

# Provedor de Dados Instalação e Configuração

---

João Lima

# Módulos

---

- Validador
    - Banco de Dados com registro de metadados
    - Programa que valida a estrutura e o perfil dos publicadores
  - Interface WEB OAI-PMH
    - Expõe os metadados utilizando o protocolo OAI-PMH
-

# Validador - Passos

---

## Instalação


1. Criar conta no banco de dados
2. Obter o arquivo de profile
3. Copiar o arquivo de configuração
4. Configurar o arquivo `lexm-db.properties`
5. Configurar os scripts de validação e instalação
6. Criar as tabelas

## Carga dos Metadados

7. Inserir metadados na base de dados

## Execução do Validador

8. Validar os metadados inseridos
-



# Instalação e Configuração do Validador

# 1. Criando conta no banco de dados e database

---

Usando o Oracle:

exemplo utilizando o sql\*plus:

```
SQL> connect system
```

\*digite a senha

```
SQL>
```

```
CREATE USER lexmluser IDENTIFIED BY lexmlpassword
```

```
DEFAULT TABLESPACE users
```

```
TEMPORARY TABLESPACE temp
```

```
QUOTA UNLIMITED ON users;
```

```
SQL> GRANT CONNECT TO lexmluser;
```

Usando o Mysql:

```
usuario@localhost>mysql -uroot -p
```

```
mysql> create database lexmldatabase
```

```
mysql> GRANT ALL PRIVILEGES ON lexmldatabase.* to 'lexmluser'@'localhost'  
IDENTIFIED BY 'lexmlpassword' WITH GRANT OPTION;
```

```
mysql> FLUSH PRIVILEGES;
```

---

## 2. Obtendo o arquivo de profile

---

- ❑ O arquivo de profile define as combinações possíveis dos elementos localidade, autoridade e tipo de documento de um publicador.
  - ❑ Obs: Na fase de implantação, este arquivo XML será enviado por e-mail a partir do preenchimento da planilha "Perfil Provedor Dados.xls"
-

	A	B	C	D	E	F	G
1	Definição de Profile de Provedor de Dados LexML						
2							
3		Obs: Enviar esta planilha para <a href="mailto:lexml@senado.gov.br">lexml@senado.gov.br</a>					
4							
5		<b>Provedor de Dados (PD)</b>					
6			Nome do Órgão:	Senado Federal			
7			Administrador do PD:	João Lima			
8			Email:	<a href="mailto:joaolima@senado.gov.br">joaolima@senado.gov.br</a>			
9			Telefone:	(61) 3311-2254			
10							
11		<b>Publisher 1</b>					
12			Nome da Seção:	Subsecretaria de Informações			
13			Sigla da Seção:	SSINF			
14			Responsável:	Antonio José Viana Filho			
15			Email:	<a href="mailto:ajosevia@senado.gov.br">ajosevia@senado.gov.br</a>			
16			Telefone:	(61) 3311-3105			
17							
18		<b>Publisher 2</b>					
19			Nome da Seção:	Secretaria Geral da Mesa			
20			Sigla da Seção:	SGM			
21			Responsável:	Flávio Heringher			
22			Email:	<a href="mailto:heringher@senado.gov.br">heringher@senado.gov.br</a>			
23			Telefone:	(61) 3311-3939			
24							

	A	B	C	D	E	F
1	Localidade		Autoridade		Tipo Documento	
2	Nome	URN	Nome	URN	Nome	URN
3	Brasil	br	Federal	federal	Lei	lei
4	Brasil	br	Federal	federal	Decreto	decreto
5	Brasil	br	Federal	federal	Lei Complementar	lei.complementar
6	Brasil	br	Federal	federal	Lei Delegada	lei.delegada
7	Brasil	br	Federal	federal	Constituição	constituicao
8	Brasil	br	Federal	federal	Emenda Constitucional	emenda.constitucional
9	Brasil	br	Federal	federal	Decreto Lei	decreto.lei
10	Brasil	br	Federal	federal	Medida Provisória	medida.provisoria
11	Brasil	br	Senado Federal	senado.federal	Resolução	resolucao
12	Brasil	br	Câmara dos Deputados	camara.deputados	Resolução	resolucao
13						
14						
15						

	A	B	C	D	E	F	G
1	Localidade		Autoridade		Tipo Documento		Subtipo
2	Nome	URN	Nome	URN	Nome	URN	Nome
3	Brasil	br	Senado Federal	senado.federal	Projeto de Lei	projeto.lei	
4	Brasil	br	Senado Federal	senado.federal	Projeto de Lei Complementar	projeto.lei.complementar	
5	Brasil	br	Senado Federal	senado.federal	Proposta de Emenda Constitucional	proposta.emenda.constitucional	
6	Brasil	br	Senado Federal	senado.federal	Projeto de Lei de Conversão	projeto.lei.conversao	
7	Brasil	br	Senado Federal	senado.federal	Projeto de Decreto Legislativo	projeto.decreto.legislativo	
8	Brasil	br	Senado Federal	senado.federal	Projeto de Lei;PLC	projeto.lei;plc	Projeto de Lei da Câmara
9							
10							
11							
12							
13							
14							
15							



### 3. Copiar o arquivo de configuração do banco de dados

---

- Se estiver utilizando Oracle:

- Copie o arquivo  
lexml-db-ORACLE.properties  
para:  
lexml-db.properties

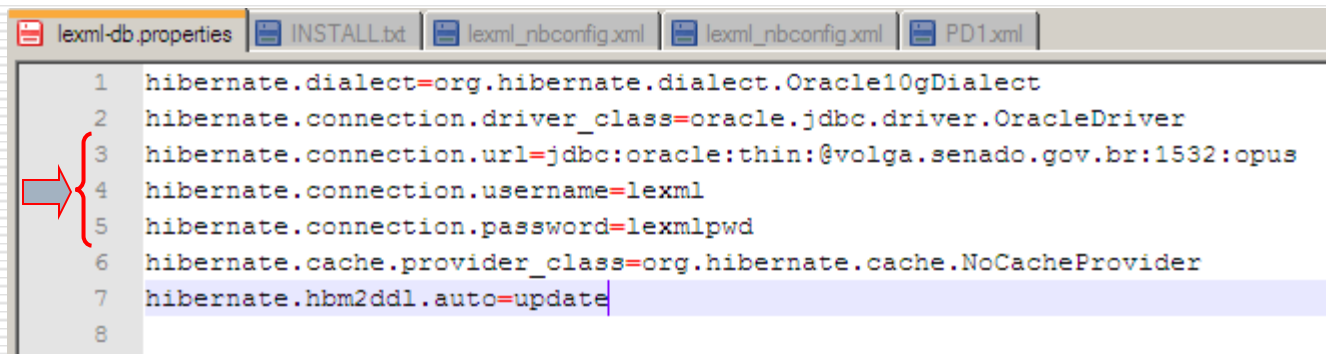
- Se estiver utilizando Mysql:

- Copie o arquivo  
lexml-db-MYSQL.properties para:  
lexml-db.properties
-

## 4. Configurar o arquivo lexm-db.properties

---

- Edite o arquivo recém criado lexml-db.properties substituindo
  - o usuario,
  - a senha do usuário e
  - a string de conexão jdbc.



```
1 hibernate.dialect=org.hibernate.dialect.Oracle10gDialect
2 hibernate.connection.driver_class=oracle.jdbc.driver.OracleDriver
3 hibernate.connection.url=jdbc:oracle:thin:@volga.senado.gov.br:1532:opus
4 hibernate.connection.username=lexml
5 hibernate.connection.password=lexmlpwd
6 hibernate.cache.provider_class=org.hibernate.cache.NoCacheProvider
7 hibernate.hbm2ddl.auto=update
8
```

## 5. Configurando os scripts

---

- ❑ Se Windows
    - Edite os arquivos install.bat e validar.bat
  - ❑ Se Unix/Linux
    - Edite os arquivos install.sh e validar.sh
  - ❑ Atualize a variável **LEXML\_HOME** com o caminho do diretório onde o validador foi instalado.
-

## 6. Criar as tabelas no banco

---

- Se Windows:
    - Execute o arquivo instalar.bat
  - Se unix/linux:
    - Execute o arquivo instalar.sh
-



Carga do banco de dados

# 7. Inserir metadados na base de dados.

---

- Se o banco é MySQL utilize a partir da sua aplicação de exportação de metadados inserts do tipo:
    - ```
INSERT INTO `registro_item` (`id_registro_item`,  
`tx_metadado_xml`, `cd_status`, `cd_validacao`,  
`ts_registro_gmt`)  
VALUES('oai:ssinf.senado.gov.br:njur/112321',  
'$OAI_LEXML', 'N', 'I', UTC_TIMESTAMP());
```
  
  - Se o banco é Oracle a instrução modifica-se pouco:
    - ```
INSERT INTO registro_item (id_registro_item,  
tx_metadado_xml, cd_status, cd_validacao ,  
ts_registro_gmt)  
VALUES('oai:ssinf.senado.gov.br:njur/112321',  
'$OAI_LEXML', 'N', 'I', sysdate);
```
-

# Sobre os dados

---

- A regra de formação da coluna `id_registro_item` é:
    - `id_registro_item := prefixo_oai + "/" + chave`
    - `prefixo_oai := 'oai:' + [orgao + "."] + dominio + ":" + sistema`
    - `chave := identificador_interno + [ ";" + detalhe]`
    - `identificador_interno :=` é a chave do sistema de origem do registro
    - EXEMPLO: `'oai:ssinf.senado.gov.br:njur/112321'`
  - O valor 'N' da coluna `cd_status` significa que o registro é novo.
  - O valor 'I' da coluna `cd_validacao` significa que o registro é pendente de validação.
  - A coluna `ts_registro_gmt` deve receber sempre a data atual UTC ('2008-08-08 08:08:08') dependendo do horário de verão são 2 ou 3 horas a mais do que o horário oficial de Brasília.
  - A coluna `tx_metadado_xml` deve receber o arquivo XML com os metadados de acordo com o schema `lexml_oai.xsd`
    - [http://www.lexml.gov.br/sch/oai\\_lexml.xsd](http://www.lexml.gov.br/sch/oai_lexml.xsd)
-

# XML de Metadados

---

```
<lexml:LexML xsi:schemaLocation="http://www.lexml.gov.br/sch/oai_lexml.xsd"
  xmlns:lexml="http://www.lexml.gov.br/oai_lexml"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
  <Item formato="text/html" idPublicador="1" tipo="conteudo">
    http://www6.senado.gov.br/legislacao/ListaPublicacoes.action?id=132554
  </Item>
  <Item formato="text/html" idPublicador="1" tipo="metadado">
    http://www6.senado.gov.br/legislacao/DetalhaDocumento.action?id=132554
  </Item>

  <DocumentoIndividual>urn:lex:br:federal:lei:1988-12-02;7682@inicio.vigenci
a;publicacao;1988-12-05~texto;pt-br </DocumentoIndividual>
<Epigrafe>Lei nº 7.682, de 02 de dezembro de 1988</Epigrafe>
<Ementa>ALTERA O DECRETO-LEI 2.406, DE 5 DE JANEIRO DE 1988, E DA
OUTRAS PROVIDENCIAS.</Ementa>
<Relacionamento
  tipo="publicacao.oficial">urn:lex:br:imprensa.nacional:publicacao.oficial;diario
.oficial.uniao:1988-12-05@1988-12-05;publicacao;1988-12-05~texto;pt-br</
Relacionamento>
</lexml:LexML>
```

---





# Execução do Validador

## 8. Validando os registros novos no banco de dados:

---

- ❑ Execute:
    - ❑ `validar.bat` (se windows)
    - ❑ `sh ./validar.sh` (se unix/linux)
  - ❑ Durante o processamento são gerados os arquivos `validacao_err.txt` e `validacao_log.txt`.
    - O primeiro é um relatório dos erros identificados durante a validação e o segundo somente um log dos registros processados.
    - Utilize o arquivo `validacao_err.txt` para identificar os registros com erros.
  - ❑ O Log e os Erros também são registrados no banco de dados.
-

# Log

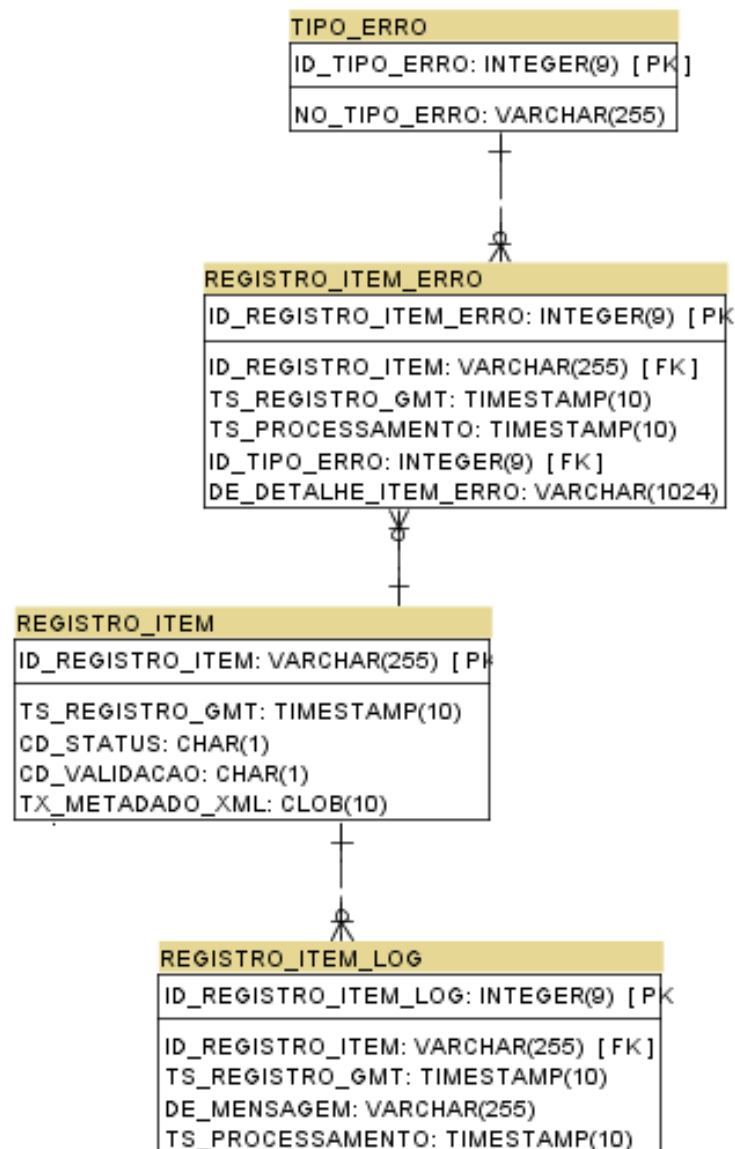
---

```
INFO : Validação com sucesso do registro :  
oai:ssinf.senado.gov.br:njur/132630  
2008-10-16 12:48:04,842 [ValidadorBO.java:204]  
INFO : Validação com sucesso do registro :  
oai:ssinf.senado.gov.br:njur/132631  
2008-10-16 12:48:05,561 [ValidadorBO.java:242]  
INFO : Total de registros processados: 525  
2008-10-16 12:48:05,561 [ValidadorBO.java:243]  
INFO : Processados com sucesso: 525  
2008-10-16 12:48:05,561 [ValidadorBO.java:244]  
INFO : Registros que apresentavam erros: 0  
2008-10-16 12:48:05,561 [ValidadorBO.java:246]  
INFO : Aproveitamento de aproximadamente: 100%
```

	ID_TIPO_ERRO	NO_TIPO_ERRO	
▶ 1	1	Erro Genérico	...
2	2	XML mal formado	...
3	3	XML inválido	...
4	4	URN mal formado	...
5	5	URN inválido	...
6	6	URN incompatível com o perfil do provedor	...

	ID_REGISTRO_ITEM	TS_REGISTRO_GMT	CD	CD	TX_METADA	
1	oai:ssinf.senado.gov.br:njur/223828	16/10/2008 12:48:01	N	O	<CLOB>	...
2	oai:ssinf.senado.gov.br:njur/132445	16/10/2008 12:48:01	N	O	<CLOB>	...
▶ 3	oai:ssinf.senado.gov.br:njur/132484	16/10/2008 12:48:01	N	O	<CLOB>	...
4	oai:ssinf.senado.gov.br:njur/132506	16/10/2008 12:48:01	N	O	<CLOB>	...
5	oai:ssinf.senado.gov.br:njur/132525	16/10/2008 12:48:01	N	O	<CLOB>	...
6	oai:ssinf.senado.gov.br:njur/132552	16/10/2008 12:48:01	N	O	<CLOB>	...
7	oai:ssinf.senado.gov.br:njur/132553	16/10/2008 12:48:01	N	O	<CLOB>	...
8	oai:ssinf.senado.gov.br:njur/132554	16/10/2008 12:48:01	N	O	<CLOB>	...
9	oai:ssinf.senado.gov.br:njur/132555	16/10/2008 12:48:01	N	O	<CLOB>	...
10	oai:ssinf.senado.gov.br:njur/132556	16/10/2008 12:48:01	N	O	<CLOB>	...
11	oai:ssinf.senado.gov.br:njur/132557	16/10/2008 12:48:01	N	O	<CLOB>	...

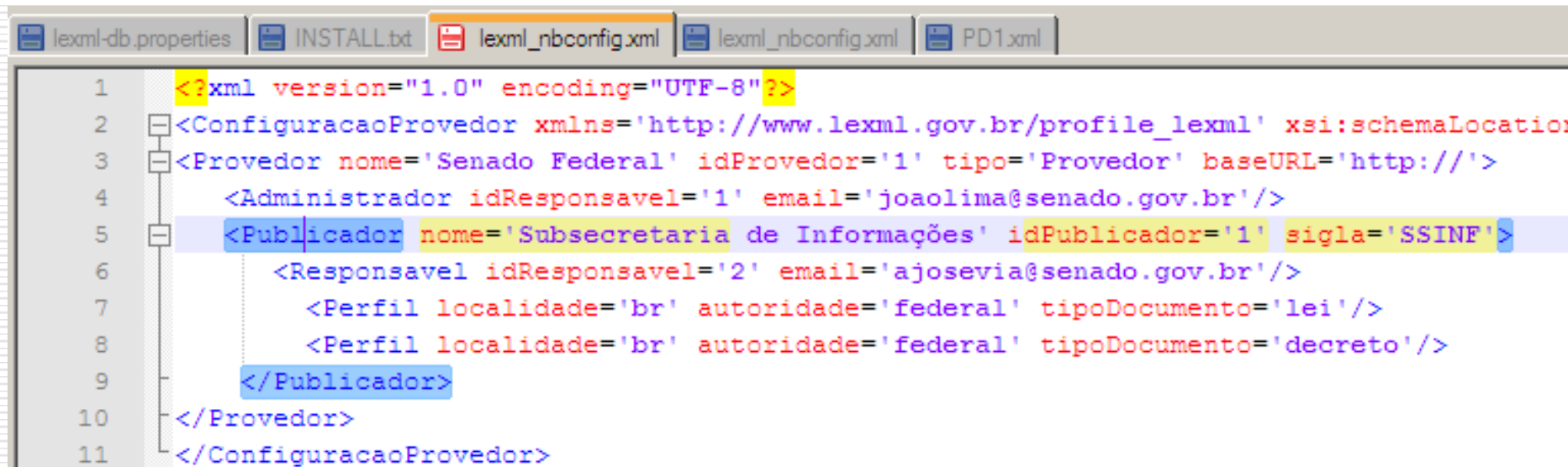
	ID_REGISTRO_ITEM_LOG	ID_REGISTRO_ITEM	TS_REGISTRO_GMT	DE_MENSAGEM	TS_PROCESSAMEN
▶ 1	225	oai:ssinf.senado.gov.br:njur/132649	16/10/2008 12:43:32	RegistroItem validado com sucesso	16/10/2008 12:48:03
2	226	oai:ssinf.senado.gov.br:njur/132650	16/10/2008 12:43:32	RegistroItem validado com sucesso	16/10/2008 12:48:03
3	227	oai:ssinf.senado.gov.br:njur/223558	16/10/2008 12:43:32	RegistroItem validado com sucesso	16/10/2008 12:48:03
4	228	oai:ssinf.senado.gov.br:njur/223559	16/10/2008 12:43:32	RegistroItem validado com sucesso	16/10/2008 12:48:03
5	229	oai:ssinf.senado.gov.br:njur/223560	16/10/2008 12:43:32	RegistroItem validado com sucesso	16/10/2008 12:48:03



# Arquivo lexml\_nbconfig.xml

## Perfis dos publicadores

---



```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <ConfiguracaoProvedor xmlns='http://www.lexml.gov.br/profile_lexml' xsi:schemaLocation='http://www.lexml.gov.br/profile_lexml http://www.lexml.gov.br/profile_lexml.xsd'>
3 <Provedor nome='Senado Federal' idProvedor='1' tipo='Provedor' baseUrl='http://www.senado.gov.br'>
4 <Administrador idResponsavel='1' email='joaolima@senado.gov.br'>
5 <Publicador nome='Subsecretaria de Informações' idPublicador='1' sigla='SSINF'>
6 <Responsavel idResponsavel='2' email='ajosevia@senado.gov.br'>
7 <Perfil localidade='br' autoridade='federal' tipoDocumento='lei'>
8 <Perfil localidade='br' autoridade='federal' tipoDocumento='decreto'>
9 </Publicador>
10 </Provedor>
11 </ConfiguracaoProvedor>
```



# Interface WEB – OAI-PMH

- 
- ❑ Criar variável de ambiente
    - LEXML\_HOME com o valor do diretório de instalação do validador
  - ❑ Copiar arquivo oai.war para um container java
    - ❑ Tomcat, Jetty
    - ❑ OC4J, em breve
-