ISAN & EIDR Metadata Schema Mapping

ISAN-IA
1A, rue du Beulet
CH-1203 Geneva
Switzerland

Tel: +41 22 545 10 00
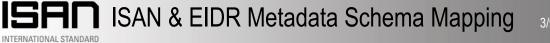Fax: +41 22 545 10 40
Email: info@isan.org

Version: 1.0, December 12, 2012

Table of contents

1   Introduction

1.1   GENERAL OVERVIEW

This document is resulting from the joint work produced by ISAN and EIDR to map their respective metadata schemas. It aims to provide guidelines to users of ISAN or EIDR who are willing to map their data from one schema to another.

The current version of the document only deals with mapping audiovisual works schemas (root objects). The mapping of versions of audiovisual works will be addressed in a future version of the document.

2 Mapping EIDR Records to ISAN

2.1 CLASSIFICATION IN WORKS AND VERSIONS

The first step when mapping EIDR data with ISAN is the classification of each EIDR record in either

- **Audiovisual Work**: Serial / Single Work / Serial Episode
- **Version**: Language / Edition / Media / Related Content / Related Item

For a given EIDR record, the combination of a "Base Object" with one or several "Derived Objects" is used to distinguish audiovisual works from versions, and to define the ISAN:VersionKind when the record is a version. The table below illustrates some examples:

AV Content	EIDR Object	ISAN Object Correspondence	
A Non Episodic Title	**Base Object**	**Single Work**	AV Works (ISAN)
A Serial	**Base Object + "Serie"** Derived Object	**Serial Header**	
A Season of a Serial	**Base Object + "Season"** Derived Object	**N/A:** ISAN does not identify seasons, instead the season information is referenced in the Episode descriptive information.	
An Episode of a Serial	**Base Object + "Episode"** Derived Object	**Serial Episode Work**	
Interactive Material	**Base Object +** "**Interactive Material**" Derived Object	**Single Work** with Work Type **Multimedia** or **Video Game** (refer to the ISAN user guide for cases where a **Related Content Version** is more appropriate)	
A composite non episodic title	**Base Object + "Composite"** Derived Object	**Single Work** with composite information	
A composite episodic title	**Base Object + "Episode" + "Composite"** Derived Objects	**Serial Episode Work** with composite information	
Changes in the audio or visual content	**Base Object + "Edit"** Derived Object	**Version** with "Version Kind" = "**EDITION**"	Versions (V-ISAN)
Changes in one or more Language track	**Base Object + "Language Variant"** Derived Object	**Version** with "Version Kind" = "**LANGUAGE**"	
Changes in the digital representation of an asset	**Base Object + "Encoding"** Derived Object	**Version** with "Version Kind" = "**MEDIA**". For technical encodings such as a DVD with multiple tracks (video, audio, text) and/or supplemental material etc… a separate version (VersionKind = "Related Item" or "Related Content") for each track can be created; these versions can then be combined in the composite field of the final "Media" Version.	
A clip taken from an identified content	**Base Object + "Clip"** Derived Object	**Version** with "Version Kind" = "**EDITION**"	
Changes the audio or visual content + inclusion of other identified material	**Base Object + "Edit" + "Composite"** Derived Objects	**Version** with "Version Kind" = "**EDITION**" + use of the **version composite** fields	
A clip taken from an Edited identified content	**Base Object +"Edit" + "Clip"** Derived Object**s**	**Version** with "Version Kind" = "**EDITION**"	
Changes the visual and audio content	**Base Object + "Edit" + "Language Variant"** Derived Objects	**Version** with a 1st "Version Kind " = "**EDITION**" & a 2nd "Version Kind"=" **LANGUAGE**"	

2.2 "EIDR TO ISAN" MAPPING FOR NON EPISODIC WORKS

The table below describes the mapping of EIDR metadata for base records (root objects) to ISAN metadata for non Episodic Works (aka: Single Works)

ISAN Field	Value to use: from EIDR metadata		Notes
ISAN:WorkType	*EIDR:ReferentType*	*ISAN:WorkType*	*Note*
	TV	TV Entertainment (TE) / Telefilm (TF)/ Documentary (DO) / Sport (SP) / News (NE) / Live Event (LV)/ Performance(PF)/ etc	Defined by the context & specific pattern that qualify a given type of content (or TV format) such as the length of the program, etc..
	Movie	**FF (Feature Film)** / Documentary (DO) / etc	Use **FF (Feature Film)** unless another work types is more appropriate (identified by the registrant or deduced from the context).
	Web	**Multimedia (MM)** / Short (SH) / Documentary (DO)/ etc	Use **Multimedia (MM)** unless another work type is more appropriate (identified by the registrant or deduced from the context).
	Short	**Short (SH)** / Video Clip (VC) / Adds (AD)/ etc	Use **Short (SH)** unless another work type is more appropriate (identified by the registrant or deduced from the context).
	Composite	**Compilation (CO)**	At least one work that compose the compilation must be registered first with ISAN
	Interactive Material	Use **Multimedia** (MM)	
	Supplemental	Register Supplemental material as "Related Content" versions of the work. If a different ISAN root is desirable, use the appropriate WorkType (Documentary, Ads, short, …)	
	Series	Registered as a "Serial Header"	
TimeValue (RunningTime)	EIDR:ApproximateLength		Translate from EIDR:xs:duration in ISAN:TimeValue and ISAN:TimeUnit
TimeUnit (RunningTime)	Typically use "MIN"		
ColorKind	N/A in EIDR		Get the color information from an external source. Sometimes it can be deduced from the EIDR:Edit(s) where the "ColorType" is mandatory. However this has to be used carefully as the purpose of the Edit can precisely identify a change in color (e.g. a colorized version from a B&W work). The following principles usually applies: - FF productions before 1960 (US) / 1970 (Europe) were generally in Black and White, and in Color after. - TV productions before 1970 were generally in Black and White, and in Color after.
WorkKind	N/A in EIDR		get this information from the context (e.g. Disney cartoons) or from an external source.
ReferenceCountryList	*The list contains an ISAN:ReferenceCountry attribute for each EIDR:CountryOfOrigin*		
ReferenceCountry/CountryCode	EIDR: CountryofOrigin		
ReferenceCountry/RelatedAction	"PRO" (PrincipalAgent/role = "producer") "DIST" (PrincipalAgent/role = "distributor" or "broadcaster")		

INTERNATIONAL STANDARD
AUDIOVISUAL NUMBER

ISAN Field	Value to use: from EIDR metadata	Notes
TitleList	*The list contains one ISAN:TitleDetail attribute created from EIDR:RessourceName*	
TitleDetail/TitleKind	"ORIGINAL"	Must be translated from RFC 4646 to ISO 639-2
TitleDetail/Language	EIDR:RessourceName/Language	
TitleDetail/Title	EIDR:RessourceName	
	The list also contains an ISAN:TitleDetail attribute for each EIDR:AlternateRessourceName	
TitleDetail/TitleKind	"ALTERNATE"	Must be translated from RFC 4646 to ISO 639-2
TitleDetail/Language	EIDR:AlternateRessourceName/Language	
TitleDetail/Title	EIDR:AlternateRessourceName	
OriginalLanguageList	*The list contains ISAN:OriginalLanguage attributes*	Define an **OriginalLanguage** for each EIDR Primary and secondary languages
OriginalLanguage	*if (EIDR:PrimaryLanguage/Manifestation=audio)* use EIDR:PrimaryLanguage/Language *else if (Mode = "Visual")* use "XMU" or "QSI" (the work is silent)	Must be translated from RFC 4646 to ISO 639-2
OriginalLanguage	*for each (EIDR:SecondaryLanguage with Manifestation = audio)* EIDR:SecondaryLanguage/Language	Must be translated from RFC 4646 to ISO 639-2
CompanyList	*The list conttains an ISAN:Company attribute for each EIDR PrincipalAgent with a role=producer*	
Company/CompanyKind	"PRO"	
Company/CompanyName	EIDR:PrincipalAgent/party	
ExternalIDList/ ExternalID/Code	"EIDR"	
ExternalIDList/ ExternalID/Id	EIDR ID of the work	
ParticipantList	FOR ALL *(EIDR:Credits/Director) AND(EIDR:Credits/Actor)*	
FirstName	EIDR:DisplayName	need to convert EIDR:DisplayName in ISAN:FirstName / ISAN:LastName
LastName	EIDR:DisplayName	
RoleCode	*if (EIDR:Credits/Director)* use "DIR" *if (EIDR:Credits/Actor)* use "ACT"	
CompositeList	ISAN like EIDR requires that the different components are previously registered. Register first the different components with ISAN and report then each resulting ISAN in the Composite list. Ignore the other EIDR:Composite object : class, start, end, duration, …	

2.3 "EIDR TO ISAN" MAPPING FOR SERIES

In ISAN, all episodes of a serial are registered under a unique Serial Header for each serial. The Serial Header is not identified with an ISAN but with a SHID (SHID = root ISAN common to all Episodes of the serial). Prior serial registration, the title of the serial must be looked up in the Serial Header list. If it is not found, register a new Serial Header for the Serial and use the resulting SHID to register the episodes of the serial.

The table below describes the mapping of EIDR metadata for serial records (*base object + serie derived object*) to ISAN metadata for Serial Headers

ISAN Field	Value to use: from EIDR metadata	Notes
TitleList	*The list contains one ISAN:TitleDetail attribute created from EIDR:RessourceName*	
TitleDetail/TitleKind	"ORIGINAL"	Must be translated from RFC 4646 to ISO 639-2
TitleDetail/Language	EIDR:RessourceName/Language	
TitleDetail/Title	EIDR:RessourceName	
	The list also contains an ISAN:TitleDetail attribute for each EIDR:AlternateRessourceName	
TitleDetail/TitleKind	"ALTERNATE"	Must be translated from RFC 4646 to ISO 639-2
TitleDetail/Language	EIDR:AlternateRessourceName/Language	
TitleDetail/Title	EIDR:AlternateRessourceName	
SerialHeaderRegistrantID	EIDR ID of the Serie Container	

2.4 "EIDR TO ISAN" MAPPING FOR EPISODES OF SERIES

In ISAN, serial works are described with the same base data as non episodic works but with an additional set of data: "EpisodeDetails".

The table below describes the mapping of EIDR metadata for episodic records (*base object + episode derived object*) to ISAN metadata for Serial Episode Works.

ISAN Field	Value to use: from EIDR metadata		Notes
EpisodeDetail	*Specific metadata for Episodic titles*		
Season	EIDR: SeasonInfo/SequenceNumber		from EIDR:Season object
Episode	EIDR: EpisodeInfo/SequenceNumber		
SerialHeaderRegistrantID	EIDR ID of the Serie Container		
SerialHeaderID	N/A		SHID of the Serial Header
ISAN:WorkType	The work Type Series (SE) exist only for non episodic works. Typically use **Series (SE)** for Fictional Episodic Works. All other work types are also possible.		
	EIDR:ReferentType	*ISAN:WorkType*	*Note*
	TV	**Series (SE)** / TV Entertainment (TE) / Telefilm (TF)/ etc	Use **SE (Serie)** unless another work type is more appropriate (identified by the registrant or deduced from the context).
	Movie	Identical as for Single Works.	
	Web	Identical as for Single Works.	
	Short	Identical as for Single Works.	
	Composite	Identical as for Single Works.	
	Interactive Material	Use **Multimedia** (MM)	
	Supplemental	Register Supplemental material as "Related Content" versions of the work. If a different ISAN root is desirable, use the appropriate WorkType (Documentary, Ads, short, …)	
	Series	*Registered as a "Serial Header"*	

Other metadata fields are identical as for non episodic Works, refer therefore to the single works mapping above.

3   Mapping ISAN Records to EIDR

3.1   "ISAN TO EIDR" MAPPING FOR NON EPISODIC WORKS

The table below describes the mapping of ISAN metadata for base records to EIDR metadata for root objects.
General Notes:

- This covers root objects only; future versions will also cover episodic works and edits/versions.
- Values to use may be by value from ISAN metadata, by algorithm or constant value.

EIDR Field	Value to use: from ISAN metadata	Notes
Structural Type	"Performance"	See *EIDR Best Practices* document for the few cases where this might not be the case.
Mode	"Visual" if ISAN language is XMU or QSI, "AudioVisual" otherwise	
ReferentType	"Movie", "TV", or "Short"	See EIDR Best Practices, but in general, use ISAN FF: EIDR Movie ISAN TF, TE, SE: EIDR TV Other ISAN types require more complex EIDR objects to map.
ResourceName	ISAN TitleDetail/Title field withTitleKind ORIGINAL	Use the first one if more than one ORIGINAL. If there is more than one ORIGINAL, use the first one, and use the remainder as values for AlternateResourceName (see below)
ResourceName language attribute	ISAN TitleDetail/Language for the title used	Requires translation from ISO 639-2 to RFC 4646
ResourceName titleClass attribute	"release"	
AlternateResourceName	ISAN TitleDetail/Title field withTitleKind ALTERNATE, ORIGNAL, or TRANSLATED	An ISAN ALTERNATE title is the release title in a non-original language.
Alternate ResourceName language attribute	ISAN TitleDetail/Language for the title used	Requires translation from ISO 639-2 to RFC 4646
AlternateResourceName titleClass attribute	"other" if ISAN TitleKind is TRANSLATED "release" if ISAN TitleKind is ORIGINAL or ALTERNATE	See comments in RegistrantExtra on ALTERNATE
PrimaryLanguage	ISAN LanguageList/Language	If the item is not silent (Mode is AudioVisual): If the ISAN item has more than one entry, use the first one. See below under "Description" for the disposition of the others. If the item is silent (Mode is Visual): If the ISAN item has no languages other than XMU or QSI, use the ISO 639-2 Code zxx to indicate no linguistic content. Otherwise, apply the rule for AudioVisual works, ignoring XMU and QSI language fields. In all cases where the language is not empty, this requires translation from ISO 639-2 to RFC 4646
PrimaryLanguage manifestaiton attribute	"audio", unless item is silent (EIDR Mode is "visual"), in which case it should be "subtitle"	See "Mode" for determining if something is silent

EIDR Field	Value to use: from ISAN metadata	Notes
PrincipalAgent	From ISAN CompanyList field	CompanyName must be looked up in EIDR party list. If it is not found, request EIDR to register a new Party and use the resulting ID
PrincipalAgent role attribute	"producer" if ISAN Company/CompanyKind is PRO "distributor" if ISAN Company/CompanyKind is PUB	
ReleaseDate	ISAN YearOfFirstPublication	EIDR allows yyyy or yyyy-mm-dd; the extra should be added only if the date is known for sure and is available to the registrant
CountryofOrigin	ISAN ReferenceCountryList/Country	Use first one with relatedActionType PRO If none present, use first with relatedActionType PUB Else, use first with relatedActionType SHO
Status	"valid"	Only register public ISAN records
ApproximateLength	ISAN Duration/TimeValue	Must be translated from MIN or SEC to xs:duration
AlternateID	ISAN id of the work	AlternateID of type "ISAN"
DisplayName	empty	
Registrant	EIDR ID of the registrant	
Credits/Director	ISAN Participant with RoleCode DIR	If more than one, choose the first
CurrentAssetHolder		
BackupContact		
Credits/Actor	ISAN Participant with RoleCode ACT	Use the first two
Description	"CROSS-REG" If ISAN had more than one primary language, ":" followed by comma-separated list of LanguageList/Language fields that were not used for the EIDR PrimaryLanguage field (excluding XMU and qsi.)	These are included to allow for improved duplicate detection, and to allow correction if the wrong primary language was chosen at registration time.
RegistrantExtra	These fields, in any order, separated by ":" "ISAN-ORIGIN-PUB" or "ISAN-ORIGIN-SHO" if relatedActionType for country of origin was not PRO "ISAN-ALT-TITLES" if ISAN alternate titles were used as ALternateResourcenames with titletype "other" "ISAN-MULTI-DIR" if ISAN provided more than one director	These are included to allow for improved de-duplication of cross-registrations, and to allow automated fix-ups if EIDR structures or controlled vocabulary change.