

ExtGASoai: OAI-PMH gateway per ExtGAS - manuale di installazione e d'uso

luglio 2007

Giuseppe A. Romano
Istituto di Scienza e Tecnologie dell'Informazione,
e.mail giuseppe.romano@isti.cnr.it

Indice

0. Premessa	pag	4
1. ExtGAS OAI-PMH gateway	pag	5
2. Il software ExtGAS OAI-PMH gateway	pag	7
3. I prerequisiti di installazione e l'installazione	pag	8
4. La security di ExtGAS OAI-PMH gateway	pag	13
5. La selezione dell'informazione	pag	14
5.0. Il modello di selezione e accesso ai dati	pag	14
5.1. L'organizzazione della lista dei risultati	pag	16
5.2. La struttura dell'OAI-PMH identifier	pag	16
5.3. Il modello di rappresentazione delle date richiesto dal gateway	pag	16
5.4. Il modello di restituzione di Identifiers/Records	pag	17
5.5. Il restart dell'accesso ai dati	pag	18
5.6. Le API CDS/ISIS	pag	23
6. L'utilizzo del proxy-server	pag	25
7. I parametri della OAI System Table	pag	27
7. 0. I parametri generali	pag	30
7. 1. I parametri del repository	pag	33
7. 2. I parametri di restituzione	pag	37
7. 3. I parametri dei record identifier	pag	40
7.4. I parametri dei Set	pag	41
7. 5. I parametri dei resumptionToken	pag	42
8 .Bibliografia	pag	44

O.Premessa

ExtGASoai è il gateway OAI-PMH per il sistema **ExtGAS/IsisGAS**. **ExtGAS/IsisGAS** è composto dal sistema **IsisGAS** descritto in "Isis/GAS : un sistema multi piattaforma per l'accesso a basi di dati testuali via Internet - ISTI-B4-28, 2003" e **ExtGAS** che accede a sistemi catalografici memorizzati su database system diversi da CDS/ISIS; vedi : " ExtGAS un sistema multipiattaforma per l'accesso ai cataloghi di sistemi biblioteconomici eterogenei - ISTI-B4-15,2005 ". Naturalmente il gateway si applica alla parte **IsisGAS** del sistema **ExtGAS/IsisGAS**.

L'implementazione del gateway si è resa necessaria dato l'utilizzo del sistema ExtGAS/IsisGAS per la selezione e l'accesso alla Digital Library delle pubblicazioni degli Istituti del CNR partecipanti al progetto Metapub. Tramite il gateway è stato/è tuttora possibile esportare i metadati che descrivono le pubblicazioni verso altri servizi integrati a livello italiano e europeo.

Nella descrizione delle procedure di installazione e configurazione del gateway gli esempi fanno riferimento all'ambiente Unix, ma sono validi anche per ambiente Windows. Non esiste un ambiente preferenziale.

Si presuppone che il lettore abbia conoscenza di OAI-PMH Protocol version 2.0 del 2002-06-14 e del sistema CDS/ISIS, delle sue procedure di generazione, indicizzazione e ordinamento, e conosca, inoltre, il linguaggio di stampa/formattazione dei dati (vedi Bibliografia).

Nel seguito si parlerà indifferentemente di:

1. OAI System File
2. OAI System Table
3. Database System Table
4. Database System File
5. Configuration File

Intendendo il file di configurazione che definisce l'applicazione **ExtGASoai** verso un particolare database CDS/Isis. Lo stesso software può gestire più di un gateway OAI/database per volta. Da ciò consegue che lo OAI System File definisce il database da utilizzare, le modalità di selezione e accesso ai dati e le modalità di restituzione secondo lo standard OAI-PMH Protocol version 2.0 del 2002-06-14, in sostanza definisce un **repository** che si appoggia su un database CDS/Isis.

1. ExtGAS OAI-PMH gateway

Il gateway sviluppato implementa i seguenti verbi del OAI-PMH Protocol version 2.0 del 2002-06-14

1. **Identify** che accede alle informazioni relative al repository di competenza.

Di default il repository gestito da gateway ha le seguenti caratteristiche:

```
<repositoryName>ExtGAS/ISIS Open Archive Initiative Repository </repositoryName>
<baseURL>http://leonardo.isti.cnr.it/oai</baseURL>
<protocolVersion>2.0</protocolVersion>
<adminEmail>giuseppe.romano@isti.cnr.it</adminEmail>
<earliestDatestamp>1977-08-15</earliestDatestamp>
<granularity>YYYY-MM-DD</granularity>
<deletedRecord>no</deletedRecord>
```

Fondamentalmente sono definite:

1. la data più bassa rispetto a tutti record che si necessita di esportare
 2. la più fine granularità supportata dal repository.
 3. Il modo con cui il sistema gestisce i record cancellati
2. **ListIdentifiers** per accedere al contenuto degli identificatori dei record e inviarli all'Harvester client. In fase di configurazione del gateway è possibile definire il massimo numero di identifier da inviare sia la prima che le successive volte per soddisfare la richiesta (il default è 100) che può selezionare un massimo numero di record (il default è 10000).

Sono accettati i seguenti parametri:

1. **metadataPrefix** che identifica il formato di restituzione dei record: è **obbligatorio**
 2. **set** che identifica, se esiste, un set o una gerarchia di set che permettono di selezionare un particolare sottoinsieme dei record
 3. **resumptionToken** che definisce il restart-download pointer: esso è esclusivo cioè non accetta alcun parametro aggiuntivo in quanto si riferisce a una selezione già di per sé definita.
 4. **from** che identifica il from-date da cui partire a selezionare i record
 5. **until** che identifica il to-date fino a cui selezionare i record
3. **ListRecords** per accedere al contenuto dei record e inviarli all'Harvester client nel **metadataPrefix** specificato. In fase di configurazione del gateway è possibile definire il massimo numero di record da inviare sia la prima che le successive volte per soddisfare la richiesta (il default è 100) che può selezionare un massimo numero di record (il default è 10000).

Sono accettati i seguenti parametri:

1. **metadataPrefix** che identifica il formato di restituzione dei record: **è obbligatorio**
 2. **set** che identifica, se esiste, un set o una gerarchia di set che permettono di selezionare un particolare sottoinsieme dei record
 3. **resumptionToken** che definisce restart-download pointer: esso è esclusivo cioè non accetta alcun parametro aggiuntivo in quanto di riferisce a una selezione già di per se definita.
 4. **from** che identifica il from-date da cui partire a selezionare i record
 5. **until** che identifica il to-date fino a cui selezionare i record
4. **GetRecord** per accedere selettivamente al contenuto di record e inviarlo all'Harvester client nel **metadataPrefix** specificato.

Sono accettati i seguenti parametri:

1. **identifier** che rappresenta l'identificatore del record: **è obbligatorio**
 2. **metadataPrefix** che identifica il formato di restituzione dei record: **è obbligatorio**
5. **ListSets** per accedere alla descrizione dei **Set** e **Subset** cui possono appartenere i record.

Di default possono essere definiti un massimo di 100 set, ma il valore può essere modificato in fase di configurazione del repository. Si tenga presente che il **resumptionToken** che definisce il restart-download pointer **non è supportato** per il verbo ListSet

6. **ListMetadataFormats** che descrive le modalità di restituzione dei metadati (formati di restituzione)

Sono accettati i seguenti parametri:

1. **identifier** che rappresenta l'identificatore del record: **è opzionale**

2. Il software ExtGAS OAI-PMH gateway

Il software ExtGASoai è costituito da un insieme di moduli java che implementano i "verbi" OAI e da una API CDS/ISIS per accedere ai dati e produrre i risultati secondo lo standard OAI:

1. **ExtGASoai.java** è il driver principale del sistema
2. **ExtGASoai_v_Id.java** gestisce il verbo Identify
3. **ExtGASoai_v_LI.java** gestisce il verbo ListIdentifiers
4. **ExtGASoai_v_LS.java** gestisce il verbo ListSets
5. **ExtGASoai_v_LMF.java** gestisce il verbo ListMetadataFormats
6. **ExtGASoai_v_LR.java** gestisce il verbo ListRecords
7. **ExtGASoai_v_GR.java** gestisce il verbo GetRecord
8. **ExtGASoaiUtil.java** è un insieme di Utility generali
9. **ExtGASoaiIsisQuery.java** costruisce le query e esegue gli accessi a CDS/ISIS
10. **ExtGASoaiWriteResults.java** produce i risultati da restituire via protocollo OAI-PMH
11. **ExtGASoaiSysTable.java** descrive e definisce i parametri generali del sistema in modo che sia possibile la personalizzazione del Gateway OAI-PMH.
12. **AccessIsis.java** accede, mediante l'API CDS/ISIS, ai dati.
13. **<wwwisis>** che è la API CDS/ISIS per accedere al database. Esso corrisponde al programma eseguibile `www-isis`, distribuito free nella versione 3.0, per piattaforme unix-linux, unix-solaris, unix-ibmaix, unix-hp e windows, da www.unesco.org e prodotto da BIREME www.bireme.br

Nel complesso il software è costituito da:

1. I driver: da 1 a 7 nella lista precedente
2. I servizi: da 9 a 10 nella lista precedente
3. Le Utility: 8 nella lista precedente
4. Le definizioni: 11 nella lista precedente
5. L'API: 13 nella lista precedente

Il software richiede l'installazione di java e di un java http server per rendere operativo il gateway

Il test del gateway è stato certificato da <http://re.cs.uct.ac.za/> : Computer Science Dpt.
- Cape Town University.

Il software può essere scaricato alla seguente URL:

<http://leonardo.isti.cnr.it/metaopac/MetaopacDirectorySaver/solaris9/ExtGASoai.src.tgz>

3. I prerequisiti di installazione e l'installazione

ExtGAS OAI-PMH gateway è multiplatforma e può essere installato:

1. Su sistemi Microsoft Windows 9x, NT, 200x, XP (per il momento non è stato verificato per il nuovo sistema Windows Vista).
2. Su sistemi Unix Linux, Sun solaris, IBM AIX, HP Unix compatibilmente con le API CDS/Isis (programma [www-isis](http://www-isis.org)).

Essendo il gateway scritto in java è necessario installare sul server prescelto almeno la Java Runtime Engine (JRE) per java :

<http://java.sun.com/>

o qualunque altro java runtime disponibile maggiore o uguale al release 1.5.0.

Inoltre è necessario disporre di un Java http server o (Jserver) che può essere prelevato da :

<http://tomcat.apache.org/>

e configurato.

E' necessario installare sul server il/i relativi database CDS/Isis unitamente al gateway [www-isis](http://www-isis.org) 3.0 che viene distribuito da

www.unesco.org/

Se è necessario ricaricare i database di test o caricare nuovi database bisogna scaricare le CISIS utility unitamente alla loro documentazione dal sito www.bireme.br (successivamente passare nell'ambiente **Isis Family**). Se il software non fosse disponibile scaricare i seguenti contenitori:

1. <http://leonardo.isti.cnr.it/metaopac/MetaopacDirectorySaver/windows2000-prof/BiremePrograms.zip>
2. <http://leonardo.isti.cnr.it/metaopac/MetaopacDirectorySaver/windows2000-prof/BiremePrograms.zip>

In essi sono contenuti le **wwwisis API** e le **CISIS Utility** unitamente alla documentazione disponibile.

Supposto di avere installato il Jserver, la Java Runtime Engine e/o il Java Development Kit (JDK) e il [wwwisis](http://www-isis.org) 3.0 si scarichi l'ExtGAS OAI gateway da:

<http://leonardo.isti.cnr.it/metaopac/MetaopacDirectorySaver/solaris9/ExtGASoai.src.tgz>

lo si decompila. Il file di distribuzione contiene:

1. La directory **ExtGASoai.src** che contiene:
 1. il source del software e le relative classi Java.
 2. La directory **ProxyDefinition** che contiene le definizioni del ProxyServer
 3. La directory **savedir** che contiene versioni obsolete dei componenti del software
 4. La directory **IsisDbdir.oai** che contiene un piccolo database di prova da poter usare con il servizio `http://<host>/oai_ienige`
 5. La directory **Docs** che contiene la documentazione del sistema:
 1. ExtGAS_OAI.pdf
 2. ExtGAS_OAI_install.pdf
 3. ExtGAS_OAI_logic.pdf
2. La directory **oai** che è la directory di configurazione del servlet **ExtGASoai**, e contiene:
 1. La directory **WEB-INF/classes** che contiene le classi java costituenti il servlet
 2. Il file **WEB-INF/web.xml** che contiene la definizione del servlet
3. La directory **oaiConf** che attiva (via Proxy-server) il servizio `http://leonardo.ist.cnr.it/oai` verso il database che contiene i dati del progetto Metapub e viene usato come sito di test.
4. La directory **oai_ienigeConf** che attiva (via Proxy-server) il servizio `http://leonardo.ist.cnr.it/oai_ienige` verso il database che contiene i dati del progetto Metapub e viene simulato come **repository** di un unico Istituto (IENIGE) via parametro **GeneralQuerySuffix..**
5. La directory **oai_cnrpubConf** che attiva (via Proxy-server) il servizio `http://leonardo.ist.cnr.it/oai_cnrpub` verso il database che contiene i dati del progetto Metapub e vengono simulati più **repository** logici di un unico database fisico via specifiche definizioni di **set** (è operativo e viene usato come repository per lo sviluppo del software BRICKS)
6. La directory **oai_cnrpubdriverConf** che attiva (via Proxy-server) il servizio `http://leonardo.ist.cnr.it/oai_cnrpubdriver` verso il database che contiene i dati del progetto Metapub e vengono simulati più **repository** logici di un unico database fisico via specifiche definizioni di **set** (è operativo e viene usato come repository per la dissemination del progetto DRIVER)
7. La directory **oai_cnrPisaLibraryConf** che attiva (via Proxy-server) il servizio `http://leonardo.ist.cnr.it/oai_cnrPisaLibrary` verso il database che contiene i metadati del database del Catalogo cumulato dell'Area di Ricerca di Pisa; vengono simulati più **repository** logici di un unico database fisico via specifiche definizioni di **set** (Monografie, Periodici, Pubblicazioni).
8. La directory **oai_iitPisaLibraryConf** che attiva (via Proxy-server) il servizio `http://leonardo.ist.cnr.it/oai_cnrPisaLibrary` verso il database che contiene i metadati del database del Catalogo cumulato dell'Area di Ricerca di Pisa e viene simulato come **repository** di un unico Istituto (IIT) via parametro **GeneralQuerySuffix** e utilizzo di **set** (Monografie, Periodici, Pubblicazioni).

Nel seguito descriveremo l'installazione riferita a un Jserver apache-tomcat installato su una macchina Unix Linux. In tutti i casi le indicazioni sono valide per ambienti Unix e Windows dove, nel caso di link simbolici si provveda a copiare il contenuto della directory cui fa riferimento il link.

Verificata l'installazione dell'apache tomcat:

1. Nella directory <apache-tomcat>/webapps si copi, per ambiente windows, la directory <distrib>oai rinominandola ExtGASoai, mentre per ambiente Unix si esegua il seguente link simbolico:

```
In -s <distrib>/oai <apache-tomcat>webapps/ExtGASoai
```

2. Si faccia restart dell'apache-tomcat
3. Si apra un browser alla seguente URL:

```
http://<host>:8080/ExtGASoai/servlet/oai?Conf=<distrib>/mpisa/oaiConf/oai.system.file&verb=Identify
```

che dovrebbe produrre il seguente risultato

```
<?xml version="1.0" encoding="UTF-8" ?>
- <!--
OAlpmh gateway versus CDS/ISIS, November 2006
By Beppe Romano giuseppe.romano@isti.cnr.it
-->
-<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2007-07-16T17:11:19+0200</responseDate>
  <request verb="Identify">http://leonardo.isti.cnr.it/oai</request>
- <Identify>
  <repositoryName>ExtGAS/ISIS Open Archive Initiative Repository</repositoryName>
  <baseUrl>http://leonardo.isti.cnr.it/oai</baseUrl>
  <protocolVersion>2.0</protocolVersion>
  <adminEmail>giuseppe.romano@isti.cnr.it</adminEmail>
  <earliestDatestamp>1977-08-15</earliestDatestamp>
  <granularity>YYYY-MM-DD</granularity>
  <deletedRecord>no</deletedRecord>
  </Identify>
</OAI-PMH>
```

Che sta a indicare che il gateway risponde e risponde anche ai verbi ListSets e ListMetadataFormats, mentre per gli altri verbi è necessario configurare correttamente l'OAI System Table.

In tutti i casi se si omettesse il file di configurazione si avrebbe il seguente messaggio di errore che anch'esso è una dimostrazione del funzionamento del gateway:

System configuration file does not exist: /oai/oaiConf/oai.system.file

OAI System File: /oai/oaiConf/oai.system.file does not exist - OAIpmh gateway versus CDS/ISIS was not started

A questo punto si deve passare alla configurazione completa dell'OAI System Table supposto di avere già disponibile un CDS/ISIS database.

Nel caso il database non esista ancora è possibile utilizzare il database di test presente nella directory <distrib>/ExtGASoai.src/IsisDbdir.oai.

Si tenga presente che il database è stato generato in ambiente Unix Linux-debian. Se fosse necessario trasportarlo in altro ambiente è indispensabile aggiornare i programmi di generazione nella procedura loadiso.cnr e genproc.cnr in accordo con il sistema operativo ospite, lo stesso deve avvenire per la directory <distrib>/ExtGASoai.src/IsisDbdir.oai/gizmodb che contiene il database adatto a costruire l'inverted file con i diacritici dell'ISO-8859-1 e per il database adatto alla traduzione dei caratteri diacritici nelle codifiche XHTML.

Ovviamente, e solo in questo caso, i database andranno rigenerati.

Si faccia riferimento alla directory <distrib>/oai_ienigeConf. Essa contiene i seguenti file:

1. cnr_li.pft è il Print Format per il verbo ListIdentifiers
2. cnr.txt è il CDS/ISIS cipar file name che contiene i riferimenti alla collocazione dei file che vengono usati dalla API wwwisis. In esso vanno aggiornati i path a:
 1. <distrib>/mpisa/oai_ienigeConf
 2. <distrib>/ExtGASoai.src/IsisDbdir.oai/cnr
3. epilog.cnr è l'epilog file
4. lat2html_extended.iso
5. lat2html_extended.mst
6. lat2html_extended_saved.iso
7. lat2html_extended.xrf
8. loadiso.lat2html_extended
9. NEVERCANCELlat2html_extended.iso
10. NEVERCANCELlat2html_extended_saved.iso
11. oai_dc.pft è il Print Format per il verbo ListRecords e GetRecords per produrre out in oai_dc format
12. oai.system.file è la OAI System Table per questa applicazione. In esso vanno aggiornati i path di:
 1. ExtGASoaiCiparFileName
 2. ExtGASoaiProgramName
 3. ResumptionTokenDir
13. prolog.cnr è il prolog file
14. Sdate.pft è il Print Format per ordinare i risultati dei verbi ListRecord e

ListIdentifiers in ordine crescente di data.

Da punto 4-10 sono i file del database di conversione da ISO8859-1 a XHTML (unitamenete al file iso di input, la procedura di caricamento e i 2 file del database).

Il cipar file è il seguente:

```
cnr.*=/export/home/metaopac/mpisa/IsisDbdir.cnr_ienige/cnr.*
cnr_li.pft=/export/home/metaopac/mpisa/oai_ienigeConf/cnr_li.pft
cnr_lr.pft=/export/home/metaopac/mpisa/oai_ienigeConf/cnr_li.pft
oai_dc.pft=/export/home/metaopac/mpisa/oai_ienigeConf/oai_dc.pft
cnroaidc_lr.pft=/export/home/metaopac/mpisa/oai_ienigeConf/oai_dc.pft
Sdate.pft=/export/home/metaopac/mpisa/oai_ienigeConf/Sdate.pft
prolog.cnr=/export/home/metaopac/mpisa/oai_ienigeConf/prolog.cnr
epilog.cnr=/export/home/metaopac/mpisa/oai_ienigeConf/epilog.cnr
lat2html.*=/export/home/metaopac/mpisa/oai_ienigeConf/lat2html_extended.*
```

Si notino la directory del database e i Print Format

L'OAI System Table in oai.system.file è la seguente:

```
# OAI-pmh gateway versus CDS/ISIS: oai.system.file
# TRACE on
MaxObjectsToBeDisplayed 10
MinObjectsToBeDisplayed 12
ServerURI http://leonardo.isti.cnr.it/oai_ienige
ExtGASoaiCiparFileName /export/home/metaopac/mpisa/oai_ienigeConf/cnr.txt
GeneralQuerySuffix AND ap/(904) and IENIGE/(900)
ExtGASoaiProgramName /export/home/metaopac/mpisa/BiremePrograms/bin/wwwisis
#ExtGASoaiProgramName /export/home/metaopac/mpisa/ExtGASoai.src/access_isis.sh
ExtGASoaiArchiveName cnr
AllowedHosts 146.48.80-87.0-255, 137.158.128.0-255, 81.189.125.130, 155.232.128.10
#DeniedHosts 146.48.82.119
#RestrictedHosts 146.48.82.119
ExtGASoaiDataRecordDelimiter ;
ExtGASoaiDateQualifier /(5)
ExtGASoaiIdentifierQualifier /(1)
ExtGASoaiSortPftFile @Sdate.pft
ExtGASoaiGizmoFile lat2html
ResumptionTokenDir /export/home/metaopac/mpisa/OAISAVE/
```

Si faccia riferimento alla descrizione dei singoli parametri per la comprensione della OAI Sytem Table.

4. La security di ExtGAS OAI-PMH gateway

Nonostante la definizione del nome del protocollo stesso "Open Access Initiative" per motivi di sicurezza e riservatezza l'accesso al repository gestito dall'ExtGAS OAI-PMH gateway è controllato e di default non è aperto a tutti ma solo agli autorizzati.

Mediante la OAI System Table è possibile definire lo status generale della security che di default è impostata a **restricted**.

I livelli di sicurezza vengono associati all'indirizzo IP dell'Harvester client definendo liste reali di HostAllowed e HostDenied e di conseguenza la lista logica di HostRestricted.

I livelli di sicurezza sono:

1. **allowed**: l'Harvester client può eseguire tutti i verbi implementati, accedere alle liste dei record e, ove ne abbia ulteriormente diritto, scaricare dalla Digital Library il documento/oggetto.
2. **denied**: l'Harvester client pur accedendo al repository non può visualizzare né record identifier né record. Può eseguire i verbi ListSets, ListMetadataFormats e Identify
3. **restricted**: l'Harvester client pur accedendo al repository può visualizzare al massimo i primi MinObjectToBeDisplayed (default 5) identifier/record che soddisfano la richiesta e vengono successivamente ordinati in ordine crescente di data, ma non può eseguire il verbo GetRecord. Può eseguire invece, i verbi ListSets, ListMetadataFormats e Identify

Il livello di sicurezza **restricted** permette di rendere pubblico il gateway e le sue funzionalità garantendo una certa sicurezza negli accessi. Si tenga presente che per default viene tracciato un log degli accessi con i diritti associati agli indirizzi IP e le operazioni eseguite.

Genericamente potrebbe essere utile implementare un repository completamente controllato definendo la default security **denied** e aprendo solo a coloro che sono menzionati nella allowed list. Contemporaneamente si potrebbe definire un repository logico implementato sullo stesso database del repository operativo che abbia (mediante la definizione di set appropriati) una versione molto parziale delle informazioni.

5. La selezione dell'informazione

5.0. Il modello di selezione e accesso ai dati

Poiché per il protocollo OAI-PMH la selezione degli oggetti può essere fatta esclusivamente applicando i limiti temporali (parametri **from** e **until**) e l'appartenenza o meno a un **Set** e la ricerca sui campi data non esiste in CDS/ISIS, è stato necessario applicare un particolare algoritmo per la selezione degli oggetti via campo data (in formato carattere) utilizzando o meno il carattere di mascheramento a destra ('\$') per ottenere la **ResultList** di tutti quegli oggetti che appartengono a una lista definita con anno di inizio e anno di fine rappresentato come carattere. Successivamente, ottenuta la lista degli oggetti selezionati, essa viene ulteriormente verificata con il valore numerico degli anni mesi e giorni limite tramite il software sviluppato. In questa maniera si ha un primo processo che seleziona gli identifier velocemente accedendo all'inverted file con chiave mascherata o meno che sia, e successivamente ogni identifier viene o meno accettato valutandone il valore numerico della data.

Si tenga presente che i parametri **from** e **until** non sono obbligatori per il protocollo OAI-PMH. Per il funzionamento del Gateway stesso, in caso di mancanza completa di entrambi i parametri o di uno di essi vengono applicati i valori limite sostitutivi che sono memorizzati nella OAI-PMH System Table nei parametri **ExtGASoaiSysTable/ExtGASoaiUpperDate** (default: 20151231) e **ExtGASoaiSysTable/ExtGASoaiLowerDate** (default: 20000101)

Nell'esempio che segue vengono selezionate le pubblicazioni del **Set** "IENIGE" con **from=20060401** e **until=20070331** e **metadataPrefix=oai_dc**.

La seguente è la URL della richiesta:

http://leonardo.isti.cnr.it/oai_cnrpub?verb=ListIdentifiers&from=2006-04-01&until=2007-03-31&set=IENIGE&metadataPrefix=oai_dc

La query CDS/Isis generata dal GateWay è la seguente:

((2006\$/(5) or 2007\$/(5)) AND ap/(904) and IENIGE/(900))

I due anni, 2006 e 2007 sono messi in or e mascherati a destra. Vengono, inoltre, qualificati come appartenenti al relativo field 5. La qualificazione si ottiene assegnando il valore relativo al parametro **ExtGASoaiSysTable/ExtGASoaiDateQualifier** (**ExtGASoaiDateQualifier** /(5))

Viene così prodotto una lista temporanea, **ResultList**, che contiene 775 oggetti:

<NumberOfSelectedRecords>
775:775

```

</NumberOfSelectedRecords>
<Start>
<Record>;20060109;Metapub:cnr.ieni/cnr.ieni.ge/1999-A0-008;000492;IENIGE,AP;</Record>
<Record>;20060109;Metapub:cnr.ieni/cnr.ieni.ge/1997-A0-008;000493;IENIGE,AP;</Record>
<Record>;20060109;Metapub:cnr.ieni/cnr.ieni.ge/1991-A0-002;000495;IENIGE,AP;</Record>
.....
.....
<Record>;20060303;Metapub:cnr.ieni/cnr.ieni.ge/1995-A0-009;000771;IENIGE,AP;</Record>
<Record>;20060317;Metapub:cnr.ieni/cnr.ieni.ge/2002-A0-017;000923;IENIGE,AP;</Record>
<Record>;20060410;Metapub:cnr.ieni/cnr.ieni.ge/1999-A2-002;000780;IENIGE,AP;</Record>
<Record>;20060410;Metapub:cnr.ieni/cnr.ieni.ge/1995-A2-001;000784;IENIGE,AP;</Record>
.....
.....
<Record>;20060728;Metapub:cnr.ieni/cnr.ieni.ge/2004-A3-022;001048;IENIGE,AP;</Record>
<Record>;20060728;Metapub:cnr.ieni/cnr.ieni.ge/2004-A3-023;001049;IENIGE,AP;</Record>
<Record>;20070426;Metapub:cnr.ieni/cnr.ieni.ge/2006-A0-023;000466;IENIGE,AP;</Record>
<Record>;20070426;Metapub:cnr.ieni/cnr.ieni.ge/2006-A0-019;000903;IENIGE,AP;</Record>
<Record>;20070426;Metapub:cnr.ieni/cnr.ieni.ge/2006-A0-008;001050;IENIGE,AP;</Record>
.....
.....
<Record>;20070606;Metapub:cnr.ieni/cnr.ieni.ge/2006-A6-017;001184;IENIGE,AP;</Record>
<Record>;20070606;Metapub:cnr.ieni/cnr.ieni.ge/2007-PR-002;001185;IENIGE,AP;</Record>
<Record>;20070606;Metapub:cnr.ieni/cnr.ieni.ge/2006-PR-013;001186;IENIGE,AP;</Record>
</Start>

```

Tale lista viene successivamente ridotta a 472 elementi mediante il test di inclusione dei valori numerici dei parametri **from** e **until** confrontati con la data di pubblicazione/aggiornamento dei metadati:

```

Resume from [cnr.1182440921475.0.RT]
Complete List Size : 472
Cursor : 0
Expiration Date : 2007-06-21T17:48:41+0200

```

e viene reso disponibile il primo identifier valido come qui in esempio.

```

identifier : Metapub:cnr.ieni/cnr.ieni.ge/1999-A2-002
datestamp : 2006-04-10
setSpec : IENIGE
setSpec : AP

```

La costruzione della lista in esempio viene realizzata utilizzando il formato:

```
<ExtGASoaiArchiveName>_li.pft
```

Il cui nome è calcolato per default dal sistema con il parametro
ExtGASoaiSysTable/ListIdentifiersExtGASoaiFormat

L'esempio che segue è il formato cnr_li.pft:

```

mhl,if val(v1001) = val(v1031) then'<NumberOfSelectedRecords>','/,F((val(v1032)-
val(v1031)+1),3,0),','v1002,','</NumberOfSelectedRecords>','/,<Start>','/, mhl,ref(['cnr']val(v1^m),if
p(v5^b) then '<Record>','v5^b,if p(v1) then ';Metapub:','v1,','mf,','v900,','v904,','</Record>','/,fi,fi)

```

Il formato suddetto utilizza un formato di ordinamento, per restituire i record ordinati dal più vecchio al più nuovo, che è definito nel parametro

ExtGASoaiSysTable/ExtGASoaiSortPftFile il cui valore volutamente assegnato è "@Sdate.pft" (mentre per default è null) che è così definito:

```
mhu,if p(v5^b) then '^t',v5^b,' ^m',mfn,fi
```

essendo il campo 5 costituito da due sottocampi: ^a che contiene la data in formato legale e ^b che contiene invece il formato numerico della stessa data:

```
^a31/12/2006^b20063112
```

5.1. L'organizzazione della lista dei risultati

La struttura dell'elemento della lista **ResultList** è costituita da 6 elementi separati dal parametro **ExtGASoaiSysTable/ExtGASoaiDataRecordDelimiter** che per default vale carattere/stringa ";" (punto_e_virgola):

1. "<Record>" stringa costante: **obbligatoria**
2. data in formato numerico del tipo `yyyymmdd`, `yyyymm`, `yyyy` : **obbligatoria**
3. OAI-PMH Identifier: **obbligatorio**
4. CDS/ISIS Master File Record Number (viene usato per il retrieval) : **obbligatorio**
5. Lista dei **Set**, separati da un **SetIdentifierSetDelimiter** che vale di default "," (virgola), cui appartiene l'OAI Identifier. Se non esistono Set inserire il valore "NullSet": **obbligatorio**
6. "</Record>" stringa costante: **obbligatoria**

L'OAI-PMH identifier è costituito da una stringa di caratteri separati da separatori specifici tale da identificare univocamente il record nel repository.

5.2. La struttura dell'OAI-PMH identifier

1. **IdentifierPrefix** che di default vale null.
2. Set o lista di Set, separati gerarchicamente dal carattere ':' (due_punti) così come specificato nel protocollo OAI-PMH, e separati dall'effettivo identifier da **SetIdentifierDelimiter** che di default vale "::". Tali Set, se esistono, vengono utilizzati per qualificare l'effettivo identifier al momento della sua selezione per ottenere i metadati.
3. Effettivo identifier: **obbligatorio**

5.3. Il modello di rappresentazione delle date richiesto dal gateway

Sono accettati come modelli di date i seguenti mapping:

1. `yyyymmdd`

2. yyyy-mm-dd
3. yyyy-mm
4. yyyy-mm
5. yyyy

Il default definito nella OAI SystemTable mediante il parametro **ExtGASoaiSysTable/DateExtGASoaiFormat** è **yyyymmddd**.

Escludendo il mapping 5 tutti i rimanenti richiedono, per poter essere selezionati, un carattere "\$" per il mascheramento a destra. In questa maniera verranno selezionati tutti gli oggetti che hanno data di modifica che inizia per "yyyy" che può determinare una espansione della query mettendo in "or" tutti gli anni compresi e uguali tra i valori di **from** e **until**.

Lo stesso si applica al caso 5 senza mascheramento

5.4. Il modello di restituzione di Identifiers/Records

Il modello di accesso e selezione descritto sopra viene utilizzato per i verbi ListIdentifiers e ListRecords. Successivamente nel caso di:

1. ListIdentifiers: leggendo la lista degli oggetti selezionati, vengono inviati gli OAI identifiers e se necessario costruito il resumptionToken restart file e inviato il ResumptionToken pointer (vedi più avanti nell'esempio)
2. ListRecords: leggendo la lista degli oggetti selezionati, mediante il CDS/ISIS Master File Record Number, vengono acceduti i record e prodotte le restituzioni, in accordo al metadataPrefix specificato, per un numero massimo di record pari a **ExtGASoaiSysTable/MaxObjectsToBeDisplayed**. Nel caso sia necessario viene costruito il resumptionToken restart file e inviato il ResumptionToken pointer.

Nel caso del verbo GetRecord l'accesso viene fatto tramite l'OAI identifier del record. L'OAI identifier viene smembrato nei suoi componenti (vedi sopra) e viene costruita una particolare query per identificare univocamente il record. Nel caso del seguente OAI identifier:

isti:CNR:oai:Genova::cnr.ieni/cnr.ieni.ge/2003-A0-001

dove:

1. **ExtGASoaiSysTable/IdentifierPrefix** è definito come **isti:CNR:oai:**
2. **ExtGASoaiSysTable/SetIdentifierDelimiter** definito come ":" da cui deriva che, dato l'**IdentifierPrefix** in 1, esiste anche, definito nell'identifier, un Set qualificato di valore **Genova**.
3. Un Set definito a livello generale con il parametro **ExtGASoaiSysTable/SetDefinition** di valore:

Genova+x_ienige+cnr Genova publications+AND ap/(904) and (IENIGE/(900) or IBFGE/(900));

Dove **InternalSetDefinitionDelimiter** vale uno dei seguenti caratteri "+%# \$" e **SetDefinitionDelimiter** vale ';' (punto_e_virgola)

4. Un effettivo identifier di valore **cnr.ieni/cnr.ieni.ge/2003-A0-001** su cui è definito il qualifier **ExtGASoaiIdentifierQualifier** (ExtGASoaiIdentifierQualifier /(1)) e l'operatore **ExtGASoaiDataIdentifierOperator** impostato a (ExtGASoaiDataIdentifierOperator " \$ ") (dollaro_spazio_dollaro)

La query risultante è:

```
( ( cnr/(1) $ ieni/(1) $ cnr/(1) $ ieni/(1) $ ge/(1) $ 2003/(1) $ A0/(1) $ 001/(1) ) AND ap/(904) and (IENIGE/(900) or IBFGE/(900) )
```

dove il Field 1 rappresenta la identificazione univoca del documento nel suo contesto:

cnr.ieni/cnr.ieni.ge/2003-A0-001

Essa viene usata per realizzare una lista di risultati come per i verbi ListIdentifiers e ListRecords, ma composta da una unica entry di tipo <Record>.....</Record>. Successivamente tramite il Master File record Number viene prodotto il record da inviare al client secondo il metadataPrefix specificato. Nell'esempio:

```
<NumberOfSelectedRecords>
1:1
</NumberOfSelectedRecords>
<Start>
<Record>;20051129;isti:CNR:oai:Genova::cnr.ieni/cnr.ieni.ge/1993-A0-001;000404;IENIGE,AP,Genova;</Record>
</Start>
```

Si tenga presente che nel formato di restituzione delle liste di OAI-PMH identifier il blocco:

```
<NumberOfSelectedRecords>
n:m
</NumberOfSelectedRecords>
```

non è assolutamente obbligatorio.

5.5. Il restart dell'accesso ai dati

Per assicurarsi la completa certezza dello scaricamento è opportuno che il server OAI invii all'Harvester un numero limitato di record/identifier per transazione. L'OAI server invierà all'Harvester piccoli blocchi di record/identifier dando sempre la possibilità di restarting, a certi livelli, della transazione. Per ottenere questo risultato ogni qualvolta vengono inviate piccole quantità di dati all'Harvester viene inviato il nome di un pointer che rende disponibile il blocco successivo di informazioni e che ha una "expiration date" tale da permettere a qualunque livello il caricamento parziale dei dati persi durante la

trasmissione.

Il restart pointer viene chiamato **ResumptionToken** (identificatore che permette il ripristino/continuazione dell'operazione di scaricamento).

Il **resumptionToken** viene identificato da un nome univoco che identifica un file memorizzato su una particolare directory del server che è definita nel parametro **ExtGASoaiSysTable/ResumptionTokenDir** che è di default caricato come:

1. /oai/RT/ in ambiente unix
2. C:\oai\RT\ in ambiente Windows

Il nome del file è costruito da **ExtGASoaiUtil/GetResumptionTokenFileName** nella seguente maniera:

1. **ResumptionTokenDir**
2. un prefisso contenuto nel parametro **ExtGASoaiSysTable/ExtGASoaiResumptionTokenFilePrefix** che di default è caricato con il valore del parametro **ExtGASoaiSysTable/ExtGASoaiArchiveName**
3. "." : il carattere punto.
4. data e ora espressa in millisecondi a partire dal 1 gennaio 1974
5. "." : il carattere punto seguito ad un valore numerico calcolato dinamicamente dal sistema per rendere univoco il nome del file
6. "." : il carattere punto.
7. un suffisso contenuto nel parametro **ExtGASoaiSysTable/ExtGASoaiResumptionTokenFileSuffix** che di default vale la stringa "RT"

Il seguente è un esempio di **resumptionToken** file/directory:

```
/home/metaopac/mpisa/OAISAVE/cnr.1182529863184.0.RT  
/home/metaopac/mpisa/OAISAVE/cnr.1182529876784.0.RT  
/home/metaopac/mpisa/OAISAVE/cnr.1182529882474.0.RT
```

dove **ExtGASoaiSysTable/ResumptionTokenDir** è /home/metaopac/mpisa/OAISAVE/

Per default il **resumptionToken** ha un **Expiration Date** espressa in millisecondi nel parametro **ExtGASoaiSysTable/ResumptionTokenTTL** che vale 3600000 (un'ora).

Il **resumptionToken** File contiene un massimo di 10 parametri così come definito nel parametro **ExtGASoaiSysTable/MaxResumptionTokenParms** separati da un separatore in scrittura che è **ExtGASoaiSysTable/ResumptionTokenWriteDelim** e vale il carattere "#" mentre in lettura sono separati da uno dei caratteri definiti in **ExtGASoaiSysTable/ResumptionTokenReadDelim** che vale uno dei seguenti caratteri "%#;:,".

La struttura del `resumptionToken` record è costituita dai seguenti parametri suddivisi dal delimiter **ExtGASoaiSysTable/ResumptionTokenWriteDelim** :

1. Numero totale di record/identifier selezionati
2. Numero di record/identifier di inizio inviato
3. Numero di record/identifier di fine inviato
4. Valore del `dateFrom`
5. Valore del `dateUntil`
6. CDS/Isis query atta a selezionare gli stessi oggetti
7. OAI-PMH verb utilizzato (`ListIdentifiers/ListRecords`)
8. Formato di restituzione corrispondente al `metadataPrefix` selezionato

Il seguente è un esempio di `resumptionToken` file:

2944#11#20#20050101#20070331#2005\$/(5) or 2006\$/(5) or 2007\$/(5)#ListRecords#@oai_dc.pft

La richiesta di esecuzione di un verbo(7) `ListIdentifiers` o `ListRecords` a partire da un `resumptionToken` identifier provoca la riesecuzione completa della richiesta CDS/Isis(6) e successivamente la selezione degli oggetti richiesti e l'invio degli oggetti stessi secondo il `metadataPrefix`(8) definito.

5.6. Le API CDS/ISIS

Le API CDS/ISIS sono utilizzate dal sistema per accedere alle informazioni contenute nel database CDS/ISIS stesso. In realtà esiste un unico applicativo che interfaccia il database ed è il programma `wwwisis` che viene distribuito free dall'Unesco (www.unesco.org) e da BIREME (www.bireme.br). Il programma può ricevere in input il nome di un file che contiene i parametri da utilizzare per l'accesso al database e produrre i relativi risultati sul file di output. L'attivazione di `wwwisis` viene eseguita dal servizio **AccessIsis.java**.

Complessivamente il programma `wwwisis` viene attivato per eseguire due tipi di operazione:

1. L'operazione **search** per selezionare i record che soddisfano la richiesta (**IsisSearchAPI**).
2. L'operazione **get** per acquisire il contenuto del record e inviarne il contenuto, in accordo con il `metadataPrefix` selezionato, all'Harvester client (**IsisGetAPI**).

In realtà il servizio **AccessIsis** attiva una chiamata a una applicazione esterna definita in **ExtGASoaiSysTable/ExtGASoaiProgramName** passando come parametro:

in=<filename>

Nello standard è:

wwwisis in=<filename>

Al ritorno dalla chiamata i risultati sono resi disponibili in un file il cui nome e la locazione sono noti al programma chiamante.

Genericamente è possibile chiamare un generico script che a sua volta può eseguire una serie di operazioni prima e dopo la chiamata alle API CDS/ISIS. Di seguito un esempio attivato dalla chiamata

access_isis.sh in=<filename>

```
#!/bin/bash
# this is a script example to change results after CDS/ISIS access
# input is access_isis.sh in=<filename>
# echoes parmvalue
echo $1
# execute ISIS query
/export/home/metaopac/mpisa/BiremePrograms/bin/wwwisis $1
# get input file name and echoes it
f=${1:3}
echo $f
# load parm with pft name and echoes it
parm=`grep pft= $f`
echo $parm
# if pft name is pft=@cnr_li.pft execute external program
$if [ $parm == pft=@cnr_li.pft ]
echo -e "executing external program"
$endif
exit 0
```

Di seguito sono analizzate le chiamate alle API generate a fronte di:

1. ListIdentifiers

verb=ListIdentifiers&from=2006-12-31&metadataPrefix=oai_dc&set=IENIGE

1. **IsisSearchAPI** per selezionare il numero massimo di record che soddisfa la richiesta:

```
db=cnr
bool=( ( 2006$/(5) or 2007$/(5) or 2008$/(5) or 2009$/(5) or 2010$/(5) or 2011$/(5) or 2012$/(5) or
2013$/(5) or 2014$/(5) or 2015$/(5) ) AND ap/(904) and IENIGE/(900) )
h1=1
h2=1
prolog=@prolog.cnr
epilog=@epilog.cnr
pft=mhl,v1002,/
cipar=/export/home/metaopac/mpisa/oai_cnrpubConf/cnr.txt
out=/tmp/ExtGASoai.1183710556093.0.query.src.out
```

che produce

```
<Prolog>
CNR/Pisa
</Prolog>
```

2. **IsisSearchAPI** per selezionare la **ResultList** dei record che soddisfano la richiesta:

```
db=cnr
bool=( ( 2006$(/5) or 2007$(/5) or 2008$(/5) or 2009$(/5) or 2010$(/5) or 2011$(/5) or 2012$(/5) or
2013$(/5) or 2014$(/5) or 2015$(/5) ) AND ap/(904) and IENIGE/(900) )
h1=1
h2=775
prolog=@prolog.cnr
epilog=@epilog.cnr
gizmo=lat2html
freq=@Sdate.pft
pft=@cnr_li.pft
cipar=/export/home/metaopac/mpisa/oai_cnrpubConf/cnr.txt
out=/tmp/ExtGASoai.1183710556093.0.out
```

che produce

```
<Prolog>
CNR/Pisa
</Prolog>
<NumberOfSelectedRecords>
775:775
</NumberOfSelectedRecords>
<Start>
<Record>;20060109;Metapub:cnr.ieni/cnr.ieni.ge/1975-A0-001;000481;IENIGE,AP;</Record>
<Record>;20060109;Metapub:cnr.ieni/cnr.ieni.ge/1991-A0-002;000482;IENIGE,AP;</Record>
.....
<Record>;20070606;Metapub:cnr.ieni/cnr.ieni.ge/2007-PR-002;000962;IENIGE,AP;</Record>
</Start>
```

da cui poi, applicando la selezione per data viene inviata la lista dei primi
(**ExtGASoaiSysTable/MaxObjectsToBeDisplayed**) identifier

2. **GetRecord**

```
verb=GetRecord&metadataPrefix=oai_dc&identifier=Metapub%3Acnr.ieni%2Fcnr.ieni.ge%2F2006-A0-021
```

1. **IsisSearchAPI** per selezionare il record che soddisfa la richiesta:

```
db=cnr
bool=cnr/(1) $ ieni/(1) $ cnr/(1) $ ieni/(1) $ ge/(1) $ 2006/(1) $ A0/(1) $ 021/(1)
h1=1
h2=1
epilog=@epilog.cnr
prolog=@epilog.cnr
gizmo=lat2html
freq=@Sdate.pft
pft=@cnr_li.pft
cipar=/export/home/metaopac/mpisa/oai_cnrpubConf/cnr.txt
out=/tmp/ExtGASoai.1183710629457.0.out
```

che produce una **ResultList** di un unico elemento

```
<NumberOfSelectedRecords>
1:1
</NumberOfSelectedRecords>
```

```
<Start>
<Record>;20070426;Metapub:cnr.ieni/cnr.ieni.ge/2006-A0-021;000444;IENIGE,AP;</Record>
</Start>
```

2. **IsisGetAPI** per accedere tramite **ResultList** via Master File Record Number al contenuto del record stesso:

```
db=cnr
from=000444
count=1
gizmo=lat2html
pft=@oai_dc.pft
cipar=/export/home/metaopac/mpisa/oai_cnrpubConf/cnr.txt
out=/tmp/ExtGASoai.1183710629576.0.out
```

che produce il record da inviare all'Harvester client in formato oai_dc

```
<oai_dc:dc
  xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
    http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
  <dc:title>Dynamic Capillary Pressure Measurements in the short Time Range by applying a fast
  growing Drop Technique</dc:title>
  <dc:creator>Makievski Alexander</dc:creator>
  <dc:creator>Kr&#228;gel J&#252;rgen</dc:creator>
  <dc:creator>Pandolfini Piero</dc:creator>
  <dc:creator>Loglio Giuseppe</dc:creator>
  .....
  <dc:identifier>http://puma.isti.cnr.it/linkdoc.php?icode=2006-A0-021&#38;authority=cnr.ieni.ge&#38;collection=cnr.ieni&#38;langver=en</dc:identifier>
</oai_dc:dc>
```

3. ListRecords

verb=ListRecords&from=2006-12-31&metadataPrefix=oai_dc&set=IENIGE

1. **IsisSearchAPI** per selezionare il numero massimo di record che soddisfa la richiesta (come per **ListIdentifier IsisSearchAPI 1**).
2. **IsisSearchAPI** per selezionare la **ResultList** dei record che soddisfano la richiesta (come per **ListIdentifier IsisSearchAPI 2**), da cui poi, applicando la selezione per data vengono accettati i primi record (**ExtGASoaiSysTable/MaxObjectsToBeDisplayed**) da inviare all'Harvester client.
3. **IsisGetAPI** per accedere tramite **ResultList** via Master File Record Number al contenuto dei (**ExtGASoaiSysTable/MaxObjectsToBeDisplayed**) record da inviare (come per **GetRecord IsisGetAPI 2**)

Per la comprensione dei parametri delle **IsisSearchAPI** e delle **IsisGetAPI** si faccia riferimento a:

1. BIREME - Centro Latino-Americano e do Caribe de Informação em Ciências da

Saúde - WWWISIS: a world-wide web server for ISIS-databases Version 3.0,
URL <http://leonardo.isti.cnr.it/metaopac/wwwisis.pdf>

2. Giuseppe A. Romano - ExtGas/IsisGAS revisione 2.0: manuale di installazione e d'uso - ISTI-B4-16, 2006 URL
<http://leonardo.isti.cnr.it/metaopac/ExtGASOmoDbInstall.pdf>

6. L'utilizzo del proxy-server

Nella directory <distrib>/ExtGASoa.src/ProxyDefinition sono contenuti due file di configurazione del proxy-server che sono adatti a pilotare i repository definiti con l'ExtGAS OAI-PMH gateway in modo da rendere trasparente all'utente sia la collocazione reale del repository che la OAI System Table che lo pilota.

Di regola l'accesso al gateway ExtGAS OAI-PMH viene fatto indicando il nome dell'applicazione oai relativa:

`http://<hostname>/<oai-application-name>`

Nell'esempio:

http://leonardo.isti.cnr.it/oai_cnrpub

Il proxy-server leonardo.isti.cnr.it possiede una regola di riscrittura del tipo:

```
RewriteCond %{REQUEST_URI} /=oai_cnrpub
RewriteCond %{QUERY_STRING} ^(.+)$
RewriteRule ^.*
http://metabsrv.isti.cnr.it:8080/ExtGASoai/servlet/oai?Raddr=%{REMOTE_ADDR}&Conf=/export/home/
metaopac/mpisa/oai_cnrpubConf/oai.system.file&%1 [P,L]
```

```
RewriteCond %{REQUEST_URI} /=oai_cnrpub
RewriteCond %{QUERY_STRING} !^(.+)$
RewriteRule ^.*
http://metabsrv.isti.cnr.it:8080/ExtGASoai/servlet/oai?Raddr=%{REMOTE_ADDR}&Conf=/export/home/
metaopac/mpisa/oai_cnrpubConf/oai.system.file [P,L]
```

che accede al servizio sul server OAI metabsrv.isti.cnr.it:8080 assegnando il relativo file di configurazione e in Raddr l'indirizzo IP del client chiamante. In questa maniera è possibile accettare o meno le richieste del client a seconda che sia o meno autorizzato l'indirizzo IP relativo.

`http://metabsrv.isti.cnr.it:8080/ExtGASoai/servlet/oai?Raddr=146.48.82.119&Conf=/export/home/metaopac/mpisa/oai_cnrpubConf/oai.system.file`

Nel caso si volesse evitare l'uso delle regole di riscrittura via proxy-server il gateway ExtGAS OAI-PMH può essere attivato nella seguente maniera:

`http://leonardo.isti.cnr.it:8080/ExtGASoai/servlet/oai?Conf=/export/home/metaopac/mpisa/oai_cnrpubConf/oai.system.file`

o ancora meglio come

<http://leonardo.isti.cnr.it:8080/ExtGASoai/servlet/oai>

dove si presuppone che il OAI System File sia /oai/oaiConf/oai.system.file per ambiente

unix e C:\oai\oaiConf\oai.system.file

Attenzione: il OAI System File **deve** esistere a meno di non attivare l'OAI gateway.

Sarà compito del driver del servlet calcolare l'indirizzo IP del client chiamante e garantire così la sicurezza dell'accesso.

7. I parametri della OAI System Table

La OAI System Table permette di costruire intorno al database CDS/ISIS il relativo repository. La OAI System Table è un file che contiene una serie di assegnazioni di valori a particolari parametri. Ogni parametro può essere definito più volte in maniera ordinata (nel senso che il secondo valore viene concatenato al primo ... e così via) ma può essere definito in qualunque sezione del file di configurazione. Ogni parametro possiede un valore default che viene assegnato e che determina particolari comportamenti del gateway.

I parametri sono classificati nei seguenti tipi:

1. Parametri generali: che definiscono le caratteristiche del sistema
2. Parametri del repository: che definiscono le caratteristiche del repository
3. Parametri di restituzione: che definiscono le strutture dati XML che contengono le informazioni
4. Parametri dei record identifier.
5. Parametri dei Set
6. parametri dei ResumptionToken

Il seguente è un esempio di OAI System Table che implementa un repository che rende disponibile i dati dell'Istituto IENI (Sezione di Genova) che si appoggia sul database del progetto Metapub: vede solo una parte del database.

```
# OAI-pmh gateway versus CDS/ISIS: oai.system.file
#TRACE on
MaxObjectsToBeDisplayed 10
MinObjectsToBeDisplayed 12
ServerURI http://leonardo.isti.cnr.it/oai_jenige
ExtGASoaiCiparFileName /export/home/metaopac/mpisa/oai_jenigeConf/cnr.txt
GeneralQuerySuffix AND ap/(904) and IENIGE/(900)
ExtGASoaiProgramName /export/home/metaopac/mpisa/BiremePrograms/bin/wwwisis
ExtGASoaiArchiveName cnr
AllowedHosts 146.48.80-87.0-255, 137.158.128.0-255, 81.189.125.130, 155.232.128.10
#DeniedHosts 146.48.82.119
#RestrictedHosts 146.48.82.119
ExtGASoaiDataRecordDelimiter ;
ExtGASoaiDateQualifier /(5)
ExtGASoaiIdentifierQualifier /(1)
ExtGASoaiSortPftFile @Sdate.pft
ExtGASoaiGizmoFile lat2html
ResumptionTokenDir /export/home/metaopac/mpisa/OAISAVE/
```

La successiva OAI System Table definisce un repository operativo da cui il progetto Driver carica i dati

```
# OAI-pmh gateway versus CDS/ISIS: oai.system.file
#TRACE on
#ResumptionTokenTTL 1000
ExtGASoaiCiparFileName /export/home/metaopac/mpisa/oai_cnrpubdriverConf/cnr.txt
#GeneralQuerySuffix AND ap/(904) and IENIGE/(900)
```

```

ExtGASoaiProgramName /export/home/metaopac/mpisa/BiremePrograms/bin/wwwisis
ExtGASoaiArchiveName cnr
AllowedHosts 146.48.80-87.0-255, 129.70.12.42, 137.158.128.0-255, 155.232.250.0-255,
137.158.96.152, 81.189.125.130, 155.232.128.0-255
#DeniedHosts 146.48.82.119
ServerURI http://leonardo.isti.cnr.it/oai_cnrpubdriver
#MaxObjectsToBeDisplayed 100
MaxObjectsToBeDisplayed 10
ExtGASoaiLowerDate 19770815
ExtGASoaiDataRecordDelimiter ;
ExtGASoaiDateQualifier /(5)
ExtGASoaiIdentifierQualifier /(1)
ExtGASoaiSortPftFile @Sdate.pft
ExtGASoaiGizmoFile lat2html
SetDefinition IAMCTA+x_iamcta+cnr IAMCTA publications+AND ap/(904) and IAMCTA/(900);
SetDefinition IBFGE+x_ibfge+cnr IBFGE publications+AND ap/(904) and IBFGE/(900);
SetDefinition IBFPI+x_ibfge+cnr IBFPI publications+AND ap/(904) and IBFPI/(900);
SetDefinition IENIGE+x_ienige+cnr IENIGE publications+AND ap/(904) and IENIGE/(900);
SetDefinition IFC+x_ifc+cnr IFC publications+AND ap/(904) and IFC/(900);
SetDefinition IGGPI+x_igmpi+cnr IGGPI publications+AND ap/(904) and IGGPI/(900);
SetDefinition IIT+x_iit+cnr IIT publications+AND ap/(904) and (IIT/(900) OR IAT/(900));
SetDefinition IMATIGE+x_imatige+cnr IMATIGE publications+AND ap/(904) and IMATIGE/(900);
SetDefinition ISEMGE+x_isemge+cnr ISEMGE publications+AND ap/(904) and ISEMGE/(900);
SetDefinition ISIBPD+x_isibpd+cnr ISIBPD publications+AND ap/(904) and ISIBPD/(900);
SetDefinition ISTI+x_isti+cnr ISTI publications+AND ap/(904) and ISTI/(900);
SetDefinition AP+x_ap+publications+AND ap/(904)
IdentifierPrefix Metapub:
ResumptionTokenDir /export/home/metaopac/mpisa/OAISAVE/
# the Identify data
Identify <Identify>
Identify <repositoryName>ExtGAS/ISIS Open Archive Initiative Repository </repositoryName>
Identify <baseURL>http://leonardo.isti.cnr.it/oai_cnrpubdriver</baseURL>
Identify <protocolVersion>2.0</protocolVersion>
Identify <adminEmail>giuseppe.romano@isti.cnr.it</adminEmail>
Identify <earliestDatestamp>1977-08-15</earliestDatestamp>
Identify <granularity>YYYY-MM-DD</granularity>
Identify <deletedRecord>no</deletedRecord>
Identify <compression>deflate</compression>
Identify <description>
Identify <oai-identifier xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
Identify <scheme>oai</scheme>
Identify <repositoryIdentifier>leonardo.isti.cnr.it</repositoryIdentifier>
Identify <delimiter>:</delimiter>
Identify <sampleIdentifier>oai:leonardo.isti.cnr.it:Metapub:cnr.isti/cnr.isti/2003-B4-28</sampleIdentifier>
Identify </oai-identifier>
Identify </description>
Identify </Identify>
# ends the Identify data
# the MetadataFormats definition
ListMetadataFormats <ListMetadataFormats>
ListMetadataFormats <metadataFormat>
ListMetadataFormats <metadataPrefix>oai_dc</metadataPrefix>
ListMetadataFormats <schema>http://www.openarchives.org/OAI/2.0/oai_dc.xsd </schema>
ListMetadataFormats <metadataNamespace>http://www.openarchives.org/OAI/2.0/oai_dc/
</metadataNamespace>
ListMetadataFormats </metadataFormat>
ListMetadataFormats <metadataFormat>

```

```
ListMetadataFormats <metadataPrefix>oai_rfc1807</metadataPrefix>
ListMetadataFormats <schema>http://www.openarchives.org/OAI/1.1/rfc1807.xsd</schema>
ListMetadataFormats
    <metadataNamespace>http://www.w3.org/2001/XMLSchemainstance</metadataNamespace>
ListMetadataFormats </metadataFormat>
ListMetadataFormats </ListMetadataFormats>
# ends the MetadataFormats definition
```

7.0. I parametri generali

I parametri generali del gateway sono utilizzati per definire l'ambiente operativo del gateway stesso.

1. **TRACE**(on/off) : abilita/disabilita la traccia delle operazioni effettuate dal gateway
 1. non cancella i file temporanei, utilizzati per l'accesso al database, dalla work directory definita con il parametro **WorkDir**
 2. traccia le operazioni commentandole in XML in modo che possano essere visualizzate senza compromettere il funzionamento del gateway.
2. **ExtGASoaiProgramName** (/tmp/wwwisis | C:\TEMP\wwwi32.exe) : definisce il programma che implementa le API CDS/ISIS. Il seguente è un esempio di definizione:

ExtGASoaiProgramName /export/home/metaopac/mpisa/BiremePrograms/bin/wwwisis

3. **ExtGASoaiCiparFileName** (WorkDir+ExtGASoaiArchiveName+".txt") definisce il file che contiene le locazioni degli oggetti manipolati dalle API CDS/Isis. Il seguente è un esempio di definizione:

ExtGASoaiCiparFileName /export/home/metaopac/mpisa/oai_cnrpubdriverConf/cnr.txt

E il successivo è un esempio di Cipar File

```
cnr.*=/export/home/metaopac/mpisa/IsisDbdir.cnr_ienige/cnr.*
cnr_li.pft=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/cnr_li.pft
cnr_lr.pft=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/cnr_li.pft
oai_dc.pft=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/oai_dc.pft
oai_rfc1807.pft=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/oai_rfc1807.pft
cnroaidc_lr.pft=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/oai_dc.pft
Sdate.pft=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/Sdate.pft
prolog.cnr=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/prolog.cnr
epilog.cnr=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/epilog.cnr
lat2html.*=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/lat2html_extended.*
```

1. La prima riga definisce la locazione del database
2. la seconda e la terza i formati di costruzione della **ResultList** per i verbi ListIdentifiers e ListRecords.
3. La quarta, la quinta e la sesta i due formati definiti per la restituzione dei metadati: oai_dc.pft e oai_rfc1807.pft
4. La settima definisce un formato di ordinamento della **ResultList** in ordine crescente di data
5. La ottava e la nona due formati di stampa per il prolog e l'epilog delle API CDS/Isis
6. La decima definisce un tabella di conversione dei caratteri diacritici e speciali compatibile con l'XML che viene usata durante lo scaricamento dei

metadati. Essa è definita nella OAI System Table con il parametro **ExtGASoaiGizmoFile**.

Maggiori informazioni su: WWWISIS: a world-wide web server for ISIS-databases Version 3.0 (BIREME - Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde) www.bireme.br alla seguente URL: <http://leonardo.isti.cnr.it/metaopac/wwwisis.pdf>.

4. **ExtGASoaiGizmoFile** (null) definisce il nome della tabella (è in realtà in CDS/Isis database di supporto) che contiene i caratteri diacritici da convertire in XML charset compatibile. Verificare in una delle directory *Conf il file **NEVERCANCELlat2html_extended.iso**: esso mappa tutti i caratteri da 128 a 255 decimale con il corrispondente XHTML &#nnn; e i caratteri [",',<,>,&]. Si tenga presente che se fosse necessario ricaricare il **ExtGASoaiGizmoFile** nella stessa directory sono presenti il file delle definizioni e la utility per ricaricarlo (loadiso.lat2html_extended).
5. **ExtGASoaiArchiveName** (CDS) definisce il nome del database da cui estrarre i metadati. Il nome del database deve essere conosciuto e contenuto nel **ExtGASoaiCiparFileName** da cui la API CDS/Isis deriva la collocazione dello stesso. Il seguente è un esempio:

ExtGASoaiArchiveName cnr

Cui corrisponde la sua definizione nel **ExtGASoaiCiparFileName** (prima riga nel punto 3.)
6. **WorkDir** (/tmp/ | C:\TEMP) definisce la work directory su cui appoggiare i file temporanei durante l'esecuzione del gateway.
7. **ExtGASoaiLogFile** (/tmp/ExtGASoaiLogFile | C:\TEMP\ExtGASoaiLogFile) definisce il nome e la collocazione del file di log del gateway. Se si volesse evitare di tracciare il log delle operazioni è sufficiente definire il file nella seguente maniera:
 1. **ExtGASoaiLogFile** /dev/null per ambiente Unix
 2. **ExtGASoaiLogFile** NUL per ambiente Windows
8. **HelpFileUrl** (/help/ExtGASoaiHelpFile.html) definisce la URL del file che contiene l'Help.
9. **MaxRecordLength** (8192) definisce la massima lunghezza del record (linea di stampa prodotta dalle API CDS/Isis tramite uno dei formati di stampa utilizzati) che viene letta dai file temporanei. Può essere modificata se la lunghezza di ogni record risulta superiore al default.
10. **TotalSelectedRecords** () contiene dinamicamente il numero totale di record

selezionabili a fronte di una particolare ricerca sul database CDS/Isis utilizzato come metadata repository. Tale numero viene successivamente utilizzato dalle API CDS/Isis per la costruzione della **ResultList** durante l'esecuzione di un verbo ListIdentifiers o ListRecords

11. **ExtGASoaiFilePrefix** (ExtGASoai) definisce il prefisso da utilizzare per tutti i file temporanei gestiti dal gateway.
12. **ExtGASoaiResultFileSuffix** (out) definisce il suffisso da utilizzare per tutti i file risultato temporanei gestiti dal gateway.
13. **ExtGASoaiQueryFileSuffix** (query) definisce il suffisso da utilizzare per tutti i file query temporanei gestiti dal gateway.
14. **ExtGASoaiPftFileSuffix** (pft) definisce il suffisso da utilizzare per tutti i file PrintFormat temporanei gestiti dal gateway.

Per la comprensione della gestione dei file temporanei si riferisca al paragrafo 5.6. Le API CDS/ISIS

7.1. I parametri del repository

I parametri del repository definiscono le modalità di gestione e manipolazione della informazione a prescindere dalla collocazione fisica del database.

1. **ServerURI** (<http://leonardo.isti.cnr.it/oai>) la Uniform Resource Identification del repository
2. **Granularity** (YYYY-MM-DD) specifica il più fine dettaglio della data che può essere utilizzato durante le operazioni di Selective Harvesting.
3. **dcMetadataPrefix** (oai_dc) descrive il nome del MetadataPrefix che riferisce al metadataFormat **dublin core** descritto dal verbo ListMetadataFormats.
4. **ListIdentifiersExtGASoaiFormat** (@+<ExtGASoaiArchiveName>+ _li.pft) definisce il default CDS/Isis Print Format per produrre la **ResultList** se il verbo è ListIdentifiers. Vedi l'associazione in **CiparFile**:

cnr_li.pft=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/cnr_li.pft

5. **ListRecordsExtGASoaiFormat** (@+<ExtGASoaiArchiveName>+ _lr.pft) definisce il default CDS/Isis Print Format per produrre la **ResultList** se il verbo è ListRecords. Vedi l'associazione in **CiparFile**:

cnr_lr.pft=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/cnr_lr.pft

6. **ExtGASoaiSortPftFile** (null) definisce il CDS/Isis Print Format File per estarre i dati dal database per utilizzarli come chiavi di ordinamento per i record selezionati. Il seguente è un esempio di ordinamento per data:

```
mhu,if p(v5^b) then '^t',v5^b,' ^m',mfn,fi
```

contenuto nel Sdate.pft. Per maggiore comprensione vedi " 5.0. Il modello di selezione e accesso ai dati "

7. **ListRecordsOAI_DCExtGASoaiFormat** (@+<ExtGASoaiArchiveName>+oaidc _lr.pft) definisce il default CDS/Isis Print Format per produrre i risultati in una **ResultList** in caso di metadataPrefix assente. Vedi l'associazione in **CiparFile**:

cnroaidc_lr.pft=/export/home/metaopac/mpisa/oai_cnrpubdriverConf/oai_dc.pft

Attenzione: questo parametro dovrebbe essere obsoleto causa controlli stretti imposti dal protocollo sui parametri dei verbi.

8. **RecordExtGASoaiFormat** (@+<ExtGASoaiArchiveName>+ _r.pft) definisce il default CDS/Isis Print Format per produrre i metadati se il verbo è GetRecord.

Attenzione questo parametro è al momento definito ad uso futuro.

9. **ExtGASoaiDefaultRights** (restricted, allowed, denied) definisce i diritti default per un client generico: se il client (IPAddress del client) non è menzionato nella AllowedList o nella DeniedList acquisisce i diritti **ExtGASoaiDefaultRights** (restricted).

10. **AllowedHosts** (null) definisce la lista degli Ipaddress, separati da virgola, dei client cui è permesso l'accesso. Il seguente ne è un esempio:

AllowedHosts 146.48.80-87.0-255, 129.70.12.42, 137.158.128.0-255, 155.232.250.0-255, 137.158.96.152, 81.189.125.130, 155.232.128.0-255

Gli ottetti 1-0 dell'IPAddress possono essere configurati con un valore minimo e un massimo in modo da definire una netmask limitata.

11. **DeniedHosts** (null) definisce la lista degli Ipaddress, separati da virgola, dei client cui è impedito l'accesso. Il seguente ne è un esempio:

DeniedHosts 146.48.82.119

Per i DeniedHosts valgono le stesse regole degli AllowedHosts. Si tenga presente che la Restricted List è una lista virtuale e non può essere definita.

12. **DateExtGASoaiFormat** (yyyymmdd, yyyy-mm-dd, yyyymm, yyyy-mm, yyyy) definisce il modello di rappresentazione delle date all'interno del CDS/Isis database. Per maggiori dettagli vedi "5.3. Il modello di rappresentazione delle date richiesto dal gateway".

13. **ExtGASoaiUpperDate** (20151231) rappresenta la massima data di aggiornamento dei record (presenti/che si vogliono rendere visibili). Attenzione essa viene utilizzata nella selezione dei record nella **ResultList** per cui deve essere compatibile con il **DateExtGASoaiFormat** secondo il seguente modello:

1. yyyymmdd e yyyy-mm-dd sono yyyymmdd
2. yyyymm e yyyy-mm sono yyyymm
3. yyyy è yyyy

14. **ExtGASoaiLowerDate** (20000101) rappresenta la minima data di aggiornamento dei record (presenti/che si vogliono rendere visibili). Per esso valgono le stesse regole di **ExtGASoaiUpperDate**

15. **GeneralQuerySuffix**(null) : è il suffisso, se esiste, da appendere alla query già tradotta nel query language del sistema CDS/Isis. In questa maniera è possibile limitare staticamente l'accesso a un sottoinsieme qualificato di record in modo da rendere accessibile solo una parte dei record memorizzati sul database.

Definita QUERY la query in ISIS query language, la traduzione avviene secondo il seguente schema:

```
if (GeneralQuerySuffix!=null) QUERY="( "+QUERY+" )" "+GeneralQuerySuffix
```

Il seguente è un esempio:

```
GeneralQuerySuffix AND ap/(904) and IENIGE/(900)
```

16. **GeneralQueryPrefix(null)** : è il prefisso, se esiste, da anteporre alla query già tradotta nel query language del sistema CDS/Isis: assume lo stesso significato di **GeneralQuerySuffix**.

Definita QUERY la query in ISIS query language, la traduzione avviene secondo il seguente schema:

```
if (GeneralQueryPrefix!=null) QUERY=GeneralQueryPrefix+"( "+QUERY+" )"
```

```
if (GeneralQueryPrefix!=null AND GeneralQuerySuffix!=null)
    QUERY=GeneralQueryPrefix+"( "+QUERY+" )" "+GeneralQuerySuffix
```

17. **ExtGASoaiIdentifierQualifier (null)** rappresenta il qualificatore degli elementi dell'identifier, da utilizzare nella CDS/Isis query, per la selezione del relativo record. Per maggiori dettagli vedi "5.4. Il modello di restituzione di Identifiers/Records ./4"

18. **ExtGASoaiDateQualifier (null)** rappresenta il qualificatore del campo tipo data che contiene la data di aggiornamento dei record. Serve per selezionare in maniera precisa i record che hanno/iniziano per la stessa data. Per maggiori dettagli vedi "5.0. Il modello di selezione e accesso ai dati"

19. **ExtGASoaiDataIdentifierOperator ("\$_")** rappresenta l'operatore da utilizzare nella CDS/Isis query prodotta dall'identifier (qualificato da un **Qualifier** o meno) di un record. Anche per esso unitamente a **ExtGASoaiIdentifierQualifier** vedi "5.4. Il modello di restituzione di Identifiers/Records ./4"

20. **MinObjectsToBeDisplayed (5)** rappresenta il numero di oggetti record/identifier che possono essere inviati all'Harvester client se l'IPAddress del client è classificato come **restricted**. Per maggiori informazioni vedi: " La security di ExtGAS OAI-PMH gateway "

21. **MaxObjectsToBeDisplayed (100)** rappresenta il numero massimo di record/identifier che possono essere inviati in una transazione del gateway. Se il numero di oggetti selezionati è maggiore viene generata una informazione specifica per l'Harvester client (**resumptionToken**) che permette di continuare a ricevere i record/identifier fino al completamento della richiesta

22. **MaxObjectsToBeSelected** (10000) rappresenta il massimo numero di record/identifier che possono essere selezionati in una richiesta.

23. **ExtGASoaiDataRecordDelimiter** (:) definisce il delimiter dei field che fanno parte della **ResultList** che viene prodotta dai formati **ListIdentifiersExtGASoaiFormat** e **ListRecordsExtGASoaiFormat**. Nell'esempio:

```
<Record>;20060109;Metapub:cnr.ieni/cnr.ieni.ge/1991-A0-002;000495;IENIGE,AP;</Record>
```

l'elemento di **ResultList** è prodotto da <istrib>/oai_ienigeConf/cnr_li.pft

```
mhl,if val(v1001) = val(v1031) then '<NumberOfSelectedRecords>',/,F((val(v1032)-  
val(v1031)+1),3,0),';',v1002,/, '</NumberOfSelectedRecords>',/, '<Start>',/,  
mhl,ref(['cnr']val(v1^m),if p(v5^b) then '<Record>',v5^b,if p(v1) then  
';Metapub:',v1,;',',mf,;',',v900,;',',v904,;', '</Record>',/,fi,fi)
```

per maggiori spiegazioni vedi " 5.1. L'organizzazione della lista dei risultati "

24. **SetIdentifierSetDelimiter** (.) identifica il separatore dei Set, se ne esiste più di uno, all'interno dell'elemento di **ResultList**. Per lo schema di riferimento vedi il punto precedente.

7.2. I parametri di restituzione

Parametri di restituzione definiscono le strutture dati XML che contengono le informazioni. Normalmente non andrebbero modificati tranne i parametri **Identify** e **ListMetadataFormats**.

1. **ListMetadataFormats** definisce i formati dei metadati supportati dal **repository**. Il default è il seguente e mappa il dublin core MetadataFormat:

```
<ListMetadataFormats>
  <metadataFormat>
    <metadataPrefix>oai_dc</metadataPrefix>
    <schema>http://www.openarchives.org/OAI/2.0/oai_dc.xsd </schema>
    <metadataNamespace>http://www.openarchives.org/OAI/2.0/oai_dc/</metadataNamespace>
  </metadataFormat>
</ListMetadataFormats>
```

In " 7. I parametri della OAI System Table " viene presentata una OAI System Table che oltre al metadataPrefix **dublin core** presenta un altro metadataPrefix **oai_rfc1807**

```
<ListMetadataFormats>
  <metadataFormat>
    <metadataPrefix>oai_dc</metadataPrefix>
    <schema>http://www.openarchives.org/OAI/2.0/oai_dc.xsd </schema>
    <metadataNamespace>http://www.openarchives.org/OAI/2.0/oai_dc/</metadataNamespace>
  </metadataFormat>
  <metadataFormat>
    <metadataPrefix>oai_rfc1807</metadataPrefix>
    <schema>http://www.openarchives.org/OAI/1.1/rfc1807.xsd</schema>
    <metadataNamespace>http://www.w3.org/2001/XMLSchemainstance</metadataNamespace>
  </metadataFormat>
</ListMetadataFormats>
```

I nomi/codici inseriti in metadataPrefix vengono utilizzati per identificare e caricare il relativo CDS/Isis Print Format secondo il seguente modello:

```
<metadataPrefix>+ ".pft"
```

Si guardi il relativo **CiparFile**

2. **Identify** che definisce le informazioni che caratterizzano il repository e vanno inviate all'Harvester client a fronte del verbo Identify. Il default è visualizzabile a " 1. ExtGAS OAI-PMH gateway /1. Identify "
3. **ListIdentifiersHeader** (<ListIdentifiers>) definisce il ListIdentifiers Header.
4. **ListIdentifiersTrailer** (</ListIdentifiers>) definisce il ListIdentifiers Trailer.
5. **GetRecordHeader** (<GetRecord>) definisce il GetRecord Header.

6. **GetRecordTrailer** (`</GetRecord>`) definisce il GetRecord Trailer.
7. **ListRecordsHeader** (`<ListRecords>`) definisce il ListRecords Header.
8. **ListRecordsTrailer** (`</ListRecords>`) definisce il ListRecords Trailer.
9. **MetadataPrefixHeader** (`<metadataPrefix>`) definisce il Metadata Prefix Header
10. **MetadataPrefixTrailer** (`</metadataPrefix>`) definisce il Metadata Prefix Trailer
11. **MetadataHeader** (`<metadata>`) definisce il Metadata Header.
12. **MetadataTrailer** (`</metadata>`) definisce il Metadata Trailer
13. **RecordHeader** (`<record>`) definisce il Record Header.
14. **RecordTrailer** (`</record>`) definisce il Record Trailer.
15. **SetNameHeader** (`<setName>`) definisce il Set Name Header.
16. **SetNameTrailer** (`</setName>`) definisce il Set Name Trailer.
17. **ListSetsHeader** (`<ListSets>`) definisce il List Sets Header.
18. **ListSetsTrailer** (`</ListSets>`) definisce il List Sets Trailer.
19. **SetHeader** (`<set>`) definisce il Set Header.
20. **SetTrailer** (`</set>`) definisce il Set Trailer.
21. **SetSpecHeader** (`<setSpec>`) definisce il Set Spec Header.
22. **SetSpecTrailer** (`</setSpec>`) definisce il Set Spec Trailer.
23. **IdentifierHeader** (`<identifier>`) definisce l'Identifier Header.
24. **IdentifierTrailer** (`</identifier>`) definisce l'Identifier Trailer.
25. **datestampHeader** (`<datestamp>`) definisce il Date Stamp Header.
26. **datestampTrailer** (`</datestamp>`) definisce il Date Stamp Trailer.
27. **HeaderHeader** (`<header>`) definisce l'Header dell'Header.

28. **HeaderTrailer** (</header>) definisce il Trailer dell' Header.
29. **ResponseDateHeader** (<responseDate>) definisce il Response Date Header.
30. **ResponseDateTrailer** (</responseDate>) definisce il Response Date Trailer.
31. **OAIPMHtrailer** (</OAI-PMH>) definisce il OAI-PMH Trailer.
32. **OAIPMHheader** definisce il OAI-PMH Header. Il suo default è:

```
<OAI-PMHxmlns="http://www.openarchives.org/OAI/2.0/"  
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"  
xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/  
http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
```

7.3. I parametri dei record identifier

I parametri dei record identifier definiscono i delimiters, i prefissi e gli eventuali suffissi che costituiscono la base per la costruzione dell'identifier e la base per la costruzione della query CDS/Isis per selezionare il relativo record del particolare identifier.

1. **IdentifierPrefix** (null) definisce un eventuale prefisso che viene collocato in testa all'identifier. Nell'esempio "Metapub:" è il prefisso dell'identifier

IdentifierPrefix Metapub:

In fase di costruzione dell'identifier viene inserito nella stringa, mentre in fase di selezione dell'identifier viene escluso dalla costruzione della CDS/Isis Query. Per maggiori dettagli vedi " 5.2. La struttura dell'OAI-PMH identifier ".

2. **SetIdentifierDelimiter** (::) Il set identifier delimiter identifica, all'interno dell'identifier, l'esistenza o meno di un set o una gerarchia di set. Nell'esempio:

isti:CNR:oai:Genova::cnr.ieni/cnr.ieni.ge/2003-A0-001

ist:CNR:oai è **IdentifierPrefix**

Genova è il **Set**

:: è **SetIdentifierDelimiter**

cnr.ieni/cnr.ieni.ge/2003-A0-001 è l'identifier effettivo del record

3. **ExtGASoaiDataIdentifierDelimiter** (./-:;) contiene i caratteri delimiter che individuano gli elementi componenti dell'identifier effettivo (realidentifier).
Facendo riferimento al precedente esempio realidentifier è:

cnr.ieni/cnr.ieni.ge/2003-A0-001

Applicando la regola di cui sopra gli elementi componenti sono:

cnr ieni cnr ieni ge 2003 A0 001

E applicando **ExtGASoaiDataIdentifierOperator**, **ExtGASoaiIdentifierQualifier** e le specifiche del Set **Genova** si ottiene la seguente query per selezionare il record identificato dall'identifier di cui sopra:

((cnr/(1) \$ ieni/(1) \$ cnr/(1) \$ ieni/(1) \$ ge/(1) \$ 2003/(1) \$ A0/(1) \$ 001/(1)) AND ap/(904) and (IENIGE/(900) or IBFGE/(900))

7.4. I parametri dei set

I parametri dei Set permettono di definire i Set fino a un massimo di **MaxSetDefined** e ne determinano le caratteristiche peculiari. I Set sono utilizzati per sezionare dinamicamente in classi il contenuto del repository.

1. **SetDefinition** (null) definisce un Set o un insieme di Set.
Ogni Set è separato dal successivo da un **SetDefinitionDelimiter**.
Ogni Set contiene quattro parametri separati tra di loro da un **InternalSetDefinitionDelimiter**
 1. Il nome del Set.
 2. Un codice di identificazione (di uso futuro).
 3. Descrizione del Set.
 4. Qualificazione del Set che corrisponde a uno o più elementi di CDS/Isis query che verranno suffissate alla CDS/Isis query che deriva dalla interpretazione degli OAI verb.

Il successivo è un esempio di definizione di Set:

```
SetDefinition ISIBPD+x_isibpd+cnr ISIBPD publications+AND ap/(904) and ISIBPD/(900);  
SetDefinition ISTI+x_isti+cnr ISTI publications+AND ap/(904) and ISTI/(900)
```

Si noti il carattere '+' che è un **InternalSetDefinitionDelimiter** e il carattere ';' che è un **SetDefinitionDelimiter**.

2. **InternalSetDefinitionDelimiter** (+%#) definisce un insieme di caratteri delimiter ognuno dei quali diversifica i 4 parametri della definizione del Set.
3. **SetDefinitionDelimiter** (:) definisce un insieme di caratteri delimiter ognuno dei quali diversifica i Set definiti. Naturalmente **SetDefinitionDelimiter** deve contenere caratteri che siano completamente diversi da **InternalSetDefinitionDelimiter**.
4. **MaxSetDefined** (100) massimo numero di Set che è possibile definire. Il valore può essere elevato senza alcun problema.
5. **MaxSetLevelHierarchy** (5) specifica il massimo livello di gerarchia definibile tra i Set.

7.5. I parametri dei resumptionToken

I parametri dei resumptionToken vengono utilizzati dal gateway per memorizzare le posizioni di restart di un insieme di identifier/record il cui numero è superiore a **MaxObjectsToBeDisplayed**. Le informazioni sono memorizzate in file univoci da cui è possibile procedere al restart. Ad ognuno dei **resumptionToken/file** viene assegnato un tempo di esistenza al termine del quale l'informazione è obsoleta. In questa maniera è possibile procedere a successive acquisizioni degli identifier/record o procedere a un eventuale recovery di informazioni perse. Per maggiori dettagli sull'organizzazione e la struttura dell'informazione dei **resumptionToken** vedi " 5.5. Il restart dell'accesso ai dati ".

1. **MaxResumptionTokenParms** (10) definisce il massimo numero di parametri che costituiscono l'informazione del **ResumptionToken** e vengono prodotti e utilizzati dal gateway per le operazioni di save/restart.
2. **ResumptionTokenWriteDelim** (#) definisce il carattere delimiter tra i parametri che costituiscono l'informazione del **ResumptionToken**.
3. **ResumptionTokenReadDelim** (%#;.,) definisce un insieme di caratteri ognuno dei quali può essere utilizzato nell'interpretazione dei parametri del **ResumptionToken**.
4. **ResumptionTokenTTL** (3600000) definisce il tempo di esistenza del **resumptionToken**. Il valore è espresso in millisecondi.
5. **ResumptionTokenDir** (/oai/RT/ [C:/oai/RT](#)) definisce la directory su cui memorizzare i **ResumptionToken** del repository.
6. **ResumptionTokenFilePrefix** (<ExtGASoaiArchiveName>) definisce il prefisso da utilizzare nella costruzione del **resumptionToken/file** del repository
7. **ResumptionTokenFileSuffix** (RT) definisce il suffisso da utilizzare nella costruzione del **resumptionToken/file** del repository

Bibliografia

1. Giuseppe A. Romano - Isis/GAS : un sistema multi piattaforma per l'accesso a basi di dati testuali via Internet - ISTI-B4-28, 2003 - URL:
<http://leonardo.isti.cnr.it/metaopac/IsisGas.pdf>
2. Giuseppe A. Romano - Isis/GAS : manuale di installazione e d'uso - ISTI-B4-31, 2003 - URL: <http://leonardo.isti.cnr.it/metaopac/IsisGasInstall.pdf>
3. Giuseppe A. Romano - Il MetaOpac Pisano: l'organizzazione logica e il mapping fisico - ISTI-B4-02, 2003 - URL:
<http://leonardo.isti.cnr.it/metaopac/MetaOpacLogic.pdf>
4. Giuseppe A. Romano - Il MetaOpac Pisano: descrizione tecnica e procedure di recovery - ISTI-B4-30, 2003 - URL:
<http://leonardo.isti.cnr.it/metaopac/MetaopacPisano.pdf>
5. Giuseppe A. Romano - Il metaopac: controllo e recovery del sistema - ISTI-B4-07, 2004 - URL: http://leonardo.isti.cnr.it/metaopac/IsisGAS_BriefMan.pdf
6. Biagioni S., Giannini S., Giuseppe A. Romano - Il Metaopac del CNR di Pisa e la sua collocazione nel Metaopac pisano - ISTI-B4-14, 2004 - URL:
<http://leonardo.isti.cnr.it/metaopac/OpacCNR.pdf>
7. Pierlugi Niccolai, Enrico Venturini - TECA: Applicazione di CDS/ISIS per i dati bibliografici - Titivillus , Firenze 1993
8. Stefania Biagioni, Carlo Carlesi, Daniela Musa - Sistema Atlas: catalogazione derivata da registrazioni in formato UNIMARC - IEI-B4-30 : Istituto di Elaborazione della Informazione, settembre 1994
9. CDS/ISIS - software, tools, descriptions - URL:
http://portal.unesco.org/ci/ev.php?URL_ID=2071&URL_DO=DO_TOPIC&URL_SECTION=201&reload=1082372136&PHPSESSID=7462f4450cf4d320877976df1ae94292
10. BIREME - CISIS reference manual - URL:
<http://productos.bvsalud.org/html/en/home.html>
11. BIREME - wwwisis version 3.0 - URL:
<http://productos.bvsalud.org/html/en/home.html>
12. LIBERO - software - URL: <http://www.infologic.it/ita/libero.htm>
13. Antonio Scolari - Unimarc, Roma, AIB, 2000 Enciclopedia Tascabile ETIS
14. Giuseppe A. Romano - Il MetaOpacPisano : organizzazione, controllo e recovery del sistema per l'anno 2005 - ISTI-B4-17,2005 URL:
http://leonardo.isti.cnr.it/metaopac/MOP_BriefMan.pdf
15. Giuseppe A. Romano - ExtGAS un sistema multipiattaforma per l'accesso ai cataloghi di sistemi biblioteconomici eterogenei - ISTI-B4-15,2005 - URL:
<http://leonardo.isti.cnr.it/metaopac/ExtGas.pdf>

16. BIREME - Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde - WWWISIS: a world-wide web server for ISIS-databases Version 3.0, URL <http://leonardo.isti.cnr.it/metaopac/wwwisis.pdf>
17. Giuseppe A. Romano - ExtGAS manuale di installazione e d'uso - ISTI-B4-16,2005 - URL: <http://leonardo.isti.cnr.it/metaopac/ExtGasInstall.pdf>
18. Giuseppe A. Romano - ExtGas/IsisGAS revisione 2.0: manuale di installazione e d'uso - ISTI-B4-16, 2006 URL: <http://leonardo.isti.cnr.it/metaopac/ExtGASOmoDbInstall.pdf>