



## Conferência 10 anos do KOHA em Portugal

18 e 19 de maio de 2017

Instituto Politécnico de Lisboa (IPL)

Lisboa, Portugal

# Conversão de Registos



UNIVERSIDADE DE ÉVORA  
SERVIÇOS DE INFORMÁTICA

Nelson Madeira | [nmadeira@uevora.pt](mailto:nmadeira@uevora.pt)

# Processo de Migração para o KOHA

- instalação de KOHA no servidor
  - criação de instâncias de testes e produção
  - configurações várias... KOHA, Auth LDAP, Vhosts, SSL
  - configuração do Zebra
- 
- **análise e migração de dados**
- 
- customização do OPAC
  - testes e produção

# Análise e Migração de Dados

- **leitores**
- **registos bibliográficos e exemplares**
- **registos de autoridades (em fase de finalização)**

## Análise de Dados...

**...com a Biblioteca da Universidade  
... e com o Eng. Rafael António**

# Migração de Leitores

- **CSV com dados da plataforma atual**

- cruzamen

- scripts (PH

- funcional

- **identificação** (id\_leitor, nome, cod\_leitor)

- **contactos** (emails agrupados, moradas, telefones agrupados)

- **datas** (nascimento, inscrição, validade)

- **tipologia e género**

# Migração de Leitores

- CSV com dados da plataforma atual
- **cruzamento de dados dos leitores com SI da Universidade**
- scripts (PHP)
- funcionalidade
  - num\_cc
  - curso, grau, escola (nos alunos)

**Integrados em campos extra do KOHA**

# Migração de Leitores

- CSV com dados da plataforma atual
- cruzamento de dados dos leitores com SI da Universidade
- **scripts para tratamento de dados e geração de CSV final**
- funcionalidades
  - limpeza de dados *dummy* nos contactos
  - conversão de tipologia
  - separação de emails e telefones (antes agrupados)
  - geração do userid com base no email institucional
  - userid's duplicados ou inexistentes
  - cardnumber's duplicados ou inexistentes

**CSV final com base no "Starter CSV"  
disponibilizado pelo KOHA**

# Migração de Leitores

- CSV com dados da plataforma atual
- cruzamento de dados dos leitores com SI da Universidade
- scripts (PHP e SQL) para tratamento de dados e geração de CSV final
- **funcionalidade “Ferramentas > Importar Leitores” do KOHA**

- **match de leitores por cardnumber e userid (únicos)**
- **acrescentar campo “flags” sem conteúdo no CSV (permissão leitor)**  
**(falta do campo “flags” atribui a permissão default “Superlibrarian”)**

# Migração de Registos Bibliográficos e Exemplares

- exportação de dados da plataforma atual para ficheiros ISO
- importação de ISO's no **J-ISIS**
- análise estatística de tags UNIMARC
- tratamento e normalização de dados com **scripts Groovy no J-ISIS**
- exportação de dados do **J-ISIS** para um ISO final
- importação no KOHA com recurso ao “**bulkmarcimport.pl**”



J-ISIS – Java CDS/ISIS

<https://github.com/J-ISIS>



<http://groovy-lang.org>

# (J-ISIS...)



J-ISIS – Java CDS/ISIS

<https://github.com/J-ISIS>

## Pontos Fortes

- aplicação Java, multi-plataforma
- ferramenta auxiliar com várias funcionalidades integradas, pronta a utilizar
- importação de BD's de dados bibliográficos, visualização e pesquisa nas tags UNIMARC
- consola de execução de scripts Groovy/Java (estatísticas e conversão de registos)
- exportação para ISO2709 e MARCXML, passíveis de serem importados no KOHA

## Pontos “Fracos”

- tempo de execução do script de conversão de registos grande quando temos um numero grande de registos
- editor de scripts pouco avançando

## Alternativas

- editar os scripts com editores externos
- realização da execução por lotes de registos
- implementação de aplicação Java, com libs disponibilizadas pelo J-ISIS, para execução e debug externo ao J-ISIS



## J-ISIS 10 March 2016 Final Release:

---

Subfield pick lists, Numerous Enhancements, and bug fixes. Details are described in the release note included with the distribution. JDK 8 is needed.

### Windows, Linux release:

[https://github.com/J-ISIS/J-ISIS/releases/download/v1.0/jisis\\_suite.10.March.2016.zip](https://github.com/J-ISIS/J-ISIS/releases/download/v1.0/jisis_suite.10.March.2016.zip)

### The Mac OS X release:

<https://github.com/J-ISIS/J-ISIS/releases/download/v1.0/jisis-mac-10-March-2016.pdf>

<https://github.com/J-ISIS/J-ISIS/wiki>

## J-ISIS 10 Ma

---

<https://github.com/J-ISIS/J-ISIS/wiki>

## J-ISIS Reference Manual:

---

New chapters on "J-ISIS History Mechanism" and "Template Databases authority and liteBibliographic". Plus corrections, some updates, and details added.

<https://github.com/J-ISIS/J-ISIS/releases/download/v1.0/J-ISIS.Reference.Manual.23.March.2015.pdf>

# Migração de Registos Bibliográficos e Exemplares

- **exportação de dados da plataforma atual para ficheiros ISO**
- **importação de ISO's no J-ISIS**
- **análise** ▪ **exportação de registos bibliográficos em lotes de 50mil registos com todos os campos disponíveis**
- **tratamento** ▪ **importação dos ISO's no J-ISIS**
  - ficheiro ISO
  - encoding do ISO
  - formato: ISO2709, MARC, MARCXML, etc
  - tipo de import: Load, Merge, Update
  - nome da BD no J-ISIS
- **exportação**
- **importação**

# Migração de Registos Bibliográficos e Exemplares

- exportação
- importação
- análise es
- tra
- ex
- im

tag	num_ocorrencias	tipo_conteudo	conversao_ou_nova_tag
966	183191	label_form_prisma	campo 995
966^0	230524	Chave do Fundo(3) (valores de 1 a 14) (1-Geral,2-BGUE,3-Monografias, etc)	
966^1	131179	codigo de barras (enviados codigos duplicados ao António para	995^f
966^2	230524	Fundo(3) - Mediateca e outros	995^a e 995^c se ESESJD, 995^r se forem valores identificados na correspo
966^3	230524	Fundo(1) - BIB	
966^4	230522	Fundo(2) - BGUE Biblioteca Principal	
966^5	230524	??? (valores 1 e 2) em testes não se conseguiu ligar isto aos estados nem a nenhum outro campo	
966^a	230523	nº registo/inventário	995^6 (manter espaços)
966^b	208022	dt criação do registo no Prisma (redundante, descartar)	
966^c	194	numero (agrupar nº, volume e tomo com prefixos Tomo:, Vol: e Nº: - v	995^h
966^d	6781	volume (agrupar nº, volume e tomo com prefixos Tomo:, Vol: e Nº: - v	995^h
966^e	30	ano (2ª parte form exemplar)	concatenar no 995^z com a forma "Ano: ano"
966^f	83	ano de publicação (redundante, descartar)	
966^g	75	mês inicial (tinham dados importados - Trimestres/Quadrimestres - q n interessam)	
966^h	4	mês final (tinham dados importados - Trimestres/Quadrimestres - q n interessam)	
966^i	9	fascículo (enviados ao António para mudança de tag)	
966^j	nunca preenchido	série	
966^k	11	parte (valores 1 a 3) (enviados ao António para mudança de tag)	
966^l	231899	localização (converter os da ESE e descartar o resto)	995^a e 995^c (codigo de escola, default CES)
966^m	217354	tipo documento (no PRISMA tem datas de catalogação) (redundante,	995^r (vem do fundo(3) se for um valor identificado, se nao for o default será
966^n	144028	notas -- no PRISMA tem fornecedor + preço (pouco normalizado, ver	se o leader3007 = "m"(Monografias) verifica-se se existem ":" e se sim 995^
966^o	nunca preenchido	secção	
966^p	44165	preço -- no PRISMA tem o proponente	995^g
966^q	nunca preenchido	dia inicial	
966^r	nunca preenchido	dia final	
966^s	233562	cota (só tem CDU, passagem direta)	995^k, acrescentar CES ao início quando a cota inicia com um numero
966^t	3	outra numeração (descartar)	
966^u	241	tomo (agrupar nº, volume e tomo com prefixos Tomo:, Vol: e Nº: - ver	995^h
966^v	19	numeração alternativa (enviado para o António para alteração)	
966^w	33725	Forma de Aquisição (1-Assinatura,2-Compra,3-Depósito Legal,4-Ofert	995^x (Assinatura converte-se em Compra, Outra ignora-se e Depósito Lega
966^x	93	titulo (encontra no registo bibliográfico)	
966^y	32	url (enviados ao António para mudança de tag)	

# Migração de Registos Bibliográficos e Exemplares

- exportação de dados da plataforma atual para ficheiros ISO
- importação de ISO's no J-ISIS
- análise estatística de tags UNIMARC
- **tratamento e normalização de dados com scripts Groovy no J-ISIS**
- e
  - limpeza de pointers "<... >" nos campos de texto
- i
  - conversão da tag 966 para a 995 (exemplares no KOHA)
    - levantamento e normalização de "valores autorizados" – cotas, país, entidades proponentes, fornecedores
    - outras conversões nos exemplares (forma de aquisição, tipo doc, biblioteca de aquisição e empréstimo, proponente, fornecedor, preço e notas internas)

# Migração de Registos Bibliográficos e Exemplares

- exportação de dados da plataforma atual para ficheiros ISO
- importação de ISO's no J-ISIS
- análise estatística de tags UNIMARC
- tratamento e normalização de dados com scripts Groovy no J-ISIS
- **exportação de dados do J-ISIS para um ISO final**
- in
  - **funcionalidade "Database > Export database"**
  - **ficheiro "Reformatting FST" (Field Select Table)**
    - **seleção dos campos a exportar**
    - **descarte do campo 966, passagem apenas do 995 utilizado pelo KOHA**

# Migração de Registos Bibliográficos e Exemplares

- exportação de dados da plataforma atual para ficheiros ISO
- importação de ISO's no J-ISIS
- análise estatística de tags UNIMARC
- tratamento e normalização de dados com scripts Groovy no J-ISIS
- exportação de dados do J-ISIS para um ISO final
- **importação no KOHA com recurso ao "bulkmarcimport.pl"**

**...precar "Configuração da ligação de subfields da tag 995 com os campos da tabela Items do KOHA" antes do import**

```
perl bulkmarcimport.pl -c=UNIMARC -m=ISO2709 -biblios -nocleanisbn -file file.iso
```

**[-c] MARC flavour**

**[-m] formato do ISO - ISO2709, ...**

**[-biblios] tipo de importação – registos bibliográficos**

**[-file] ficheiro a importar**

# ... J-ISIS na prática

# Principais Desafios

(...para um técnico não bibliotecário)

- **aquisição de know-how na “representação” UNIMARC** (só possível com o contacto direto e próximo com os bibliotecários)
- **aquisição de know-how na plataforma KOHA** (conjunto vasto de funcionalidades e configurações)
- **análise de dados e esforço de normalização/conversão de dados**

# Obrigado!

Nelson Madeira  
nmadeira@uevora.pt