisation: submission from Alan Poulter**

Finally they emphasized the importance of maintaining the distinction between the ‘subject’ and ‘concept’ entities, as they had defined them, although they note a potential issue with the former. Their analysis did not give any attention to citation order within ‘subjects’, which would be essential for the coherence and readability of the strings of ‘concepts’ used in subjects. They conclude that their proposal:

“demonstrates a greater affinity with systems based on logical analysis and synthesis techniques, rather than those systems based on lists of pre-constituted headings”( Buizza and Guerrini 2002, p44)

The second attempt at expounding a subject extension for FRBR/FRAD came from Tom Delsey, who, as the chief author of RDA, recognised it as the next hurdle. In Delsey (2005) he stated that neither FRBR or FRAD were complete in their conceptual analysis of data relevant to subject access as performed by bibliographic and authority records. Refining and extending their models to reflect subject access fully would require a significant re-examination of the entities in those models and their attributes and relationships. The new entities when defined would have to completely cover the range of topics that would be required for subjects as understood by library users. Also needed would be all the attributes for the construction and use of subject access points and subject authority records. Finally, there would be need for a model to provide a clear and robust representation of the range of subject access tools – thesauri, subject headings, classification schemes and the syntactic structures used in indexing strings, as these would all be needed. Major expansions of the FRBR and FRAD models would be required:

“In examining the entities in the existing models, we need to check whether they cover the whole ‘subject universe’ and whether they can forge the range of tools used to implement the subject universe.” (Delsey 2005, p52)

For each Group 1 entity in FRBR, an identifier (one or more attributes) and other appropriate attributes, are defined. In FRBR, the entities ‘work’, ‘expression’, ‘manifestation’ and ‘item’, get attributes ‘title’ and ‘identifier’, as well as additional attributes that may be needed for clarification in entries, for example ‘form’, ‘date’ and ‘language’. Again, for the FRBR entities ‘person’ and ‘corporate body’, the identifying attribute is ‘name’, which can be supplemented by, for example, ‘date’, ‘number’ and ‘place’. This is not the case for each of the ‘concept’, ‘object’, ‘event’ and ‘place’ entities, for which only one attribute was currently defined, ‘term’ for use as an entry element in a subject access point and for all other roles needed in subject access. Delsey felt that this was not enough and that there was a need to define additional attributes for ‘concept’, ‘object’, ‘event’ and ‘place’ so that they could be used in subject access points and authority records.

In FRAD the attributes for FRBR access roles, ‘name’, ‘title’ and ‘term’ become entities in themselves with sets of attributes for types and their identifiers. For example, ‘name’ has attributes ‘title’, ‘corporate name’, ‘identifier’ etc, elements like ‘forename’, ‘surname’, etc and additional elements, like ‘scope’, ‘language’, ‘dates of usage’ etc. Also, in FRAD the attributes for each of the FRBR entities was expanded by additional attributes which were needed for confirming the identity of the entity represented by the access point. So, for example, a work might need a ‘place of origin’ or a manifestation a ‘sequence number’. For the entities ‘person’, ‘corporate body’ and ‘family’, corresponding attributes would be ‘place of birth’, ‘gender’, ‘citizenship’, ‘location of head office’ etc. In FRAD for ‘concept’ only ‘type’ is given as an attribute, while ‘object’ has ‘type’, ‘date of production’ etc, The entity

## **New Directions in Information Organisation: submission from Alan Poulter**

‘event’ had ‘date’ and ‘place’ as attributes while ‘place’ had the attribute ‘co-ordinates’ and other geographic terms. Thus, only the ‘type’ attribute of ‘concept’ and the ‘type’ attribute of ‘object’ could be useful in implementing the categorizations that are reflected in the facets and hierarchies defined in thesauri and classification schemes.

Relationships would also need extending. In FRBR there were two levels of relationships, those that worked at the highest level on down – work ‘is realised by’ expression, person ‘is known by name’ etc and those that operated between specific instances of the same or different entity type – e.g. work ‘has supplement’. The relationship ‘has a subject’ would have to encompass not just the expected features (like subject headings) but also links by genre, form and possibly geographic and temporal categories. Also provision for semantic relationships would be needed, between subject terms, narrower and broader, equivalent and related, associative, and chronological/geographical ranges. Delsey noted that associative relationships (‘see also’) would be the hardest to accommodate, as they were neither equivalent nor hierarchical but simply what did not fit into those two groups. There was a need to establish whether associative relationships only operated between instances of ‘concept’ or did they operate as well between ‘place’, ‘event’ and ‘object’ as defined in FRBR?

Delsey also attempted to check the FRBR/FRAD models at a high level, to determine whether they encompassed all possible subjects by comparing them against a recognised universal model, Indecs. Indecs was the outcome of a project funded by the European Community Info 2000 initiative and commercial rights organisations (Rust and Bide 2000). It defined ‘percepts’ (things that the senses perceive), ‘concepts’ (things that the mind perceives), and ‘relations’, which are composed of two or more percepts and objects. At a lower level, percepts divided into animates, ‘beings’, and inanimates, ‘things’, and relations into dynamic ‘events’ and static ‘situations’. The FRBR entity, ‘object’ equated to Indecs ‘percepts’, and ‘concept’ is in both FRBR and Indecs. However, the FRBR entity ‘event’ equated to a subclass of ‘relation’, while FRBR’s ‘place’ in Indecs was paired with ‘time’ as in Indecs these two concepts together were needed to fix an ‘event’ or ‘situation’. ‘Person’ in FRBR was a problem as it needed a subset of the Indecs ‘beings’, while FRBR’s ‘corporate body’ was a special instance of ‘group’ (which included family, societies etc) which would go under either ‘object’ or ‘concept’ in Indecs. These were problems chiefly caused by FRBR’s need to focus on distinct entities needed for bibliographic purposes, but the mismatch in the high level classification of reality in the two models did raise serious doubt on the viability of the FRBR Group 3 entities.

Delsey also noted Buizza and Guerrini’s approach, in creating a new entity to represent the entire string or indexing terms forming a topic. He agreed that syntactic priorities for ordering the terms would still need to be applied within the string, so some system of assigning string roles and ordering was required. The challenge in creating such a system:

“lies in the wide and diverse range of such relationships...Ideally the relationship types would be the same range of relationships but would do so at a higher level of generalization to which specific types in indexing languages could be mapped...On a practical level it would also provide the basis for mapping syntactic relationships to generic categories to support subject access across databases containing index strings constructed using different thesauri and subject heading lists” (Delsey 2005, p52)

## **The arrival of FRSAD**

## **New Directions in Information Organisation: submission from Alan Poulter**

The Working Group on the Functional Requirements for Subject Authority Data (FRSAD) was the third IFLA working group of the FRBR family. Formed in April 2005, it was charged with the task of developing a conceptual model of FRBR Group 3 entities within the FRBR framework as they relate to the "aboutness" of works.

It began by conducting two user studies. The first was a study of attendees at the 2006 Semantic Technologies Conference (San Jose, California, USA). The second was an international survey sent to information professionals throughout the world during the months of May-September 2007. In both, participants were asked to describe their work and their use of subject authority data in different contexts. The FRSAR five user tasks were based on the results. Another objective was to redefine the FRBR/FRAD user-tasks towards "aboutness", so a new set was produced:

“Find one or more subjects and/or their appellations, that correspond(s) to the user’s stated criteria, using attributes and relationships;

Identify a subject and/or its appellation based on its attributes or relationships (i.e. to distinguish between two or more subjects or appellations with similar characteristics and to confirm that the appropriate subject or appellation has been found);

Select a subject and/or its appellation appropriate to the user’s needs (i.e., to choose or reject based on the user's requirements and needs);

Explore relationships between subjects and/or their appellations (e.g., to explore relationships in order to understand the structure of a subject domain and its terminology).” (FRSAD, 2010 p9)

The last one, ‘explore’ is a new task not in FRBR/FRAD, to enable users to browse subject resources

Although ‘aboutness’ is the focus, FRSAD also considers ‘of-ness’ in terms of form, genre and target audience as this concept overlaps that of the pure subject search,

There seems to have been a general agreement that Group 3 entities should be “revisited”. Alternative models, including the one discussed previously from Buizza and Guerrini, were considered. Delsey’s approach of using other general models to examine the Group 3 entities was copied, and indecs, and other general models, like Ranganathan’s, were examined.

By 2007, focus had shifted towards the development of a different conceptual model of Group 3 entities.. What was proposed was a very new general model, based on ‘thema’ and ‘nomen’, whereby the former, defined as ‘any entity used as the subject of a work’, was represented by the latter, defined as ‘any sign or sequence of signs’. In general a ‘thema’ could have many ‘nomens’ and vice versa, while ‘works’ could have many ‘thema’ and one ‘thema’ could apply to many works. A ‘nomen’ was defined as any sign or sequence of signs (alphanumeric characters, symbols, sound, etc.) by which a thema was known by, referred to, or addressed as. For example, "indexing", or "025.4". These two entities enabled the task “to build a conceptual model of Group 3 entities within the FRBR framework as they relate to the aboutness of works” to be fulfilled, and the model resulting was very compact and generic. Any existing subject access scheme could be ‘represented’ and examples were given in Appendices.

## **New Directions in Information Organisation: submission from Alan Poulter**

Themas could vary substantially in complexity or simplicity. Depending on the circumstances (the subject authority system, user needs, the nature of the work, etc.) the aboutness of a work could be expressed as a one-to-one relationship between the work and the thema. In an implementation themas could be organised based on category, kind, or type. The report did not suggest specific types etc, because they may differ depending on implementations.

Thema attributes were 'type', the category to which a thema belonged in the context of a particular subject organisation system and 'scope note', text describing and/or defining the thema or specifying its scope within a particular subject organisation system.

Nomen attributes were 'type' (e.g. identifier, controlled term), 'scheme', reference source, representation (e.g. ASCII), 'language', 'script', 'script conversion', 'form' (additional information, 'time of validity' (of the nomen not the subject), 'audience' and 'status'.

Finally, the 'thema' and 'nomen' conceptual model also matches well with schemas such as SKOS (Simple Knowledge Organisation System), OWL (Web Ontology Language), and the DCMI Abstract Model, making it ideal for resource the sharing and reuse of subject authority data.

Although produced by IFLA, the reports have come from different groups over a long period of time which has meant that their approaches and outcomes have differed. There is a significant conceptual mis-match between the reports in how far to go when proposing a new conceptual model. The FRSAD report is also different in that it reads more like an academic paper than a structure that lays the foundations for practical developments, which the earlier reports do.

However, by using such a simple model the aim "to provide a clearly-defined structured frame of reference for relating the data that are recorded in subject authority records to the needs of the users of that data" is fulfilled on paper and in theory. What is needed is bridge into being able to apply FRSAD's abstract model using a tried and tested tool.

To try and move on, without re-visiting work on FRSAD, it seems prudent to adopt the general model it proposes but actually use an existing system that is based on solid theory, congruent with that in FRSAD, that has been tried and tested and possesses the ability to form a structure that can both exist on its own and also can serve to interlink between other existing schemes, especially the dominant ones, LCSH (Library of Congress Subject Headings) and DDC (Dewey Decimal Classification). PRECIS is proposed for this role.

### **Implementing FRSAD with PRECIS**

PRECIS (Preserved Context Indexing System) is not a list of terms/codes. It is two sets of procedures, one syntactic using a general 'grammar' of roles to generate one or more terms (a 'string') to unambiguously represent a topic, the other semantic setting up permanent thesaural connections between terms where needed. It does not prescribe terms. PRECIS grew out of research into classification which produced its set of syntactic codes, known as 'role operators' (Austen 1974).

Implemented first by the British National Bibliography, to streamline subject operations, each PRECIS string was given a unique Subject Indicator Number (SIN). Added to the SIN were equivalents in DDC and LCSH. Once SINs were created, their re-use would save time and

## **New Directions in Information Organisation: submission from Alan Poulter**

effort. Reference Indicator Numbers (RINs) performed a similar role for thesaural aspects (Austin, D. 1984). In its heyday, PRECIS was being used in bilingual Canada and its use in a number of languages was being investigated (Detemple, S. 1982, Assuncao, J.B. 1989). It was even given a trial at the Library of Congress (Dykstra, M. 1978). Subject data can be seen as more crucial to the growth of Semantic Web than descriptive data. Austen (1982) attacked the early claims of machine retrieval. It is surely prudent to equip cataloguers as soon as possible with the tools to mount one more offensive.

Derek Austin joined the BNB (British National Bibliography) in 1963 as a Subject Editor, after having worked as a reference librarian for many years. He says in his memoirs (Austin 1998) that:

“A hard pressed reference librarian quickly learns to distinguish among and evaluate everyday working tools such as indexes and bibliographies, and tends as a matter of course, to identify, possibly at a sub-conscious level, those features which mark one index, say, as more or less successful than another.”

This practical experience was crucial to his utilitarian, rather than philosophical approach to subject retrieval. His job at BNB was checking the appropriateness and accuracy of DDC (Dewey Decimal Classification) numbers. In 1967 he was seconded to research work for the Classification Research Group (CRG).

At the time there was general dissatisfaction with the two main schemes used for subject access, DDC and the Library of Congress Classification (LCC) as their lack of a well-explained logical structure and inconsistencies in their sub-division, made it hard to accommodate new subjects. However, critiques of existing schemes in themselves did not solve these issues nor give a basis for more solid approach. One potential route was offered by S.R. Ranganathan, whose facet analysis approach, based on the universal facets of place, material, energy, space and time (PMEST).

At a conference of the Classification Research Group (CRG) in London in 1963, as well as investigating the design of a new systematic arrangement of main classes within a new classification scheme, the citation order of components of compound subjects was also discussed. This was proposed as the basis for a ‘freely-faceted’ scheme, initially intended to provide open-ended extension capabilities for classification schemes.

Later work was funded by BNB and NATO. A general system of categories based on fundamental classes of ideas was produced. Things were distinct from Actions. Concrete things were different from Ideas. Concrete things were divided into Naturally-occurring and Artificial. Types of relationship between categories were also defined: whole/part, genus/species etc. Categories and types were to supply the semantics of the subject representation scheme. No notation was added, in order to avoid traps set by its form, e.g. decimal numbers only allowing up to ten choices.

As well, work proceeded on handling compound topics:

‘for example, a topic such as training of supervisors in Californian industries involves an *action/patient* relationship linking ‘training’ to ‘supervisors’, a whole/part relationship between ‘supervisors’ and ‘industries’ and a ‘space/location’ relationship which links ‘industries’ to California. A basic set of these *syntactical* relations was implicit in

## **New Directions in Information Organisation: submission from Alan Poulter**

Ranganathan's PMEST and this had been expanded and modified by Vickery as the sequence: Things (Products), Kinds, Parts, Materials, Properties, Operations, Agents' (Austen 1998, p31).

Using this sequence would not remove all ambiguity however. The CRG had tried to address this problem by using a set of role operators, single digit numbers in brackets, which not only determined the citation order of elements but also indicated their roles.

Also at this time the automated production of BNB was being upgraded and a project was set up to create a new indexing system for it, the existing alternatives all being ruled out. The job of generating this index was to be automated, so a system that created strings of terms for each index entry, with lead term(s) indicated and the appropriate formatting and display of other terms. Unlike the previous chain indexing system, each entry would display the full set of terms in the entry. As well as index entries, *See* and *See Also* references would also be automatically generated. Finally, unlike the old chain index system, which was bound to a classification system, the new system would use a set of role operators to identify and order concepts in an index entry and that the set of role operators and index terms used should be able represent any subject.

To achieve this novel last goal, two innovations were made. One was the development of a generic set of role operators, that were not tied to any existing scheme. They were to provide complete disambiguation of meaning in any string of indexing terms. To aid in this disambiguation, a new form for index entries was required.

Terms were ordered by the principle of context dependency, in which terms set the context for following terms. Thus in the topic "training of supervisors in Californian industries", 'California' would come first to set the location for the remaining terms. In California are located 'Industries', so this is the second term. In those industries are supervisors who are being trained, so 'Supervisors' provides the context for 'Training', the last term. So the final string of index terms would be:

California – Industries – Supervisors – Training

The above string is unambiguous, but if it shunted around to create entries for the other terms as in a KWIC index, then ambiguity re-appears, for example in:

Training - California – Industries – Supervisors

it is not clear whether the supervisors are being trained or giving the training. To solve this issue a multi-line entry format was developed, a lead term, followed by terms in a 'qualifier' and under this line of terms were the remaining terms in a 'display', for example:

California

    Training - Industries – Supervisors

Industries – California

    Supervisors - Training

Supervisors - Industries – - California

    Training

## **New Directions in Information Organisation: submission from Alan Poulter**

Training - Supervisors - California – Industries

This ‘shunting’ process produces a lead term set in its wider context (if any) by the ‘qualifier’ and given more detail by the ‘display’. To compress the index display, if different strings have the same lead and qualifier, then only their displays need to be shown, For example, suppose another string is:

Industries – California  
Technicians - Salaries

then combining its display with the previous example string would give:

Industries – California  
Supervisors - Training  
Technicians - Salaries

The driver of string creation was a set of primary operators, denoting roles and identified by numbers, the most important being:

- 0 - Location
- 1 – Key concept
- 2 – Action/Effect of action
- 3 – Performer/Agent/Instrument

There were also secondary operators, the most commonly used being ‘p’ for part or property. To code the example string would produce:

- 0 - California
- 1 – Industries
- P - Supervisors
- 2 – Training

Note that in the above string, ‘Supervisors’ are considered a part of ‘Industries’. Strings had to contain a Key Concept and an Action, else they would be rejected as being invalid. The best to build a string was to work out first the activity involved (the ‘2’ Action) and then what the target of the action was (the ‘1’ Key concept).

PRECIS was taken up by the Australian National Library and the National Film Board of Canada. It was used for back of the book indexing including the final edition of the PRECIS manual (Austen 1984) and the IFLA UNIMARC Manual (Holt 1987). The first edition of the manual has trials of PRECIS in other languages and suggests that PRECIS follows an underlying grammar (BNB 1974). This grammar is not language itself, as attempts to teach PRECIS as a grammar failed. There is some similarity between the roles in PRECIS and grammatical categories, but there are significant differences. For example, sentences have verbs, PRECIS strings contain only nouns or noun phrases. PRECIS seemed to work well in related languages like French and German as well as in different languages like Tamil and Telegu (Vencatachari,1982).

## **New Directions in Information Organisation: submission from Alan Poulter**

Austin (1998) suggests that this generality in indexing capability comes from Chomsky's theory of transformational generative grammar (1965). He posits that there is a deep structure underlying language which is understood only innately and a surface structure, which is comprehended by speakers. The same deep structure is common across languages, which accounts for their common form and functions, while their surface structures seemingly differ. People can innately understand deep grammar, which enables them to learn surface languages easily, since language acquisition and use is vital for human society. Other theorists support this approach and Longacre (1976) lists four basic elements common across different theorists: locative, agentive, instrumental and patient/object. There is an obvious similarity between these and the role operators in PRECIS. PRECIS was tested for its application across languages, and while many trials were successful, there was pressure to expand the set of role operators to address particular issues with certain languages. For example, codes to handle Komposita in German were devised but never added to the core set. However, even if extra codes for special situations with certain languages had been added to PRECIS, these would never have complicated the majority of indexing which would have used the core operators.

### **Conclusion: what future for FRSAD?**

This paper has traced the development of the FRSAD model and suggested a mechanism, based on PRECIS, for putting into practice this model. Yet there seems to be a general denial of the FRSAD model. Rather than being incorporated into RDA, at the most recent meeting (November 2011) of the RDA's JSC, its existence appears not to have been mentioned.

According to a blog post by the ALA's JSC representative (Attig 2011) there was a suggestion to:

“consider the "subject" entities [Concept, Object, Event, and Place] independent of their grouping in FRBR as Group 3 "subject" entities, but rather consider them as bibliographic entities and define whatever attributes and relationships seem appropriate to each entity. One implication of this is that entities should not be limited to the subject relationship, but considered more broadly within the context of bibliographic information. The JSC accepted this as a basis for further development and discussion.”

which could be interpreted as a rethink leading up to the recognition and incorporation of FRSAD. However, one proposal which was passed seems to completely ignore FRSAD:

“there was tentative consensus that there should be a very general definition of the subject relationship; that the Concept and Object entities should be defined in RDA; and that further discussion was needed about the Event/Time/Place entities.”

The JSC is not an organisation tied to IFLA so it is not bound to recognise IFLA standards. However it is strange that it is planning a revision of a now superseded structure. The literature review for this paper found no fundamental criticisms of FRSAD, and its gestation seems to have been open and informed by the same processes that FRBR and FRAD went through. Its lineage back to work from Buizza and Guerina, and from Tom Delsey, is clear. Yet it is almost as though FRSAD itself has never appeared. The blanks in RDA will go though. From the same blog post:

## **New Directions in Information Organisation: submission from Alan Poulter**

“The suggestion was made that we delete the "placeholder" chapters from RDA outline -- because they are so closely related to Group 3/Subject concepts -- and rethink how we wish to define and document additional entities”

FRSAD seems to have come and gone in the night: a strange case indeed!

Word count: 5648

### **References:**

Assuncao, J.B. (1989)

PRECIS em portugues: em busca uma adaptacao.

Revista da Escola Biblioteconomia da UFMG; 18 (2), p153-365

Attig, J. (2011)

Report of the Meeting of the Joint Steering Committee, 1 November 2011

[http://www.personal.psu.edu/jxa16/blogs/resource\\_description\\_and\\_access\\_ala\\_rep\\_notes/2011/11/report-of-the-meeting-of-the-joint-steering-committee-1-november-2011.html](http://www.personal.psu.edu/jxa16/blogs/resource_description_and_access_ala_rep_notes/2011/11/report-of-the-meeting-of-the-joint-steering-committee-1-november-2011.html)

Austin, D. (1974).

The development of PRECIS: a theoretical and technical history. *Journal of Documentation* 30(1): p47-102.

Austin (1982)

Basis concept classes and primitive relations. In: *Universal classification: proceedings of the Fourth International Study Conference on Classification Research, Augsburg, June 1982.*

Index-Verlag.

Austin, D. (1984)

PRECIS: a manual of concept analysis and indexing. London : British Library. p397

Austen, D. (1986)

Vocabulary control and information control. *Aslib Proceedings*. 38(1), p1-15

Austin, D. (1998)

Developing PRECIS, Preserved Context Index System. *Cataloging and Classification Quarterly*; 25 (2/3): p.23-66

British National Bibliography (1974)

PRECIS: a manual of content analysis and indexing.

Buizza, P. and Guerrini, M. A. (2002)

Conceptual model for the new "Soggettario": Subject indexing in the light of FRBR. *Cataloging & Classification Quarterly*, 34(4): 31-45.

Chomsky, N. (1965)

Aspects of the theory of syntax. MIT Press.

Delsey, T. (2005).

## **New Directions in Information Organisation: submission from Alan Poulter**

Modeling subject access: Extending the FRBR and FRANAR conceptual models.  
Cataloging & Classification Quarterly 39(3/3), p49-61

Detemple, S. (1982)  
PRECIS. Bibliothek: Forschung und Praxis; 6 (1/2), p4-46

Dykstra, M. (1978)  
The lion that squeaked. Library Journal; 103 (15) 1 Sept, p1570-1572

Holt. B.P. (1987)  
UNIMARC Manual. British Library for IFLA.

IFLA Study Group on the Functional Requirements for Bibliographic Records (2009)  
Functional Requirements for Bibliographic Records: final report.  
[http://www.ifla.org/files/cataloguing/frbr/frbr\\_2008.pdf](http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf).

IFLA Working Group on Functional Requirements and Numbering of Authority Records (2009)  
Functional Requirements for Authority Data – a conceptual model. München: K.G. Saur.

IFLA Working Group on the Functional Requirements for Subject Authority Records (2010)  
Functional Requirements for Subject Authority Data (FRSAD) : a Conceptual Model.  
<http://www.ifla.org/files/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>.

International Organization for Standardization (1986)  
ISO 27788 : Documentation: Guidelines for the establishment and development of monolingual thesauri. 2<sup>nd</sup> ed.

International Organization for Standardization (1985)  
ISO 5964 : Documentation: Guidelines for the establishment and development of multi-lingual thesauri. 2<sup>nd</sup> ed.

Longacre (1976)  
An anatomy of speech notions. Peter De Ridder Press.

Patton, G, (2005)  
FRAR: Extending FRBR Concepts to Authority Data.  
<http://archive.ifla.org/IV/ifla71/papers/014e-Patton.pdf>

Rust, G and Bide, M. (2000)  
The <indec> metadata framework: principles, model and data dictionary  
[http://www.doi.org/topics/indec/indec\\_framework\\_2000.pdf](http://www.doi.org/topics/indec/indec_framework_2000.pdf)

U.S. RDA Test Coordinating Committee (2011)  
Report and Recommendations of the U.S. RDA Test Coordinating Committee  
<http://www.loc.gov/bibliographic-future/rda/rdatesting-finalreport-20june2011.pdf>.

Vencatachari, P.N. (1982)

## **New Directions in Information Organisation: submission from Alan Poulter**

Application of PRECIS to Indian languages : a case study. In: Perspectives in library and information science /edited by S.N.Agawhal. Lucknow: Printhouse.

Zeng, M.L., Zumer, M.  
Introducing FRSAD and mapping it with SKOS and other models  
<http://www.ifla.org/files/hq/papers/ifla75/200-zeng-en.pdf>.

Zumer, M.,Salaba, A. and Zeng, M.  
Functional Requirements for Subject Authority Records (FRSAR):  
a conceptual model of aboutness

**Alan Poulter** (Lecturer, past CILIP Rep on the RDA JSC, past British Library Cataloguer/PRECIS indexer)  
Strathclyde University  
Department of Computer and Information Sciences  
Livingstone Tower  
24 Richmond Street  
Glasgow, **Scotland**, G11 1XH