



5to COLOQUIO SOBRE RDA EN AMÉRICA LATINA Y EL CARIBE

Implementación de RDA en los sistemas integrados de nueva generación: el caso de Koha

Presenta

Dr. Gerardo Belmont Luna

en

26, 27 y 28 de octubre de 2022

Presentación

- Biblioteca Francisco Xavier Clavigero (BFXC),
 - Es una biblioteca con 488,411 registros y casi un millón de ejemplares,
 - se organizan en cuatro bibliotecas: BFXC, Libros Antiguos y Raros (LAYR), Biblioteca Eusebio Kino (Kino) y Biblioteca Digital BFXC (BiDiBFXC).
 - En ella laboran 65 personas: 16 académicos, 5 coordinaciones, 3 secretarías, una dirección, 15 bibliotecarios profesionales y 35 personas de apoyo al servicio.
 - Pertenece a los grupos Sistema Universitario Jesuita (SUJ) y Asociación de Universidades de la Compañía de Jesús en América Latina (AUSJAL).
 - A mediados del 2019, la dirección de la BFXC propuso a Koha como opción para remplazar a Aleph 500;
 - El cambio tecnológico que se dio entre los años 2020 y 2021;
 - Los resultados no son los esperados.
-

Diagnóstico de la base de datos

Conceptos	Casos	Causa	Efectos que se produce
Registros bibliográficos	488,411	Registros migrados	Total de registros migrados en Koha.
Número de control	701,914 variables	Se identifican infinidad de formas y número de sistema	Ingobernabilidad para el sistema.
Llaves del sistema:	Sin definir en Koha	Sin control	Recupera de todo
Bibliotecas	342	Se desconocieron las políticas y la supervisión se relajó	Redundancia y ruido en la búsqueda y recuperación de información
Colecciones	218	En la práctica se van a acumulando las variables por ocurrencia.	No se distinguen las colecciones debidamente.
Títulos inconsistentes	132,000	Puede ser defecto de migración, no hay claridad en las etiquetas MARC.	Se indiza datos inconexos, irrecuperables, sin inventariar.
Indicadores dudosos	29,594	Índices cuyo contenido está fragmentado.	Indicadores inconexos.
Títulos fragmentados	145,000	Puede ser defecto de migración, no hay claridad en las etiquetas MARC.	Sin recuperación ni edición para su corrección.
Títulos vacíos	259	Puede ser por malas prácticas en la migración, desconocimiento del sistema o dudas en el proceso al momento de la catalogación por descarga.	Relleno sin utilidad.
Autores	25%	Se mezclan con autores corporativos.	Ausencia de capacitación, supervisión y políticas por parte de la coordinación de Organización de la Información.
Autores	90%	El ejercicio de investigación es ausente para el registro de autoridades, no se mira el propio catálogo.	Los índices se elevan y las variables son se controlan, Selección sin perspectiva.
Registros sin índice	1,083	Esto ocurre cuando se quita gobernabilidad al sistema.	El sistema se vuelve vulnerable y poco confiable.
Notas 5XX	Incontables	Sin parámetros	No hay políticas para su asiento. Aunque son libres, deben establecerse las formas de asentar.

.. Además

- En la configuración Koha
 - Las Plantillas MARC están indefinidas;
 - Se incluyen todos los campos para la indización y recuperación;
 - Dos OPAC, uno para Catalogadores y otro para usuarios normales;
 - OPAC configurado para nivel experto de catalogación;
 - 240 variables de bibliotecas,
 - 280 colecciones
-

Koha



- A Koha es un software de licencia tipo GNU (General Public License) creado por Katipo Communications en el año de 1999;
 - Por sus características lo ubicamos como un SIGB de nueva generación;
 - Su base tecnológica es Linux;
 - Incluye a los manejadores de bases de datos son MariaDB y MySQL;
 - Cumple con las normas y estándares de la bibliotecología como son: MARC21, RCA, ISDB, SCHEMA.ORG, OpenURL, RSS/Atom, Z39.50, SIP2 y, además nosotros queremos que cumpla con RDA-FRBR y la representación genealógica de la información.
 - Garantiza la interoperabilidad de los registros bibliográficos con base a la Norma ISO 2709 y se armoniza con lagunas tecnologías de LC MARCEdit.
 - Ver Official Website of Koha Library Software. <https://koha-community.org/about/history/>
-

Sistemas Integrales Para la Gestión de Bibliotecas (SIGB)

“Son una herramienta tecnológica para automatizar las operaciones bibliotecarias más comunes como lo son: la catalogación, circulación, consulta y adquisición de materiales. Y, más recientemente: sistemas de resolución de enlaces, metabuscadores, gestores de recursos electrónicos, nuevos interfaces o portales web de consulta.” (Martín, 2008:2)



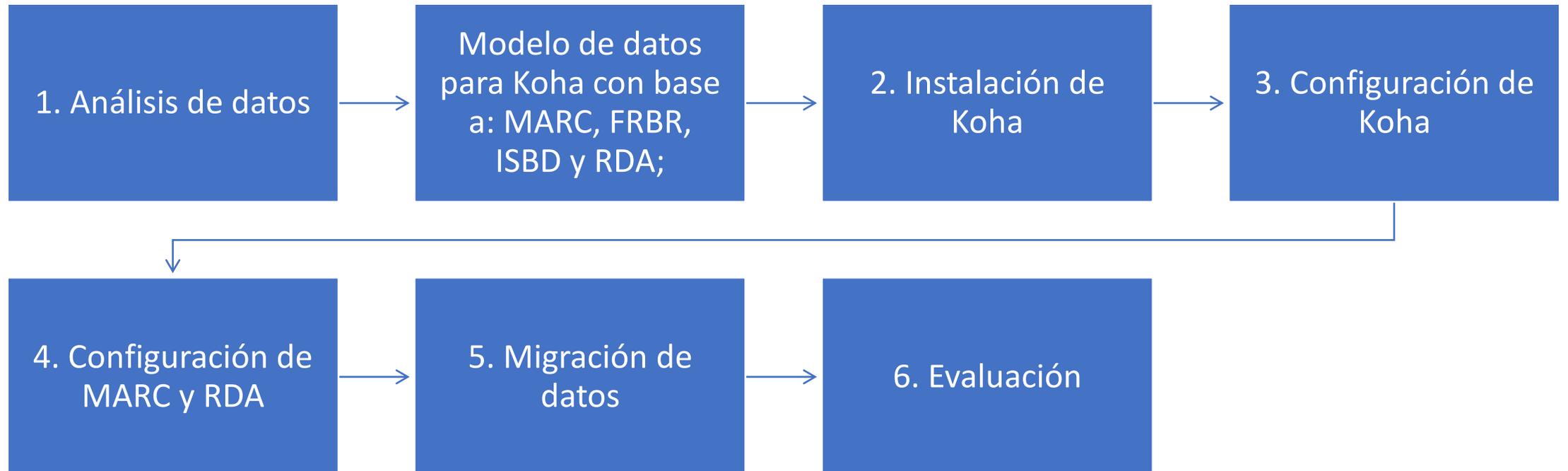
- Gisber y Esengeldiev (2014), debe “cumplir con todos los requisitos de cada una de las normas, con un conjunto de políticas y procedimientos; Mejorar la eficiencia general mediante la eliminación de las tareas duplicadas; Definir funciones y responsabilidades de forma clara frente a objetivos comunes; y, que sea más fácil y simple para mejorar continuamente la gestión.” (Gisbert y Esengeldiev, 2014:248)

Nombre	Versión	Lenguaje de programación	Manejador de base de datos	Sistema operativo	Licencia	Dirección electrónica
Koha	20.11 (2022)	Perl	MariaDB, MySQL,	Multiplataforma	GPLv2	http://koha-community.org/
Koha-Kobli	1.12.4 (2013)	Perl	MySQL ; PostgreSQL	Multiplataforma	GPLv2	http://kobli.bage.es/
PHPMylibrary	2.3.1.1	PHP	MySQL	Multiplataforma	GPLv2	http://code.google.com/p/phpmylibrary/
PMB (PhpMyBibli)	4.0 (2012)	PHP	MySQL	Multiplataforma	GPLv2	http://www.pmbservices.fr
OpenBiblio	0.7.1 (17 marzo 2012)	PHP	MySQL	Multiplataforma	GPLv2	Obsoleto
Evergreen	2.3 (2012)	PHP	PostgreSQL	Multiplataforma	GPLv2	http://open-ils.org/
NewGenLib	3.0.4 (agosto 2012)	Java	PostgreSQL	Multiplataforma	GPLv2	http://www.verussolutions.biz/web/content/download
Invenio	1.1.0 (21 oct. 2012)	Java	MySQL, PostgreSQL, etc.	Multiplataforma	GPLv2	https://inveniosoftware.org/

SIGB libres

Tabla 1. Principales SIGB. (elaboración propia con base a análisis en la Web)

Metodología



1. Análisis de datos

Herramientas:

- MARCEdit 7
- C# MARC
- Excel
- NotePad++
- Pivot
- Koha

- diccionario de datos;
- Eliminaron campos cuyo contenido no tenía estructura;
- Se observaron campos con letras y se modificaron para engrosar el propio;
- Campos que se utilizaron una sola vez
- Se identificó información sin correspondencia en campo como 3XX, 4XX, 5XX, 6XX, 7XX, 8XX y 9XX.

Conceptos	Casos
Registros	488,411
Número de control	701,914 variables
Bibliotecas	342
Colecciones	218
Títulos inconsistentes	132,000
Indicadores alterados	29,594
Títulos fragmentados	145,000
Títulos vacíos	259
Autores	25%
Registros sin índice	1,083
Notas 5XX	

- 363 campos MARC;
- 1993 Subcampos;
- Se encontraron elementos RCA, RCA2 y RDA (se entiende que se han acumulado 3 migraciones)
- Más de 17 millones de líneas en texto plano para revisar.

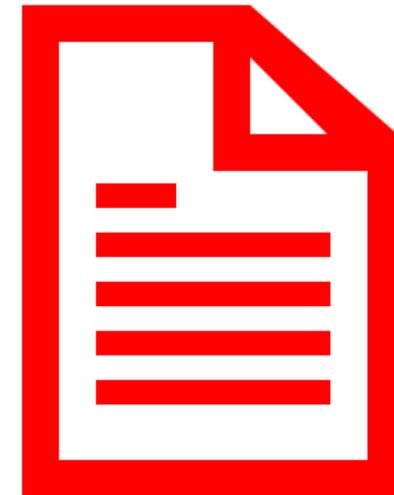
2. Modelo de datos para Koha con base a: MARC, FRBR, ISBD y RDA;

- Herramientas:
- Ubuntu Server
- Koha 20.11
- Configuración
 - Declarar las bibliotecas
 - Declarar las colecciones
 - Crear las plantillas para la descripción MARC



3. Configuración Koha

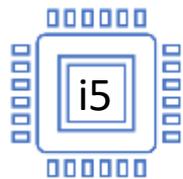
- Con base a:
- diccionario de datos,
- políticas de la biblioteca para la circulación y
- Documentos R-5, ISBD 2011, AACR2 - REGLAS DE CATALOGACIÓN – RDA: FORMA DE PUNTOS DE ACCESO (2010)



4. Configuración con base a Koha

- Plantilla de configuración

🔗 C# MARC Editor



esquemas	FRBR/RDA	MARC	No. De casos	Total de registros	RDA
Obra / Expresión		006	81	81	x
		008	488358	488360	x
Libro		020	196153	285656	x
PP		022	2891	2920	x
Otros materiales		024	2682	2807	x
Otros materiales		028	1012	1102	x
		030	32	32	x
Manifestación / coordinación		034	504	505	x
		048	3	3	x
Libro, PP, otros materiales		074	1	1	x
Obra-creador		100	360418	360435	X
Obra-creador		110	17792	17793	X
Obra-creador		111	6710	6711	X
Obra-título / medio de acceso		130	3369	3371	X
Obra-título / medio de acceso		240	39242	39277	x
Obra-título / medio de acceso		243	19	19	x
Obra-título		245	488410	488625	x
		250	115183	115389	x
Tipo de puntuación / Musica		254	72	72	x
Coordnación / escala		255	6798	6802	x
		260	424722	424772	x
		300	483809	483924	x
Obra-expresión		362	2105	2125	x
		533	210	229	x
Obra-creador		700	201810	308392	X
Obra-creador		710	59821	68789	X
Obra-creador		711	2077	2160	X
		730	1419	1590	X

5. Migración de datos

- Con base a MARC-RDA-FRBR, se configuro la base de datos y se desarrollan las platillas MARC para 22 colecciones.
 - Se ejecuta y se salva el archivo con formato txt . Tiene problemas de codificación: se abre con block de notas y se guarda con codificación utf8, se cierra y se abre con Excel: archivo delimitado por comas y se asigna a cada campo el valor texto.
 - Se procesa un lote de 488 mil registros en MARCEdit7.
-

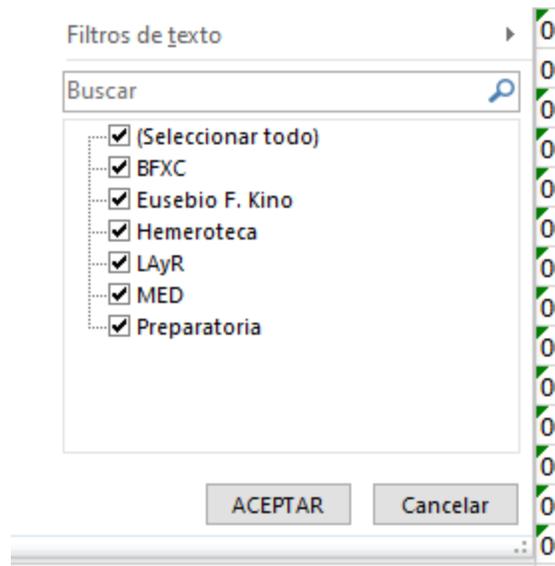
6. Evaluación

- Estamos evaluando la base de datos migrada y su configuración para el servicio y circulación de los recursos bibliográficos.
 - Se están revisando la asignación de tareas: buscar, encontrar, seleccionar y obtener.
 - Se asignaron relaciones con base a estructura de Koha en las entradas principales y \$9 sin embargo aun no se sabe como se relacionan en Zebra –motor de búsqueda. Se insertó el subcampo \$8 para mantener relaciones para campos 1xx, 2xx, 4xx y 6xx.
-

Resultados

Síntoma	Acción	Proceso	Resultado
Registros sin uniformidad	Normalización por campos MARC	488,411	Se normalizo conforma a MARC.
Número de control diferentes	Se asignaron por orden consecutivo, a partir del campo 999 y 001	Todos los registros con índice, único e irrepetible.	Se concluyó la unificación.
Llaves del sistema	39 llaves	39 llaves bien definidas.	Efectos en los campos: 006,008, 020, 022, 024, 028, 030, 034, 048, 074,100, 110, 111, 130, 240, 243, 245, 250, 254, 255, 260, 300, 362, 533, 700, 710, 711, 730, 900, 902, 903, 904, 905, 906, 913, 915, 916, 942, 952, 999
Definición de Bibliotecas	4	Definidas	Cargadas en el campo correspondiente.
Definición de Colecciones	22	Definidas	Falta traducción en Koha.
Títulos inconsistentes	Revisión de las tablas de etiqueta 245	Se revisarán a final del proceso	Eliminar títulos en blanco.
Indicadores	Se dejan sin asignar	Se requiere conocer a Zebra	Indizar todo lo necesario
Títulos fragmentados	Títulos que no se escribieron completos y el registro quedó fracturado.	Se revisarán al final	Eliminar fragmentación
Títulos vacíos	259	Se eliminarán caso por caso	Por revisar
Autores	Revisión de las tablas de etiqueta 1XX	En proceso de normalización	Se busca minimizar las variables de nombre.
Autores	Acomodar los indicadores (sirven para que el sistema despliegue el índice correctamente)	Rectificar por tipo de autor, personal, corporativo o evento.	100%
Registros sin índice	Revisión de las tablas en excel	Se asignó por sistema	Resuelto
Notas 5XX	Se usan indistintamente.	OI debe revisarlas.	Se recomienda usar un manual específico.

Bibliotecas y colecciones



- CD de Audio
- CD de Música
- Colección Enrique Dussel
- Colección Manuale Arango
- Colección Masonería
- Colección Nicaragua
- Colección Pérez Alonso
- Colecciones Especiales
- Consulta
- Depósito Oficial
- Diapositivas I
- EA
- EC
- ECE
- Ekino
- ER
- ES
- Eusebio F. Kino: AC y C
- Eusebio F. Kino: Fondo Reservado
- FILA
- FKINO
- FMC
- Folletos
- GEN
- Hemeroteca
- I
- INEGI y/o Mapas
- JC
- Mapas
- Mapas (Exepto INEGI)
- Material complementario
- Preparatoria
- Rincón Chino
- Teología
- Teología Consulta
- Teología Fondo Reservado
- Tesis Digital: AC, DO (2 discos)
- Tesis Impresas

La base paso de 544 Mb a 286 Mb.

Archivos que no se migran

Tipos de registro	Registros que se migra	Registros que no se migran	Campos utilizados en la selección	Fecha de ejercicio	Logro
Bibliográficos	Registros bibliográficos visibles en OPAC	Registros bibliográficos MARC ocultos en OPAC	035, 9450, subcampo 0	otoño 2022	Garantizar la consistencia de la información
Ejemplares	Registros visibles en OPAC y enlazados a registros visibles OPAC	Registros de ejemplares ocultos en OPAC o enlazados a registros bibliográficos MARC ocultos	035, 9450	otoño 2022	Garantizar la consistencia de la información
Seriadas	Registros bibliográficos MARC de publicaciones periódicas de fondos descritos	Registros bibliográficos MARC de publicaciones periódicas sin fondos descritos	035, 9450	otoño 2022	Garantizar la consistencia de la información
Autoridades	No se tiene información	Se asigna un índice con el subcampo \$9 al final de cada entrada (1XX, 2XX, 4XX, 6XX, 7XX)	Koha \$9	otoño 2022	Garantizar la consistencia de la información
Inventario	Acervo general y consulta	Registros no allados en inventario	Código de barras y acervo general y consulta.	otoño 2022	Garantizar la consistencia de la información

Diccionario de datos

Aplica a libros, Seriadadas, Archivo de computadora Mapas música, materiales gráficos, materiales mixtos, código de registro lingüístico, lingüístico manuscrito, Archivos de computadora, Mapas impresos, Mapas manuscritos, Registro sonoro no musical, Registro sonoro musical

A	B	C	D	E	F	G	H	I
Esquemas	FRBR/R	MAF	Subcampos	No. De casos	Total de registros	RCA2	RDA	ISBD
Libro / Expresión		006		81	81		x	Título prop
		008		488358	488360		x	Información
Libro		020		196153	285656		x	
PP		022		2891	2920		x	
Otros materiales		024		2682	2807		x	
Otros materiales		028		1012	1102		x	
		030		32	32		x	
Manifestación / coordinación		034		504	505		x	
		048		3	3		x	
Libro, PP, otros materiales		074		1	1		x	
Libro-creador		100		360418	360435	X	X	
Libro-creador		110		17792	17793	X	X	
Libro-creador		111		6710	6711	X	X	
Libro-título / medio de acceso		130		3369	3371	X	X	
Libro-título / medio de acceso		240		39242	39277		x	
Libro-título / medio de acceso		243		19	19		x	
Libro-título		245		488410	488625		x	
		250		115183	115389	x	x	
Tipo de puntuación / Música		254		72	72		x	
Coordinación / escala		255		6798	6802		x	
		260		424722	424772		x	
		300		483809	483924		x	
Libro-expresión		362		2105	2125		x	
		533		210	229		x	
Libro-creador		700		201810	308392	X	X	
Libro-creador		710		59821	68789	x	X	
Libro-creador		711		2077	2160	X	X	
		730		1419	1590	X	X	
		900		1385	1567		x	
		902		26895	26946		x	
		903		22	22		x	
		904		16	16		x	
		905		481723	481804		x	
		906		183	183		x	
		913		32	32		x	
		915		482528	509349		x	
		916		481153	481376		x	
		942		488390	488606		x	
		952		488411	721061		x	
		999		488411	488411		x	

Koha

Campos para Ítem en Koha

Tag/subfield	Data element	Description	Notes
952\$0	Estado retirado	Estado del artículo, conéctese con la lista de valores autorizados 'RETIRO'	Los valores por defecto de la lista son: 0 = "" [No retirado]; 1 = 'Retirado'
952\$1	estado perdido	Estado del artículo, conectar con la lista de valores autorizados 'PERDIDO'	Los valores por defecto de la lista son: 0 = "" [Not Lost]; 1 = 'Perdido'; 2= 'Mucho retraso (perdido)'; 3 = 'Perdido y Pagado'; 4 = 'Falta'
952\$2	Fuente de clasificación o esquema de estantería	Esquema de clasificación que define las reglas de archivo utilizadas para ordenar las signaturas topográficas.	Si no se ingresa un esquema de clasificación, el sistema utilizará el valor ingresado en la preferencia de catalogación "DefaultClassificationSource".
952\$3	Materiales especificados (volumen encuadernado u otra parte)	Emisiones específicas de publicaciones seriadas o multipartes.	Los datos de copia/volumen se almacenan en 952\$h y \$t para su visualización en vistas normales. \$3 solo se muestra de forma predeterminada en la interfaz del personal, pero se puede hacer visible en el OPAC mediante los ajustes de configuración de la columna (Error 13547). La nota de materiales también se mostrará en el momento de la compra y el registro.
952\$4	Estado dañado	Estado del artículo, conéctese con la lista de valores autorizados 'DAÑADO'	Los valores por defecto de la lista son: 0 = "" [No dañado]; 1 = 'Dañado'
952\$5	Restricciones de uso	Estado del artículo, conectar con la lista de valores autorizados 'RESTRINGIDO'	Los valores por defecto de la lista son: 0 = "" [No restringido]; 1 = 'Restringido'
952\$6	Clasificación normalizada de Koha para la clasificación		
952\$7	no para préstamo	Estado del artículo, conéctese con la lista de valores autorizados 'NOT_LOAN'	Los valores por defecto de la lista son: -1 = Ordenado; 0 = "" [Disponible para préstamos]; 1 = No para préstamo; 2 = Colección de personal
952\$8	Código de colección	Estado del artículo, conectar con la lista de valores autorizados 'CCODE'	El código de la colección proporciona otro medio para dividir la colección de la biblioteca en categorías (Ejemplo: Ficción, Misterios, etc.) para los informes. También se utiliza para catalogación y estadísticas de circulación.
952\$9	Número de artículo de Koha (generado automáticamente)		Número de artículo generado por el sistema. No se muestra en el registro del artículo.

Tag/subfield	Data element	Description	Notes
952\$a	código de sucursal	biblioteca propietaria	El código debe ser definido en Administración del Sistema → Bibliotecas, Sucursales y Grupos
952\$b	código de sucursal	holding biblioteca (usu. lo mismo que 952\$a)	Misma nota del subcampo anterior
952\$c	Código de ubicación de estantería	Valor codificado, que coincide con la lista de valores autorizados 'LOC'	
952\$d	Fecha de adquisición	AAAA-MM-DD	La fecha en que se compró o adquirió el artículo.
952\$e	Fuente de adquisición	valor codificado o cadena de proveedor	Nombre o abreviatura del editor, distribuidor o vendedor. Este también puede ser el código de proveedor del módulo de adquisiciones, pero actualmente hay pocas funciones que utilicen este campo.
952\$f	Calificador de ubicación codificada	Código de tres caracteres que identifica las emisiones específicas del artículo que se encuentran separadas de las existencias principales del mismo artículo.	Esto no tiene ninguna función en Koha.
952\$g	Precio de compra	Precio de compra original o precio de lista	número decimal, sin símbolo de moneda
952\$h	Leyenda de enumeración en serie	Información de volumen y emisión para artículos en serie.	
952\$i	Numero de inventario	El número de inventario.	No es repetible y no es obligatorio. Agregado con 3.4.0
952\$j	Número de control de estantería	El número de pila del artículo, código de inventario.	
952\$l	Pagos totales	Número total de pagos. Mostrar solo campo.	
952\$m	Renovaciones totales	Número total de renovaciones. Mostrar solo campo.	Este es el total acumulado. El número total de renovaciones no se restablece a 0 cuando se registra el elemento.
952\$n	Retenciones totales	Número total de retenciones. Mostrar solo campo..	
952\$o	Número completo de Koha		
952\$p *	código de barras	máximo 20 caracteres	
952\$q	Fecha de vencimiento	Fecha de vencimiento del artículo prestado. AAAA-MM-DD.	Cuando el artículo necesita devolverse de un préstamo. NULL si no se emite.
952\$r	Fecha de la última vez que se vio	La fecha en que el artículo se vio por última vez en la biblioteca (registrado/desprotegido/inventariado).	
952\$s	Fecha de última salida	Última fecha de pago del artículo. Mostrar solo campo.	
952\$t	número de copia	Copie el número de artículos que tienen la misma ubicación y número de clasificación.	
952\$u	Identificador uniforme de recursos	Una URL o URN, que proporciona datos de acceso electrónico en una sintaxis estándar.	
952\$v	Precio de reemplazo	número decimal, sin símbolo de moneda	Si un artículo se marca como perdido, este valor se carga al usuario.
952\$w	Precio efectivo desde	AAAA-MM-DD	Esta fecha se refiere al 'precio de compra normal'.
952\$x	Nota no pública	Nota del personal interno.	
952\$y *	Tipo de elemento Koha	valor codificado, campo obligatorio para la circulación	Valor codificado, debe ser definido en Administración del Sistema → Tipos de artículos y Códigos de Circulación
952\$z	Nota pública		

¿Qué nos falta?

- Generar el Manual de procedimientos;
 - Capacitar a los catalogadores a la nueva práctica;
 - Establecer los nuevos indicadores para el control de calidad;
 - Revisar el programa NACO para las autoridades;
-

Bibliografía

- Martín Gavilán, César. (2008). Catalogos y Gestión de Autoridades: Diseño y presentaciones de OPACs. [en temas de Bibliotecología]. 110 p. [en línea] <http://eprints.rclis.org/13188/1/sigb.pdf>
-