



UNIMARC et *Linked Data* (ou Web de données)

Gordon Dunsire

Consultant indépendant

Edimbourg, GB

Mél. : gordon@gordondunsire.com

Mirna Willer

Université de Zadar

Département des Sciences de l'information et des bibliothèques

Zadar, Croatie

Mél. : mwiller@unizd.hr

Traduit de l'anglais par :

Françoise Leresche

Bibliothèque nationale de France

Département de l'Information bibliographique et numérique

Meeting:

187 — *Advancing UNIMARC: alignment and innovation* — IFLA UNIMARC Programme (UNIMARC)

Résumé :

Le principal objectif de cette communication est de présenter des arguments et de faire des recommandations pour représenter les formats UNIMARC pour les données bibliographiques et d'autorité en RDF (Resource Description Framework = Cadre de référence pour la description de ressources), la norme du W3C pour structurer les données dans l'environnement du Web sémantique et du Linked Data (ou Web de données). C'est une continuation du travail déjà entrepris par les groupes respectifs de l'IFLA pour représenter en RDF l'ISBD et les modèles conceptuels FRBR, FRAD et FRASAD. Les auteurs recommandent vivement que le PUC propose à l'IFLA le financement du développement de la représentation d'UNIMARC en RDF comme un projet de recherche et développement.

Introduction et arrière-plan

« Le terme *Linked Data* (ou Web de données) renvoie à un ensemble de bonnes pratiques pour publier et interconnecter des données structurées sur le Web »¹. Dans cette approche, les données sont exprimées sous la forme de déclarations simples utilisant RDF (*Resource Description Framework* = Cadre de référence pour la description de ressources) et interconnectées au moyen d'identifiants exploitables par les machines et conformes à la syntaxe définie pour les URI (*Uniform Resource Identifier* = Identifiant unique de ressource). Les déclarations RDF se présentent sous la forme d'une

structure en trois parties « sujet-prédicat-objet » où le sujet identifie à propos de quoi la déclaration est faite, le prédicat identifie l'aspect spécifique du sujet qui est décrit et l'objet identifie ou présente la valeur de cet aspect. Une déclaration RDF est de ce fait communément connue comme un « triplet ». La base d'un triplet est son prédicat qui est représenté comme une propriété RDF tandis que le sujet et l'objet particuliers d'un triplet sont représentés comme des membres de classes RDF. Les classes décrivent des choses et les propriétés décrivent les relations entre ces choses ; les classes et les propriétés sont les types d'éléments fondamentaux en RDF. La chose décrite comme une classe peut être tout type de ressource ou d'entité à propos de laquelle on veut faire une déclaration ; elle est utilisée comme le sujet d'un triplet. Les terminologies contrôlées utilisées comme objets des triplets peuvent être représentées comme des « vocabulaires de valeurs » en utilisant SKOS (*Simple Knowledge Organization System* = Système simple d'organisation des connaissances)² qui est un ensemble particulier d'éléments RDF défini pour des thésaurus et des taxonomies simples. L'objet peut aussi être représenté par une chaîne de données littérale, comme un nom de personne, une mention d'édition, etc., ne faisant pas référence à un vocabulaire ou une terminologie contrôlée.

Un triplet, ce sont essentiellement des métadonnées, c'est-à-dire des données sur des données ; dans ce cas, des données sur le sujet du triplet. Le *Linked Data* (ou Web de données) devrait donc intéresser particulièrement la communauté des bibliothèques qui a élaboré des modes d'accès aux métadonnées bibliographiques, sophistiqués et centrés sur l'utilisateur, sous la forme de catalogues régis par des normes internationalement reconnues. Une caractéristique du *Linked Data* (ou Web de données) est qu'il est à l'échelle du Web, le Web sémantique, ce qui permet le partage des données à un niveau global entre des sources multiples et hétérogènes. À nouveau, ceci devrait être d'une utilité significative pour les bibliothèques qui échangent des notices MARC (*MACHine-Readable Cataloguing* = Catalogage lisible par machine) depuis les années 1960.

Les données liées des bibliothèques, issues des notices existantes fondées sur des normes internationales, seront d'une grande qualité et seront disponibles en grande quantité, couvrant un grand nombre des ressources d'information disponibles qui sont susceptibles d'intéresser les utilisateurs du Web sémantique. À lui seul, WorldCat d'OCLC contient plus de 230 millions de notices bibliographiques³. Une analyse du contenu des notices MARC⁴ a trouvé plus de 13 millions de sous-zones dans environ 420 000 notices ; en supposant que chaque sous-zone puisse générer un triplet, cela suggère une moyenne de 31 triplets potentiels par notice. Ce chiffre n'est pas réduit par les effets de duplications à l'intérieur de WorldCat, dans la mesure où ceux-ci sont facilement compensés par les notices qui ne sont pas rassemblées dans WorldCat, ce qui indique qu'il doit y avoir au moins des milliards de triplets renfermés dans les notices existantes. D'une égale importance sont les données créées par les bibliothèques pour le contrôle des points d'accès par des notices d'autorité qui couvrent les personnes, les collectivités, les lieux, les mots-matière et d'autres choses qui sont susceptibles d'intéresser une audience plus large que les utilisateurs traditionnels des bibliothèques.

Utiliser les normes en vigueur dans les bibliothèques comme base pour de nouveaux triplets et extraire des triplets des notices existantes nécessite la représentation de ces normes en RDF, soit en créant les éléments RDF appropriés, soit en établissant des correspondances avec des éléments existants. Ceci ne permettra pas seulement au Web sémantique de bénéficier des métadonnées des bibliothèques ; ce devrait aussi améliorer l'interopérabilité entre les entités bibliographiques, attributs et relations, décrits dans des normes différentes mais apparentées. Les propriétés RDF peuvent être choisies à partir de différentes normes et mélangées à l'intérieur d'une même application pour atteindre ses fonctionnalités requises, en utilisant un profil d'application du Dublin Core⁵ ou une ontologie exprimée en RDF/OWL⁶.

L'IFLA, en tant qu'organisme de normalisation, devrait être particulièrement intéressée à entrer dans le *Linked Data* (ou Web de données) et dans l'environnement du Web sémantique car elle a pour mandat de développer et de maintenir les modèles et les normes de l'information bibliographique, et par là, de permettre à la communauté des bibliothèques de mieux servir ses utilisateurs dans un environnement technologique en mutation. En outre, en soutenant les développements conduisant à la présentation en RDF de ses normes internationalement reconnues, l'IFLA fournit une garantie d'authenticité et de fiabilité dans les métadonnées créées par les bibliothèques, ce qui est d'une importance exceptionnelle dans un environnement qui laisse « n'importe qui dire n'importe quoi sur n'importe quelle ressource », tout en promouvant en même temps sa marque au-delà des frontières de la communauté des bibliothèques. En utilisant des relations explicitement définies, « il est possible de créer un web de confiance par des calculs informatiques [Godlbeck and Parsia]. Établir un système de confiance dans le Web sémantique rendra plus facile aux ordinateurs la tâche de déterminer quelles informations proviennent d'une source autorisée et quelles autres non »⁷.

On peut faire remonter à 2006 la première initiative pour revoir les normes de l'IFLA dans le contexte des technologies et des services du Web lorsque le Groupe de révision de l'ISBD (*ISBD Review Group*) de la Section de Catalogage de l'IFLA a décidé de prendre acte de la recommandation de son Groupe d'étude sur l'indication des types de ressources (*Material Designations Study Group*) et de développer un schéma XML pour l'ISBD. Le Groupe d'étude sur l'ISBD en XML (*ISBD/XML Study Group*)⁸ a été mis en place en 2008 avec cet objectif ; toutefois, comme le travail du Groupe de révision du modèle FRBR (*FRBR Review Group*)⁹ pour mettre en relation le modèle FRBR¹⁰ avec RDF avait commencé l'année précédente, le Groupe d'étude sur l'ISBD en XML a décidé de court-circuiter un balisage général en XML et d'examiner la représentation de l'ISBD lui-même en RDF. Le projet sur trois ans du Groupe d'étude sur l'ISBD en XML est aujourd'hui dans sa phase finale et sera terminé en décembre 2011^{11, 12}. Le Groupe de révision du modèle FRBR a poursuivi son travail de représentation en RDF des modèles de l'IFLA, en l'étendant aux modèles pour les données d'autorité (FRAD)¹³ et pour les données d'autorité matière (FRSAD)¹⁴. Les représentations en RDF des trois modèles comme de l'ISBD ont été créées en utilisant l'OMR (*Open Metadata Registry* = Bureau d'enregistrement des métadonnées ouvertes)¹⁵.

Il faut toutefois mentionner que toutes ces activités ont été menées en concertation avec des recherches similaires dans le domaine faites par d'autres parties intéressées, alimentant en retour le développement des représentations en RDF des normes de l'IFLA¹. Il faut aussi noter une recherche faite pour tester la possibilité d'appliquer RDA (*Resource Description and Access* = Ressources : Description et Accès) comme norme de contenu pour UNIMARC, en complément et en correspondance avec l'ISBD, dans le contexte du Web sémantique¹⁶.

Les dernières éditions des formats UNIMARC, troisièmes éditions du format bibliographique comme du format pour les données d'autorité, ont été publiées respectivement en 2007¹⁷ et en 2009¹⁸ ; les mises à jour suivantes sont en préparation par le PUC (*Permanent UNIMARC Committee* = Comité permanent UNIMARC). Dans sa 3^e édition, le format UNIMARC pour les données d'autorité a déjà implémenté des caractéristiques particulières du modèle FRAD afin d'être plus étroitement aligné sur ce modèle¹⁹, tandis que son alignement sur le modèle FRSAD n'est toujours pas réalisé. L'alignement du format UNIMARC pour les données bibliographiques avec le modèle FRBR et aussi avec la nouvelle édition intégrée de l'ISBD²⁰, est en cours d'approbation. Il va sans dire que les formats UNIMARC suivent étroitement les autres normes de l'IFLA ; dans le cas des données bibliographiques, il s'agit plus particulièrement de l'ISBD. Ainsi, le travail pour représenter UNIMARC en RDF est, en ce qui concerne spécifiquement les données bibliographiques, une extension du

ⁱ Pour plus d'information, voir : Dunsire, G ; M. Willer. *Ibidem*.

travail du Groupe d'étude sur l'ISBD en XML qui constitue le point de départ de cette communication.

Des espaces de noms pour UNIMARC

RDF exige que les classes et les propriétés soient dotées d'identifiants exploitables par les machines et conformes à la syntaxe définie pour les URI (*Uniform Resource Identifier* = Identifiant unique de ressource)²¹. Un ensemble d'URI donnant les informations essentielles sur les classes et les propriétés correspondantes, publié et géré dans un contexte particulier, est connu comme un espace de noms. Tous les URI d'un espace de noms seront généralement construits à partir d'une chaîne de caractères commune, désignée comme le domaine de base, à laquelle est ajoutée une chaîne de caractères distinctive, désignée comme la partie locale. Un avantage de cette méthode réside dans le fait que le domaine de base peut être abrégé pour raccourcir les URI pour l'affichage à des humains ; il est développé automatiquement pour le traitement par les machines. Les éléments et les vocabulaires d'UNIMARC en RDF seront initialement créés et maintenus en utilisant l'OMR, selon la même démarche que celle utilisée pour l'ISBD et pour la famille des modèles de métadonnées sur les Fonctionnalités requises (FRBR, FRAD et FRSD). En particulier, l'OMR supporte des intitulés et d'autres annotations dans plusieurs langues. C'est une exigence importante pour les normes de l'IFLA qui sont conçues pour une application dans un environnement international et prévues pour une interopérabilité multilingue. UNIMARC a été traduit de l'anglais en chinois, croate, français, italien, lituanien, portugais, russe, etc. RDF est fondamentalement neutre vis-à-vis des langues car il est conçu pour le traitement par les machines, mais il permet que des intitulés, définitions, notes d'application et autres annotations en différentes langues soient associés au même élément. L'infrastructure à plus long terme exigée pour gérer les espaces de noms de l'IFLA sera étudiée et développée par le Groupe de travail spécial sur les espaces de noms de l'IFLA (*IFLA Namespaces Task Group*), selon les recommandations faites dans le rapport établi par ce Groupe²².

Identifier les éléments d'UNIMARC par étiquette/sous-zone/position codée

Dans cette communication, des codes abrégés sont utilisés pour identifier les éléments d'UNIMARC définis par des étiquettes (zones), des sous-zones, des indicateurs et des positions codées dans le Bloc des informations codées 1XX, en utilisant le système suivant :

Étiquette + indicateur 1 + indicateur 2 + code de sous-zone ; « b » est utilisé pour désigner un indicateur à blanc et remplace la valeur dièse (#) signalant un espace. Le premier caractère du code de sous-zone (\$) n'est pas nécessaire car chaque code abrégé s'applique à une seule sous-zone et la sixième position y indique le code de sous-zone.

Exemples :

010bba = Numéro (ISBN)

2001ba = Titre propre (le titre est significatif)

2000ba = Titre propre (le titre n'est pas significatif)

Pour le Bloc des informations codées, la position codée est ajoutée pour former le code abrégé.

Exemples :

100bba8 = Type de date de publication

100bba17-19 = Type de public

100bba34-35 = Écriture du titre

Domaines d'espaces de noms

Un ou plusieurs espaces de noms avec les domaines de base correspondants seront nécessaires pour représenter les éléments et vocabulaires d'UNIMARC en RDF.

Réutilisation d'espaces de noms existants

C'est une bonne pratique de réutiliser des éléments et vocabulaires RDF provenant d'espaces de noms existants, lorsque c'est approprié : cela épargne du temps et évite l'effort de développer les éléments et les vocabulaires, puis de les maintenir ; c'est plus simple pour développer des applications et des services à partir des métadonnées ; cela encourage une approche combinée des applications ; cela favorise le web d'éléments interconnectés et de données liées. Il est toutefois extrêmement important de garantir que les éléments réutilisés sont rigoureusement associés à la norme qui les utilise, de manière à ce que tout changement dans leur signification directe ou indirecte (voisinage sémantique) soit immédiatement reflété dans l'espace de noms correspondant afin de prévenir des « dérives » sémantiques entre les deux espaces de noms.

Le format UNIMARC pour les données bibliographiques est aligné sur l'ISBD, lequel a déjà un espace de noms publié pour son ensemble d'éléments et ses vocabulaires (Zone 0 – Zone du contenu et du type de médiation). Il y a donc un choix à faire pour UNIMARC :

1. Réutiliser les classes et propriétés de l'ISBD partout où c'est approprié, au lieu d'en créer des distinctes dans l'espace de noms UNIMARC Bibliographique. Cette option n'est appropriée que si les changements proposés pour l'une ou l'autre norme prennent en considération l'impact sur l'autre norme ; c'est-à-dire, si les deux normes sont gérées et maintenues comme « une seule ».
2. Représenter tous les éléments du format UNIMARC pour les données bibliographiques dans un espace de noms UNIMARC spécifique et établir des liens vers les classes et propriétés équivalentes dans l'espace de noms de l'ISBD. C'est cette option qui doit être choisie si l'ISBD et l'UNIMARC continuent à être développés séparément, même s'il existe une liaison étroite entre eux.

Ce choix doit être fait avant d'engager tout développement substantiel d'un espace de noms pour le format UNIMARC pour les données bibliographiques.

Le Tableau 1 montre un exemple de mise en correspondance entre les propriétés potentielles d'UNIMARC/B et les propriétés existantes de l'ISBD. Il convient de noter que le domaine de l'espace de noms des URI pour chacune des propriétés d'UNIMARC et de l'ISBD n'est pas inclus par souci de concision.

UNIMARC	Intitulé en anglais	Intitulé en français	Propriété de l'ISBD	Intitulé en anglais	Intitulé en français
P205bba	has edition statement	a pour mention d'édition	P1008	has edition statement	a pour mention d'édition
P205bbb	has issue statement*	a pour autre mention d'édition	P1011	has additional edition statement	a pour autre mention d'édition
P205bbd	has parallel edition statement	a pour mention parallèle d'édition	P1009	has parallel edition statement	a pour mention parallèle d'édition
P205bbf	has statement of responsibility relating to edition	a pour mention de responsabilité relative à l'édition	P1010	has statement of responsibility relating to edition	a pour mention de responsabilité relative à l'édition
P205bbg	has subsequent statement of responsibility**	a pour mention de responsabilité suivante *	P1010	has statement of responsibility relating to edition	a pour mention de responsabilité relative à l'édition

Tableau 1. Exemple de mise en correspondance entre les propriétés d'UNIMARC/B et les propriétés existantes de l'ISBD en RDF : Zone 205 du format UNIMARC Bibliographique (Mention d'édition)

* Des différences dans les intitulés, par exemple pour P205bbb, peuvent être gérées en utilisant la propriété de SKOS pour une variante d'intitulé. C'est ce qui a été fait pour certaines propriétés FRBR qui ont des intitulés différents dans FRAD.

** Voici un exemple où l'élément UNIMARC est plus spécifique que celui de l'ISBD. Pour conserver l'élément UNIMARC et sa granularité plus fine, il est nécessaire de déclarer une propriété UNIMARC dans l'espace de noms UNIMARC Bibliographique.

Comme cela a déjà été mentionné, le format UNIMARC pour les données d'autorité « prend en compte les attributs des entités et les relations entre entités tels que les définit le modèle FRAD (*Functional Requirements for Authority Data: a Conceptual Model* = Fonctionnalités requises des données d'autorité : un modèle conceptuel) »ⁱⁱ, pour les aspects suivants : « changement de vocabulaire, définition des zones, évolution de la sous-zone de contrôle \$5 (Données codées sur les relations) [...]. Les blocs ont été renommés en 2XX Point d'accès autorisé, 4XX Variante de point d'accès, 5XX Point d'accès en relation et 7XX Point d'accès autorisé dans une autre langue et/ou écriture, tandis que les étiquettes désignent les noms des entités représentées par les points d'accès contrôlés, telles que Nom de personne, Nom de collectivité, Titre »ⁱⁱⁱ. Toutefois, FRAD est un modèle, alors qu'UNIMARC/A est un véhicule de contenu au niveau applicatif, aussi l'équivalence entre les définitions de FRAD et celles d'UNIMARC/A devra être vérifiée. En outre, contrairement à l'ISBD, FRAD est une extension du modèle FRBR et lui-même réutilise les éléments appropriés de l'espace de noms du modèle FRBR. Il faut aussi noter qu'il existe un alignement intrinsèque entre les formats UNIMARC pour les données bibliographiques et pour les données d'autorité, ce qui signifie qu'UNIMARC/A répercute les changements et les additions apportées au format bibliographique. La correspondance entre UNIMARC/A et UNIMARC/B est exprimée par la sous-zone \$3 dans les zones de points d'accès en UNIMARC/B et expressément expliquée dans le Guide d'utilisation du format

ⁱⁱ UNIMARC Manual. Authorities Format. *Ibidem*, p. 14

ⁱⁱⁱ Willer, M. Foreword to the third edition. *Ibidem*, p. 8

UNIMARC/A. En conséquence, il sera nécessaire d'étudier comment positionner UNIMARC/A par rapport au bloc 7XX Bloc des responsabilités d'UNIMARC/B d'une part et à FRAD d'autre part pour déterminer des relations entre éléments pour des instances particulières de données liées. Par exemple, si à un point d'accès /élément d'entrée particulier dans une notice en UNIMARC/B correspond une notice d'autorité en UNIMARC/A qui a aussi été publiée en RDF, une chaîne de données liées peut alors être établie entre elles. Sinon, le point d'accès en UNIMARC/B, par exemple 700 Nom de personne – Responsabilité principale, utilisé sans renvoyer à la zone correspondante d'une notice d'autorité en UNIMARC/A par l'utilisation de la sous-zone \$3 peut être représenté comme une chaîne littérale.

Des espaces de noms particuliers pour UNIMARC

L'ISBD ne couvre pas les points d'accès (ou vedettes) et les notices d'autorité qui leur correspondent, aussi le format UNIMARC pour les données d'autorité n'aura pas de correspondance avec les classes ou propriétés de l'ISBD. Un espace de noms est absolument nécessaire pour le format UNIMARC pour les données d'autorité.

L'exemple présenté dans le Tableau 1 montre que les éléments de l'ISBD ne présentent pas le même niveau de granularité que le format UNIMARC pour les données bibliographiques ; aussi un espace de noms est-il nécessaire pour le format UNIMARC pour les données bibliographiques, afin de rendre compte des éléments qui ne sont pas couverts par l'ISBD, sans tenir compte de la réutilisation des éléments de l'ISBD. Les formats UNIMARC pour les données bibliographiques et pour les données d'autorité doivent avoir des espaces de noms distincts pour refléter la publication séparée des deux textes et distinguer entre les encodages utilisant la même zone/sous-zone.

On propose que les espaces de noms pour UNIMARC suivent le schéma déjà établi par le Groupe de travail spécial sur les espaces de noms de l'IFLA (*IFLA Namespaces Task Group*) pour l'ISBD ainsi que pour les modèles FRBR, FRAD et FRASD.

Pour les éléments du format UNIMARC pour les données d'autorité, UNIMARC/A, le domaine de base de l'espace de noms est :

<http://iflstandards.info/ns/unimarc/unimarca/elements/>

Il peut être abrégé en « unimarca » pour l'affichage. Il faut noter que ce n'est pas une URL ; c'est un URI pérenne (« cool » URI).

Pour les éléments du format UNIMARC pour les données bibliographiques, UNIMARC/B, le domaine de base de l'espace de noms est :

<http://iflstandards.info/ns/unimarc/unimarcb/elements/>

Il peut être abrégé en « unimarcb ».

Pour les vocabulaires UNIMARC (ou listes de valeurs), un nom de domaine distinct est utilisé pour chaque vocabulaire, selon la pratique utilisée pour l'ISBD. Le domaine de base est composé par une chaîne de caractères identifiant un vocabulaire particulier qui est ajoutée à un nom de domaine général pour les vocabulaires UNIMARC :

<http://iflstandards.info/ns/unimarc/terms/> + chaîne de caractères identifiant le vocabulaire

La chaîne de caractères identifiant un vocabulaire particulier ne peut pas être fondée sur le code composé par la chaîne étiquette/indicateurs/sous-zone/position codée, car certains vocabulaires, comme celui pour les Écritures, sont utilisés dans plusieurs zones à l'intérieur des formats UNIMARC/A et UNIMARC/B :

UNIMARC/A: 100bba21-22

UNIMARC/B: 100bba34-35

En revanche, on peut utiliser une abréviation du titre du vocabulaire :

Exemple :

<http://iflstandards.info/ns/unimarc/terms/graphicssmd>

C'est l'espace de noms défini pour le vocabulaire correspondant à l'indication spécifique du type de ressource pour les ressources graphiques, utilisé en 116bba1.

Cet exemple peut être abrégé en « unimarcgsmd » pour l'affichage.

Il faut noter qu'il n'y a nul besoin d'utiliser le nom de domaine pour indiquer si le vocabulaire provient du format pour les données bibliographiques ou du format pour les données d'autorité. Les vocabulaires peuvent être utilisés comme l'objet de toute propriété appropriée utilisée dans un triplet et l'utilisation du vocabulaire sera décrite dans un profil d'application comme un schéma d'encodage du vocabulaire associé aux propriétés RDF pertinentes. Cette méthode détache un vocabulaire de son utilisation particulière en UNIMARC et rend plus aisée sa réutilisation par d'autres communautés dans des applications non-UNIMARC.

Profils d'application

Un ou plusieurs profils d'application du Dublin Core seront nécessaires pour UNIMARC, afin de définir la réutilisation, s'il y a lieu, des classes et propriétés de l'ISBD et de FRAD, l'utilisation de déclarations agrégées composées de deux propriétés ou plus (comme en ISBD), l'utilisation de vocabulaires précis comme schémas d'encodage du vocabulaire et toute autre contrainte sur l'utilisation des éléments dans une notice UNIMARC (notice bibliographique ou notice d'autorité) bien formée, comme le caractère obligatoire ou non, répétable ou non, d'un élément.

Les deux formats, UNIMARC/B comme UNIMARC/A, définissent des éléments obligatoires, ainsi que des éléments répétables et non-répétables, dans les notices UNIMARC. Les zones obligatoires dans les deux formats sont les zones 001 Identifiant de la notice, 100 Données générales de traitement (seulement certains éléments de données, identiques dans les deux formats) et 801 Source de catalogage, tandis que les zones obligatoires propres à un format sont en UNIMARC/B, 200\$a Titre propre (plus certaines zones particulières au type de ressource), et en UNIMARC/A, le point d'accès autorisé du bloc 2XX. Les deux formats définissent des éléments répétables et non répétables au niveau des zones et des codes de sous-zones. Ainsi, en UNIMARC/B, la zone 010 ISBN est répétable, tandis que la sous-zone 010\$a Numéro (ISBN) ne l'est pas. En UNIMARC/A, la zone 220 Point d'accès autorisé – Nom de famille est répétable, mais seulement pour des formes dans des écritures différentes, tandis que la sous-zone 220\$a Élément d'entrée ne l'est pas. Dans une notice UNIMARC, l'ordre des codes de sous-zones n'est pas précisé, dans la mesure où l'ordre est déterminé par les données.

Méta-métadonnées

Les données sur une notice UNIMARC particulière sont contenues dans le Label de notice et dans le bloc 1XX Bloc des informations codées. Ce sont des méta-métadonnées, c'est-à-dire des données sur des métadonnées. En RDF, il y a un certain nombre de techniques qui peuvent être utilisées pour représenter ces données, comme un qualificatif de langue qui peut être ajouté à une chaîne littérale,

par exemple un titre, utilisée comme l'objet d'un triplet ; par exemple, « @en » indique que la chaîne de caractères est en anglais, « @fr » qu'elle est en français, etc. Ces techniques n'exigent pas des éléments UNIMARC particuliers et ne seront donc pas davantage abordées dans cette communication.

Il y a aussi des méta-métadonnées qui sont propres à une notice UNIMARC en tant qu'une instance du format d'échange structuré de notices défini par la norme ISO 2709, comme la longueur de la notice, les codes d'application, la longueur de l'indicateur, etc. Ces éléments ne sont pas pertinents quand les métadonnées sont exprimées sous la forme de triplets RDF.

Vocabulaires du bloc des informations codées

Pour les codes, avec les valeurs et les définitions qui leur sont associées, qui sont utilisés dans le bloc 1XX Bloc des informations codées pour décrire une ressource (plutôt que la notice UNIMARC), la meilleure représentation est celle d'un vocabulaire SKOS, comme cela a été fait pour les vocabulaires de la zone 0 de l'ISBD.

Le code UNIMARC pour un terme du vocabulaire peut être utilisé comme la partie locale de son URI.

Exemple :

`http://iflastandards.info/ns/unimarc/terms/graphicsmd#a`

Cet URI désigne le terme « collage » dans le vocabulaire correspondant à l'indication spécifique du type de ressource pour les ressources graphiques

Quand le code UNIMARC pour un terme du vocabulaire est un chiffre, il doit être préfixé avec une lettre comme « T » (pour terme), afin d'éviter des problèmes en XML avec les parties locales qui commencent par un caractère numérique. Ceci suit la règle appliquée pour l'ISBD.

Traiter ainsi le code UNIMARC pour un terme du vocabulaire préserve l'indépendance de l'URI par rapport au langage utilisé, évite de surcharger les URI avec de la sémantique et évite aussi toute confusion si le terme (anglais) est modifié dans le futur (qu'il passe de « collage » à « sculpture en deux dimensions utilisant plusieurs matériaux »).

Le code UNIMARC lui-même peut être explicitement représenté en utilisant la propriété `skos:notation`. Dans les exemples qui suivent, les triplets utilisant cette propriété ont comme sujet l'URI pour un terme d'un vocabulaire et le code associé comme valeur de l'objet.

Exemples :

`<http://iflastandards.info/ns/unimarc/terms/graphicsmd#a> skos:notation "a".`

(ou en utilisant la forme abrégée de l'espace de noms : `unimarcgsmd#a skos:notation "a".`)

`<http://iflastandards.info/ns/unimarc/terms/publicationdatatype#f> skos:notation "f".`

`<http://iflastandards.info/ns/unimarc/terms/titlescript#ca> skos:notation "ca".`

De même, le terme lui-même peut être représenté en utilisant la propriété `skos:prefLabel` avec un qualificatif de langue.

Exemples :

`<http://iflastandards.info/ns/unimarc/terms/graphicsmd#a> skos:prefLabel "collage"@en.`

`<http://iflastandards.info/ns/unimarc/terms/graphicsmd#a> skos:prefLabel "collage"@fr.`

<<http://iflastandards.info/ns/unimarc/terms/publicationdatetype#f>> skos:prefLabel
"monograph, date of publication uncertain"@en.

<<http://iflastandards.info/ns/unimarc/terms/publicationdatetype#f>> skos:prefLabel
"monographie, date de publication incertain"@fr.

<<http://iflastandards.info/ns/unimarc/terms/titlescript#ca>> skos:prefLabel "Cyrillic"@en.

<<http://iflastandards.info/ns/unimarc/terms/titlescript#ca>> skos:prefLabel "Cyrillique"@fr.

Le Tableau 2 donne un exemple complet d'un vocabulaire du bloc 1XX Bloc des informations codées avec les intitulés en anglais, en italien et en portugais pris dans les traductions officielles d'UNIMARC.

N	PrefLabel@en	PrefLabel@fr	PrefLabel@it	PrefLabel@pt	Definition@en	Definition@fr	ScopeNote@en	ScopeNote@fr
a	collage	collage	collage	colagem	An original work created by affixing various materials (paper, wood, newspaper, cloth, etc.) to a surface.	Œuvre originale créée en fixant des matériaux variés (papier, bois, journal, tissu, etc.) sur une surface.		
b	drawing	dessin	disegno	desenho	An original visual representation (other than a print or painting) made with pencil, pen, chalk, or other writing instrument on paper or similar non-rigid support.	Représentation visuelle originale (qui n'est ni une estampe ni une peinture) réalisée au moyen d'un crayon, d'une plume, d'une craie ou de tout autre instrument permettant d'écrire sur du papier ou un support souple similaire.		
c	painting	peinture	pittura	pintura	An original visual representation produced by applying paint to a surface.	Représentation visuelle originale créée en appliquant de la peinture sur une surface.		

d	Photomechanical reproduction	reproduction photo-mécanique	riproduzione fotomeccanica	reprodução fotomecânica	Any picture produced in imitation of another picture through the use of a photographic process to transfer the image to a printing surface.	Toute image reproduisant exactement une autre image par l'utilisation d'un procédé photographique permettant de transférer cette image sur une surface d'impression.	Hence, a snapshot made to document a painting or a Xerox copy of a print are considered photomechanical reproductions. Art reproductions, postcards, posters, and study prints are included here.	Par conséquent, un instantané pris pour documenter une peinture ou la photocopie d'une estampe sont considérés comme des reproductions photomécaniques. Les reproductions d'art, les cartes postales, les affiches et les images à usage pédagogique sont incluses ici.
e	photonegative	négatif photographique	fotonegativo	negativo fotográfico	A piece of film, a glass plate, or paper on which appears a "negative" image, i.e. directly opposite to a "positive" image (photoprint), slide, or transparency. Used to produce a positive print.	Partie de film, plaque de verre, ou papier sur lequel apparaît une image « négative », c'est-à-dire directement inversée, d'une image « positive » (épreuve photographique), d'une diapositive ou d'un transparent. Utilisé pour produire une épreuve positive.	Does not include negative photoprints, photoprints that are a combination of negative and positive images, photographs or solarized prints, all of which are considered to be techniques used when making photoprints.	Ce code ne couvre pas les épreuves photographiques négatives, celles qui combinent des images négatives et positives, les photographies ou les tirages solarisés ni tout ce qui relève des techniques utilisées pour réaliser des épreuves photographiques.

f	photoprint	épreuve photographique	riproduzione fotografica	positivo fotográfico	A positive image made either directly or indirectly on a sensitised surface by the action of light or other radiant energy.	Image positive produite soit directement soit indirectement sur une surface sensible par l'action de la lumière ou de toute autre énergie rayonnante.	The term "photoprint" is used here as a more precise term than "photograph", which technically can cover both the print and the negative. Radiographs and opaque stereographs are included here.	Le terme « épreuve photographique » est utilisé ici comme étant plus précis que « photographie », qui techniquement peut couvrir à la fois le tirage et le film négatif. Les radiographies et les vues stéréoscopiques sont incluses ici.
h	picture	image	immagine	imagem	A twodimensional visual representation accessible to the naked eye and generally on an opaque backing.	Représentation visuelle à deux dimensions visible à l'œil nu et généralement placée sur un fond opaque.	This term is used when a more specific designation is unknown or not desired.	Ce terme est utilisé lorsqu'il n'existe pas de désignation plus précise ou lorsque celle-ci n'est pas nécessaire.
i	print	estampe	stampa	gravura	A design or picture transferred from an engraved plate, wood block, lithographic stone, or other medium.	Tracé ou image que l'on réalise à partir d'une plaque gravée, d'un bloc de bois, d'une pierre lithographique ou tout autre support intermédiaire.	Generally, there are four types: planographic print, relief print, intaglio print, and stencil print.	On distingue en général quatre types d'estampes : à plat, en relief, en creux (taille-douce) et au pochoir.

k	technical drawing	dessin technique	disegno tecnico	desenho técnico	A cross section, detail, diagram, elevation, perspective, plan, working plan, etc., made for use in an engineering or other technical context.	Coupe transversale, détail, diagramme, élévation, perspective, plan, plan de travail, etc., utilisé en ingénierie ou dans tout autre contexte technique.		
m	master	matrice	master	matriz	Any plate, mould, matrix, die etc. which allows the reproduction of the same impression.	Plaque, moule, matrice, cliché etc. qui permet la reproduction d'une même impression.		
z	other nonprojected graphic type	autre type de ressource graphique non-projetée	altro tipo di documento grafico non proiettabile	outro material gráfico não-projectável	Other types not included in the above. [Types other than collage, drawing, painting, photomechanical reproduction, photonegative, photoprint, picture, print, technical drawing, master.]	Tout autre type de ressource non spécifié ci-dessus. [Type de ressource autre que collage, dessin, peinture, reproduction photomécanique, négatif photographique, épreuve photographique, image, estampe, dessin technique, matrice.]	Includes mixed media productions made by a combination of freehand and printing techniques when one or the other does not predominate. In some cases, where mixed media are applied, one must decide whether the creator intends the item to be a photoprint (even though it is painted	Cela comprend les productions qui combinent des techniques manuelles avec des techniques d'impression sans qu'aucune ne prédomine. Dans certains cas, lorsque le créateur utilise plusieurs techniques, il faut déterminer si son intention était de réaliser une épreuve photographique

							<p>over the photographic image). Hand colouring is considered a technique applied to a printing process; this aspect is covered by a character position 3. Computer produced graphics and the various duplication masters (including spirit masters and transparency masters) are included here.</p>	<p>(même s'il a peint ensuite sur l'image photographique). Le coloriage à la main est considéré comme une technique appliquée à une procédure d'impression ; cet aspect est exprimé par un code dans la position 3. Les images graphiques produites par ordinateur sont incluses ici, ainsi que les différentes matrices pour la duplication (y compris les matrices à alcool et les matrices de transparents).</p>
--	--	--	--	--	--	--	--	---

Tableau 2 : Exemple complet d'un vocabulaire du bloc 1XX Bloc des informations codées : Vocabulaire de la position codée 116bba0 = Données codées pour les ressources graphiques : Indication spécifique du type de ressource
(colonne N = Notation ; c'est aussi la partie locale de l'URI)

Les termes, définitions et notes d'application sont extraits des textes suivants, utilisés comme sources :

@en : <http://archive.ifla.org/VI/8/unimarc-concise-bibliographic-format-2008.pdf>

@it : <http://unimarc-it.wikidot.com/116>

@pt : http://www.ifla.org/files/uca/Unimarc_bib_3%C2%AAed_abrev.pdf

NdT : Les termes, définitions et notes d'application en français ont été ajoutés à partir de la traduction officielle en français de la zone 116 du format UNIMARC/B, disponible en ligne :

@fr : http://multimedia.bnf.fr/unimarc_b_trad/B116-6-2010.pdf

Les définitions et notes d'application comportent des marques de correction pour montrer comment elles dérivent du texte anglais (et du texte français).

Utilisation de vocabulaires SKOS externes

Certains ensembles de valeurs codées en UNIMARC sont explicitement fondés sur un vocabulaire ou une terminologie externe. Par exemple, la langue de l'incipit dans la sous-zone 036bbz et la langue de la ressource à plusieurs endroits de la zone 101 utilise un code à trois lettres, pris dans l'Annexe A et qui est le même que celui de la Liste MARC pour les langues (*MARC List for Languages*). Cette liste est disponible comme un vocabulaire SKOS²³, qui peut être utilisé directement dans tout triplet UNIMARC. De même, les codes d'aires géographiques pour la sous-zone 660bba sont aussi disponibles en SKOS²⁴.

Le pays de publication dans la sous-zone 102\$a utilise un code à deux lettres, pris dans l'Annexe B et qui correspond à la norme ISO 3166-1. Ce n'est pas le même que la Liste MARC pour les pays (*MARC List for Countries*) que la Bibliothèque du Congrès a également publié comme un vocabulaire SKOS²⁵. Toutefois, une représentation en RDF de la norme ISO 3166-1 est disponible²⁶, même si un examen plus approfondi de sa pertinence pour UNIMARC est nécessaire.

La disponibilité de représentations en SKOS d'autres vocabulaires externes utilisés en UNIMARC doit être vérifiée et validée. Si aucune traduction en SKOS ne peut être trouvée, le Comité permanent UNIMARC devrait prendre contact avec le propriétaire du vocabulaire pour discuter le développement d'une représentation appropriée en RDF.

Certains vocabulaires internes peuvent déjà avoir des représentations appropriées en SKOS, même s'ils ne sont pas explicitement fondés sur un vocabulaire ou une terminologie externe. On peut en trouver un exemple avec les codes utilisés dans les positions codées 100bba26-29 (Jeux de caractères utilisés). En effet, une autre communauté peut avoir développé une représentation en SKOS d'un vocabulaire similaire qui contient tous les termes et codes utilisés en UNIMARC ; si tel est le cas, ces URI en SKOS peuvent être réutilisées par UNIMARC. Ceci aussi nécessite un examen plus approfondi et des vérifications, par exemple en utilisant un moteur de recherche d'ontologies tel que Swoogle.

Valeurs de date

Les éléments codés qui sont des dates dans un format défini, comme l'année, peuvent être représentés en utilisant la propriété `rdfs:range`. Par exemple, les dates de publication 1 et 2 dans la zone 100 Données générales de traitement du format UNIMARC pour les données bibliographiques peuvent être exprimées sous la forme de triplets (en suivant la convention pour les URI donnée plus bas) :

```
unimarc:100bba09-12 rdfs:range xsd:gYear.
```

```
unimarc:100bba13-16 rdfs:range xsd:gYear.
```

Classes UNIMARC

Comme pour l'ISBD, il n'y a qu'une seule classe à prendre en compte pour le format UNIMARC pour les données bibliographiques, à l'exception des classes pour les schémas d'encodage de la syntaxe qui sont nécessaires pour les éléments agrégés, comme noté plus bas. Il s'agit de la classe ISBD Ressource qui peut être utilisée comme le domaine de définition pour toutes les propriétés RDF du format UNIMARC pour les données bibliographiques. Il n'y a pas lieu de créer une classe UNIMARC pour Ressource.

Les classes pour le format UNIMARC pour les données d'autorité nécessitent un examen plus approfondi, surtout en relation avec FRAD/FRSAD. L'espace de noms de FRAD a des classes pour toutes les entités de FRAD, comme Entité bibliographique, Nom, Identifiant, Point d'accès contrôlé, Règles et Agence et des sous-classes pour Nom d'une personne, Nom d'une collectivité, Nom d'une famille et Nom d'une œuvre ; il a aussi une sous-classe pour Collectivité parce que sa définition a été modifiée par rapport à celle du modèle FRBR et une classe pour Famille qui n'est pas définie comme une entité dans le modèle FRBR tel que publié. L'espace de noms de FRAD ne comprend toutefois pas de classes pour d'autres entités telles que Personne, Œuvre, Expression, Manifestation, etc., car elles sont déjà publiées dans l'espace de noms du modèle FRBR. À l'analyse, il devient évident que les classes FRBR ne répondent pas à tous les types de classes possibles d'UNIMARC/A : Personne, Collectivité, Œuvre peuvent globalement être considérées comme alignées, mais il y a des cas où certains types d'entités d'UNIMARC/A ou candidats à devenir des classes requièrent une analyse particulière. On en a un exemple avec Lieu qui est simplement défini dans le modèle FRBR comme « des coordonnées topographiques » ; il peut être aligné avec le Nom de territoire ou nom géographique d'UNIMARC/A, mais pas vraiment avec Accès par lieu – cette zone de point d'accès avait été conçue à l'origine pour enregistrer le lieu (pays et ville) d'impression des livres anciens, mais elle a été considérablement étendue pour couvrir les Lieu et date de publication, de représentation ou d'enregistrement, etc. ; en outre, comme UNIMARC/A est un format intégré pour les données d'autorité qui couvre à la fois les points d'accès par nom et l'indexation-matière, la classe « Lieu » doit aussi être examinée dans le contexte de FRSAD. Un autre exemple est Œuvre/Expression : parmi ses types d'entités, UNIMARC/A définit les entités suivantes : titre, rubrique de classement, nom/titre et nom/rubrique de classement, mais dans sa 3^e édition, il ne distingue pas si le type d'entité est l'Œuvre ou l'Expression^{iv}. FRAD a une sous-classe pour le Nom d'une œuvre, mais si UNIMARC définit un nouveau type d'entité pour l'Expression, il devrait ajouter à son espace de noms une sous-classe UNIMARC pour le Nom d'une expression. Des classes qui peuvent être considérées hors du champ d'application de FRBR et FRAD sont les types d'entités d'UNIMARC/A Marque et Forme, genre ou caractéristiques physiques que FRSAD avait exclus volontairement de son étude.

Propriétés UNIMARC

En général, les éléments étiquette/indicateur/sous-zone d'UNIMARC seront représentés comme des propriétés RDF, comme cela a été fait pour l'ISBD.

Tous les éléments UNIMARC ne sont pas propices ou appropriés à une représentation sous la forme de propriétés RDF. Cela inclut les éléments de méta-métadonnées qui ont été traités plus haut. Toutefois, d'autres données dans le Label de notice, telles que le type de notice (dans les deux formats) sont en fait des métadonnées sur la ressource ; bien que la plupart de ces informations doivent se retrouver dans le corps de la notice, ce n'est pas toujours le cas et certains éléments du Label de notice nécessiteront une représentation en RDF. Parmi ces éléments, on peut identifier, par exemple, le niveau bibliographique et le niveau hiérarchique en UNIMARC/B et le type d'entité en UNIMARC/A. Cela exige une plus ample instruction.

Les codes abrégés étiquette/indicateur/sous-zone utilisés dans cette communication peuvent former la partie locale de l'URI. Les URI qui utilisent une barre oblique comme dernier délimiteur hiérarchique (« *Slash* » URI) sont recommandées de préférence aux URI qui utilisent un dièse (« *Hash* » URI) lorsqu'il y a un grand nombre de propriétés. La partie locale doit être préfixée avec une lettre pour éviter des problèmes en XML avec les parties locales qui commencent par un

^{iv} Un mécanisme particulier pour distinguer entre l'Œuvre ou l'Expression est en cours d'approbation par le PUC et sera publié dans la prochaine mise à jour des formats UNIMARC.

caractère numérique ; l'ISBD utilise la lettre « P » (pour propriété) et UNIMARC peut suivre cette convention. Par exemple, l'URI pour l'élément Mention d'édition pourrait être :

`http://iflastandards.info/ns/unimarc/unimarc/elements/P205bba`

(ou, en utilisant la forme abrégée de l'espace de noms : `unimarc:P205bba`)

Cette façon de faire est indépendante de la langue, évite de surcharger les URI avec de la sémantique et évite aussi toute confusion si le libellé associé à l'étiquette, à l'indicateur ou au code de sous-zone change.

Le libellé lui-même peut être représenté en utilisant la propriété `rdfs:label` et un qualificatif de langue en RDF. Ce libellé peut exiger de faire une synthèse des libellés associés respectivement à l'étiquette, à l'indicateur et/ou à la sous-zone. Selon la pratique suivie pour l'ISBD, les intitulés des propriétés RDF peuvent être exprimés sous une forme verbale en ajoutant « a pour » devant le libellé de l'élément en UNIMARC

Exemple :

`unimarc:P205bba rdfs:label "has edition statement"@en.`

`unimarc:P205bba rdfs:label "a pour mention d'édition"@fr.`

L'OMR exige un nom d'enregistrement différent utilisant la propriété `reg:name` ; celui-ci peut être construit selon la méthode habituelle, comme une version écrite en « casse de chameau » (« *CamelCase* ») du `rdfs:label`.

Exemple :

`unimarc:P205bba reg:name "hasEditionStatement".`

Le Tableau 3 donne un exemple complet des propriétés RDF issues d'une seule étiquette sans indicateurs définis.

URI	Label@en	Label@fr	Name
P205bba	has edition statement	a pour mention d'édition	hasEditionStatement
P205bbb	has issue statement	a pour autre mention d'édition	hasIssueStatement
P205bbd	has parallel edition statement	a pour mention parallèle d'édition	hasParallelEditionStatement
P205bbf	has statement of responsibility relating to edition	a pour mention de responsabilité relative à l'édition	hasStatementOfResponsibilityRelatingToEdition
P205bbg	has subsequent statement of responsibility	a pour mention de responsabilité suivante	hasSubsequentStatementOfResponsibility

Tableau 3 : Exemple complet de propriétés RDF représentant une zone UNIMARC dont les indicateurs ne sont pas définis : Mention d'édition (zone 205 d'UNIMARC/B)

Comme cela a déjà été expliqué, chaque combinaison différente des indicateurs et d'une sous-zone à l'intérieur d'une étiquette constitue potentiellement une propriété RDF distincte, avec un intitulé approprié distinct. Une méthode pour y parvenir consiste à qualifier le libellé de la sous-zone avec le « libellé » de l'indicateur (i.e. le sens de la valeur de l'indicateur), comme le montre le Tableau 4.

URI	Label@en	Label@fr	Name
P206bba	has mathematical data statement (unstructured)	a pour mention des données mathématiques (non structurée)	hasMathematicalDataStatementUnstructured
P206bbb	has statement of scale (unstructured)	a pour mention de l'échelle (non structurée)	hasStatementOfScaleUnstructured
P206bbc	has statement of projection (unstructured)	a pour mention de la projection (non structurée)	hasStatementOfProjectionUnstructured
P206bbd	has statement of coordinates (unstructured)	a pour mention des coordonnées (non structurée)	hasStatementOfCoordinatesUnstructured
P206bbe	has statement of zone (unstructured)	a pour mention de la zone (non structurée)	hasStatementOfZoneUnstructured
P206bbf	has statement of equinox (unstructured)	a pour mention de l'équinoxe (non structurée)	hasStatementOfEquinoxUnstructured
P2060ba	has mathematical data statement (structured)	a pour mention des données mathématiques (structurée)	hasMathematicalDataStatementStructured
P2060bb	has statement of scale (structured)	a pour mention de l'échelle (structurée)	hasStatementOfScaleStructured
P2060bc	has statement of projection (structured)	a pour mention de la projection (structurée)	hasStatementOfProjectionStructured
P2060bd	has statement of coordinates (structured)	a pour mention des coordonnées (structurée)	hasStatementOfCoordinatesStructured
P2060be	has statement of zone (structured)	a pour mention de la zone (structurée)	hasStatementOfZoneStructured
P2060bf	has statement of equinox (structured)	a pour mention de l'équinoxe (structurée)	hasStatementOfEquinoxStructured

Tableau 4 : Exemple complet de propriétés RDF représentant une zone UNIMARC avec un seul indicateur pouvant prendre une valeur binaire : Zone particulière à certains types de ressources : Ressources cartographiques – Données mathématiques (zone 206 d'UNIMARC/B)

Quand, dans une étiquette, les deux indicateurs sont définis, avec plusieurs valeurs pour chaque indicateur, le nombre de propriétés RDF potentielles est multiplié du fait de l'explosion des combinaisons possibles, comme le Tableau 5 en fait la démonstration.

URI	Label@en	Label@fr
P210bba	has place of publication, distribution, etc. (sequence of publication data not applicable or earliest available publisher; produced in multiple copies, usually published or publically distributed)	a pour lieu de publication, de production, de diffusion, etc. (historique des changements dans la publication non applicable ou premier éditeur connu ; produit en nombre, pour l'édition ou la diffusion publique)
P210b1a	has place of publication, distribution, etc. (sequence of publication data not applicable or earliest available publisher; not published or publically distributed)	a pour lieu de publication, de production, de diffusion, etc. (historique des changements dans la publication non applicable ou premier éditeur connu ; ni publié, ni diffusé publiquement)
P2100ba	has place of publication, distribution, etc. (intervening publisher; produced in multiple copies, usually published or publically distributed)	a pour lieu de publication, de production, de diffusion, etc. (éditeur intermédiaire ; produit en nombre, pour l'édition ou la diffusion publique)
P21001a	has place of publication, distribution, etc. (intervening publisher; not published or publically distributed)	a pour lieu de publication, de production, de diffusion, etc. (éditeur intermédiaire ; ni publié, ni diffusé publiquement)
P2101ba	has place of publication, distribution, etc. (current or latest publisher; produced in multiple copies, usually published or publically distributed)	a pour lieu de publication, de production, de diffusion, etc. (éditeur courant ou dernier connu ; produit en nombre, pour l'édition ou la diffusion publique)
P21011a	has place of publication, distribution, etc. (current or latest publisher; not published or publically distributed)	a pour lieu de publication, de production, de diffusion, etc. (éditeur courant ou dernier connu ; ni publié, ni diffusé publiquement)

Tableau 5 : Exemple partiel de propriétés RDF représentant une zone UNIMARC avec deux indicateurs pouvant prendre des valeurs multiples : Publication, production, diffusion, etc. (zone 210 d'UNIMARC/B)

Le nombre total des propriétés potentielles pour cet exemple est 48 : 3 (valeurs du 1^{er} indicateur) fois 2 (valeurs du 2^e indicateur) fois 8 (sous-zones).

Il y a des étiquettes UNIMARC avec un nombre bien plus grand de combinaisons possibles :

327 Note de contenu : $4 \times 2 \times 12 = 96$

620 Lieu et date de publication, de représentation ou d'enregistrement, etc. : $7 \times 3 \times 15 = 315$

621 Lieu et date de provenance : $7 \times 3 \times 16 = 336$

852 Localisation et cote : $7 \times 4 \times 16 = 448$

Il est nécessaire de procéder à un examen plus approfondi de ces cas pour déterminer si certaines combinaisons ne sont pas pertinentes et ne nécessitent pas de propriété distincte.

Déclarations agrégées

Toutes les étiquettes répétables contenant plusieurs sous-zones constituent une déclaration agrégée. Il est nécessaire de maintenir les sous-zones ensemble pour chaque répétition de la zone, de manière à ce qu'elles ne se mélangent pas.

Exemple

010bba Numéro ISBN + 010bbb Qualificatif de l'ISBN

Les déclarations agrégées sont représentées en RDF en utilisant des schémas d'encodage de la syntaxe (*Syntax encoding schemes*, SES), et il faudrait suivre la pratique de l'ISBD. Réutiliser des éléments de l'ISBD qui sont eux-mêmes des déclarations agrégées évitera d'avoir à développer des équivalents en UNIMARC.

Conclusion et recommandations

L'investissement de l'IFLA dans l'activité de publication en RDF de ses normes et modèles internationalement reconnus, comme une première étape pour marquer les métadonnées des bibliothèques comme fiables et faisant autorité dans le Web sémantique, a déjà été fait. Toutefois, bien que cette première étape concerne les trois modèles conceptuels, FRBR, FRAD et FRSAD, ainsi que la norme bibliographique qu'est l'ISBD, un travail complémentaire est nécessaire. En présentant certaines solutions et en soulevant des questions qui appellent une analyse plus poussée, cette communication plaide en faveur du besoin de représenter de la même manière les formats UNIMARC de l'IFLA, le format pour les données bibliographiques comme le format pour les données d'autorité. Les auteurs soutiennent aussi que la coordination du travail pour représenter la documentation des normes de l'IFLA devrait être menée de manière beaucoup plus étroite, car en pratique, ces normes sont envisagées et utilisées à l'unisson, et aussi, parce que leur développement ultérieur y gagnera beaucoup en efficacité et en économie. Un autre aspect du travail pour représenter les normes en RDF est qu'il offre des remontées d'information sur les normes elles-mêmes, leur structure, leur précision dans l'expression des concepts et des définitions, leur cohérence, l'interopérabilité avec d'autres normes connexes de métadonnées à l'usage des bibliothèques ou d'autres communautés, etc. – ce qui est désormais nécessaire dans le nouveau paradigme technologique du Web sémantique.

Recommandations au Comité permanent UNIMARC (PUC) pour discussion complémentaire et approbation :

- Approuver la méthode pour identifier les éléments et vocabulaires de l'UNIMARC.
- Prendre une décision sur la création initiale et la maintenance des éléments et vocabulaires de l'UNIMARC dans l'OMR (*Open Metadata Registry* = Bureau d'enregistrement des métadonnées ouvertes).
- Soutenir et promouvoir la traduction des classes et propriétés de l'UNIMARC dans les langues nationales.
- Prendre une décision entre la réutilisation pour UNIMARC/B de l'espace de noms existant pour l'ISBD ou la représentation de tous les éléments de l'UNIMARC/B et la création de liens vers les classes et les propriétés existant pour l'ISBD lorsque c'est approprié.
- Étudier en outre la réutilisation pour UNIMARC/A des espaces de noms existant pour FRAD/FRBR et FRSAD ou la représentation de tous les éléments de l'UNIMARC/A et la création de liens vers les classes/sous-classes et les propriétés existant pour FRAD/FRBR/FRSAD lorsque c'est approprié.
- Approuver le schéma proposé pour les espaces de noms pour les éléments et les vocabulaires d'UNIMARC/A et d'UNIMARC/B.
- Discuter et examiner les besoins de profils d'applications pour UNIMARC.
- Contrôler et vérifier la disponibilité de représentations en SKOS d'autres vocabulaires externes qui sont utilisés en UNIMARC.

- Examiner et vérifier les vocabulaires internes d'UNIMARC pour en donner des représentations adéquates en SKOS ; étudier la possibilité d'entrer en contact avec les propriétaires de vocabulaires externes pour assurer la liaison sur le développement de représentations en SKOS.
- Examiner plus avant les classes appropriées pour UNIMARC/A en relation avec UNIMARC/B, FRAD/FRBR et FRSAD.
- Examiner plus avant « l'explosion combinatoire » des propriétés d'UNIMARC ; déterminer si certaines combinaisons ne sont pas pertinentes et ne nécessitent pas de propriété distincte.
- Étudier et approuver la réutilisation d'éléments agrégés de l'ISBD qui sont représentés en RDF en utilisant des schémas d'encodage de la syntaxe (*Syntax encoding schemes*, SES), ce qui évitera d'avoir besoin de développer des équivalents pour UNIMARC.
- Suivre les développements de MARC21 susceptibles de le concerner, en particulier l'annonce récente par la Bibliothèque du Congrès de l'Initiative pour une transition du cadre de l'information bibliographique (*Bibliographic Framework Transition Initiative*²⁷).

Les auteurs de cette communication recommandent vivement que le PUC propose à l'IFLA le financement du développement de la représentation d'UNIMARC en RDF comme un projet de recherche et développement.

Références bibliographiques

- ¹ Bizer, Christian ; Tom Heath, Tim Berners-Lee. Linked data – The story so far. // International Journal on Semantic Web and Information Systems (IJSWIS), vol.5, issue 3. (2009). Pré-publication disponible à l'adresse suivante : <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf>
- ² W3C. SKOS Simple Knowledge Organization System – Home page. 2010. Disponible à l'adresse suivante : <http://www.w3.org/2004/02/skos/>
- ³ OCLC. WorldCat facts and statistics. 2011. Disponible à l'adresse suivante : <http://www.oclc.org/worldcat/statistics/default.htm>
- ⁴ Moen, William E. ; Penelope Benardino. Assessing metadata utilization : an analysis of MARC content designation use. // International Conference on Dublin Core and Metadata Applications, DC-2003--Seattle Proceedings. Disponible à l'adresse suivante : <http://dcpapers.dublincore.org/ojs/pubs/article/view/745/741>
- ⁵ Coyle, Karen ; Thomas Baker. Guidelines for Dublin Core application profiles. 2009. Disponible à l'adresse suivante : <http://dublincore.org/documents/profile-guidelines/index.shtml>
- ⁶ W3C OWL Working Group. OWL 2 Web Ontology Language: document overview. 2009. Disponible à l'adresse suivante : <http://www.w3.org/TR/owl2-overview/>
- ⁷ Graves, Mike ; Adam Constarbaris, Dan Brickley. FOAF : connecting people on the Semantic Web. // Knitting the Semantic Web / Jane Greenberg, Eva Méndez, editors. Binghampton, NY : The Howarth Information Press, 2007. P. 196.
- ⁸ IFLA. Cataloguing Section. ISBD Review Group, ISBD/XML Study Group. 2008. Disponible à l'adresse suivante : www.ifla.org/en/node/1795
- ⁹ IFLA. Cataloguing Section. FRBR Review Group. Meeting Report Durban, August 21, 2007. Disponible à l'adresse suivante : www.ifla.org/files/cataloguing/frbrrg/meeting_2007.pdf
- ¹⁰ Functional requirements for bibliographic records : final report / IFLA Study Group on the Functional Requirements for Bibliographic Records. München : Saur, 1998. Amended 2009. Disponible à l'adresse suivante : www.ifla.org/en/publications/functional-requirements-for-bibliographic-records

Traduction française de l'édition originale disponible à l'adresse suivante :

http://www.bnf.fr/documents/frbr_rapport_final.pdf

¹¹ Dunsire, Gordon ; Mirna Willer. Standard library metadata models and structures for the Semantic Web. // Library hi tech news, vol. 28, no. 3. (2011), pp. 1-12. Disponible à l'adresse suivante :

<http://dx.doi.org/10.1108/07419051111145118>

¹² Willer, Mirna ; Gordon Dunsire, and Boris Bosančić. ISBD and the Semantic Web. // JLIS.it Journal of Library and Information Science. Italy, vol. 1, no. 2, (2010), pp. 213-236. Disponible à l'adresse suivante : <http://dx.doi.org/10.4403/jlis.it-4536>

¹³ Functional requirements for authority data : a conceptual model / edited by Glenn E. Patton ; IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). Final report, December 2008 / approved by the Standing Committees of the IFLA Cataloguing Section and IFLA Classification and Indexing Section, March 2009. München : K. G. Saur, 2009.

Également disponible à l'adresse suivante : www.ifla.org/publications/functional-requirements-for-authority-data

Traduction française disponible à l'adresse suivante :

http://www.bnf.fr/documents/frad_rapport_final.pdf

¹⁴ Functional requirements for subject authority data (FRSAD) : a conceptual model / edited by Marcia Lei Zeng, Maja Žumer and Athena Salaba ; IFLA Working Group on Functional Requirements for Subject Authority Records (FRSAR). Berlin ; München : De Gruyter Saur, 2011. Rapport final disponible à l'adresse suivante : www.ifla.org/files/classification-and-indexing/functional-requirements-for-subjectauthority-data/frsad-final-report.pdf

¹⁵ Open Metadata Registry. Sans date. Disponible à l'adresse suivante : <http://metadataregistry.org/>

¹⁶ Dunsire, Gordon. UNIMARC, RDA and the Semantic Web. In : International Cataloguing and Bibliographic Control (ICBC), vol. 39, no. 2 (April/June 2010).

Fondé sur une communication présentée au 75^e congrès de l'IFLA : World Library and Information Congress : 75th IFLA General Conference and Assembly, 23-27 August 2009, Milan, Italy. Disponible à l'adresse suivante : <http://www.ifla.org/files/hq/papers/ifla75/135-dunsire-en.pdf>

Traduction française de la communication présentée au 75^e congrès de l'IFLA disponible à l'adresse suivante : <http://www.ifla.org/files/hq/papers/ifla75/135-dunsire-fr.pdf>

¹⁷ UNIMARC Manual. Bibliographic Format / edited by Alan Hopkinson. 3rd edition. München : Saur, 2007.

Traduction française disponible en ligne à l'adresse suivante :

http://www.bnf.fr/fr/professionnels/anx_formats/a.unimarc_manuel_format_bibliographique.html

¹⁸ UNIMARC Manual. Authorities Format / edited by Mirna Willer. 3rd edition. München : Saur, 2009.

¹⁹ Willer, Mirna. Foreword to the third edition. // UNIMARC Manual. Authorities Format / edited by Mirna Willer. 3rd edition. München : Saur, 2009. Pp. 7-9.

²⁰ International standard bibliographic description (ISBD). Consolidated edition. Projet à la date du 2010-05-10. Disponible à l'adresse suivante :

www.ifla.org/files/cataloguing/isbd/isbd_wvr_20100510_clean.pdf.

La publication de l'édition normalisée est attendue pour le congrès de l'IFLA de 2011.

²¹ Semantic Web Education and Outreach (SWEO) Interest Group. Cool URIs for the Semantic Web. 2008. Disponible à l'adresse suivante: <http://www.w3.org/TR/cooluris/>

²² IFLA Namespaces Task Group. IFLA namespaces - requirements and options. 2010. Disponible à l'adresse suivante : http://www.ifla.org/files/classification-and-indexing/ifla-namespaces-requirements-optionsreport_corrected.pdf

²³ Library of Congress. Sans date. MARC list for languages. Disponible à l'adresse suivante : <http://id.loc.gov/vocabulary/languages.html>

²⁴ Library of Congress. Sans date. MARC list for geographic areas. Disponible à l'adresse suivante : <http://id.loc.gov/vocabulary/geographicAreas.html>

²⁵ Library of Congress. Sans date. MARC list for countries. Disponible à l'adresse suivante :

<http://id.loc.gov/vocabulary/countries.html>

²⁶ Martin, Earle. ISO 3166 RDF representation. 2005. Disponible à l'adresse suivante :

<http://downlode.org/Code/RDF/ISO-3166/>

²⁷ Library of Congress. Bibliographic Framework Transition Initiative. 2011. Disponible à l'adresse suivante: <http://www.loc.gov/marc/transition/>