CNS 國家標準草案：開放典藏推動之詮釋資料擷取協定

(The Open Archives Initiative Protocol for Metadata Harvesting, 2004)

中央研究院. 數位典藏國家型科技計畫後設資料工作組

與

國立台灣師範大學圖書資訊學研究所

研譯

經濟部標準檢驗局 96 年度 委辦計畫

規劃與建置數位內容與數位生活應用之技術標準環境

(案號：1D15960125-20)

The Open Archives Initiative Protocol for Metadata Harvesting

目次

1.適用範圍

Open Archives Initiative Protocol for Metadata Harvesting(OAI-PMH)協定係基於詮釋資料擷取, 提供與應用無關的互運框架. 在這框架下有二種類別之參與者:

(1) 資料提供者為一管理系統，係支援以 OAI-PMH 對外提供詮釋資料；

(2) 服務提供者，係透過使用 OAI-PMH 標準擷取詮釋資料，作為建構加值服務之基礎。

本文件中提到的“必須(must)”、“不須(must not) ”、“需要(required) ”、“應(shall) ”、“不應(shall not) ”、“宜(should) ”、“不宜(should not) ”、“建議(recommended) ”、“可(may) ”以及“選用(optional) ”等黑體字元是參照 RFC2119 格式表示。若一程式未滿足一至多個“必須”或“需要”等級需求，則此實作結果將無法相容協定。

The Open Archives Initiative Protocol for Metadata Harvesting (referred to as the OAI-PMH in the remainder of this document) provides an application-independent interoperability framework based on metadata harvesting. There are two classes of participants in the OAI-PMH framework:

(1) Data Providers administer systems that support the OAI-PMH as a means of exposing metadata; and

(2) Service Providers use metadata harvested via the OAI-PMH as a basis for building value-added services.

In this document the key words "must", "must not", " required", "shall", "shall not", "should", " should not", "recommended", "may", and "optional " in bold face are to be interpreted as described in RFC 2119 . An implementation is not conformant if it fails to satisfy one or more of the "must" or "required" level requirements for the protocols it implements.

2.用語釋義

2.1 擷取器

擷取器是發出 OAI-PMH 請求的客戶端應用程式, 通常受服務提供者所驅動運作，用以由儲存庫收集詮釋資料。

2.1 Harvester

A harvester is a client application that issues OAI-PMH requests. A harvester is operated by a service provider as a means of collecting metadata from repositories.

2.2 儲存庫

儲存庫即為可以進行網路存取的伺服器, 且此伺服器可以處理本文中所描述的六種 OAI-PMH 請求。儲存庫由資料提供者進行管理, 以提供詮釋資料讓擷取器使用。為了讓多種組態的儲存庫可以進行運作, OAI-PMH 將對外提供詮釋資料的方式分為三種不同層次：

(1) 資源：詮釋資料所關連到的物件, 至於資源本身的特性, 如實體的或是數位的, 或是儲存於儲存庫之中或是連到到外界其他的資料庫, 並不在 OAI-PMH 標準的討論範圍中。

(2) 資料項：儲存庫中用以描述資源的詮釋資料, 且可以進行散佈。詮釋資料可以用特定的表示形式, 透過其關連到儲存庫中的資源進行散佈。

(3) 紀錄：使用特定格式表達的詮釋資料。在本標準中, 使用特定格式的詮釋資料, 必須使用 XML 編碼成為資料流的形式, 來作為某協定請求的回應。

2.2 Repository

A repository is a network accessible server that can process the 6 OAI-PMH requests in the manner described in this document. A repository is managed by a data provider to expose metadata to harvesters. To allow various repository configurations, the OAI-PMH distinguishes between three distinct entities related to the metadata made accessible by the OAI-PMH:

(1) resource - A resource is the object or "stuff" that metadata is "about". The nature of a resource, whether it is physical or digital, or whether it is stored in the repository or is a constituent of another database, is outside the scope of the OAI-PMH.

(2) item - An item is a constituent of a repository from which metadata about a resource can be disseminated. That metadata may be disseminated

on-the-fly from the associated resource, cross-walked from some canonical form, actually stored in the repository, etc.

(3) record - A record is metadata in a specific metadata format. A record is returned as an XML-encoded byte stream in response to a protocol request to disseminate a specific metadata format from a constituent item.

2.3 資料項

資料項是儲存庫的基本組成單位, 其內容為描述用以傳佈資源的詮釋資料。資料項在概念上可以視為儲存或是動態產生多種格式詮釋資料的容器(container), 各項目可透過 OAI-PMH 來進行紀錄的擷取動作。各資料項在儲存庫中皆會擁有唯一的識別符。

2.3 Item

An item is a constituent of a repository from which metadata about a resource can be disseminated. An item is conceptually a container that stores or dynamically generates metadata about a single resource in multiple formats, each of which can be harvested as records via the OAI-PMH. Each item has an identifier that is unique within the scope of the repository of which it is a constituent.

2.4 唯一識別符

唯一的識別符為不會讓資料項在儲存庫之中產生識別混淆的代碼;此識別符是在 OAI-PMH 取用資料項的詮釋資料請求下使用。資料項可包含多種格式的詮釋資料。此唯一的識別符會對應到資料項以及所有使用相同識別符的可能紀錄。

唯一識別符的格式必須參考 URI (統一資源識別符, Uniform Resource Identifier)語法進行表達。不同的社群可以針對定義出該社群特定的 URI 方案, 以利未來在不同的儲存庫之間進行運作。唯一識別符的方案基本上不可以與現有的 URI 方案相同, 除非該識別符與這些方案相容。儲存庫則可以實作出在實作指導綱要(Implementation Guidelines document)所描述的 OAI 識別符語法。

2.4 Unique Identifier

A unique identifier unambigiously identifies an item within a repository; the unique identifier is used in OAI-PMH requests for extracting metadata from the item. Items may contain metadata in multiple formats. The unique identifier

maps to the item, and all possible records available from a single item share the same unique identifier.
The format of the unique identifier must correspond to that of the URI (Uniform Resource Identifier) syntax. Individual communities may develop community-specific URI schemes for coordinated use across repositories. The scheme component of the unique identifiers must not correspond to that of a recognized URI scheme unless the identifiers conform to that scheme. Repositories may implement the oai-identifier syntax described in the accompanying Implementation Guidelines document.

唯一識別符在協定中扮演的角色有二:

(1) 回應：識別符會在 ListIdentifiers 以及 ListRecords 兩項請求中作為回應內容。

(2) 請求：結合詮釋資料前綴屬性(metadataPrefix)使用的識別符，可以用在 GetRecord 請求中，作為請求特定詮釋資料格式紀錄的方法。

Unique identifiers play two roles in the protocol:

(1) Response: Identifiers are returned by both the ListIdentifiers and ListRecords requests.

(2) Request: An identifier, in combination with a metadataPrefix, is used in the GetRecord request as a means of requesting a record in a specific metadata format from an item.

備考：在這裡所描述的識別符並非指資源識別符。資源識別符的特性並不在 OAI-PMH 討論的範圍內。為了便利那些擁有詮釋資料資源的存取，儲存庫必須使用詮釋資料紀錄中的某個元件來建立紀錄(以及該紀錄的識別符)以及資源識別符(如 URL、URN、DOI 等等)之間的連結關係。都柏林核心集格式中必備的欄位中就有識別符元件可以供上述需要使用。

Note that the identifier described here is not that of a resource. The nature of a resource identifier is outside the scope of the OAI-PMH. To facilitate access to the resource associated with harvested metadata, repositories should use an element in metadata records to establish a linkage between the record (and the identifier of its item) and the identifier (URL, URN, DOI, etc.) of the associated

resource. The mandatory Dublin Core format provides the identifier element that should be used for this purpose.

2.5 紀錄

紀錄是以單一格式表示的詮釋資料。一筆紀錄可以使用 XML 編碼成為資料流的形式，以作為 OAI-PMH 請求某資料項詮釋資料的回應。紀錄透過結合資料項的唯一識別符(在此指那些擁有紀錄的項目)，以及使用資料前綴屬性(metadataPrefix，該筆紀錄藉以進行詮釋資料格式識別)以提供明確的識別。使用 XML 編碼的紀錄在組織上擁有以下幾個部份：

2.5 Record

A record is metadata expressed in a single format. A record is returned in an XML-encoded byte stream in response to an OAI-PMH request for metadata from an item. A record is identified unambigiously by the combination of the unique identifier of the item from which the record is available, the metadataPrefix identifying the metadata format of the record, and the datestamp of the record. The XML-encoding of records is organized into the following parts:

(1) 標頭 (header)：此部份包含該資料項的唯一識別符，以及必須用於選擇性擷取之性質。標頭是由以下部份所組成：

(a) 唯一識別符：於儲存庫某資料項之唯一識別符；

(b) 日戳(datestamp)：提供用於選擇性擷取動作之紀錄建立、修改以及刪除的日期。

(c) 0 或多個 setSpec 元件：提供用於選擇性擷取動作之資料項集合成員屬性。

(d) 值為”已刪除”之選擇性狀態屬性係表示某資料項的特定詮釋資料格式已遭到刪除不再供存取之用。此部份與儲存庫支援刪除的能力有關。

(2) 詮釋資料 (metadata)：指資料項中單一的詮釋資料呈現。OAI-PMH 目前支援資料項擁有多種的詮釋資料呈現(也就是格式)。在此儲存庫所提供最低限度的支援，就是使用都柏林核心集格式描述的紀錄詮釋資料

(且不包含任何的自訂屬性)。除此之外，儲存庫也可以傳佈其他格式的詮釋資料。如果要使用特定的詮釋資料格式，就必須透過指定詮釋資料前綴(metadataPrefix)，作為 GetRecord 或是 ListRecord 請求用的參數，藉以產生紀錄。而使用 ListMetadataFormats 請求則可以取得儲存庫或是針對某資料項所有可使用的詮釋資料格式列表(這可以透過指定 ListMetadataFormats 請求參數來達到目的)。

(3) 其他相關 (about)：此部份係為選項且可重複之容器，容器持有之資料係用以表達該紀錄的詮釋資料部份，且其內容必須符合 XML 綱要(XML Schema)。實作出本標準的社群可建立定義此容器的 XML 綱要。此相關部份的容器最常見的使用方式為：

(a) 權利說明(rights statements)：部份儲存庫可能會需要在外界透過 OAI-PMH 使用詮釋資料時附帶描述這類說明。在 OAI-PMH 中，並沒有特定針對權利描述訂定任何特定的 XML 標籤，因此此容器可以允許使用社群自訂的權利描述標籤。

(b) 出處說明(provenance statements)：另外一種建議使用相關部份容器的方式，是用以描述詮釋紀錄的出處，如這項紀錄是如何被擷取的，從哪個儲存庫，以及何時被擷取。在本標準實作指導綱要文件(Implementation Guidelines document)中有針對出處說明容器以及其他相關支援資料設計出 XML 綱要可供使用。

(1) header -- contains the unique identifier of the item and properties necessary for selective harvesting. The header consists of the following parts:

(a) the unique identifier -- the unique identifier of an item in a repository;

(b) the datestamp -- the date of creation, modification or deletion of the record for the purpose of selective harvesting.

(c) zero or more setSpec elements -- the set membership of the item for the purpose of selective harvesting.

(d) an optional status attribute with a value of deleted indicates the withdrawal of availability of the specified metadata format for the item, dependent on the repository support for deletions.

(2) metadata -- a single manifestation of the metadata from an item. The OAI-PMH supports items with multiple manifestations (formats) of metadata. At a minimum, repositories must be able to return records with metadata expressed in the Dublin Core format, without any qualification. Optionally, a repository may also disseminate other formats of metadata. The specific metadata format of the record to be disseminated is specified by means of an argument -- the metadataPrefix -- in the GetRecord or ListRecords request that produces the record. The ListMetadataFormats request returns the list of all metadata formats available from a repository, or for a specific item (which can be specified as an argument to the ListMetadataFormats request).

(3) about -- an optional and repeatable container to hold data about the metadata part of the record. The contents of an about container must conform to an XML Schema. Individual implementation communities may create XML Schema that define specific uses for the contents of about containers. Two common uses of about containers are:

(a) rights statements: some repositories may find it desirable to attach terms of use to the metadata they make available through the OAI-PMH. No specific set of XML tags for rights expression is defined by OAI-PMH, but the about container is provided to allow for encapsulating community-defined rights tags.

(b) provenance statements: One suggested use of the about container is to indicate the provenance of a metadata record, e.g. whether it has been harvested itself and if so from which repository, and when. An XML Schema for such a provenance container, as well as some supporting information is available from the accompanying Implementation Guidelines document.

以下是使用 XML 編碼過的紀錄以及其組成：

(1) 標題：

(a) 資料項之唯一識別符，紀錄係傳播自該資料項，其範例內容為 oai:arXiv.org:cs/0112017。

(b) 該紀錄的日戳，內容為 2002-02-28。

(c) 兩個 setSpecs 元件，分別為 cs 以及 math，用以說明此紀錄分屬於儲存庫中兩個的集合成員。

(2) 詮釋資料：此部份是由單一的根標籤(root tag)組成，例如 oai_dc:dc，其下則由相對應的詮釋資料格式所屬的標籤構成巢狀的資料結構，都柏林核心集的元件 dc:title 即為一例。請注意在此詮釋資料中根標籤用法與一般 XML 文件相同，皆會涵括一些屬性，用以標示使用到的命名空間以及確認用的格式綱要：

(2.1)命名空間宣告(namespace declarations)：詮釋資料中會使用到命名空間的宣告，其中每個宣告皆會使用 xmlns 作為其前綴。命名空間的宣告在使用上有兩大類別：

(a) 與詮釋資料格式相關的特定命名空間(metadata format specific namespace)：每個詮釋資料的部份內容皆必須包含一或多個以 xmlns 為前綴的屬性，用以定義所使用的詮釋資料格式前綴(如 dc)以及相對的詮釋資料格式命名空間 URI (使用 XML 命名空間規格進行定

義)。部份詮釋資料格式可能會使用多個命名空間, 因此需要使用多個 xmlns 前綴屬性—在本例中就宣告了 oai_dc 以及 dc。

(b) XML 綱要命名空間(xml schema namespace)：每個詮釋資料的部份內容皆必須使用 xmlns:xsi 屬性, 在本例中, 其資料值必須為 XML 綱要的命名空間。

(2.2)xsi:schemaLocation：本標籤的資料值為 URI, URL 配對；第一部份為本部份詮釋資料的命名空間 URI (使用 XML 命名空間規格進行定義), 第二部份則為用以確認詮釋資料的 XML 綱要 URL。

(3) 該紀錄的相關部份(其中使用了 oai_provenance.xsd 綱要), 在實作指導綱要文件中有所說明, 這是用以提供紀錄值中該部份詮釋資料出處的方式。請注意每個說明部份(about part)的根元件皆必須和其詮釋資料的根元件具有相同的結構。

The following example shows an XML-encoding of a record and its components:

(1) the header part with:

(a) a unique identifier of the item from which the record was disseminated, equal to oai:arXiv.org:cs/0112017;

(b) the datestamp of the record equal to 2002-02-28;

(c) two setSpecs, respectively cs and math, indicating that the item from which the record was disseminated belongs to two sets of the repository;

(2) the metadata part. This consists of a single root tag - in the example the tag oai_dc:dc - with the nested tags belonging to the corresponding metadata format - in the example, Dublin Core elements such as dc:title. Note that the root tag within the metadata part includes a number of attributes that are common to all XML documents that use namespaces and schema validity:

(2.1)namespace declarations -- the declarations of the namespaces used within the metadata part, each of which is prefixed with xmlns.

Namespace declarations within the metadata part fall into two categories:

(a) metadata format specific namespace(s) - every metadata part must include one or more xmlns prefixed attributes that define the correspondence between a metadata format prefix -- e.g. dc -- and the namespace URI (as defined by the XML namespace specification) of the respective metadata format. Some metadata formats employ tags from multiple namespaces, requiring multiple xmlns prefixed attributes -- in the example, there are declarations for both oai_dc and dc.

(b) xml schema namespace - every metadata part must include the attribute xmlns:xsi, the value of which must always be the URI shown in the example, which is the namespace URI for XML schema.

(2.2)xsi:schemaLocation -- the value of which is a URI, URL pair; the first is the namespace URI (as defined by the XML namespace specification) of the metadata that follows in this part, and the second is the URL of the XML schema for validation of the metadata that follows.

(3) one about part of the record which uses the oai_provenance.xsd schema, described in the accompanying Implementation Guidelines document, as a means to provide information regarding the origins of the metadata part of the record. Note that the root element within each about part has the same structure as the root element in the metadata part.

```
<header>
  <identifier>oai:arXiv:cs/0112017</identifier>
  <datestamp>2002-02-28</datestamp>
  <setSpec>cs</setSpec>
  <setSpec>math</setSpec>
```

```
</header>
<metadata>
  <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>Using Structural Metadata to Localize Experience of Digital
                    Content</dc:title>
     <dc:creator>Dushay, Naomi</dc:creator>
     <dc:subject>Digital Libraries</dc:subject>
     <dc:description>With the increasing technical sophistication of both
       information consumers and providers, there is increasing demand for
       more meaningful experiences of digital information. We present a
       framework that separates digital object experience, or rendering,
       from digital object storage and manipulation, so the
       rendering can be tailored to particular communities of users.
     </dc:description>
     <dc:description>Comment: 23 pages including 2 appendices,
                          8 figures</dc:description>
     <dc:date>2001-12-14</dc:date>
     <dc:type>e-print</dc:type>
     <dc:identifier>http://arXiv.org/abs/cs/0112017</dc:identifier>
  </oai_dc:dc>
</metadata>
<about>
   <provenance
         xmlns="http://www.openarchives.org/OAI/2.0/provenance"
         xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
         xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/provenance
         http://www.openarchives.org/OAI/2.0/provenance.xsd">
      <originDescription harvestDate="2002-02-02T14:10:02Z" altered="true">
         <baseURL>http://the.oa.org</baseURL>
         <identifier>oai:r2:klik001</identifier>
         <datestamp>2002-01-01</datestamp>

<metadataNamespace>http://www.openarchives.org/OAI/2.0/oai_dc/</metadataNamespa
```

```
ce>
      </originDescription>
   </provenance>
</about>
```

2.5.1 已刪除紀錄(Deleted records)

如果一筆紀錄已無法再取得的話，就稱之為已刪除。儲存庫必須針對 Identify 回應中的 deletedRecord 元件，提出對已刪除紀錄的三個層級之一的支援：

(1) 不支援(no)：儲存庫並不針對刪除動作做任何資料維護。使用這種層級的儲存庫不可以在任何回應中揭露已刪除的狀態訊息。

(2) 永續的(persistent)：儲存庫針對刪除動作所做任何資料維護並沒有時間限制。使用這種層級的儲存庫必須持續的追蹤刪除動作的資訊，同時也必須常態的揭露已刪除紀錄的狀態訊息。

(3) 暫時的(transient)：儲存庫並不保證刪除列表會受到持續以及常態的維護。使用這種層級的儲存庫可以揭露已刪除紀錄的狀態訊息。

如果儲存庫並不去追蹤刪除動作，各項紀錄就會在回應中消失，同時也無法讓擷取器在長期的擷取動作中察覺刪除的現象。如果儲存庫針對刪除動作進行追蹤，則已刪除紀錄中日戳的內容就必須是遭到刪除的日期時間。針對已刪除紀錄的 GetRecord 請求的回應內容中，標頭部份必須有一狀態屬性(status)被設為‘deleted’，同時內容也不能包含詮釋資料或是相關說明部份。同理，針對選擇性擷取動作的請求，其請求內容中就必須有集合成員及日期範圍，其中已刪除紀錄也必須內含標頭資訊。因此透過長期的擷取動作就可以發覺並追蹤那些刪除動作.

已刪除狀態對每筆紀錄皆是實例。已刪除的紀錄如同正常紀錄，可使用唯一識別符，metadataPrefix 以及日戳來進行識別。至於那些擁有不同 metadataPrefix 卻使用相同唯一識別符的紀錄，該資料項仍然可以使用。

2.5.1 Deleted records

If a record is no longer available then it is said to be deleted. Repositories must declare one of three levels of support for deleted records in the deletedRecord element of the Identify response:

(1) no - the repository does not maintain information about deletions. A repository that indicates this level of support must not reveal a deleted status in any response.

(2) persistent - the repository maintains information about deletions with no time limit. A repository that indicates this level of support must persistently keep track of the full history of deletions and consistently reveal the status of a deleted record over time.

(3) transient - the repository does not guarantee that a list of deletions is maintained persistently or consistently. A repository that indicates this level of support may reveal a deleted status for records.

If a repository does not keep track of deletions then such records will simply vanish from responses and there will be no way for a harvester to discover deletions through continued incremental harvesting. If a repository does keep track of deletions then the datestamp of the deleted record must be the date and time that it was deleted. Responses to GetRecord request for a deleted record must then include a header with the attribute status="deleted", and must not include metadata or about parts. Similarly, responses to selective harvesting requests with set membership and date range criteria that include deleted records must include the headers of these records. Incremental harvesting will thus discover deletions from repositories that keep track of them.

Deleted status is a property of individual records. Like a normal record, a deleted record is identified by a unique identifier, a metadataPrefix and a datestamp. Other records, with different metadataPrefix but the same unique identifier, may remain available for the item.

2.6 集合(Set)

為了選擇性擷取目的而用以將資料項予以分群的選項。資料端可以將資料項組織成不同的集合。集合的組織可以是單層次的形式，如單一的列表。同時使用不同且獨立上層節點的多個階層架構也是被允許的。如果要將集合透過階層式架構的方式進行組織，可以使用 setSpec 參數語法進行描述，請參見以下說明。當儲存庫定義了一集合組織之後，其必須在回應的資料

項的標頭中包含集合成員的相關資訊，包括 ListIdentifiers ， ListRecords 以及 GetRecord 請求。
各儲存庫中集合組織的節點皆包含以下內容：

(1) setSpec：使用冒號[:]作為標示集合架構的自根節點以下至特定節點路徑的分隔符號。各列表中的元件皆代表由 URI 允許使用正常字元所構成的字串，且一定不能包含冒號[：]。由於 setSpec 可以在儲存庫中代表唯一識別符，因此各集合皆必須是唯一的。單層次的集合組織只能允許那些 setSpec 中不包含冒號[：]的資料。

(2) setName：此為針對集合所命名的可讀(human-readable)短字串。

(3) setDescription：此為選擇性且可以重複的容器元件，其可以擁有針對集合所定義的社群特定資料(經 XML 編碼過)；本標準的實作指導綱要中提供了使用這類容器的建議說明。

2.6 Set

A set is an optional construct for grouping items for the purpose of selective harvesting. Repositories may organize items into sets. Set organization may be flat, i.e. a simple list, or hierarchical. Multiple hierarchies with distinct, independent top-level nodes are allowed. Hierarchical organization of sets is expressed in the syntax of the setSpec parameter as described below. When a repository defines a set organization it must include set membership information in the headers of items returned in response to the ListIdentifiers , ListRecords and GetRecord requests.
Each node in a set organization of a repository has:

(1) a setSpec -- a colon [:] separated list indicating the path from the root of the set hierarchy to the respective node. Each element in the list is a string consisting of any valid URI unreserved characters, which must not contain any colons [:]. Since a setSpec forms a unique identifier for the set within the repository, it must be unique for each set. Flat set organizations have only sets with setSpec that do not contain any colons [:].

(2) a setName -- a short human-readable string naming the set.

(3) a setDescription -- an optional and repeatable container that may hold community-specific XML-encoded data about the set; the accompanying Implementation Guidelines document provides suggestions regarding the usage of this container.

以下為在儲存庫中集合階層架構的範例：

(1) 機構(Institutions)

(a) Oceanside University of Nebraska。

(b) Valley View University of Florida。

(2) 主題(Subjects)

(a) Existential Kenesiology。

(b) Quantum Psychology。

The following is an example of a possible set hierarchy in a repository:

(1) Institutions

(c) Oceanside University of Nebraska

(d) Valley View University of Florida

(2) Subjects

(c) Existential Kenesiology

(d) Quantum Psychology

下表說明了上述集合階層架構可能的資料表示，其中使用了 setName 以及相對應的 setSpec 資料值來進行表達。

The following table shows a possible representation of the above set hierarchy by means of setName and respective setSpec values.

setName	setSpec
Institutions	institution
Oceanside University of Nebraska	institution:nebraska

Valley View University of Florida	institution:florida
Subjects	subject
Existential Kenesiology	subject:kenesiology
Quantum Psychology	subject:quantum

一個資料項可以被組織在一集合中，也可以是數個集合，或者根本不屬於任一集合。在上例中，可以看見單一的資料項被同時組織在主題以及 institution:florida 中。擷取器不應該假設當擷取所有儲存庫集合內的資料時，就擷取到所有的資料項，因為資料項也可能只被組織在集合階層架構中的內部節點下。

An item may be organized in one set, several sets, or no sets at all. In the example above, it is conceivable that an individual item is organized in both subject and institution:florida. A harvester should not assume that harvesting every set in a repository will retrieve metadata from all items in the repository. Items may also be assigned to interior nodes in the set hierarchy.

集合或是儲存庫內對集合的管理真正的意義並不在此協定內定義，而只應該用以讓特定社群針對集合定義良好的組態，其中可能內含供 setNames 以及 setSpec 使用的控制辭彙，以及開發出將資料提供給擷取器的方式。舉例來說，特定領域的電子列印典藏庫可以同意使用儲存庫特定的集合來管理詮釋資料組態，且這些集合使用了控制的主題分類。

The actual meaning of a set or of the arrangement of sets in a repository is not defined in the protocol. It is expected that individual communities may formulate well-defined set configurations with perhaps a controlled vocabulary for setNames and setSpec , and may even develop mechanisms for exposing these to harvesters. For example, a group of cooperating e-print archives in a specific discipline may agree on sets that arrange metadata in their repositories based on a controlled subject classification.

儲存庫的集合階層架構在本協定中可以透過 setSpec 表示。ListSets 請求會傳回儲存庫所含集合的組態列表。各列表中的成員皆必須內含 setSpec 以及 setName，同時也可以內含 setDescription。ListRecords 請求以及 ListIdentifiers 可以內含選擇性的集合參數，內含資料為 setSpec 的資料值，可以用以指定選擇性擷取的特定標的集合。在上例中的集合階層架構，setSpec institution:nebraska 可以用在請求的傳回資料中，用以傳回那些紀錄在 setSpec 中的資料集合。在此必須注意以下五個重點：

(1) 如果儲存庫支援集合的話，則 ListIdentifiers，ListRecords 以及 GetRecords 請求必須包含集合成員的資訊。而 setSpec 列表應該要包含指稱集合成員的最少 setSpec 數目。以上例的集合階層架構來說，在 institution:florida 中的資料項的標頭就不應該包含 setSpec 的機構，這是因為其已經隱性的表達在 institution:florida 之中。

(2) 各資料項可以被組織在一或多個集合中；此意味著不同的 setSpec 參數可能會傳回相同的紀錄。

(3) 各資料項不需要隸屬於任一集合；此意味著即使針對所有可能的 setSpec，使用 ListRecords 請求，也不一定保證會傳回儲存庫內的所有紀錄。唯一可以確保擷取所有紀錄或標頭的方法，是使用沒有帶 setSpec 參數的 ListRecords 或 ListIdentifiers 請求。

(4) 當使用 setSpec 作為參數時，回應內容必須包含所有 setSpec 中指定的紀錄或是標頭，以及所有該集合下的所有子結構中的所有紀錄或標頭。以上例的集合階層架構來說，ListRecords 請求中的機構 setSpec 參數，將會導致所有 setSpec 資料值中與該機構相符以及其下所有子結構(institution:florida 與 institution:nebraska)的詮釋紀錄被傳回。

(5) 儲存庫中集合階層架構可以包含空集合。

A repository's set hierarchy is represented in the protocol via setSpecs. ListSets returns a list indicating the configuration of sets in a repository. Each member of this list must include a setSpec and a setName and may include a setDescription. ListRecords and ListIdentifiers requests may include an optional set argument, the value of which is a setSpec, to specify the target set for selective harvesting. In the previous example of a set hierarchy, the setSpec institution:nebraska could be used in a request to return only those records that are disseminated from items organized in the set represented by this setSpec. Five issues should be noted here:

(1) If a repository supports sets then it must include set membership information in response to ListIdentifiers, ListRecords and GetRecord

requests. The list of setSpec should include only the minimum number of setSpec required to specify the set membership. Using the previous example of a set hierarchy, the header for an item organized in set institution:florida should not include setSpec institution since that is implied by the setSpec institution:florida.

(2) An item may be organized in more than one set; meaning that different setSpec arguments may return the same record(s).

(3) An item need not be organized in any set; meaning that an exhaustive repetition of ListRecords requests with all possible setSpecs is not guaranteed to return all records in the repository. The only guaranteed methods of harvesting all records or headers are ListRecords or ListIdentifiers requests with no setSpec argument.

(4) When a setSpec is used as an argument, the response must include records or headers from all items in the set specified by the setSpec , and all records or headers from items in sets that are descendant from the specified set. Using the previous example of a set hierarchy, a setSpec of institution to the ListRecords request will return all records from metadata organized within the set with a setSpec value equal to institution and within the descendent sets with setSpec values equal to institution:florida and institution:nebraska.

(5) The set hierarchy of a repository may include sets that are empty.

2.7 選擇性擷取

選擇性擷取允許擷取器限制其擷取範圍為儲存庫部份的詮釋資料。OAI-PMH 會以兩種擷取準則來支援選擇性擷取，該等準則為 datestamps 以及集合成員，可於一個 OAI-PMH 請求中合併使用。

2.7 Selective Harvesting

Selective harvesting allows harvesters to limit harvest requests to portions of the metadata available from a repository. The OAI-PMH supports selective harvesting with two types of harvesting criteria that may be combined in an OAI-PMH request: datestamps and set membership.

2.7.1 選擇性擷取與日戳(Datestamp)

擷取器可以使用日戳來指定擷取的日期範圍，包含紀錄的建立時間、刪除時間、以及修改時間。如果要針對選擇性擷取指定日期範圍，就必須將日期列入 ListRecords 以及 ListIdentifiers 中的選擇性參數，分別指定起訖日期。擷取動作將會以指定的日期範圍參數進行運作，如果開始日期不存在(忽略)，則開始時間會延展到最早的日戳，如果結束時間不存在，則結束時間會延展到最近的日戳。範圍指定是包含端點的，此意味著開始日期必須解釋為“此日期之後或等於此日期”，而結束日期則必須解釋為“此日期之前或等於此日期”。因此開始時間必須與結束時間相等或是在其之前，否則儲存庫就必須發出 badArgument 的錯誤訊息。

2.7.1 Selective Harvesting and Datestamps

Harvesters may use datestamps to harvest only those records that were created, deleted, or modified within a specified date range. To specify datestamp-based selective harvesting, datestamps are included as values of the optional arguments, from and until, in the ListRecords and ListIdentifiers requests. Harvesting is restricted to the range specified by the from and until arguments, extending back to the earliest datestamp if from is omitted, and forward to the most recent datestamp if until is omitted. Range limits are inclusive: from specifies a bound that must be interpreted as "greater than or equal to", until specifies a bound that must be interpreted as "less than or equal to". Therefore, the from arugment must be less than or equal to the until argument. Otherwise, a repository must issue a badArgument error.

儲存庫對於選擇性擷取的支援可以到以日為單位的程度，同時對 Identify 請求則可以選擇性的支援到以秒計的程度。無論是在請求或是回應中，時間戳記的表達皆必須與 UTCdatetime 規格相容。儲存庫在紀錄改變時必須更新其日戳，間接也會影響到該紀錄的詮釋資料部份(以 XML 編碼過)。這樣的變更最少包含了以下幾件事情：變更紀錄的詮釋資料，變更紀錄的詮釋資料格式，引進新的詮釋資料格式，停止支援某種詮釋資料格式等等。

Repositories must support selective harvesting with the from and until arguments expressed at day granularity. Optional support for seconds granularity is indicated in the response to the Identify request. The value of datestamps in both requests and responses must comply to the specifications for UTCdatetime in this document. A repository must update the datestamp of a record if a change occurs, the result of which would be a change to the metadata part of the XML-encoding of the record. Such changes include, but are not limited to, changes to the metadata of the record, changes to the metadata format of the record, introduction of a new metadata format, termination of support for a metadata format, etc.

選擇性擷取的日戳範圍通常是使用在傳送 ListRecords 與 ListIdentifiers 請求中之 from 及 until 引數上。儲存庫必須針對特定變更的型式，對應此日戳範圍，並使用以下的規則來建立 ListRecords 的回應。至於 ListIdentifiers 請求的回應部份，也必須依循相同的規則，但是簡化為標頭部份而非紀錄。

(1) 修改日期(modification)：本回應必須包含針對特定 metadataPrefix 參數符合起訖日期，且在儲存庫中改變過的紀錄。

(2) 建立日期(creation)：本回應必須包含針對特定 metadataPrefix 參數符合起訖日期，且在儲存庫中可以被使用的紀錄。

(3) 刪除日期(deletion)：本項回應必須根據儲存庫針對刪除紀錄追蹤的支援程度而定，回應內容可以包含紀錄的標頭，針對特定 metadataPrefix 參數符合起訖日期，且在儲存庫中遭到刪除的部份。透過標頭中的狀態屬性表示已刪除之狀態，且不包含任何詮釋資料。

Datestamp ranges for selective harvesting are expressed in the from and until arguments that may be submitted in the ListRecords and ListIdentifiers requests. Repositories must use the following rules to create a ListRecords response matching the specified datestamp range according to the type of change that occured within the repository. The response to a ListIdentifiers request follows the same rules but is abbreviated to include only headers rather than records.

(1) modification - the response must include records, corresponding to the metadataPrefix argument, which have changed within the bounds of the from and until arguments.

(2) creation - the response must include records, corresponding to the metadataPrefix argument, that have become available from the repository within the bounds of the from and until arguments.

(3) deletion - depending on the level at which a repository keeps track of deleted records, the response may include headers of records, corresponding to the metadataPrefix argument, which have been withdrawn from the repository within the bounds of the from and until arguments. Deleted status is indicated via the status attribute of the header element and no metedata is included.

經由 GetRecord，ListRecords 或是 ListIdentifiers 請求傳回的標頭皆包含了日戳，用以反映出依上述規則所定義的建立、修改、或是刪除時間。

Every header returned by the GetRecord, ListRecords or ListIdentifiers requests contains a datestamp, which reflects the most recent date and time of the creation, modification, or deletion according to the rules defined above.

2.7.2 選擇性擷取與集合(Selective Harvesting and Sets)

擷取器可以透過指定集合成員來進行選擇性的擷取。如果要指定以集合為基礎的選擇性擷取，則必須在 ListRecords 以及 ListIdentifiers 請求中使用 setSpec 作為選擇性的集合參數，這樣一來便可以針對特定集合執行選擇性的擷取動作。

當 setSpec 作為參數時，其相對應的回應必須包含：

(1) 透過 setSpec 指定的項目，包括針對特定 metadataPrefix 的資料項，或是相符但已刪除的標頭。

(2) 針對特定 metadataPrefix 的資料項，或是相符但已刪除的標頭，此二者以下子結構中的所有資料或是標頭紀錄。

2.7.2 Selective Harvesting and Sets

Harvesters may specify set membership as a criteria for selective harvesting. To specify set-based selective harvesting, a setSpec is included as the value of the optional set argument to the ListRecords and ListIdentifiers requests, thereby specifying selective harvesting of records from items within the respective set.

When a setSpec is used as an argument, the response must include:

(3) the records corresponding to the metadataPrefix argument, or headers thereof in the case of deleted records, available from those items in the set specified by the setSpec;

(4) the records corresponding to the metadataPrefix argument, or headers thereof in the case of deleted records, available from those items in sets that are descendant from the specified set.

3.參考標準

[OAI-PMH] Open Archives Initiative Protocol for Metadata Harvesting （OAI-PMH）2.0, 2004. < http://www.openarchives.org/OAI/openarchivesprotocol.html >

4.協定特點

4.1 HTTP 內嵌的 OAI-PMH 請求(HTTP Embedding of OAI-PMH requests)

OAI-PMH 請求是以表達成 HTTP 請求的形式。常見的標準實作方式，便是使用標準的 Web 伺服器, 將其設定為可以分派 OAI-PMH 請求至相關的處理軟體的狀態。本段其他部份將描述與 HTTP 內嵌相關的協定層面。

4.1 HTTP Embedding of OAI-PMH requests

OAI-PMH requests are expressed as HTTP requests. A typical implementation uses a standard Web server that is configured to dispatch OAI-PMH requests to the software handling these requests. The remainder of this section describes the aspects of the protocol that are specific to the HTTP embedding.

4.1.1 HTTP 請求格式(HTTP Request Format)

OAI-PMH 的請求必須使用 HTTP GET 或是 POST 方法來送出，其中使用 POST 方法較具優勢，這是因為該方法對參數沒有長度的限制。儲存庫因

此必須支援 GET 與 POST 方法。對所有的請求來說，有共同單一的基礎 URL，此基礎 URL 指定了網際網路的主機(host)以及通訊埠(port)，以及選擇性的路徑名稱，用以將 HTTP 伺服器當作儲存庫。儲存庫會將基礎 URL 以 Identify 回應中的 baseURL 元件進行表達。請注意任何路徑的組合皆是由儲存庫的 HTTP 伺服器組態所決定。

除了基礎 URL 之外，所有的請求皆會包含一組關鍵詞參數，參數是以 key=value 的配對形式表達。多個參數可以以任意順序出現，且參數之間必須以[&]符號分開。各 OAI-PMH 請求皆必須包含最少一個 key=value 配對，用以指定該 OAI-PMH 請求是由擷取器所發出：

(1) 鍵值為‘verb’字串；

(2) 資料值則定義了其中一種 OAI-PMH 請求。

至於其餘的 key=value 配對的數量與內涵則依個別請求的參數而有所不同。

4.1.1 HTTP Request Format

OAI-PMH requests must be submitted using either the HTTP GET or POST methods. POST has the advantage of imposing no limitations on the length of arguments. Repositories must support both the GET and POST methods. There is a single base URL for all requests. The base URL specifies the Internet host and port, and optionally a path, of an HTTP server acting as a repository. Repositories expose their base URL as the value of the baseURL element in the Identify response. Note that the composition of any path is determined by the configuration of the repository's HTTP server.

In addition to the base URL, all requests consist of a list of keyword arguments, which take the form of key=value pairs. Arguments may appear in any order and multiple arguments must be separated by ampersands [&]. Each OAI-PMH request must have at least one key=value pair that specifies the OAI-PMH request issued by the harvester:

(1) key is the string 'verb';

(2) value is one of the defined OAI-PMH requests.

The number and nature of additional key=value pairs depends on the arguments for the individual request.

4.1.1.1 針對 HTTP GET 方法將 OAI-PMH 請求進行編碼

GET 請求是將關鍵詞參數串接到基礎 URL 之後，並透過問號[?]隔開。舉例來說，以 GetRecord 請求的 URL 為例，儲存庫基礎 URL 為 http://an.oa.org/OAI-script 時，就應該成為：

http://an.oa.org/OAI-script?verb=GetRecord&identifier=oai:arXiv.org:hep-th/9901001&metadataPrefix=oai_dc

但無論如何，由於 URI 中特殊的字元必須經過編碼，因此上述的 GET 請求 URL 應該寫成：

http://an.oa.org/OAI-script?verb=GetRecord&identifier=oai%3AarXiv.org%3Ahep-th%2F9901001&metadataPrefix=oai_dc

4.1.1.1 Encoding an OAI-PMH request in a URL for an HTTP GET

URLs for GET requests have keyword arguments appended to the base URL, separated from it by a question mark [?]. For example, the URL of a GetRecord request to a repository with base URL that is http://an.oa.org/OAI-script might be:

http://an.oa.org/OAI-script?verb=GetRecord&identifier=oai:arXiv.org:hep-th/9901001&metadataPrefix=oai_dc

However, since special characters in URIs must be encoded, the correct form of the above GET request URL is:

http://an.oa.org/OAI-script?verb=GetRecord&identifier=oai%3AarXiv.org%3Ahep-th%2F9901001&metadataPrefix=oai_dc

4.1.1.2 針對 HTTP POST 方法將 OAI-PMH 請求進行編碼

HTTP POST 方法是將關鍵詞參數透過訊息本體(message body)進行傳遞。請求的 Content-Type 部份必須定為 application/x-www-form-urlencoded。舉例來說，上例的請求如果透過 POST 方法傳遞時，URL 部份為基礎 URL，POST 的其餘部份則為：

POST http://an.oa.org/OAI-script HTTP/1.0
Content-Length: 82
Content-Type: application/x-www-form-urlencoded

verb=GetRecord&identifier=oai%3AarXiv.org%3Ahep-th%2F9901001&metadataPrefix=oai_dc

4.1.1.2 Encoding an OAI-PMH request in an HTTP POST

Keyword arguments are carried in the message body of the HTTP POST. The Content-Type of the request must be application/x-www-form-urlencoded. For example, submitting the same request as above using the POST method would use just the base URL as the URL, with the format of the POST being:

POST http://an.oa.org/OAI-script HTTP/1.0
Content-Length: 82
Content-Type: application/x-www-form-urlencoded

verb=GetRecord&identifier=oai%3AarXiv.org%3Ahep-th%2F9901001&metadataPrefix=oai_dc

4.1.1.3 將 OAI-PMH 請求中的特殊字元進行編碼

URI 的語法規則限制了部份特殊字元的使用，如果需要使用這些字元，必須透過逸出序列(escape sequence)的方式表達，即用百分符號加上該字元的十六進位碼。這些保留的字元包括：

4.1.1.3 Encoding of special characters in keyword arguments of OAI-PMH requests

The syntax rules for URIs restrict a few characters to special roles in certain contexts, and require that if these characters are used in any other way that they must be written as an escape sequence, i.e. a percent sign followed by the character code in hexadecimal. The reserved characters include:

字元	URI 角色	逸出序列
/	路徑區隔符號	%2F
?	查詢字串區隔符號	%3F
#	片段識別符(Fragment Identifier)	%23
=	名稱/資料值區隔符號	%3D
&	查詢字串中的參數區隔符號	%26
:	主機與連接埠區隔符號	%3A
;	授權命名空間區隔符號	%3B
	空白字元	%20
%	逸出字元(Escape Indicator)	%25
+	逸出空白字元(Escaped Space)	%2B

Character	URI Role	Escape Sequence
/	Path Component Separator	%2F
?	Query Component Separator	%3F
#	Fragment Identifier	%23
=	Name/Value Separator	%3D
&	Argument Separator in Query Component	%26
:	Host Port Separator	%3A
;	Authority Namespace Separator	%3B

	Space Character	%20
%	Escape Indicator	%25
+	Escaped Space	%2B

因此, 以上這些字元必須在URI中表示成相對應的逸出序列, 在OAI-PMH的場合下，此意味著這些保留字元當出現在 key=value 配對中的資料值時，就必須進行編碼，這對不論 OAI-PMH 請求使用 GET 或是 POST 編碼時，皆是一樣的。

As a result, these characters must be represented by their respective escape sequence if their use does not correspond to their established URI role. In case of the OAI-PMH, this means that the reserved characters must be encoded when they appear in the value part of the key=value pairs of the request. This applies for both the GET and POST encoding of the OAI-PMH requests.

4.1.2 HTTP 回應格式(HTTP Response Format)

本標準中對於請求的回應是透過 HTTP 回應的方式，配合適當的 HTTP 標頭欄位來進行。

4.1.2 HTTP Response Format

Responses to requests are formatted as HTTP responses, with appropriate HTTP header fields.

4.1.2.1 內容型式(Content-Type)

所有 OAI-PMH 請求的傳回內容皆必須是 text/xml 型式.

4.1.2.1 Content-Type

The Content-Type returned for all OAI-PMH requests must be text/xml.

4.1.2.2 狀態碼(Status-Code)

OAI-PMH 的錯誤碼與 HTTP 的狀態碼並不相同。由於 OAI-PMH 使用 HTTP 作為其傳輸層，因此實作 OAI-PMH 的伺服器也必須配合使用 HTTP 狀態碼的定義，運用狀態碼來回報相關 HTTP 傳輸層的狀態。OAI-PMH 儲存庫可以使用除了‘200 OK’以外的 HTTP 狀態碼，舉例來說，以下的狀態碼對於 OAI 儲存庫的負載平衡就很有幫助：(1) 302：允許某儲存庫暫時將 OAI-PMH 請求重新導向至另一儲存庫。該暫時使用儲存庫的 URI 將會在 HTTP 回應中的 Location 欄位予以登錄。

(2) 503：服務無法使用，同時將會提供 Retry-After 時間區段。此時擷取器應該要等待這段時間過後才可以重新發出另一個 OAI-PMH 請求。

4.1.2.2 Status-Code

OAI-PMH errors are distinguished from HTTP Status-Codes. Since OAI-PMH uses HTTP as a transport layer, servers implementing OAI-PMH must conform to HTTP status code definitions and report relevant HTTP transport layer status via those Status-Codes. OAI-PMH repositories may employ HTTP Status-Codes in addition to "200 OK". For instance, the following Status-Codes may be useful for load balancing in OAI repositories:
(1) 302 - Allows the repository to temporarily redirect an OAI-PMH request to another repository. The URI of the temporary repository should be given by the Location field in the HTTP response.
(2) 503 - Service unavailable, a Retry-After period is specified. Harvesters
should wait this period before attempting another OAI-PMH request.

4.1.3 回應的壓縮(Response Compression)

回應的壓縮在 OAI-PMH 中是選擇性的。針對 OAI-PMH 請求的回應所進行的壓縮處理是發生在 HTTP 層次的事情，但是有以下的限制：

(1) 擷取器可以在 OAI-PMH 請求中內含 Accept-Encoding 標頭，用以指定回應所偏好使用的壓縮方式。

(2) 如果擷取器沒有在 OAI-PMH 請求中內含 Accept-Encoding 標頭，則其永遠皆會收到未經壓縮的回應內容。

(3) 當請求內含 Accept-Encoding 標頭時，編碼列表必須包含識別(未經壓縮)編碼(唯一非零的 qvalue)。

(4) 儲存庫必須支援 HTTP 識別編碼。

(5) 儲存庫除了識別之外，應該要提供所支援的編碼，這些是透過使用 Identify 回應中的壓縮元件中進行表達。

4.1.3 Response Compression

Response compression is optional in OAI-PMH. Compression of responses to OAI-PMH requests is handled at the level of HTTP, with the following restrictions:

(1) Harvesters may include an Accept-Encoding header in their OAI-PMH requests to specify response compression preferences.

(2) Harvesters that do not include an Accept-Encoding header in their requests will always receive uncompressed responses.

(3) When a request includes an Accept-Encoding header the list of encodings must include the Identify (no compression) encoding (with a non-zero qvalue).

(4) Repositories must support the HTTP Identify encoding.

(5) Repositories should express the encodings they support in addition to Identify by including compression elements in the Identify response.

4.2 XML 回應格式(XML Response Format)

所有 OAI-PMH 請求的回應皆必須是結構完整(well-formed)的 XML 文件。XML 在編碼部份必須使用 UTF-8 的方式來呈現 Unicode。字元的參考(Character references)必須被使用，而非單元的參考。字元參考允許 XML 回應被當作獨立的文件，此文件可以不需要仰賴其他項外參考的單元宣告。

4.2 XML Response Format

All responses to OAI-PMH requests must be well-formed XML instance documents. Encoding of the XML must use the UTF-8 representation of Unicode. Character references, rather than entity references, must be used. Character references allow XML responses to be treated as stand-alone documents that can be manipulated without dependency on entity declarations external to the document.

所有 OAI-PMH 請求的回應中，XML 資料部份必須經過 XML 綱要的確認(validate)，此部份在本節最後會進行說明。由於綱要中的內容的關係，OAI-PMH 請求的回應中必須內含以下幾個共同的標示：

The XML data for all responses to OAI-PMH requests must validate against the XML Schema shown at the end of this section . As can be seen from that schema, responses to OAI-PMH requests have the following common markup:

(1) 第一個輸出的標籤為 XML 宣告部份，XML 版本必須設定為 1.0，且其編碼必須設定為 UTF-8，例如：<?xml version="1.0" encoding="UTF-8" ?>

(2) 其他部份的內容將會由根元件所包含，名稱為 OAI-PMH。各個元件必須擁有三個屬性，定義回應的其餘部份也使用到的 XML 命名空間，以及用以確認的綱要位置：

(a) xmlns：此屬性的值必須為 OAI-PMH 的 URI (http://www.openarchives.org/OAI/2.0/)。

(b) xmlns:xsi：此屬性的值必須為 XML 綱要的命名空間 URI (http://www.w3.org/2001/XMLSchema-instance)。

(c) xsi:schemaLocation：此為配對，其中第一部份為 OAI-PMH 的命名空間 URI (這是由 XML 命名空間規範所定義) (http://www.openarchives.org/OAI/2.0/)，而第二部份則為用以確認回應內容的 XML 綱要 URL (http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd)。

(3) 對所有的回應而言，根目錄下的頭兩個元件為：

(a) 回應日期(responseDate)：此為 UTCdatetime 資料，用以表達回應送出的日期與時間。此部份資料必須以 UTC 形式表達。

(b) 請求(request)：用以表達產生本回應的請求。產生此請求元件的規則如下所示：

(b-1)請求元件的內容必須為協定請求的基礎 URL；

(b-2)請求元件中唯一合法的屬性為協定中所定義的 key=value 配對，屬性質必須反應出那些 key=value 的配對；

(b-3)如果這些特定請求所產生的回應皆沒有錯誤或是例外狀況時，屬性以及屬性值必須與請求中的 key=value 配對相符；

(b-4) 如果這些特定請求所產生的回應是 badVerb 或是 badArgument 等錯誤的情況時，儲存庫必須只傳回請求的基礎 URL。此時任一屬性皆不應該提供為回應之用。

(4) 根元件下的第三個子元件可以是下列其中之一：

(a) 表示錯誤的元件(error element)，這是在錯誤或是例外狀況發生時使用；

(b) 使用與 OAI-PMH 請求相同 verb 的元件。

(1) The first tag output is an XML declaration where the version is always 1.0 and the encoding is always UTF-8, eg: <?xml version="1.0" encoding="UTF-8" ?>

(2) The remaining content is enclosed in a root element with the name OAI-PMH. This element must have three attributes that define the XML namespaces used in the remainder of the response and the location of the validating schema:

(a) xmlns -- the value of which must be the namespace URI of the OAI-PMH (http://www.openarchives.org/OAI/2.0/).

(b) xmlns:xsi -- the value of which must be the namespace URI for XML schema (http://www.w3.org/2001/XMLSchema-instance).

(c) xsi:schemaLocation -- is a pair, the first part of which is the namespace URI (as defined by the XML namespace specification) of the OAI-PMH (http://www.openarchives.org/OAI/2.0/), and the second part is the URL of the XML schema for validation of the response (http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd).

(3) For all responses, the first two children of the root element are:

(a) responseDate -- a UTCdatetime indicating the time and date that the response was sent. This must be expressed in UTC.

(b) request -- indicating the protocol request that generated this response.

The rules for generating the request element are as follows:

(b-1)The content of the request element must always be the base URL of the protocol request;

(b-2)The only valid attributes for the request element are the keys of the key=value pairs of protocol request. The attribute values must be the corresponding values of those key=value pairs;

(b-3)In cases where the request that generated this response did not result in an error or exception condition, the attributes and attribute values of the request element must match the key=value pairs of the protocol request;

(b-4)In cases where the request that generated this response resulted in a badVerb or badArgument error condition, the repository must return the base URL of the protocol request only. Attributes must not be provided in these cases.

(4) The third child of the root element is either:

(a) an error element that must be used in case of an error or exception condition;

(b) an element with the same name as the verb of the respective OAI-PMH request.

下例為 GetRecord 請求成功的回應，其內容形式為:

An example of a successful reply to the GetRecord request shown above is of the form:

```
<?xml version="1.0" encoding="UTF-8" ?>
<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>2002-05-01T19:20:30Z</responseDate>
 <request verb="GetRecord" identifier="oai:arXiv.org:hep-th/9901001"
          metadataPrefix="oai_dc">http://an.oa.org/OAI-script</request>
 <GetRecord>
  <record>
      ...
```

```
  </record>
 </GetRecord>
</OAI-PMH>
```

4.2.1 供 OAI-PMH 請求確認回應的 XML 綱要(XML Schema for Validating Responses to OAI-PMH Requests)

4.2.1 XML Schema for Validating Responses to OAI-PMH Requests

```
<schema targetNamespace="http://www.openarchives.org/OAI/2.0/"
        xmlns="http://www.w3.org/2001/XMLSchema"
        xmlns:oai="http://www.openarchives.org/OAI/2.0/"
        elementFormDefault="qualified"
        attributeFormDefault="unqualified">

  <annotation>
    <documentation>
    XML Schema which can be used to validate replies to all OAI-PMH
    v2.0 requests. Herbert Van de Sompel, 2002-05-13.
    Validated with XML Spy v.4.3 on 2002-05-13.
    Validated with XSV 1.203.2.45/1.106.2.22 on 2002-05-13.
    Added definition of protocolVersionType instead of using anonymous
    type. No change of function. Simeon Warner, 2004-03-29.
    Tightened definition of UTCdatetimeType to enforce the restriction
    to UTC Z notation. Simeon Warner, 2004-09-14.
    Corrected pattern matches for setSpecType and metedataPrefixType
    to agree with protocol specification. Simeon Warner, 2004-10-12.
    $Date: 2004/10/12 15:20:29 $
    </documentation>
  </annotation>

  <element name="OAI-PMH" type="oai:OAI-PMHtype"/>

  <complexType name="OAI-PMHtype">
    <sequence>
      <element name="responseDate" type="dateTime"/>
      <element name="request" type="oai:requestType"/>
      <choice>
        <element name="error" type="oai:OAI-PMHerrorType"
maxOccurs="unbounded"/>
```

```
            <element name="Identify" type="oai:IdentifyType"/>
            <element name="ListMetadataFormats"
type="oai:ListMetadataFormatsType"/>
            <element name="ListSets" type="oai:ListSetsType"/>
            <element name="GetRecord" type="oai:GetRecordType"/>
            <element name="ListIdentifiers" type="oai:ListIdentifiersType"/>
            <element name="ListRecords" type="oai:ListRecordsType"/>
         </choice>
      </sequence>
   </complexType>

   <complexType name="requestType">
      <annotation>
         <documentation>Define requestType, indicating the protocol request that
         led to the response. Element content is BASE-URL, attributes are arguments
         of protocol request, attribute-values are values of arguments of protocol
         request</documentation>
      </annotation>
      <simpleContent>
         <extension base="anyURI">
            <attribute name="verb" type="oai:verbType" use="optional"/>
            <attribute name="identifier" type="oai:identifierType" use="optional"/>
            <attribute name="metadataPrefix" type="oai:metadataPrefixType"
use="optional"/>
            <attribute name="from" type="oai:UTCdatetimeType" use="optional"/>
            <attribute name="until" type="oai:UTCdatetimeType" use="optional"/>
            <attribute name="set" type="oai:setSpecType" use="optional"/>
            <attribute name="resumptionToken" type="string" use="optional"/>
         </extension>
      </simpleContent>
   </complexType>

   <simpleType name="verbType">
      <restriction base="string">
         <enumeration value="Identify"/>
         <enumeration value="ListMetadataFormats"/>
         <enumeration value="ListSets"/>
         <enumeration value="GetRecord"/>
```

```
        <enumeration value="ListIdentifiers"/>
        <enumeration value="ListRecords"/>
      </restriction>
    </simpleType>

    <!-- define OAI-PMH error conditions -->
    <!-- ============================== -->

    <complexType name="OAI-PMHerrorType">
      <simpleContent>
        <extension base="string">
          <attribute name="code" type="oai:OAI-PMHerrorcodeType"
use="required"/>
        </extension>
      </simpleContent>
    </complexType>

    <simpleType name="OAI-PMHerrorcodeType">
      <restriction base="string">
        <enumeration value="cannotDisseminateFormat"/>
        <enumeration value="idDoesNotExist"/>
        <enumeration value="badArgument"/>
        <enumeration value="badVerb"/>
        <enumeration value="noMetadataFormats"/>
        <enumeration value="noRecordsMatch"/>
        <enumeration value="badResumptionToken"/>
        <enumeration value="noSetHierarchy"/>
      </restriction>
    </simpleType>

    <!-- define OAI-PMH verb containers -->
    <!-- ============================== -->

    <complexType name="IdentifyType">
      <sequence>
        <element name="repositoryName" type="string"/>
        <element name="baseURL" type="anyURI"/>
        <element name="protocolVersion" type="oai:protocolVersionType"/>
```

```
        <element name="adminEmail" type="oai:emailType"
maxOccurs="unbounded"/>
        <element name="earliestDatestamp" type="oai:UTCdatetimeType"/>
        <element name="deletedRecord" type="oai:deletedRecordType"/>
        <element name="granularity" type="oai:granularityType"/>
        <element name="compression" type="string" minOccurs="0"
maxOccurs="unbounded"/>
        <element name="description" type="oai:descriptionType"
                    minOccurs="0" maxOccurs="unbounded"/>
      </sequence>
    </complexType>

    <complexType name="ListMetadataFormatsType">
      <sequence>
        <element name="metadataFormat" type="oai:metadataFormatType"
maxOccurs="unbounded"/>
      </sequence>
    </complexType>

    <complexType name="ListSetsType">
      <sequence>
        <element name="set" type="oai:setType" maxOccurs="unbounded"/>
        <element name="resumptionToken" type="oai:resumptionTokenType"
minOccurs="0"/>
      </sequence>
    </complexType>

    <complexType name="GetRecordType">
      <sequence>
        <element name="record" type="oai:recordType"/>
      </sequence>
    </complexType>

    <complexType name="ListRecordsType">
      <sequence>
        <element name="record" type="oai:recordType" maxOccurs="unbounded"/>
        <element name="resumptionToken" type="oai:resumptionTokenType"
minOccurs="0"/>
```

```
      </sequence>
   </complexType>

   <complexType name="ListIdentifiersType">
      <sequence>
         <element name="header" type="oai:headerType" maxOccurs="unbounded"/>
         <element name="resumptionToken" type="oai:resumptionTokenType"
minOccurs="0"/>
      </sequence>
   </complexType>

   <!-- define basic types used in replies to
           GetRecord, ListRecords, ListIdentifiers -->
   <!-- ======================================== -->

   <complexType name="recordType">
      <annotation>
         <documentation>A record has a header, a metadata part, and
            an optional about container</documentation>
      </annotation>
      <sequence>
         <element name="header" type="oai:headerType"/>
         <element name="metadata" type="oai:metadataType" minOccurs="0"/>
         <element name="about" type="oai:aboutType" minOccurs="0"
maxOccurs="unbounded"/>
      </sequence>
   </complexType>

   <complexType name="headerType">
      <annotation>
         <documentation>A header has a unique identifier, a datestamp,
            and setSpec(s) in case the item from which
            the record is disseminated belongs to set(s).
            the header can carry a deleted status indicating
            that the record is deleted.</documentation>
      </annotation>
      <sequence>
         <element name="identifier" type="oai:identifierType"/>
```

```
        <element name="datestamp" type="oai:UTCdatetimeType"/>
        <element name="setSpec" type="oai:setSpecType" minOccurs="0"
maxOccurs="unbounded"/>
      </sequence>
      <attribute name="status" type="oai:statusType" use="optional"/>
   </complexType>

   <simpleType name="identifierType">
      <restriction base="anyURI"/>
   </simpleType>

   <simpleType name="statusType">
      <restriction base="string">
        <enumeration value="deleted"/>
      </restriction>
   </simpleType>

   <complexType name="metadataType">
      <annotation>
        <documentation>Metadata must be expressed in XML that complies
          with another XML Schema (namespace=#other). Metadata must be
          explicitly qualified in the response.</documentation>
      </annotation>
      <sequence>
        <any namespace="##other" processContents="strict"/>
      </sequence>
   </complexType>

   <complexType name="aboutType">
      <annotation>
        <documentation>Data "about" the record must be expressed in XML
        that is compliant with an XML Schema defined by a
community.</documentation>
      </annotation>
      <sequence>
        <any namespace="##other" processContents="strict"/>
      </sequence>
   </complexType>
```

```
  <complexType name="resumptionTokenType">
    <annotation>
      <documentation>A resumptionToken may have 3 optional attributes
        and can be used in ListSets, ListIdentifiers, ListRecords
        responses.</documentation>
    </annotation>
    <simpleContent>
      <extension base="string">
        <attribute name="expirationDate" type="dateTime" use="optional"/>
        <attribute name="completeListSize" type="positiveInteger"
use="optional"/>
        <attribute name="cursor" type="nonNegativeInteger" use="optional"/>
      </extension>
    </simpleContent>
  </complexType>

  <complexType name="descriptionType">
    <annotation>
      <documentation>The descriptionType is used for the description
      element in Identify and for setDescription element in ListSets.
      Content must be compliant with an XML Schema defined by a
      community.</documentation>
    </annotation>
    <sequence>
      <any namespace="##other" processContents="strict"/>
    </sequence>
  </complexType>

  <simpleType name="UTCdatetimeType">
    <annotation>
      <documentation>Datestamps are to either day (type date)
      or to seconds granularity (type oai:UTCdateTimeZType)</documentation>
    </annotation>
    <union memberTypes="date oai:UTCdateTimeZType"/>
  </simpleType>

  <simpleType name="UTCdateTimeZType">
```

```
    <restriction base="dateTime">
      <pattern value=".*Z"/>
    </restriction>
  </simpleType>

  <!-- define types used for Identify verb only -->
  <!-- ======================================== -->

  <simpleType name="protocolVersionType">
    <restriction base="string">
      <enumeration value="2.0"/>
    </restriction>
  </simpleType>

  <simpleType name="emailType">
    <restriction base="string">
      <pattern value="\S+@(\S+\.)+\S+"/>
    </restriction>
  </simpleType>

  <simpleType name="deletedRecordType">
    <restriction base="string">
      <enumeration value="no"/>
      <enumeration value="persistent"/>
      <enumeration value="transient"/>
    </restriction>
  </simpleType>

  <simpleType name="granularityType">
    <restriction base="string">
      <enumeration value="YYYY-MM-DD"/>
      <enumeration value="YYYY-MM-DDThh:mm:ssZ"/>
    </restriction>
  </simpleType>

  <!-- define types used for ListMetadataFormats verb only -->
  <!-- ================================================== -->
```

```
  <complexType name="metadataFormatType">
    <sequence>
      <element name="metadataPrefix" type="oai:metadataPrefixType"/>
      <element name="schema" type="anyURI"/>
      <element name="metadataNamespace" type="anyURI"/>
    </sequence>
  </complexType>

  <simpleType name="metadataPrefixType">
    <restriction base="string">
      <pattern value="[A-Za-z0-9\-_\.!~\*'\(\)]+"/>
    </restriction>
  </simpleType>

  <!-- define types used for ListSets verb -->
  <!-- ================================== -->

  <complexType name="setType">
    <sequence>
      <element name="setSpec" type="oai:setSpecType"/>
      <element name="setName" type="string"/>
      <element name="setDescription" type="oai:descriptionType"
                  minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
  </complexType>

  <simpleType name="setSpecType">
    <restriction base="string">
      <pattern value="([A-Za-z0-9\-_\.!~\*'\(\)])+(:[A-Za-z0-9\-_\.!~\*'\(\)]+)*"/>
    </restriction>
  </simpleType>

</schema>
```

This Schema is available at http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd

==========

4.3 UTCdatetime

本標準協定中，日期與時間資料皆一致使用 ISO8601 進行編碼，以 UTC 格式表示。當需要表達時間資料時，就必須使用特殊的 UTC 指定字元(‘Z’)。即使在時區未經指定的情形下，UTC 即為預設的時區。舉例來說，1957-03-20T20:30:00Z 就是指 UTC 8:30:00 PM on March 20th 1957。本協定中的請求和回應皆使用了 UTCdatetime，以下段落將繼續說明。

4.3 UTCdatetime

Dates and times are uniformly encoded using ISO8601 and are expressed in UTC throughout the protocol. When time is included, the special UTC designator ("Z") must be used. UTC is implied for dates although no timezone designator is specified. For example, 1957-03-20T20:30:00Z is UTC 8:30:00 PM on March 20th 1957. UTCdatetime is used in both protocol requests and protocol replies, in the way described in the following sections.

4.3.1 在協定請求中使用的 UTCdatetime(UTCdatetime in Protocol Requests)

日戳在 ListIdentifiers 以及 ListRecords 請求中是以選擇性的參數角色存在，用以紀錄起訖時間，編碼方面則是使用 ISO8601，表示內容則為 UTC。這些參數皆是用在選擇性擷取的場合下指定日戳之用。此外，這些參數也支援 ISO8601 中所謂的“完整日期”(Complete date)以及“完整日期加上時/分/秒”(Complete date plus hours, minutes and seconds)的形式。合法的資料格式為 YYYY-MM-DD 以及 YYYY-MM-DDThh:mm:ssZ，且這兩種參數形式的時間精細度皆是相同的。所有的儲存庫皆必須支援 YYYY-MM-DD 格式，如果儲存庫支援 YYYY-MM-DDThh:mm:ssZ 格式，則必須在識別回應(Identify response)中使用。如果由擷取器使用的請求其時間精細度超過儲存庫支援的精細度時，就會產生錯誤。

4.3.1 UTCdatetime in Protocol Requests

Datestamps used as values of the optional arguments from and until in the ListIdentifiers and ListRecords requests are encoded using ISO8601 and are expressed in UTC. These arguments are used to specify datestamp-based selective harvesting. These arguments support the "Complete date" and the "Complete date plus hours, minutes and seconds" granularities defined in ISO8601. The legitimate formats are YYYY-MM-DD and YYYY-MM-DDThh:mm:ssZ. Both arguments must have the same granularity. All repositories must support YYYY-MM-DD. A repository that supports YYYY-MM-DDThh:mm:ssZ should indicate so in the Identify response. A request by a harvester with finer granularity than that supported by a repository must produce an error.

4.3.2 在協定回應中使用的 UTCdatetime(UTCdatetime in Protocol Responses)

日戳通常在 ListIdentifiers、GetRecord、以及 ListRecords 的回應裡出現在紀錄的標頭中。這些日戳皆必須使用 ISO8601 進行編碼，表示內容則為 UTC；這些資料皆必須以儲存庫所支援最精細的程度進行表達。這些日戳資料皆是用在以日戳為基礎的選擇性擷取的場合下使用，同時也必須符合相對應的規則。

各協定回應皆包含 responseDate 元件，此元件是使用 UTC 標示的回應日期與時間，同時此資料也必須使用 ISO8601 中所謂的“完整日期加上時/分/秒”(Complete date plus hours, minutes and seconds)的形式進行編碼。

在協定回覆中使用 resumptionToken 也可以包含選擇性的參數，稱為 expirationDate，此參數也是以 UTC 格式表示。該也是使用 ISO8601 中所謂的”完整日期加上時/分/秒”(Complete date plus hours, minutes and seconds)的形式進行編碼，格式為 YYYY-MM-DDThh:mm:ssZ。

4.3.2 UTCdatetime in Protocol Responses

Datestamps appear in the headers of records that are returned in response to ListIdentifiers , GetRecord and ListRecords requests. These datestamps are encoded using ISO8601 and are expressed in UTC; they must be expressed in the finest granularity supported by the repository. The value of the datestamp must correspond to the rules for datestamp-based selective harvesting.

Each protocol response includes a responseDate element, which must be the time and date of the response in UTC. This is encoded using the "Complete date plus hours, minutes, and seconds" variant of ISO8601. This format is YYYY-MM-DDThh:mm:ssZ.

A resumptionToken in a protocol reply may include an optional argument expirationDate, which is expressed in UTC. This is encoded using the "Complete date plus hours, minutes, and seconds" variant of ISO8601. This format is YYYY-MM-DDThh:mm:ssZ.

4.4 metadataPrefix 與詮釋資料綱要(metadataPrefix and Metadata Schema)

OAI-PMH 在儲存庫針對紀錄提供多種詮釋資料格式的支援，而 ListMetadataFormats 請求則會傳回所有該儲存庫可使用的詮釋資料格式，其包含以下屬性：

(1) metadataPrefix 元件：用以在 OAI-PMH 請求中指定所使用詮釋資料格式的字串。metadataPrefix 可以由任何合法的 URI 未保留字元進行表達。metadataPrefix 參數可以用在 ListRecords、ListIdentifiers、以及 GetRecord

請求中，用以擷取紀錄，或是取出由 metadataPrefix 所指定的詮釋資料格式標頭紀錄；

(2) 詮釋資料綱要 URL(metadata schema URL)：此為表示 XML 綱要的 URL，此綱要是用以確認特定格是詮釋資料的正確性；

(3) XML 命名空間 URI，這是用以表達該命名空間的全域性識別符。

4.4 metadataPrefix and Metadata Schema

OAI-PMH supports the dissemination of records in multiple metadata formats from a repository. The ListMetadataFormats request returns the list of all metadata formatsavailable from a repository, each of which has the following properities:

(1) The metadataPrefix - a string to specify the metadata format in OAI-PMH requests issued to the repository. metadataPrefix consists of any valid URI unreserved characters. metadataPrefix arguments are used in ListRecords, ListIdentifiers, and GetRecord requests to retrieve records, or the headers of records that include metadata in the format specified by the metadataPrefix;

(2) The metadata schema URL - the URL of an XML schema to test validity of metadata expressed according to the format;

(3) The XML namespace URI that is a global identifier of the metadata format.

由 ListRecords 以及 GetRecord 請求傳回的每一筆紀錄，其詮釋資料皆必須與 XML 命名空間規範的習慣用法相容。此意味著詮釋資料部份的根元件必須包含 xmlns 屬性，其屬性資料值則為該詮釋資料格式的 XML 命名空間 URI。根元件同時也必須包含 xsi:schemaLocation 屬性，其屬性資料值包含了用以確認詮釋資料用的 XML 綱要的 URL。此 URL 必須與 metadataPrefix 所指的詮釋資料綱要 URL 相符，並作為 ListRecords 以及 GetRecord 請求的參數(此由 metadataPrefix 對應到詮釋資料綱要的部份，是由儲存庫定義 ListMetadataFormats 的回應所決定)。

The metadata in each record returned by ListRecords and GetRecord must comply with the conventions of the XML namespace specification. This means that the root element of the metadata part must contain an xmlns attribute, the

value of which is the XML namespace URI of the metadata format. The root element must also contain an xsi:schemaLocation attribute that has a value that includes the URL of the XML schema for validation of the metadata. This URL must match the URL of the metadata schema for the metadataPrefix included as an argument to the ListRecords or GetRecord request (the mapping from metadataPrefix to metadata schema is defined by the repository's response to the ListMetadataFormats request).

為了交互操作的目的，儲存庫必須使用都柏林核心集進行資料傳佈。因此，協定中必須保留 metadataPrefix `oai_dc'，同時定義都柏林核心集的詮釋資料綱要則為 http://www.openarchives.org/OAI/2.0/oai_dc.xsd。此外相對應的 XML 命名空間 URI 則為 http://www.openarchives.org/OAI/2.0/oai_dc/。

For purposes of interoperability, repositories must disseminate Dublin Core, without any qualification. Therefore, the protocol reserves the metadataPrefix `oai_dc', and the URL of a metadata schema for unqualified Dublin Core, which is http://www.openarchives.org/OAI/2.0/oai_dc.xsd. The corresponding XML namespace URI is http://www.openarchives.org/OAI/2.0/oai_dc/.

metadataPrefix 為‘all’是保留給未來之用。目前的實作不應該使用這樣的 metadataPrefix。
The metadataPrefix `all' is reserved for future use. Implementations should not use this metadataPrefix.

社群之間可以透過應用指導綱要來分享不同的 metadataPrefix，詮釋資料綱要以及詮釋資料格式的 XML 命名空間 URI，但這種應用指導綱要並不在 OAI-PMH 的討論範圍之內。本標準的實作指導綱要文件提供了一些 XML 綱要樣本以及部份實例文件作為參考，這些參考皆是常見的詮釋資料格式，如 MARC 以及 RFC1807 等。

Communities should adopt guidelines for sharing of metadataPrefixes, metadata schema and XML namespace URI's of metadata formats. Such guidelines are outside of the scope of the OAI-PMH. The accompanying Implementation Guidelines document provides some sample XML Schema and instance documents for common metadata formats such as MARC and RFC 1807.

4.5 流程控制(Flow Control)

許多 OAI-PMH 請求皆會傳回單元列表：如 ListRecords 傳回紀錄的列表, ListIdentifiers 傳回標頭列表，ListSets 則會傳回集合列表。這些請求我們通稱為列表式的請求(list requests)。在某些場合下，這些列表可能會相當的大, 因此較實際的做法是分成一系列的請求與回應來取得資料。這種分割方式步驟來完成:

(1) 儲存庫針對請求回應不完整的列表，同時附帶繼續的標籤，稱為 resumptionToken；

(2) 為了將由回應中取得完整的資料列表，擷取器需要發出一或多個帶有以 resumptionTokens 作為參數的請求。完整的資料列表便可以透過一系列的請求來逐步取得，即所謂的列表請求序列(list request sequence)。

4.5 Flow Control

A number of OAI-PMH requests return a list of discrete entities: ListRecords returns a list of records, ListIdentifiers returns a list of headers, and ListSets returns a list of sets. Collectively these requests are called list requests. In some cases, these lists may be large and it may be practical to partition them among a series of requests and responses. This partitioning is accomplished as follows:

(1) A repository replies to a request with an incomplete list and a resumptionToken;

(2) In order to make the response a complete list, the harvester will need to issue one or more requests with resumptionTokens as arguments. The complete list then consists of the concatenation of the incomplete lists from the sequence of requests, known as a list request sequence.

以下為使用 resumptionToken 的詳細流程控制：

(1) 目前定義出使用 resumptionToken 的唯一方式如下：

(a) 儲存庫必須內含 resumptionToken 元件，作為傳回不完整資料列表回應的一部份；

(b) 為了要取回完整資料列表中的下一個部份，下一個請求必須使用上一個回應中 resumptionToken 元件的資料值，作為請求中 resumptionToken 元件的資料值；

(c) 如果傳回的不完整資料列表為最後一部份，則必須內含空的 resumptionToken 元件；

所有其他由擷取器所使用的 resumptionToken 方式皆會被視為不合標準，因此會傳回錯誤訊息。

(2) 不論是什麼情形下，當 resumptionToken 發出之後，傳回的不完整資料列表必須可以構成完整的單元；舉例來說，所有由經 ListRecords 請求所傳回的不完整列表中的個別紀錄皆必須是完整的。

(3) OAI-PMH 並沒有定義 resumptionToken 的格式，因此對擷取器來說必須視為黑盒子使用。

(4) 本協定中並沒有針對所謂的不完整進行定義，因此，擷取器不應該假設不完整資料列表中的成員是在符合某些選擇標準的狀態下(如依日期排列)。

(5) 在後續請求的 URL 中內含 resumptionToken 之前，擷取器必須將其中任何的特殊字元進行編碼。

Details of flow control and the resumptionToken are as follows:

(1) The only defined use of resumptionToken is as follows:

(d) a repository must include a resumptionToken element as part of each response that includes an incomplete list;

(e) in order to retrieve the next portion of the complete list, the next request must use the value of that resumptionToken element as the value of the resumptionToken argument of the request;

(f) the response containing the incomplete list that completes the list must include an empty resumptionToken element;

All other uses of resumptionToken by a harvester are illegal and must return an error.

(2) In all cases when a resumptionToken is issued, the incomplete list must consist of complete entities; e.g., all individual records returned in an incomplete record list from a ListRecords request must be intact.

(3) The format of the resumptionToken is not defined by the OAI-PMH and should be considered opaque by the harvester.

(4) The protocol does not define the semantics of incompleteness. Therefore, a harvester should not assume that the members in an incomplete list conform to some selection criteria (e.g., date ordering).

(5) Before including a resumptionToken in the URL of a subsequent request, a harvester must encode any special characters in it.

以下列出的選擇性屬性，皆可以內含在 resumptionToken 元件之中：

(1) 有效日期(expirationDate)：此為 UTCdatetime 資料，用以表示 resumptionToken 可以正常使用的範圍。

(2) 完整列表大小(completeListSize)：此為整數數值，用以表示完整的資料列表大小(也就是所有不完整資料列表大小的總和)。由於儲存庫有可能在列表請求序列過程中發生資料變動，因此 completeListSize 只能作為真實參考列表大小的預估參考之用，且在後續的列表參考序列中有可能會變動。

(3) 指標(cursor)：這適用以表示直至目前為止所傳輸的完整資料列表的元件數量(指標是由 0 開始計數)。

The following optional attributes may be included as part of the resumptionToken element along with the resumptionToken itself:

(1) expirationDate -- a UTCdatetime indicating when the resumptionToken ceases to be valid.

(2) completeListSize -- an integer indicating the cardinality of the complete list (i.e., the sum of the cardinalities of the incomplete lists). Because there may be changes in a repository during a list request sequence, as described under Idempotency of resumptionTokens, the value of completeListSize may be only an estimate of the actual cardinality of the complete list and may be revised during the list request sequence.

(3) cursor -- a count of the number of elements of the complete list thus far returned (i.e. cursor starts at 0).

以下的例子說明了一系列的 ListRecords 請求，其完整資料列表為 175 筆紀錄，且儲存庫每次回應只傳回 100 筆紀錄。

(1) 擷取器發出 ListRecords 請求。

(2) 儲存庫的回應包含 100 筆紀錄的不完整資料列表。儲存庫將此列表透過內含非空的 resumptionToken 元件，將其標示為不完整的列表，內含兩個屬性：一是 completeListSize，其值為 175，以及 cursor 屬性，其值為 0。

(3) 擷取器接著發出後續的 ListRecords 請求，內含之前收到的 resumptionToken 元件。

(4) 接著儲存庫在回應中內含 75 筆紀錄的不完整資料列表。同時儲存庫將此列表利用空的 resumptionToken 元件，將其標示為最後的不完整資料列表，內含兩個屬性：一是 completeListSize，其值為 175，以及 cursor 屬性，其值為 100。

The following example is a series of ListRecords requests where the complete list consists of 175 records and the repository only returns 100 records per response.

(1) The harvester issues a ListRecords request.

(2) The repository responds with an incomplete list of 100 records. The repository marks this list as incomplete by including in the response a non-empty resumptionToken element, with two attributes: a completeListSize of 175, and a cursor of 0.

(3) The harvester issues a subsequent ListRecords request that includes the resumptionToken that it received in the previous response.

(4) The repository responds with an incomplete list of 75 records. The repository marks this list as the final incomplete list by including in the response an empty resumptionToken element with two attributes: a completeListSize of 175, and a cursor of 100.

這種流程控制的方式, 結合了 HTTP 傳送層所提供的功能, 為儲存庫提供了部份基本的工具, 使其在擷取界面上的提供上得以建立可接受的使用方針。實作 OAI-PMH 的社群有可能需要更完整的工具來提供儲存庫的擷取界面, 或是擷取詮釋資料的使用方針。這些額外方針所提供的強化功能並不在 OAI-PMH 的討論範圍之內。

This flow control mechanism, in combination with HTTP transport layer facilities, provides some basic tools with which a repository can enforce an acceptable use policy for its harvesting interface. Communities implementing the OAI-PMH may need more extensive tools to enforce acceptable use policies for either the harvesting interface of their repositories or for the metadata harvested from those repositories. The enforcement of such additional policies is outside of the scope of the OAI-PMH.

4.5.1 resumptionTokens 的等價關係(Idempotency of resumptionTokens)

實作 resumptionTokens 的儲存庫必須允許擷取器重複發出列表請求, 為的是讓擷取器使用最近的 resumptionToken, 透過序列的請求來持續取得不完整資料列表。此流程的目的是允許使用者得以從網路或是其他型式的錯誤中恢復回來, 否則此意味著列表序列請求需要重新開始。這種使用 resumptionToken 重複發出列表請求會發生在以下兩種情境中：

(1) 儲存庫中沒有任何變更。此時針對列表請求序列所必須傳回的完整資料列表內容並不會有所改變。在這種情形下, 儲存庫必須針對那

些請求中內含仍在有效期中的 resumptionToken，傳回相同的不完整資料列表。

(2) 當儲存庫資料有變動時。這種情形下有可能會發生完整資料列表的內容有所變更，這有可能是因為要用以傳送的紀錄已經超過了請求中所標示的日戳範圍，這些紀錄有可能在儲存庫中進行了變更，修改，或是刪除動作。在這種情況下，就不需要針對不完整資料列表進行嚴格的等價限制，取而代之的是針對重複發出的請求傳回相對應的不完整資料列表，其中包含那些符合最初請求日期範圍的紀錄(日戳未經改變的)。這些針對重複發出的請求傳回相對應的不完整資料列表也有可能內含那些超出日期範圍的紀錄，因此如果儲存庫的內容有後續的更動，儲存庫將可以發出 badResumptionToken 錯誤，用以通知擷取器必須要重新開始列表請求序列。

4.5.1 Idempotency of resumptionTokens

Repositories that implement resumptionTokens must do so in a manner that allows harvesters to resume a sequence of requests for incomplete lists by re-issuing a list request with the most recent resumptionToken. The purpose of this is to allow harvesters to recover from network or other errors that would otherwise mean that the list request sequence would have to be started again. A re-issue of a list request with a resumptionToken occurs in two contexts:

(1) When there are no changes in the repository. There are no changes to the complete list returned by the list request sequence. In this case, the repository must return the same incomplete list when the most recent list request, i.e. the one with the most recent non-expired resumptionToken, is re-issued.

(2) When there are changes in the repository. There may be changes to the complete list returned by the list request sequence. These changes occur when the records disseminated in the list move in or out of the

datestamp range of the request because of changes, modifications, or deletions in the repository. In this case, strict idempotency of the incomplete-list requests using resumptionToken values is not required. Instead, the incomplete list returned in response to a re-issued request must include all records with unchanged datestamps within the range of the initial list request. The incomplete list returned in response to a re-issued request may contain records with datestamps that either moved into or out of the range of the initial request. In cases where there are substantial changes to the repository, it may be appropriate for a repository to return a badResumptionToken error, signaling that the harvester should restart the list request sequence.

4.6 錯誤與例外狀況(Error and Exception Conditions)

當錯誤或是例外狀況發生時，儲存庫必須透過在回應中內含一或多個錯誤元件發出 OAI-PMH 錯誤，並與 HTTP 狀態碼進行區別。當錯誤元件足以表達目前的錯誤或是例外狀況時，儲存庫應該要回報所有處理請求時所發生的錯誤與例外。各錯誤元件皆必須有錯誤代碼屬性(請見下表)；並包含自定的文字字串，用以提供使用者可讀的相關錯誤訊息，這些字串並不在 OAI-PMH 的定義之中。

4.6 Error and Exception Conditions

In event of an error or exception condition, repositories must indicate OAI-PMH errors, distinguished from HTTP Status-Codes, by including one or more error elements in the response. While one error element is sufficient to indicate the presence of the error or exception condition, repositories should report all errors or exceptions that arise from processing the request. Each error element must have a code attribute that must be from the following table; each error element may also have a free text string value to provide information about the error that is useful to a human reader. These strings are not defined by the OAI-PMH.

錯誤代碼(Error Codes)	說明	適用指令 (Applicable Verbs)

badArgument	說明請求中使用了不合標準的參數, 包括使用重複參數, 或是參數值使用了不合規定的語法。	所有指令
badResumptionToken	resumptionToken 參數不正確或不再有效。	ListIdentifiers ListRecords ListSets
badVerb	所使用的並不是合法的 OAI-PMH 指令, 或是指令所需的參數不存在, 或是指令使用了重複的參數。	無
cannotDisseminateFormat	透過 metadataPrefix 所指定的詮釋資料格式儲存庫並不支援。	GetRecord ListIdentifiers ListRecords
idDoesNotExist	所適用的識別符參數在儲存庫中並不存在或是不合標準。	GetRecord ListMetadataFormats
noRecordsMatch	任何指定起訖時間, 集合, 或是 metadataPrefix 參數所取得的資料列表是空的。	ListIdentifiers ListRecords
noMetadataFormats	指定的資料項沒有任何可用的詮釋資料格式。	ListMetadataFormats
noSetHierarchy	該儲存庫並不支援集合。	ListSets ListIdentifiers ListRecords

Error Codes	Description	Applicable Verbs
badArgument	The request includes illegal arguments, is missing required arguments, includes a repeated argument, or values for arguments have an illegal syntax.	all verbs
badResumptionToken	The value of the resumptionToken argument is invalid or expired.	ListIdentifiers ListRecords ListSets
badVerb	Value of the verb argument is not a legal OAI-PMH verb,	N/A

	the verb argument is missing, or the verb argument is repeated.	
cannotDisseminateFormat	The metadata format identified by the value given for the metadataPrefix argument is not supported by the item or by the repository.	GetRecord ListIdentifiers ListRecords
idDoesNotExist	The value of the identifier argument is unknown or illegal in this repository.	GetRecord ListMetadataFormats
noRecordsMatch	The combination of the values of the from, until, set and metadataPrefix arguments results in an empty list.	ListIdentifiers ListRecords
noMetadataFormats	There are no metadata formats available for the specified item.	ListMetadataFormats
noSetHierarchy	The repository does not support sets.	ListSets ListIdentifiers ListRecords

以下的範例說明了使用錯誤指令的錯誤處理情形。以下所有的請求 URL 皆經過轉換以維持較高的可讀性。

The following example demonstrates error handling in the case of an illegal verb argument. All request URLs shown from now on will be wrapped to make them more readable.

請求
http://arXiv.org/oai2?
verb=nastyVerb

Request
http://arXiv.org/oai2?
verb=nastyVerb

回應

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-05-01T09:18:29Z</responseDate>
  <request>http://arXiv.org/oai2</request>
  <error code="badVerb">Illegal OAI verb</error>
</OAI-PMH>
```

以下的範例說明了針對不支援集合的儲存庫使用 ListSets 請求的錯誤處理情形。

The following example demonstrates error handling in the case of a ListSets request to a repository that does not handle sets.

請求

http://arXiv.org/oai2?
verb=ListSets

Request

http://arXiv.org/oai2?
verb=ListSets

回應

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-05-01T09:18:29Z</responseDate>
  <request verb="ListSets">http://arXiv.org/oai2</request>
  <error code="noSetHierarchy">This repository does not
    support sets</error>
</OAI-PMH>
```

5.協定的請求與回應

本段將列出並說明定義在 OAI-PMH 中的所有請求(或稱指令, verb)。本文件針對每一種請求皆會以如下的組織來進行說明：

(1) 與特定請求相關的段落標題, 此標題文字就是用在表達 HTTP 請求中的指令參數。

(2) 一段簡要的摘要說明指令的意義以及用法。

(3) 列出請求所使用的其他相關參數。參數根據其特性分為三種：

(a) 必要的(required), 意指這項參數必須內含在該請求之中(其中指令參數一定是必要的, 如同 HTTP 請求格式上的說明)。

(b) 選擇性的(optional), 意指這項參數可以內含在該請求之中。

(c) 互斥的(exclusive), 意指這項參數可以內含在該請求之中, 但是必須是唯一的參數(在此指除了指令參數之外)。

(4) 該請求所屬特定的錯誤以及例外狀況。

(5) 一或多個範例請求以及相關的回應, 以及適當的重點說明。

另外尚有 XML 綱要, 用以定義所有 OAI-PMH 請求回應格式的正確性。

This section lists the requests, or verbs, defined in the OAI-PMH. The documentation for each request is organized as follows:

(1) A section title corresponding to the token used to specify the request as the required verb argument to an HTTP request.

(2) A brief summary of the meaning of the verb and notes on its usage.

(3) The list of additional arguments for the request. Arguments are of three types:

- required, the argument must be included with the request (the verb argument is always required, as described in HTTP Request Format).
- optional, the argument may be included with the request.
- exclusive, the argument may be included with request, but must be the only argument (in addition to the verb argument).

(4) Error and exception conditions specific to the protocol request.

(5) One or more example requests and corresponding responses, with explanatory notes if appropriate.

An XML Schema defines the format of valid replies to all OAI-PMH requests.

5.1 GetRecord

摘要以及使用方式說明

這項指令是用以擷取來自儲存庫的單一詮釋資料紀錄。其中所需的參數包括指定的識別符，以及該筆紀錄中會使用的詮釋資料格式。此外還定義了與儲存庫相關的刪除追蹤層次，也就是在回應的標頭中，伴隨著狀態屬性中使用“deleted”資料值，以防由 metadataPrefix 所指定的詮釋資料格式在儲存庫或是特定資料項不再可供使用時之用。

5.1 GetRecord

Summary and Usage Notes

This verb is used to retrieve an individual metadata record from a repository. Required arguments specify the identifier of the item from which the record is requested and the format of the metadata that should be included in the record. Depending on the level at which a repository tracks deletions, a header with a "deleted" value for the status attribute may be returned, in case the metadata format specified by the metadataPrefix is no longer available from the repository or from the specified item.

參數

(1) 識別符是必須使用的參數，用以指定儲存庫中資料項的唯一識別符，該項目所屬的紀錄隨後將被傳佈出去。

(2) metadataPrefix 是必須使用的參數，用以指定欲傳回紀錄中部份詮釋資料的格式。紀錄只有在與傳遞的格式有出現在 metadataPrefix 時，才可以透過識別符參數的指定進行傳佈。至於儲存庫可以使用的詮釋資料格式，或是欲存取特定紀錄的詮釋資料格式，則應該透過使用 ListMetadataFormats 請求來獲得。

Arguments

(1) identifier a required argument that specifies the unique identifier of the item in the repository from which the record must be disseminated.

(2) metadataPrefix a required argument that specifies the metadataPrefix of the format that should be included in the metadata part of the returned record . A record should only be returned if the format specified by the metadataPrefix can be disseminated from the item identified by the value of the identifier argument. The metadata formats supported by a repository and for a particular record can be retrieved using the ListMetadataFormats request.

錯誤以及例外狀況

(1) badArgument：是指該請求包含了不合標準的參數, 或是沒有使用必要的參數。

(2) cannotDisseminateFormat：是指透過 identifier 參數所參考的資料項並無法支援 metadataPrefix 參數中的資料值所指定的詮釋資料格式。

(3) idDoesNotExist：是指識別符參數中所登記的代碼, 儲存庫無法辨識或認為不合標準。

Error and Exception Conditions

(1) badArgument - The request includes illegal arguments or is missing required arguments.

(2) cannotDisseminateFormat - The value of the metadataPrefix argument is not supported by the item identified by the value of the identifier argument.

(3) idDoesNotExist - The value of the identifier argument is unknown or illegal in this repository.

範例

請求

此例為請求都柏林核心集詮釋資料格式的紀錄[為了增加可讀性，URL 部份並沒有經過編碼]。

Examples

Request

Request a record in the Dublin Core metadata format [URL shown without encoding to be more readable].

http://arXiv.org/oai2?
verb=GetRecord&identifier=oai:arXiv.org:cs/0112017&metadataPrefix=oai_dc

回應

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T08:55:46Z</responseDate>
  <request verb="GetRecord" identifier="oai:arXiv.org:cs/0112017"
           metadataPrefix="oai_dc">http://arXiv.org/oai2</request>
  <GetRecord>
   <record>
    <header>
      <identifier>oai:arXiv.org:cs/0112017</identifier>
      <datestamp>2001-12-14</datestamp>
      <setSpec>cs</setSpec>
      <setSpec>math</setSpec>
    </header>
    <metadata>
      <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>Using Structural Metadata to Localize Experience of
                  Digital Content</dc:title>
```

```
        <dc:creator>Dushay, Naomi</dc:creator>
        <dc:subject>Digital Libraries</dc:subject>
        <dc:description>With the increasing technical sophistication of
            both information consumers and providers, there is
            increasing demand for more meaningful experiences of digital
            information. We present a framework that separates digital
            object experience, or rendering, from digital object storage
            and manipulation, so the rendering can be tailored to
            particular communities of users.
        </dc:description>
        <dc:description>Comment: 23 pages including 2 appendices,
            8 figures</dc:description>
        <dc:date>2001-12-14</dc:date>
      </oai_dc:dc>
    </metadata>
  </record>
 </GetRecord>
</OAI-PMH>
```

請求

此例用以取得一筆都柏林核心集詮釋資料格式的紀錄。但是請求中所指定的紀錄不一定會傳回，這是因為該識別符可能不存在。此時回應終將不會包含紀錄的容器，同時也會擁有錯誤元件，代碼部份的資料值將會是 idDoesNotExist。[為了增加可讀性，URL 部份並沒有經過編碼]。

Request

Request a record in the Dublin Core metadata format. The requested record, however, can not be returned because the identifier does not exist. Therefore, the response does not contain a record container. It does have an error element with a code attribute that has the value idDoesNotExist. [URL shown without encoding for better readability].

http://arXiv.org/oai2?verb=GetRecord&identifier=oai:arXiv.org:quant-ph/02131001&metadataPrefix=oai_dc

回應

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T08:55:46Z</responseDate>
  <request verb="GetRecord" identifier="oai:arXiv.org:quant-ph/0213001"
           metadataPrefix="oai_dc">http://arXiv.org/oai2</request>
  <error code="idDoesNotExist">No matching identifier in arXiv</error>
</OAI-PMH>
```

請求

透過指定 oai_marc 格式來抓取 oai_marc 詮釋資料的流程。但最終此識別符皆無法取得指定格式的詮釋資料。因此回應將會沒有任何紀錄。這種情形的處理方式是使用錯誤元件，內含代碼屬性，內嵌在 cannotDisseminateFormat 中。[為了增加可讀性，URL 部份並沒有經過編碼]。

Request

Request a record in the oai_marc metadata format. However, the requested metadata format can not be disseminated for this identifier. Therefore, the response contains no record. It does contain an error element with a code attribute that has the value cannotDisseminateFormat. [URL shown without encoding for better readability].

http://arXiv.org/oai2?
verb=GetRecord&identifier=oai:arXiv.org:quant-ph/9901001&metadataPrefix=oai_marc

回應

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T08:55:46Z</responseDate>
  <request verb="GetRecord" identifier="oai:arXiv.org:quant-ph/9901001"
           metadataPrefix="oai_marc">http://arXiv.org/oai1</request>
```

```
  <error code="cannotDisseminateFormat"/>
</OAI-PMH>
```

5.2 識別(Identify)

摘要以及使用方式說明

此指令是用以擷取儲存庫的相關資訊，其中部份傳回內容對 OAI-PMH 的運作來說是必須的。此外，儲存庫也可以利用識別指令來提供額外描述性的資訊。

5.2 Identify

Summary and Usage Notes

This verb is used to retrieve information about a repository. Some of the information returned is required as part of the OAI-PMH. Repositories may also employ the Identify verb to return additional descriptive information.

參數

無

Arguments

None

錯誤以及例外狀況

(1) badArgument：請求中包含不合標準的參數錯誤以及例外狀況。

Error and Exception Conditions

(1)badArgument - The request includes illegal arguments.

回應格式(Response Format)

回應內容必須包含以下元件的實例：

(1) repositoryName：是指人可以讀取的儲存庫名稱；

(2) baseURL：是指儲存庫的基礎 URL；

(3) protocolVersion：是指儲存庫所支援的 OAI-PMH 版本；

(4) earliestDatestamp：此為 UTCdatetime 格式的資料，用以表示所有變更，修改，以及刪除日戳的下限。儲存庫不可以使用早於 earliestDatestamp 元件中所描述的日戳。同時 earliestDatestamp 必須表示成儲存庫所支援最精細的資料。

(5) deletedRecord：是指儲存庫針對已刪除紀錄所支援的標示方式。合法的資料值為 no；transient；persistent，其意義在討論刪除的段落有所說明。

(6) granularity：這是指儲存庫所支援最精細的日期格式。合法的資料值為 YYYY-MM-DD 以及 YYYY-MM-DDThh:mm:ssZ，其意義在 ISO8601 中有定義。

Response Format

The response must include one instance of the following elements:

(1) repositoryName : a human readable name for the repository;

(2) baseURL : the base URL of the repository;

(3) protocolVersion : the version of the OAI-PMH supported by the repository;

(4) earliestDatestamp : a UTCdatetime that is the guaranteed lower limit of all datestamps recording changes, modifications, or deletions in the repository. A repository must not use datestamps lower than the one specified by the content of the earliestDatestamp element. earliestDatestamp must be expressed at the finest granularity supported by the repository.

(5) deletedRecord : the manner in which the repository supports the notion of deleted records. Legitimate values are no ; transient ; persistent with meanings defined in the section on deletion.

(6) granularity: the finest harvesting granularity supported by the repository. The legitimate values are YYYY-MM-DD and YYYY-MM-DDThh:mm:ssZ with meanings as defined in ISO8601.

此外，回應內容必須也包含以下元件的一或多個實例：

(1) adminEmail：是指儲存庫管理者的 e-mail 帳號。

The response must include one or more instances of the following element:

(1) adminEmail : the e-mail address of an administrator of the repository.

另外，回應內容也可以包含以下選擇性元件的多個實例：

(1) compression：這是指儲存庫所支援的壓縮編碼，建議的資料值為定義在描述 HTTP 1.1 的 RFC 2616 中的 14.11 節，也就是說名內容編碼(Content-Encoding)標頭部份。此壓縮元件不應該涵括在識別編碼中，因為識別編碼有其預設編碼方式。

(2) description：此為讓社群可以描述其儲存庫的彈性方式。舉例來說，描述容器可以用以內含集叢層次的詮釋資料，此資料通常是用在識別請求的回應中。實作指導綱要文件中有針對此方面提出建議做法。各描述容器皆必須內含描述內含資料結構的 XML 綱要，以 URL 表示。

The response may include multiple instances of the following optional elements:

(1) compression : a compression encoding supported by the repository. The recommended values are those defined for the Content-Encoding header in Section 14.11 of RFC 2616 describing HTTP 1.1. A compression element should not be included for the Identify encoding, which is implied.

(2) description : an extensible mechanism for communities to describe their repositories. For example, the description container could be used to include collection-level metadata in the response to the Identify request. Implementation Guidelines are available to give directions with this respect. Each description container must be accompanied by the URL of an XML schema describing the structure of the description container.

範例
請求
http://memory.loc.gov/cgi-bin/oai?verb=Identify

Examples
Request

http://memory.loc.gov/cgi-bin/oai?verb=Identify

回應
Response
下例中說明識別請求的回應包含了三個描述容器的情形：

(1) oai-identifier 容器使用 XML 綱要，其位址為 http://www.openarchives.org/OAI/2.0/oai-identifier.xsd。此在實作指導綱要文件中所提供的綱要，是用以作為儲存庫將資料項的識別符指定其格式之用。此識別符的格式可以在 oai-identifier.xsd 綱要的註解中找到其說明。

(2) eprints 容器使用 XML 綱要，其位址為 http://www.openarchives.org/OAI/1.1/eprints.xsd。此在實作指導綱要文件中所提供的綱要，目前由 OAI e-print 社群同意使用，是用以提供 e-print 社群中所使用的特殊儲存庫相關訊息之用。

(3) friends 容器使用 XML 綱要，其位址為 http://www.openarchives.org/OAI/2.0/friends.xsd。此在實作指導綱要文件中所提供的綱要，目前是用以導引擷取器到其他儲存庫之用，做法是透過列出這些 URL 來達到其目的。friends 容器的使用方式上是較受推薦的，因為其可以支援擷取器發覺儲存庫的網路位址。

The below example of a response to the Identify request contains three description containers:

(1) The oai-identifier container complies to an XML Schema, which is available at http://www.openarchives.org/OAI/2.0/oai-identifier.xsd. This schema, provided in the accompanying Implementation Guidelines document, is used by repositories that choose to comply with a specific format of unique identifiers for items. The format of that identifier is explained by means of comments in the oai-identifier.xsd XML Schema.

(2) The eprints container complies to an XML Schema, which is available at http://www.openarchives.org/OAI/1.1/eprints.xsd. This schema, provided in the accompanying Implementation Guidelines document, has been agreed upon by the OAI e-print community, and contains information specific to repositories in that community.

(3) The friends container complies to an XML Schema, which is available at http://www.openarchives.org/OAI/2.0/friends.xsd. This schema, provided in the accompanying Implementation Guidelines document, is used by repositories that want to point harvesters to other repositories, by listing their base URLs. Usage of the friends container is recommended; it may support harvesters in discovering the network-location of repositories.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T12:00:01Z</responseDate>
  <request verb="Identify">http://memory.loc.gov/cgi-bin/oai</request>
  <Identify>
    <repositoryName>Library of Congress Open Archive Initiative
                    Repository 1</repositoryName>
    <baseURL>http://memory.loc.gov/cgi-bin/oai</baseURL>
    <protocolVersion>2.0</protocolVersion>
    <adminEmail>somebody@loc.gov</adminEmail>
    <adminEmail>anybody@loc.gov</adminEmail>
    <earliestDatestamp>1990-02-01T12:00:00Z</earliestDatestamp>
    <deletedRecord>transient</deletedRecord>
    <granularity>YYYY-MM-DDThh:mm:ssZ</granularity>
    <compression>deflate</compression>
    <description>
      <oai-identifier
        xmlns="http://www.openarchives.org/OAI/2.0/oai-identifier"
```

```
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            xsi:schemaLocation=
                  "http://www.openarchives.org/OAI/2.0/oai-identifier
            http://www.openarchives.org/OAI/2.0/oai-identifier.xsd">
            <scheme>oai</scheme>
            <repositoryIdentifier>lcoa1.loc.gov</repositoryIdentifier>
            <delimiter>:</delimiter>

<sampleIdentifier>oai:lcoa1.loc.gov:loc.music/musdi.002</sampleIdentifier>
          </oai-identifier>
        </description>
        <description>
          <eprints
              xmlns="http://www.openarchives.org/OAI/1.1/eprints"
              xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
              xsi:schemaLocation="http://www.openarchives.org/OAI/1.1/eprints
              http://www.openarchives.org/OAI/1.1/eprints.xsd">
            <content>

<URL>http://memory.loc.gov/ammem/oamh/lcoa1_content.html</URL>
              <text>Selected collections from American Memory at the Library
                    of Congress</text>
            </content>
            <metadataPolicy/>
            <dataPolicy/>
          </eprints>
        </description>
        <description>
          <friends
              xmlns="http://www.openarchives.org/OAI/2.0/friends/"
              xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
              xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/friends/
            http://www.openarchives.org/OAI/2.0/friends.xsd">
           <baseURL>http://oai.east.org/foo/</baseURL>
           <baseURL>http://oai.hq.org/bar/</baseURL>
           <baseURL>http://oai.south.org/repo.cgi</baseURL>
         </friends>
      </description>
```

```
  </Identify>
</OAI-PMH>
```

5.3 ListIdentifiers

摘要以及使用方式說明

此指令是 ListRecords 的簡要版本，目的是用以取出標頭資料，而非紀錄。可選擇的參數包括允許透過集合成員或是日戳等方式來進行選擇性的擷取。如果符合指定參數是已經刪除的資料時，將會傳回設定狀態屬性為”deleted” 的標頭，但是這和儲存庫對刪除功能的支援相關。

5.3 ListIdentifiers

Summary and Usage Notes

This verb is an abbreviated form of ListRecords, retrieving only headers rather than records. Optional arguments permit selective harvesting of headers based on set membership and/or datestamp. Depending on the repository's support for deletions, a returned header may have a status attribute of "deleted" if a record matching the arguments specified in the request has been deleted.

參數

(1) from 參數，這是選擇性的參數，使用 UTCdatetime 資料值，是用以指定已日戳為基礎的選擇性擷取的下限值(lower bound)。

(2) until 參數，這是選擇性的參數，使用 UTCdatetime 資料值，是用以指定已日戳為基礎的選擇性擷取的上限值(upper bound)。

(3) metadataPrefix 參數，此為必要的參數，這是當 metadataPrefix 所指的詮釋資料格式存在時(包含儲存庫所支援的已刪除紀錄)，那些指定的資料標投將會傳回。至於儲存庫或是特定資料項所支援的詮釋資料格式則可以利用 ListMetadataFormats 來取得。

(4) set 參數，這是選擇性的參數，這是參數與 setSpec 資料值配合使用，用以指定選擇性擷取的集合過濾條件。

(5) resumptionToken 參數，此為互斥的參數，此為用以作流程控制的標籤，是透過前一個 ListIdentifiers 請求(傳回不完整資料列表)傳回。

Arguments

(1) from an optional argument with a UTCdatetime value, which specifies a lower bound for datestamp-based selective harvesting.

(2) until an optional argument with a UTCdatetime value, which specifies a upper bound for datestamp-based selective harvesting.

(3) metadataPrefix a required argument, which specifies that headers should be returned only if the metadata format matching the supplied metadataPrefix is available or, depending on the repository's support for deletions, has been deleted. The metadata formats supported by a repository and for a particular item can be retrieved using the ListMetadataFormats request.

(4) set an optional argument with a setSpec value , which specifies set criteria for selective harvesting.

(5) resumptionToken an exclusive argument with a value that is the flow control token returned by a previous ListIdentifiers request that issued an incomplete list.

錯誤以及例外狀況

(1) badArgument：是指該請求包含了不合標準的參數，或是沒有使用必要的參數。

(2) badResumptionToken：是指 resumptionToken 參數的資料值錯誤或是已經失效。

(3) cannotDisseminateFormat：是指透過 identifier 參數所參考的資料項並無法支援 metadataPrefix 參數中的資料值所指定的詮釋資料格式。

(4) noRecordsMatch：是指 from、util、以及 set 參數資料值結合處理的結果唯一空的資料列表時發生。

(5) noSetHierarchy：是指儲存庫並不支援集合階層架構。

Error and Exception Conditions

(1) badArgument - The request includes illegal arguments or is missing required arguments.

(2) badResumptionToken - The value of the resumptionToken argument is invalid or expired.

(3) cannotDisseminateFormat - The value of the metadataPrefix argument is not supported by the repository.

(4) noRecordsMatch- The combination of the values of the from, until, and set arguments results in an empty list.

(5) noSetHierarchy - The repository does not support sets.

範例

請求

本例中試圖將紀錄的列表以多個條件列出，包括使用 oldarXiv 詮釋資料格式，資料的加入，修改以及刪除動作發生自 1998 年 1 月 15 日以後，以及在 physics:hep 集合中。[為了增加可讀性，URL 部份並沒有經過編碼]。

http://an.oa.org/OAI-script?verb=ListIdentifiers&from=1998-01-15&metadataPrefix=oldArXiv&set=physics:hep

Examples

Request

List the headers of records in the oldarXiv metadata format that are added, modified or deleted since January 15, 1998 in the set physics:hep. [URL shown without encoding for better readability].

http://an.oa.org/OAI-script?verb=ListIdentifiers&from=1998-01-15&metadataPrefix=oldArXiv&set=physics:hep

回應

回應的內容是傳回含有四個標頭的資料列表。標頭中包含了已刪除狀態，表示有一筆紀錄在 metadataPrefix 指定的詮釋資料格式並不存在。此外也會

傳回 resumptionToken (非空值，資料值為 xxx45abttyz)，用以說明標頭列表並不完整，需要發出一或多個後續的請求來取出完整的資料列表。在本例中，此 resumptionToken 內含有三個選擇性的屬性，分別為：expirationDate，說明此 resumptionToken 在 2002 年 6 月 1 日 11:20 PM UTC 之後將會失效；completeListSize，說明完整的資料列表是由六個識別符所組成；指標為零的屬性代表本回應之前並沒有任何標頭被傳回。

Response

A list of four headers is returned. One header has a deleted status, indicating that a record in the metadata format specified by the metadataPrefix is no longer available. In addition, a resumptionToken (non-empty, value xxx45abttyz) has been returned, indicating that the list of headers is incomplete and that one or more subsequent requests will need to be issued to retrieve a complete list. In the example, the resumptionToken comes with all of the 3 optional attributes: expirationDate indicates that the resumptionToken will become unusable after 11:20 PM UTC on June 1st 2002; completeListSize indicates that the complete list consists of 6 identifiers; the zero-value for cursor indicates that no headers have been returned previous to this reply.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-01T19:20:30Z</responseDate>
  <request verb="ListIdentifiers" from="1998-01-15"
           metadataPrefix="oldarXiv"
           set="physics:hep">http://an.oa.org/OAI-script</request>
  <ListIdentifiers>
   <header>
    <identifier>oai:arXiv.org:hep-th/9801001</identifier>
    <datestamp>1999-02-23</datestamp>
    <setSpec>physic:hep</setSpec>
   </header>
   <header>
    <identifier>oai:arXiv.org:hep-th/9801002</identifier>
    <datestamp>1999-03-20</datestamp>
    <setSpec>physic:hep</setSpec>
```

```
      <setSpec>physic:exp</setSpec>
    </header>
    <header>
      <identifier>oai:arXiv.org:hep-th/9801005</identifier>
      <datestamp>2000-01-18</datestamp>
      <setSpec>physic:hep</setSpec>
    </header>
    <header status="deleted">
      <identifier>oai:arXiv.org:hep-th/9801010</identifier>
      <datestamp>1999-02-23</datestamp>
      <setSpec>physic:hep</setSpec>
      <setSpec>math</setSpec>
    </header>
    <resumptionToken expirationDate="2002-06-01T23:20:00Z"
        completeListSize="6"
        cursor="0">xxx45abttyz</resumptionToken>
 </ListIdentifiers>
</OAI-PMH>
```

請求

針對以上的範例發出後續的請求。其中包含單一的 resumptionToken 參數，內容為前一個回應所傳回的資料值。[為了增加可讀性，URL 部份並沒有經過編碼]。

http://an.oa.org/OAI-script?
verb=ListIdentifiers&resumptionToken=xxx45abttyz

Request

Issue a subsequent request to the one issued above. The single resumptionToken argument has the value returned in the previous response. [URL shown without encoding for better readability].

http://an.oa.org/OAI-script?
verb=ListIdentifiers&resumptionToken=xxx45abttyz

回應

本次傳回了兩筆標頭資料。其中在本列表完畢後的 resumptionToken 元件沒有資料值，這是指此資料列表已經傳遞完畢(已成為完整資料列表)。其中

completeListSize 屬性的資料值仍維持為 6，但是指標屬性的資料值變成 4，說明了上一個(或前幾個)回應已經傳回了四筆識別符。

Response

Two more headers are returned. The resumptionToken element at the end of the list has no value, indicating that the list is now complete. The value of the completeListSize attribute remains 6, while the value of the cursor attribute has changed to 4, indicating that a previous reply has (or previous replies have) already delivered 4 identifiers.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-01T19:30:00Z</responseDate>
  <request verb="ListIdentifiers"
           resumptionToken="xxx45abttyz">http://an.oa.org/OAI-script</request>
  <ListIdentifiers>
    <header>
      <identifier>oai:arXiv.org:hep-th/9801020</identifier>
      <datestamp>1999-02-23</datestamp>
      <setSpec>physic:hep</setSpec>
    </header>
    <header>
    <identifier>oai:arXiv.org:hep-th/9801060</identifier>
      <datestamp>1999-02-23</datestamp>
      <setSpec>physic:hep</setSpec>
    </header>
    <resumptionToken completeListSize="6" cursor="4"/>
  </ListIdentifiers>
</OAI-PMH>
```

請求

列出 olac-formatted 紀錄的標頭，日期訂為在 2001 年 1 月 1 日進行新增或修改，同時在 Perseus:collection:PersInfo 集合中。此請求將無法對應到任何紀錄，因此，回應內容包括錯誤標籤，同時也不含有任何的標頭元件 [為了增加可讀性，URL 部份並沒有經過編碼]。

http://www.perseus.tufts.edu/cgi-bin/pdataprov?
verb=ListIdentifiers&metadataPrefix=olac&from=2001-01-01&until=2001-01-01&set=Perseus:collection:PersInfo

Request
List the headers of olac-formatted records, added or modified on January 1, 2001 in the set Perseus:collection:PersInfo. There are no matches for this request, hence, the response contains an error tag and does not contain any header elements [URL shown without encoding for better readability].

http://www.perseus.tufts.edu/cgi-bin/pdataprov?
verb=ListIdentifiers&metadataPrefix=olac&from=2001-01-01&until=2001-01-01&set=Perseus:collection:PersInfo

回應
Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T14:27:19Z</responseDate>
  <request verb="ListIdentifiers" metadataPrefix="olac"
    from="2001-01-01" until="2001-01-01"
    set="Perseus:collection:PersInfo">
    http://www.perseus.tufts.edu/cgi-bin/pdataprov</request>
  <error code="noRecordsMatch"/>
</OAI-PMH>
```

5.4 ListMetadataFormats
摘要以及使用方式說明
此指令是用以取出儲存庫中可以使用的詮釋資料格式。其中可以使用的選擇性參數透過指定特定紀錄來限制請求的適用範圍。

5.4 ListMetadataFormats
Summary and Usage Notes

This verb is used to retrieve the metadata formats available from a repository. An optional argument restricts the request to the formats available for a specific item.

參數

選擇性參數內容為唯一識別符，用以請求特定的資料項可用的詮釋資料格式。如果忽略掉此參數，則回應內容將會包含所有儲存庫所支援的詮釋資料格式。請注意，即使儲存庫支援某個詮釋資料格式，也不意味著所有的資料項皆可以透過這種詮釋資料格式進行傳佈。

Arguments

identifier an optional argument that specifies the unique identifier of the item for which available metadata formats are being requested. If this argument is omitted, then the response includes all metadata formats supported by this repository. Note that the fact that a metadata format is supported by a repository does not mean that it can be disseminated from all items in the repository.

錯誤以及例外狀況

(1) badArgument：是指該請求包含了不合標準的參數，或是沒有使用必要的參數。

(2) idDoesNotExist：是指識別符參數中所登記的代碼，儲存庫無法辨識或認為不合標準。

(3) noMetadataFormats：指定的資料項沒有任何可使用的詮釋資料格式。

Error and Exception Conditions

(1) badArgument - The request includes illegal arguments or is missing required arguments.

(2) idDoesNotExist - The value of the identifier argument is unknown or illegal in this repository.

(3) noMetadataFormats - There are no metadata formats available for the specified item.

範例

請求
列出儲存庫(位於 http://www.perseus.tufts.edu/cgi-bin/pdataprov)中可以用以傳佈的所有詮釋資料格式，同時使用唯一識別符(oai:perseus.tufts.edu:Perseus:text:1999.02.0119)來指定資料項[為了增加可讀性，URL 部份並沒有經過編碼]。

http://www.perseus.tufts.edu/cgi-bin/pdataprov?verb=ListMetadataFormats&identifier=oai:perseus.tufts.edu:Perseus:text:1999.02.0119

Examples
Request
List the metadata formats that can be disseminated from the repository http://www.perseus.tufts.edu/cgi-bin/pdataprov for the item with unique identifier oai:perseus.tufts.edu:Perseus:text:1999.02.0119 [URL shown without encoding for better readability].

http://www.perseus.tufts.edu/cgi-bin/pdataprov?verb=ListMetadataFormats&identifier=oai:perseus.tufts.edu:Perseus:text:1999.02.0119

回應
此回應列出了三個可支援使用的詮釋資料格式，分別針對識別符 oai_dc，olac 以及 perseus。其中每一種格式皆會傳回該種格式所使用的 XML 綱要的位址，以及 XML 命名空間的 URI。

Response
The response shows that 3 metadata formats are supported for the given identifier: oai_dc, olac and perseus. For each of the formats, the location of an XML Schema describing the format, as well as the XML Namespace URI is given.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T14:27:19Z</responseDate>
```

```
<request verb="ListMetadataFormats"
  identifier="oai:perseus.tufts.edu:Perseus:text:1999.02.0119">
  http://www.perseus.tufts.edu/cgi-bin/pdataprov</request>
<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>olac</metadataPrefix>
   <schema>http://www.language-archives.org/OLAC/olac-0.2.xsd</schema>
   <metadataNamespace>http://www.language-archives.org/OLAC/0.2/
    </metadataNamespace>
 </metadataFormat>
 <metadataFormat>
   <metadataPrefix>perseus</metadataPrefix>
   <schema>http://www.perseus.tufts.edu/persmeta.xsd</schema>
   <metadataNamespace>http://www.perseus.tufts.edu/persmeta.dtd
     </metadataNamespace>
 </metadataFormat>
</ListMetadataFormats>
</OAI-PMH>
```

請求

列出儲存庫(http://memory.loc.gov/cgi-bin/oai)可以用以傳佈的詮釋資料格式。

http://memory.loc.gov/cgi-bin/oai?verb=ListMetadataFormats

Request

List the metadata formats that can be disseminated from the repository http://memory.loc.gov/cgi-bin/oai.

http://memory.loc.gov/cgi-bin/oai?verb=ListMetadataFormats

回應

回應內容列出了儲存庫支援兩種詮釋資料格式：oai_dc 以及 oai_marc。對每一種格式來說，該種格式所使用的 XML 綱要的位址皆會傳回。但是在儲存庫層級支援這些格式並不意味著各儲存庫中的資料項皆可以使用這些格式。

Response

The response shows that the repository supports two metadata formats: oai_dc, and oai_marc. For each of the formats, the location of an XML Schema describing the format is given. The support of these formats at the repository-level does not imply support of each format for each item of the repository.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-08T15:19:13Z</responseDate>
  <request verb="ListMetadataFormats">
           http://memory.loc.gov/cgi-bin/oai</request>
  <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_marc</metadataPrefix>
    <schema>http://www.openarchives.org/OAI/1.1/oai_marc.xsd</schema>
    <metadataNamespace>http://www.openarchives.org/OAI/1.1/oai_marc
      </metadataNamespace>
   </metadataFormat>
  </ListMetadataFormats>
</OAI-PMH>
```

請求

列出儲存庫(http://memory.loc.gov/cgi-bin/oai)可以用以傳佈的詮釋料格式, 限定同時唯一識別符為 oai:lcoa1.loc.gov:loc.rbc/rbpe.00000111。但是此識別符並不存在, 因此回應中將含有錯誤元件, 同時也沒有內含 metadataFormat 容器 [為了增加可讀性, URL 部份並沒有經過編碼]。

http://memory.loc.gov/cgi-bin/oai?
verb=ListMetadataFormats&identifier=oai:lcoa1.loc.gov:loc.rbc/rbpe.00000111

Request
List the metadata formats that can be disseminated for the unique identifier oai:lcoa1.loc.gov:loc.rbc/rbpe.00000111 in the repository http://memory.loc.gov/cgi-bin/oai. The identifier, however, does not exist and therefore, the response contains an error element and no metadataFormat container. [URL shown without encoding for better readability].

http://memory.loc.gov/cgi-bin/oai?
verb=ListMetadataFormats&identifier=oai:lcoa1.loc.gov:loc.rbc/rbpe.00000111

回應
Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-08T15:19:13Z</responseDate>
  <request verb="ListMetadataFormats"
           identifier="oai:lcoa1.loc.gov:loc.rbc/rbpe.00000111">
           http://memory.loc.gov/cgi-bin/oai</request>
  <error code="idDoesNotExist">oai:lcoa1.loc.gov:loc.rbc/rbpe.00000111 has the
      structure of a valid LOC identifier, but it maps to no known
      item</error>
</OAI-PMH>
```

5.5 ListRecords

摘要以及使用方式說明

此指令是用以自儲存庫中擷取紀錄。可選擇的參數包括允許透過集合成員或是日戳等方式來進行選擇性的擷取。根據儲存庫對刪除功能的支援而

言，如果符合指定參數是已經刪除的資料時，將會傳回設定狀態屬性為”deleted”的標頭，但是標示為已刪除狀態的詮釋資料將不會傳回。

5.5 ListRecords

Summary and Usage Notes

This verb is used to harvest records from a repository. Optional arguments permit selective harvesting of records based on set membership and/or datestamp. Depending on the repository's support for deletions, a returned header may have a status attribute of "deleted" if a record matching the arguments specified in the request has been deleted. No metadata will be present for records with deleted status.

參數

(1) from 參數，這是選擇性的參數，使用 UTCdatetime 資料值，是用以指定以日戳為基礎的選擇性擷取的下限值(lower bound)。

(2) until 參數，這是選擇性的參數，使用 UTCdatetime 資料值，是用以指定以日戳為基礎的選擇性擷取的上限值(upper bound)。

(3) set 參數，這是選擇性的參數，這是參數與 setSpec 資料值配合使用，用以指定選擇性擷取的集合過濾條件。

(4) resumptionToken 參數，此參數不能與其他參數共同使用，此為用以作流程控制的標籤，是透過前一個 ListIdentifiers 請求(傳回不完整資料列表)傳回。

(5) metadataPrefix 參數(除非使用了 resumptionToken 參數)，此為必要的參數，這是當 metadataPrefix 所指的詮釋資料格式存在時(包含儲存庫所支援的已刪除紀錄)，那些指定的資料標頭將會傳回。至於儲存庫或是特定資料項所支援的詮釋資料格式則可以利用 ListMetadataFormats 來取得。

Arguments

(1) from an optional argument with a UTCdatetime value, which specifies a lower bound for datestamp-based selective harvesting.

(2) until an optional argument with a UTCdatetime value, which specifies a upper bound for datestamp-based selective harvesting.

(3) set an optional argument with a setSpec value , which specifies set criteria for selective harvesting.

(4) resumptionToken an exclusive argument with a value that is the flow control token returned by a previous ListRecords request that issued an incomplete list.

(5) metadataPrefix a required argument (unless the exclusive argument resumptionToken is used) that specifies the metadataPrefix of the format that should be included in the metadata part of the returned records. Records should be included only for items from which the metadata format matching the metadataPrefix can be disseminated. The metadata formats supported by a repository and for a particular item can be retrieved using the ListMetadataFormats request.

錯誤以及例外狀況

(1) badArgument：是指該請求包含了不合標準的參數, 或是沒有使用必要的參數。

(2) badResumptionToken：是指 resumptionToken 參數的資料值錯誤或是已經失效。

(3) cannotDisseminateFormat：是指透過 identifier 參數所參考的資料項並無法支援 metadataPrefix 參數中的資料值所指定的詮釋資料格式。

(4) noRecordsMatch：是指 from、until、以及 set 參數資料值結合處理的結果為一空的資料列表。

(5) noSetHierarchy：是指儲存庫並不支援集合階層架構。

範例

請求
列出使用 oai_rfc1807 詮釋資料格式進行表達的紀錄，日期範圍為在 1998 年 1 月 15 日以後進行新增或修改，且位於 physics 集合中的 hep 子集合下[為了增加可讀性，URL 部份並沒有經過編碼]。

http://an.oa.org/OAI-script?
verb=ListRecords&from=1998-01-15&set=physics:hep&metadataPrefix=oai_rfc1807

Examples
Request
List the records expressed in oai_rfc1807 metadata format, that have been added or modified since January 15, 1998 in the hep subset of the physics set [URL shown without encoding for better readability].

http://an.oa.org/OAI-script?
verb=ListRecords&from=1998-01-15&set=physics:hep&metadataPrefix=oai_rfc1807

回應
兩筆紀錄將傳回：

(1) 第一筆資料使用了 oai_rfc1807 詮釋資料表達。這筆紀錄有說明部份，該資料項隸屬於兩個集合(physics:hep 和 math)。

(2) 第二筆資料的標頭內含了狀態為已刪除的屬性(status='deleted'，因此沒有詮釋資料的部份)。

請注意：傳回的內容只能為那些以 oai_rfc1807 詮釋資料表達的紀錄。如果符合了起訖時間戳記以及集合等參數但是並沒有使用指定的詮釋資料格式時，紀錄仍不會傳回。

Response
Two records are returned:

(1) The first record is expressed in the oai_rfc1807 metadata. This record also has an about part, and the item from which it was disseminated belongs to two sets (physics:hep and math).

(2) The second has a header with a status="deleted" attribute (and therefore no metadata part).

Note: The reply only includes records for those items from which metadata in oai_rfc1807 can be disseminated. No records are returned for those items that fit the from, until, and set arguments but from which the specified format can not be disseminated.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-01T19:20:30Z</responseDate>
 <request verb="ListRecords" from="1998-01-15"
          set="physics:hep"
          metadataPrefix="oai_rfc1807">
          http://an.oa.org/OAI-script</request>
 <ListRecords>
  <record>
    <header>
      <identifier>oai:arXiv.org:hep-th/9901001</identifier>
      <datestamp>1999-12-25</datestamp>
      <setSpec>physics:hep</setSpec>
      <setSpec>math</setSpec>
    </header>
    <metadata>
     <rfc1807 xmlns=
        "http://info.internet.isi.edu:80/in-notes/rfc/files/rfc1807.txt"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation=
       "http://info.internet.isi.edu:80/in-notes/rfc/files/rfc1807.txt
        http://www.openarchives.org/OAI/1.1/rfc1807.xsd">
        <bib-version>v2</bib-version>
        <id>hep-th/9901001</id>
        <entry>January 1, 1999</entry>
        <title>Investigations of Radioactivity</title>
        <author>Ernest Rutherford</author>
```

```
            <date>March 30, 1999</date>
          </rfc1807>
        </metadata>
        <about>
          <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:publisher>Los Alamos arXiv</dc:publisher>
              <dc:rights>Metadata may be used without restrictions as long as
                  the oai identifier remains attached to it.</dc:rights>
          </oai_dc:dc>
        </about>
    </record>
    <record>
        <header status="deleted">
          <identifier>oai:arXiv.org:hep-th/9901007</identifier>
          <datestamp>1999-12-21</datestamp>
        </header>
    </record>
  </ListRecords>
</OAI-PMH>
```

請求

請求使用 oai_dc 詮釋資料格式的紀錄，日期範圍為在 2002 年 5 月 1 日 2:15pm 到 2:20pm 之間進行新增或修改[為了增加可讀性，URL 部份並沒有經過編碼]。

http://www.perseus.tufts.edu/cgi-b:in/pdataprov?verb=ListRecords&from=2002-05-01T14:15:00Z&until=2002-05-01T14:20:00Z&metadataPrefix=oai_dc

Request

Request records in the oai_dc metadata format, modified or added between 2:15pm and 2:20pm UTC on May 1st 2002. [URL shown without encoding for better readability].

http://www.perseus.tufts.edu/cgi-b:in/pdataprov?verb=ListRecords&from=2002-05-01T14:15:00Z&until=2002-05-01T14:20:00Z&metadataPrefix=oai_dc

回應

兩筆紀錄將會傳回, 其中第二筆紀錄在說明部份含有說明出處的容器, 可以說明更詳細的出處資訊。

Response

Two records are returned. The second one has a provenance container in its about element, giving an insight in its chain of provenance.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-01T19:20:30Z</responseDate>
 <request verb="ListRecords" from="2002-05-01T14:15:00Z"
          until="2002-05-01T14:20:00Z" metadataPrefix="oai_dc">
          http://www.perseus.tufts.edu/cgi-bin/pdataprov</request>
 <ListRecords>
  <record>
    <header>
      <identifier>oai:perseus:Perseus:text:1999.02.0084</identifier>
      <datestamp>2002-05-01T14:16:12Z</datestamp>
    </header>
    <metadata>
      <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>Opera Minora</dc:title>
        <dc:creator>Cornelius Tacitus</dc:creator>
        <dc:type>text</dc:type>
        <dc:source>Opera Minora. Cornelius Tacitus. Henry Furneaux.
         Clarendon Press. Oxford. 1900.</dc:source>
```

```
            <dc:language>latin</dc:language>
            <dc:identifier>http://www.perseus.tufts.edu/cgi-bin/ptext?
              doc=Perseus:text:1999.02.0084</dc:identifier>
          </oai_dc:dc>
        </metadata>
    </record>
    <record>
        <header>
          <identifier>oai:perseus:Perseus:text:1999.02.0083</identifier>
          <datestamp>2002-05-01T14:20:55Z</datestamp>
        </header>
        <metadata>
          <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>Germany and its Tribes</dc:title>
            <dc:creator>Tacitus</dc:creator>
            <dc:type>text</dc:type>
            <dc:source>Complete Works of Tacitus. Tacitus. Alfred John Church.
             William Jackson Brodribb. Lisa Cerrato. edited for Perseus.
             New York: Random House, Inc. Random House, Inc. reprinted 1942.
              </dc:source>
            <dc:language>english</dc:language>
            <dc:identifier>http://www.perseus.tufts.edu/cgi-bin/ptext?
             doc=Perseus:text:1999.02.0083</dc:identifier>
          </oai_dc:dc>
          </metadata>
        <about>
          <provenance
           xmlns="http://www.openarchives.org/OAI/2.0/provenance"
           xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
           xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/provenance
           http://www.openarchives.org/OAI/2.0/provenance.xsd">
           <originDescription harvestDate="2002-01-01T11:10:01Z"
altered="true">
```

```
            <baseURL>http://some.oa.org</baseURL>
            <identifier>oai:r2.org:klik001</identifier>
            <datestamp>2001-01-01</datestamp>

<metadataNamespace>http://www.openarchives.org/OAI/2.0/oai_dc/</metadata
Namespace>
            </originDescription>
          </provenance>
       </about>
    </record>
  </ListRecords>
</OAI-PMH>
```

請求

請求使用 oai_marc 詮釋資料格式的紀錄，日期範圍為在 2002 年 6 月 1 日 2:00am 到 3:00am 之間進行新增或修改，但是由於此日期精細度儲存庫並無法支援，因此將會傳回 badArgument，其中內含錯誤的代碼屬性[為了增加可讀性，URL 部份並沒有經過編碼]。

http://memory.loc.gov/cgi-bin/oai?verb=ListRecords&from=2002-06-01T02:00:00Z&until=2002-06-01T03:00:00Z&metadataPrefix=oai_marc

Request

Request records in the the oai_marc metadata format, modified or added between 2:00am and 3:00am UTC on June 1st 2002. The specified granularity is not supported by the repository and therefore, an error with code attribute of badArgument is returned. [URL shown without encoding for better readability].

http://memory.loc.gov/cgi-bin/oai?verb=ListRecords&from=2002-06-01T02:00:00Z&until=2002-06-01T03:00:00Z&metadataPrefix=oai_marc

回應

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-01T19:20:30Z</responseDate>
  <request verb="ListRecords" from="2002-06-01T02:00:00Z"
           until="2002-06-01T03:020:00Z"
           metadataPrefix="oai_marc">
           http://memory.loc.gov/cgi-bin/oai</request>
  <error code="badArgument"/>
</OAI-PMH>
```

5.6 ListSets

摘要以及使用方式說明

此指令是用以取出儲存庫的集合階層架構，對選擇性擷取來說相當有用。

5.6 ListSets

Summary and Usage Notes

This verb is used to retrieve the set structure of a repository, useful for selective harvesting.

參數

唯一的參數為互斥性的 resumptionToken 參數，內含的資料值可以用以進行流程控制，透過之前 ListSets 所傳回的不完整資料列表繼續傳輸。

Arguments

resumptionToken an exclusive argument with a value that is the flow control token returned by a previous ListSets request that issued an incomplete list.

錯誤以及例外狀況

(1) badArgument：是指該請求包含了不合標準的參數，或是沒有使用必要的參數。

(2) badResumptionToken：是指 resumptionToken 參數的資料值錯誤或是已經失效。

(3) noSetHierarchy：是指儲存庫並不支援集合階層架構。

Error and Exception Conditions

(1) badArgument - The request includes illegal arguments or is missing required arguments.

(2) badResumptionToken - The value of the resumptionToken argument is invalid or expired.

(3) noSetHierarchy - The repository does not support sets.

範例
請求
http://an.oa.org/OAI-script?verb=ListSets

Examples
Request
http://an.oa.org/OAI-script?verb=ListSets

回應
以下的回應內容說明擁有兩個最上層的集合, 其 setSpec 資料值分別為 music 和 video。Music 集合擁有兩個子集合，其 setSpec 資料值分別為 setSpec music:(muzak)和 music:(elec)，其中 music:(elec)子集合內含 setDescription 元件，資料值為都柏林核心集的容器，用以描述其內容。

Response
The following response indicates a set hierarchy with two top level sets with respective setSpec music and video. The music set has two subsets, with setSpec music:(muzak) and music:(elec). The subsets identified by setSpec music:(elec), has a setDescription element which holds a Dublin Core container, used to describe its contents.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-08-11T07:21:33Z</responseDate>
 <request verb="ListSets">http://an.oa.org/OAI-script</request>
 <ListSets>
  <set>
    <setSpec>music</setSpec>
    <setName>Music collection</setName>
  </set>
```

```
  <set>
    <setSpec>music:(muzak)</setSpec>
    <setName>Muzak collection</setName>
  </set>
  <set>
    <setSpec>music:(elec)</setSpec>
    <setName>Electronic Music Collection</setName>
    <setDescription>
      <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:description>This set contains metadata describing
              electronic music recordings made during the 1950ies
              </dc:description>
       </oai_dc:dc>
    </setDescription>
  </set>
  <set>
    <setSpec>video</setSpec>
    <setName>Video Collection</setName>
  </set>
 </ListSets>
</OAI-PMH>
```

請求

http://purl.org/alcme/etdcat/servlet/OAIHandler?verb=ListSets

Request

http://purl.org/alcme/etdcat/servlet/OAIHandler?verb=ListSets

回應

本回應內容說明了儲存庫中並沒有集合階層架構。

Response

The response shows that the repository does not have a set hierarchy.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2001-06-01T19:20:30Z</responseDate>
  <request verb="ListSets">
           http://purl.org/alcme/etdcat/servlet/OAIHandler</request>
  <error code="noSetHierarchy">This repository does not
         support sets</error>
</OAI-PMH>
```

6.都柏林核心集

下表說明了都柏林核心集所使用的 XML 綱要，此綱要在 OAI-PMH 中是透過使用 metadataPrefix(定為 oai_dc)來進行。本文件中所有的範例皆含有都柏林核心集詮釋資料，可以透過 XML 綱要進行確認。至於其他詮釋資料格式的綱要則在實作指導綱要中有所說明。

The following table shows the XML Schema for Dublin Core without qualification, which is associated with the reserved metadataPrefix oai_dc in the OAI-PMH. All examples in this document that include Dublin Core metadata, validate against this XML schema. Schema for other metadata formats are provided in the accompanying Implementation Guidelines document.

以下為用以確認都柏林核心集的 XML 綱要範例，其中有使用到內容值為 oai_dc 的 metadataPrefix。
A XML schema for validating Unqualified Dublin Core metadata associated with the reserved oai_dc metadataPrefix

```
<schema targetNamespace="http://www.openarchives.org/OAI/2.0/oai_dc/"
        xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
        xmlns:dc="http://purl.org/dc/elements/1.1/"
        xmlns="http://www.w3.org/2001/XMLSchema"
        elementFormDefault="qualified" attributeFormDefault="unqualified">

<annotation>
  <documentation>
      XML Schema 2002-03-18 by Pete Johnston.
```

```
        Adjusted for usage in the OAI-PMH.
        Schema imports the Dublin Core elements from the DCMI schema for
unqualified Dublin Core.
        2002-12-19 updated to use simpledc20021212.xsd (instead of
simpledc20020312.xsd)
   </documentation>
</annotation>

<import namespace="http://purl.org/dc/elements/1.1/"

schemaLocation="http://dublincore.org/schemas/xmls/simpledc20021212.xsd"/>

<element name="dc" type="oai_dc:oai_dcType"/>

<complexType name="oai_dcType">
   <choice minOccurs="0" maxOccurs="unbounded">
      <element ref="dc:title"/>
      <element ref="dc:creator"/>
      <element ref="dc:subject"/>
      <element ref="dc:description"/>
      <element ref="dc:publisher"/>
      <element ref="dc:contributor"/>
      <element ref="dc:date"/>
      <element ref="dc:type"/>
      <element ref="dc:format"/>
      <element ref="dc:identifier"/>
      <element ref="dc:source"/>
      <element ref="dc:language"/>
      <element ref="dc:relation"/>
      <element ref="dc:coverage"/>
      <element ref="dc:rights"/>
   </choice>
</complexType>

</schema>
```

This Schema is available at http://www.openarchives.org/OAI/2.0/oai_dc.xsd

範例

Examples

```
<?xml version="1.0" encoding="UTF-8"?>
<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 xml:lang="en">The Cornell Law Quarterly</dc:title>
   <dc:date>1915-1916</dc:date>
   <dc:identifier>http://heinonline.org/HeinOnline/show.pl?
         handle=hein.journals/clqv1%26id=1%26size=4</dc:identifier>
   <dc:rights>Available by Subscription.
                  See http://www.wshein.com</dc:rights>
</oai_dc:dc>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<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 xml:lang="en">Grassmann's space analysis</dc:title>
   <dc:creator>Hyde, E. W. (Edward Wyllys)</dc:creator>
   <dc:subject>LCSH:Ausdehnungslehre; LCCN QA205.H99</dc:subject>
   <dc:publisher>J. Wiley &amp;amp; Sons</dc:publisher>
   <dc:date>Created: 1906; Available: 1991</dc:date>
   <dc:type>text</dc:type>
   <dc:identifier>http://resolver.library.cornell.edu/math/1796949
        </dc:identifier>
   <dc:language>english</dc:language>
   <dc:rights xml:lang="en">Public Domain</dc:rights>
</oai_dc:dc>
```

英中名詞對照表

-A-

-B-

-C-

-D-

Datestamp *日戳*

-E-

-F-

-G-

-H-

Harvester *擷取器*

Header 標頭

-I-

-J-

-K-

-L-

-M-

-N-

-O-

-P-

-Q-

-R-

Record 紀錄

Request 請求

Repository *儲存庫*

-S-

Service Provider 服務提供者

Status-code *狀態碼*

-T-

-U-

Unique identifier 唯一識別符

-V-

-W-

-X-

-Y-

-Z-

The Open Archives Initiative Protocol for Metadata Harvesting

Contents

1.Scope

The Open Archives Initiative Protocol for Metadata Harvesting (referred to as the OAI-PMH in the remainder of this document) provides an application-independent interoperability framework based on metadata harvesting. There are two classes of participants in the OAI-PMH framework:

(3) Data Providers administer systems that support the OAI-PMH as a means of exposing metadata; and

(4) Service Providers use metadata harvested via the OAI-PMH as a basis for building value-added services.

In this document the key words "must", "must not", " required", "shall", "shall not", "should", " should not", "recommended", "may", and "optional " in bold face are to be interpreted as described in RFC 2119 . An implementation is not conformant if it fails to satisfy one or more of the "must" or "required" level requirements for the protocols it implements.

2.Terms and definitions

2.1 Harvester

A harvester is a client application that issues OAI-PMH requests. A harvester is operated by a service provider as a means of collecting metadata from repositories.

2.2 Repository

A repository is a network accessible server that can process the 6 OAI-PMH requests in the manner described in this document. A repository is managed by a data provider to expose metadata to harvesters. To allow various repository configurations, the OAI-PMH distinguishes between three distinct entities related to the metadata made accessible by the OAI-PMH:

(4) resource - A resource is the object or "stuff" that metadata is "about". The nature of a resource, whether it is physical or digital, or whether it is stored in the repository or is a constituent of another database, is outside the scope of the OAI-PMH.

(5) item - An item is a constituent of a repository from which metadata about a resource can be disseminated. That metadata may be disseminated

on-the-fly from the associated resource, cross-walked from some canonical form, actually stored in the repository, etc.

(6) record - A record is metadata in a specific metadata format. A record is returned as an XML-encoded byte stream in response to a protocol request to disseminate a specific metadata format from a constituent item.

2.3 Item

An item is a constituent of a repository from which metadata about a resource can be disseminated. An item is conceptually a container that stores or dynamically generates metadata about a single resource in multiple formats, each of which can be harvested as records via the OAI-PMH. Each item has an identifier that is unique within the scope of the repository of which it is a constituent.

2.4 Unique Identifier

A unique identifier unambigiously identifies an item within a repository; the unique identifier is used in OAI-PMH requests for extracting metadata from the item. Items may contain metadata in multiple formats. The unique identifier maps to the item, and all possible records available from a single item share the same unique identifier.

The format of the unique identifier must correspond to that of the URI (Uniform Resource Identifier) syntax. Individual communities may develop community-specific URI schemes for coordinated use across repositories. The scheme component of the unique identifiers must not correspond to that of a recognized URI scheme unless the identifiers conform to that scheme. Repositories may implement the oai-identifier syntax described in the accompanying Implementation Guidelines document.

Unique identifiers play two roles in the protocol:

(3) Response: Identifiers are returned by both the ListIdentifiers and ListRecords requests.

(4) Request: An identifier, in combination with a metadataPrefix, is used in the GetRecord request as a means of requesting a record in a specific metadata format from an item.

Note that the identifier described here is not that of a resource. The nature of a resource identifier is outside the scope of the OAI-PMH. To facilitate access to the resource associated with harvested metadata, repositories should use an element in metadata records to establish a linkage between the record (and the identifier of its item) and the identifier (URL, URN, DOI, etc.) of the associated resource. The mandatory Dublin Core format provides the identifier element that should be used for this purpose.

2.5 Record

A record is metadata expressed in a single format. A record is returned in an XML-encoded byte stream in response to an OAI-PMH request for metadata from an item. A record is identified unambigiously by the combination of the unique identifier of the item from which the record is available, the metadataPrefix identifying the metadata format of the record, and the datestamp of the record. The XML-encoding of records is organized into the following parts:

(4) header -- contains the unique identifier of the item and properties necessary for selective harvesting. The header consists of the following parts:

  (a) the unique identifier -- the unique identifier of an item in a repository;

  (b) the datestamp -- the date of creation, modification or deletion of the record for the purpose of selective harvesting.

  (c) zero or more setSpec elements -- the set membership of the item for the purpose of selective harvesting.

  (d) an optional status attribute with a value of deleted indicates the withdrawal of availability of the specified metadata format for the item, dependent on the repository support for deletions.

(5) metadata -- a single manifestation of the metadata from an item. The OAI-PMH supports items with multiple manifestations (formats) of metadata. At a minimum, repositories must be able to return records with metadata expressed in the Dublin Core format, without any qualification. Optionally, a repository may also disseminate other formats of metadata. The specific metadata format of the record to be disseminated is specified

by means of an argument -- the metadataPrefix -- in the GetRecord or ListRecords request that produces the record. The ListMetadataFormats request returns the list of all metadata formats available from a repository, or for a specific item (which can be specified as an argument to the ListMetadataFormats request).

(6) about -- an optional and repeatable container to hold data about the metadata part of the record. The contents of an about container must conform to an XML Schema. Individual implementation communities may create XML Schema that define specific uses for the contents of about containers. Two common uses of about containers are:

(a) rights statements: some repositories may find it desirable to attach terms of use to the metadata they make available through the OAI-PMH. No specific set of XML tags for rights expression is defined by OAI-PMH, but the about container is provided to allow for encapsulating community-defined rights tags.

(b) provenance statements: One suggested use of the about container is to indicate the provenance of a metadata record, e.g. whether it has been harvested itself and if so from which repository, and when. An XML Schema for such a provenance container, as well as some supporting information is available from the accompanying Implementation Guidelines document.

The following example shows an XML-encoding of a record and its components:

(4) the header part with:

(a) a unique identifier of the item from which the record was disseminated, equal to oai:arXiv.org:cs/0112017;

(b) the datestamp of the record equal to 2002-02-28;

(c) two setSpecs, respectively cs and math, indicating that the item from which the record was disseminated belongs to two sets of the repository;

(5) the metadata part. This consists of a single root tag - in the example the tag oai_dc:dc - with the nested tags belonging to the corresponding metadata format - in the example, Dublin Core elements such as dc:title. Note that the root tag within the metadata part includes a number of attributes that are common to all XML documents that use namespaces and schema validity:

(2.1) namespace declarations -- the declarations of the namespaces used within the metadata part, each of which is prefixed with xmlns. Namespace declarations within the metadata part fall into two categories:

(a) metadata format specific namespace(s) - every metadata part must include one or more xmlns prefixed attributes that define the correspondence between a metadata format prefix -- e.g. dc -- and the namespace URI (as defined by the XML namespace specification) of the respective metadata format. Some metadata formats employ tags from multiple namespaces, requiring multiple xmlns prefixed attributes -- in the example, there are declarations for both oai_dc and dc.

(b) xml schema namespace - every metadata part must include the attribute xmlns:xsi, the value of which must always be the URI shown in the example, which is the namespace URI for XML schema.

(2.2) xsi:schemaLocation -- the value of which is a URI, URL pair; the first is the namespace URI (as defined by the XML namespace specification) of the metadata that follows in this part, and the second is the URL of the XML schema for validation of the metadata that follows.

(6) one about part of the record which uses the oai_provenance.xsd schema, described in the accompanying Implementation Guidelines document, as a means to provide information regarding the origins of the metadata part of

the record. Note that the root element within each about part has the same structure as the root element in the metadata part.

```
<header>
   <identifier>oai:arXiv:cs/0112017</identifier>
   <datestamp>2002-02-28</datestamp>
   <setSpec>cs</setSpec>
   <setSpec>math</setSpec>
</header>
<metadata>
 <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>Using Structural Metadata to Localize Experience of Digital
                  Content</dc:title>
    <dc:creator>Dushay, Naomi</dc:creator>
    <dc:subject>Digital Libraries</dc:subject>
    <dc:description>With the increasing technical sophistication of both
     information consumers and providers, there is increasing demand for
     more meaningful experiences of digital information. We present a
     framework that separates digital object experience, or rendering,
     from digital object storage and manipulation, so the
     rendering can be tailored to particular communities of users.
    </dc:description>
    <dc:description>Comment: 23 pages including 2 appendices,
                          8 figures</dc:description>
    <dc:date>2001-12-14</dc:date>
    <dc:type>e-print</dc:type>
    <dc:identifier>http://arXiv.org/abs/cs/0112017</dc:identifier>
 </oai_dc:dc>
</metadata>
<about>
   <provenance
        xmlns="http://www.openarchives.org/OAI/2.0/provenance"
```

```
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/provenance
        http://www.openarchives.org/OAI/2.0/provenance.xsd">
      <originDescription harvestDate="2002-02-02T14:10:02Z" altered="true">
        <baseURL>http://the.oa.org</baseURL>
        <identifier>oai:r2:klik001</identifier>
        <datestamp>2002-01-01</datestamp>

<metadataNamespace>http://www.openarchives.org/OAI/2.0/oai_dc/</metadataNamespa
ce>
      </originDescription>
   </provenance>
</about>
```

2.5.1 Deleted records

If a record is no longer available then it is said to be deleted. Repositories must declare one of three levels of support for deleted records in the deletedRecord element of the Identify response:

(4) no - the repository does not maintain information about deletions. A repository that indicates this level of support must not reveal a deleted status in any response.

(5) persistent - the repository maintains information about deletions with no time limit. A repository that indicates this level of support must persistently keep track of the full history of deletions and consistently reveal the status of a deleted record over time.

(6) transient - the repository does not guarantee that a list of deletions is maintained persistently or consistently. A repository that indicates this level of support may reveal a deleted status for records.

If a repository does not keep track of deletions then such records will simply vanish from responses and there will be no way for a harvester to discover deletions through continued incremental harvesting. If a repository does keep track of deletions then the datestamp of the deleted record must be the date

and time that it was deleted. Responses to GetRecord request for a deleted record must then include a header with the attribute status="deleted", and must not include metadata or about parts. Similarly, responses to selective harvesting requests with set membership and date range criteria that include deleted records must include the headers of these records. Incremental harvesting will thus discover deletions from repositories that keep track of them.

Deleted status is a property of individual records. Like a normal record, a deleted record is identified by a unique identifier, a metadataPrefix and a datestamp. Other records, with different metadataPrefix but the same unique identifier, may remain available for the item.

2.6 Set

A set is an optional construct for grouping items for the purpose of selective harvesting. Repositories may organize items into sets. Set organization may be flat, i.e. a simple list, or hierarchical. Multiple hierarchies with distinct, independent top-level nodes are allowed. Hierarchical organization of sets is expressed in the syntax of the setSpec parameter as described below. When a repository defines a set organization it must include set membership information in the headers of items returned in response to the ListIdentifiers , ListRecords and GetRecord requests.

Each node in a set organization of a repository has:

(4) a setSpec -- a colon [:] separated list indicating the path from the root of the set hierarchy to the respective node. Each element in the list is a string consisting of any valid URI unreserved characters, which must not contain any colons [:]. Since a setSpec forms a unique identifier for the set within the repository, it must be unique for each set. Flat set organizations have only sets with setSpec that do not contain any colons [:].

(5) a setName -- a short human-readable string naming the set.

(6) a setDescription -- an optional and repeatable container that may hold community-specific XML-encoded data about the set; the accompanying Implementation Guidelines document provides suggestions regarding the usage of this container.

The following is an example of a possible set hierarchy in a repository:

(3) Institutions

  (e) Oceanside University of Nebraska

  (f) Valley View University of Florida

(4) Subjects

  (e) Existential Kenesiology

  (f) Quantum Psychology

The following table shows a possible representation of the above set hierarchy by means of setName and respective setSpec values.

setName	setSpec
Institutions	institution
Oceanside University of Nebraska	institution:nebraska
Valley View University of Florida	institution:florida
Subjects	subject
Existential Kenesiology	subject:kenesiology
Quantum Psychology	subject:quantum

An item may be organized in one set, several sets, or no sets at all. In the example above, it is conceivable that an individual item is organized in both subject and institution:florida. A harvester should not assume that harvesting every set in a repository will retrieve metadata from all items in the repository. Items may also be assigned to interior nodes in the set hierarchy.

The actual meaning of a set or of the arrangement of sets in a repository is not defined in the protocol. It is expected that individual communities may formulate well-defined set configurations with perhaps a controlled vocabulary for setNames and setSpec , and may even develop mechanisms for exposing these to harvesters. For example, a group of cooperating e-print archives in a specific discipline may agree on sets that arrange metadata in their repositories based on a controlled subject classification.

A repository's set hierarchy is represented in the protocol via setSpecs. ListSets returns a list indicating the configuration of sets in a repository. Each member of this list must include a setSpec and a setName and may include a setDescription. ListRecords and ListIdentifiers requests may include an optional set argument, the value of which is a setSpec, to specify the target set for selective harvesting. In the previous example of a set hierarchy, the setSpec institution:nebraska could be used in a request to return only those records that are disseminated from items organized in the set represented by this setSpec. Five issues should be noted here:

(6) If a repository supports sets then it must include set membership information in response to ListIdentifiers, ListRecords and GetRecord requests. The list of setSpec should include only the minimum number of setSpec required to specify the set membership. Using the previous example of a set hierarchy, the header for an item organized in set institution:florida should not include setSpec institution since that is implied by the setSpec institution:florida.

(7) An item may be organized in more than one set; meaning that different setSpec arguments may return the same record(s).

(8) An item need not be organized in any set; meaning that an exhaustive repetition of ListRecords requests with all possible setSpecs is not guaranteed to return all records in the repository. The only guaranteed methods of harvesting all records or headers are ListRecords or ListIdentifiers requests with no setSpec argument.

(9) When a setSpec is used as an argument, the response must include records or headers from all items in the set specified by the setSpec , and all records or headers from items in sets that are descendant from the specified set. Using the previous example of a set hierarchy, a setSpec of institution to the ListRecords request will return all records from metadata organized within

the set with a setSpec value equal to institution and within the descendent sets with setSpec values equal to institution:florida and institution:nebraska.

(10) The set hierarchy of a repository may include sets that are empty.

2.7 Selective Harvesting

Selective harvesting allows harvesters to limit harvest requests to portions of the metadata available from a repository. The OAI-PMH supports selective harvesting with two types of harvesting criteria that may be combined in an OAI-PMH request: datestamps and set membership.

2.7.1 Selective Harvesting and Datestamps

Harvesters may use datestamps to harvest only those records that were created, deleted, or modified within a specified date range. To specify datestamp-based selective harvesting, datestamps are included as values of the optional arguments, from and until, in the ListRecords and ListIdentifiers requests. Harvesting is restricted to the range specified by the from and until arguments, extending back to the earliest datestamp if from is omitted, and forward to the most recent datestamp if until is omitted. Range limits are inclusive: from specifies a bound that must be interpreted as "greater than or equal to", until specifies a bound that must be interpreted as "less than or equal to". Therefore, the from arugment must be less than or equal to the until argument. Otherwise, a repository must issue a badArgument error.

Repositories must support selective harvesting with the from and until arguments expressed at day granularity. Optional support for seconds granularity is indicated in the response to the Identify request. The value of datestamps in both requests and responses must comply to the specifications for UTCdatetime in this document. A repository must update the datestamp of a record if a change occurs, the result of which would be a change to the metadata part of the XML-encoding of the record. Such changes include, but are not limited to, changes to the metadata of the record, changes to the metadata format of the record, introduction of a new metadata format, termination of support for a metadata format, etc.

Datestamp ranges for selective harvesting are expressed in the from and until arguments that may be submitted in the ListRecords and ListIdentifiers requests. Repositories must use the following rules to create a ListRecords response matching the specified datestamp range according to the type of

change that occured within the repository. The response to a ListIdentifiers request follows the same rules but is abbreviated to include only headers rather than records.

(4) modification - the response must include records, corresponding to the metadataPrefix argument, which have changed within the bounds of the from and until arguments.

(5) creation - the response must include records, corresponding to the metadataPrefix argument, that have become available from the repository within the bounds of the from and until arguments.

(6) deletion - depending on the level at which a repository keeps track of deleted records, the response may include headers of records, corresponding to the metadataPrefix argument, which have been withdrawn from the repository within the bounds of the from and until arguments. Deleted status is indicated via the status attribute of the header element and no metedata is included.

Every header returned by the GetRecord, ListRecords or ListIdentifiers requests contains a datestamp, which reflects the most recent date and time of the creation, modification, or deletion according to the rules defined above.

2.7.2 Selective Harvesting and Sets

Harvesters may specify set membership as a criteria for selective harvesting. To specify set-based selective harvesting, a setSpec is included as the value of the optional set argument to the ListRecords and ListIdentifiers requests, thereby specifying selective harvesting of records from items within the respective set.

When a setSpec is used as an argument, the response must include:

(5) the records corresponding to the metadataPrefix argument, or headers thereof in the case of deleted records, available from those items in the set specified by the setSpec;

(6) the records corresponding to the metadataPrefix argument, or headers thereof in the case of deleted records, available from those items in sets that are descendant from the specified set.

3.Normative reference

[OAI-PMH] Open Archives Initiative Protocol for Metadata Harvesting （OAI-PMH）2.0, 2004. < http://www.openarchives.org/OAI/openarchivesprotocol.html >

4.Protocol Features

4.1 HTTP Embedding of OAI-PMH requests

OAI-PMH requests are expressed as HTTP requests. A typical implementation uses a standard Web server that is configured to dispatch OAI-PMH requests to the software handling these requests. The remainder of this section describes the aspects of the protocol that are specific to the HTTP embedding.

4.1.1 HTTP Request Format

OAI-PMH requests must be submitted using either the HTTP GET or POST methods. POST has the advantage of imposing no limitations on the length of arguments. Repositories must support both the GET and POST methods. There is a single base URL for all requests. The base URL specifies the Internet host and port, and optionally a path, of an HTTP server acting as a repository. Repositories expose their base URL as the value of the baseURL element in the Identify response. Note that the composition of any path is determined by the configuration of the repository's HTTP server.

In addition to the base URL, all requests consist of a list of keyword arguments, which take the form of key=value pairs. Arguments may appear in any order and multiple arguments must be separated by ampersands [&]. Each OAI-PMH request must have at least one key=value pair that specifies the OAI-PMH request issued by the harvester:

(3) key is the string 'verb';

(4) value is one of the defined OAI-PMH requests.

The number and nature of additional key=value pairs depends on the arguments for the individual request.

4.1.1.1 Encoding an OAI-PMH request in a URL for an HTTP GET

URLs for GET requests have keyword arguments appended to the base URL, separated from it by a question mark [?]. For example, the URL of a GetRecord request to a repository with base URL that is http://an.oa.org/OAI-script might be:

http://an.oa.org/OAI-script?verb=GetRecord&identifier=oai:arXiv.org:hep-th/9901001&metadataPrefix=oai_dc

However, since special characters in URIs must be encoded, the correct form of the above GET request URL is:

http://an.oa.org/OAI-script?verb=GetRecord&identifier=oai%3AarXiv.org%3Ahep-th%2F9901001&metadataPrefix=oai_dc

4.1.1.2 Encoding an OAI-PMH request in an HTTP POST

Keyword arguments are carried in the message body of the HTTP POST. The Content-Type of the request must be application/x-www-form-urlencoded. For example, submitting the same request as above using the POST method would use just the base URL as the URL, with the format of the POST being:

```
POST http://an.oa.org/OAI-script HTTP/1.0
Content-Length: 82
Content-Type: application/x-www-form-urlencoded

verb=GetRecord&identifier=oai%3AarXiv.org%3Ahep-th%2F9901001&metadataPrefix=oai_dc
```

4.1.1.3 Encoding of special characters in keyword arguments of OAI-PMH requests

The syntax rules for URIs restrict a few characters to special roles in certain contexts, and require that if these characters are used in any other way that they must be written as an escape sequence, i.e. a percent sign followed by the character code in hexadecimal. The reserved characters include:

Character	URI Role	Escape Sequence
/	Path Component Separator	%2F
?	Query Component Separator	%3F
#	Fragment Identifier	%23
=	Name/Value Separator	%3D
&	Argument Separator in Query Component	%26
:	Host Port Separator	%3A
;	Authority Namespace Separator	%3B
	Space Character	%20
%	Escape Indicator	%25

+	Escaped Space	%2B

As a result, these characters must be represented by their respective escape sequence if their use does not correspond to their established URI role. In case of the OAI-PMH, this means that the reserved characters must be encoded when they appear in the value part of the key=value pairs of the request. This applies for both the GET and POST encoding of the OAI-PMH requests.

4.1.2 HTTP Response Format

Responses to requests are formatted as HTTP responses, with appropriate HTTP header fields.

4.1.2.1 Content-Type

The Content-Type returned for all OAI-PMH requests must be text/xml.

4.1.2.2 Status-Code

OAI-PMH errors are distinguished from HTTP Status-Codes. Since OAI-PMH uses HTTP as a transport layer, servers implementing OAI-PMH must conform to HTTP status code definitions and report relevant HTTP transport layer status via those Status-Codes. OAI-PMH repositories may employ HTTP Status-Codes in addition to "200 OK". For instance, the following Status-Codes may be useful for load balancing in OAI repositories:

(1) 302 - Allows the repository to temporarily redirect an OAI-PMH request

to another repository. The URI of the temporary repository should be given by the Location field in the HTTP response.

(2) 503 - Service unavailable, a Retry-After period is specified. Harvesters

should wait this period before attempting another OAI-PMH request.

4.1.3 Response Compression

Response compression is optional in OAI-PMH. Compression of responses to OAI-PMH requests is handled at the level of HTTP, with the following restrictions:

(6) Harvesters may include an Accept-Encoding header in their OAI-PMH requests to specify response compression preferences.

(7) Harvesters that do not include an Accept-Encoding header in their requests will always receive uncompressed responses.

(8) When a request includes an Accept-Encoding header the list of encodings must include the identity (no compression) encoding (with a non-zero qvalue).

(9) Repositories must support the HTTP identity encoding.

(10) Repositories should express the encodings they support in addition to identity by including compression elements in the Identify response.

4.2 XML Response Format

All responses to OAI-PMH requests must be well-formed XML instance documents. Encoding of the XML must use the UTF-8 representation of Unicode. Character references, rather than entity references, must be used. Character references allow XML responses to be treated as stand-alone documents that can be manipulated without dependency on entity declarations external to the document.

The XML data for all responses to OAI-PMH requests must validate against the XML Schema shown at the end of this section . As can be seen from that schema, responses to OAI-PMH requests have the following common markup:

(5) The first tag output is an XML declaration where the version is always 1.0 and the encoding is always UTF-8, eg: <?xml version="1.0" encoding="UTF-8" ?>

(6) The remaining content is enclosed in a root element with the name OAI-PMH. This element must have three attributes that define the XML namespaces used in the remainder of the response and the location of the validating schema:

(d) xmlns -- the value of which must be the namespace URI of the OAI-PMH (http://www.openarchives.org/OAI/2.0/).

(e) xmlns:xsi -- the value of which must be the namespace URI for XML schema (http://www.w3.org/2001/XMLSchema-instance).

(f) xsi:schemaLocation -- is a pair, the first part of which is the namespace URI (as defined by the XML namespace specification) of the OAI-PMH (http://www.openarchives.org/OAI/2.0/), and the second part is the URL of the XML schema for validation of the response (http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd).

(7) For all responses, the first two children of the root element are:

(c) responseDate -- a UTCdatetime indicating the time and date that the response was sent. This must be expressed in UTC.

(d) request -- indicating the protocol request that generated this response. The rules for generating the request element are as follows:

(b-1)The content of the request element must always be the base URL of the protocol request;

(b-2)The only valid attributes for the request element are the keys of the key=value pairs of protocol request. The attribute values must be the corresponding values of those key=value pairs;

(b-3)In cases where the request that generated this response did not result in an error or exception condition, the attributes and attribute values of the request element must match the key=value pairs of the protocol request;

(b-4)In cases where the request that generated this response resulted in a badVerb or badArgument error condition, the repository must return the base URL of the protocol request only. Attributes must not be provided in these cases.

(8) The third child of the root element is either:

(c) an error element that must be used in case of an error or exception condition;

(d) an element with the same name as the verb of the respective OAI-PMH request.

An example of a successful reply to the GetRecord request shown above is of the form:

```
<?xml version="1.0" encoding="UTF-8" ?>
<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>2002-05-01T19:20:30Z</responseDate>
  <request verb="GetRecord" identifier="oai:arXiv.org:hep-th/9901001"
           metadataPrefix="oai_dc">http://an.oa.org/OAI-script</request>
  <GetRecord>
    <record>
      ...
    </record>
  </GetRecord>
</OAI-PMH>
```

4.2.1 XML Schema for Validating Responses to OAI-PMH Requests

```
<schema targetNamespace="http://www.openarchives.org/OAI/2.0/"
        xmlns="http://www.w3.org/2001/XMLSchema"
        xmlns:oai="http://www.openarchives.org/OAI/2.0/"
        elementFormDefault="qualified"
        attributeFormDefault="unqualified">

  <annotation>
    <documentation>
    XML Schema which can be used to validate replies to all OAI-PMH
    v2.0 requests. Herbert Van de Sompel, 2002-05-13.
    Validated with XML Spy v.4.3 on 2002-05-13.
    Validated with XSV 1.203.2.45/1.106.2.22 on 2002-05-13.
    Added definition of protocolVersionType instead of using anonymous
    type. No change of function. Simeon Warner, 2004-03-29.
    Tightened definition of UTCdatetimeType to enforce the restriction
    to UTC Z notation. Simeon Warner, 2004-09-14.
    Corrected pattern matches for setSpecType and metedataPrefixType
    to agree with protocol specification. Simeon Warner, 2004-10-12.
    $Date: 2004/10/12 15:20:29 $
```

```
      </documentation>
   </annotation>

   <element name="OAI-PMH" type="oai:OAI-PMHtype"/>

   <complexType name="OAI-PMHtype">
      <sequence>
         <element name="responseDate" type="dateTime"/>
         <element name="request" type="oai:requestType"/>
         <choice>
            <element name="error" type="oai:OAI-PMHerrorType"
maxOccurs="unbounded"/>
            <element name="Identify" type="oai:IdentifyType"/>
            <element name="ListMetadataFormats"
type="oai:ListMetadataFormatsType"/>
            <element name="ListSets" type="oai:ListSetsType"/>
            <element name="GetRecord" type="oai:GetRecordType"/>
            <element name="ListIdentifiers" type="oai:ListIdentifiersType"/>
            <element name="ListRecords" type="oai:ListRecordsType"/>
         </choice>
      </sequence>
   </complexType>

   <complexType name="requestType">
      <annotation>
         <documentation>Define requestType, indicating the protocol request that
         led to the response. Element content is BASE-URL, attributes are arguments
         of protocol request, attribute-values are values of arguments of protocol
         request</documentation>
      </annotation>
      <simpleContent>
         <extension base="anyURI">
            <attribute name="verb" type="oai:verbType" use="optional"/>
            <attribute name="identifier" type="oai:identifierType" use="optional"/>
            <attribute name="metadataPrefix" type="oai:metadataPrefixType"
use="optional"/>
            <attribute name="from" type="oai:UTCdatetimeType" use="optional"/>
            <attribute name="until" type="oai:UTCdatetimeType" use="optional"/>
```

```
            <attribute name="set" type="oai:setSpecType" use="optional"/>
            <attribute name="resumptionToken" type="string" use="optional"/>
          </extension>
        </simpleContent>
    </complexType>

    <simpleType name="verbType">
        <restriction base="string">
          <enumeration value="Identify"/>
          <enumeration value="ListMetadataFormats"/>
          <enumeration value="ListSets"/>
          <enumeration value="GetRecord"/>
          <enumeration value="ListIdentifiers"/>
          <enumeration value="ListRecords"/>
        </restriction>
    </simpleType>

    <!-- define OAI-PMH error conditions -->
    <!-- ============================== -->

    <complexType name="OAI-PMHerrorType">
        <simpleContent>
          <extension base="string">
            <attribute name="code" type="oai:OAI-PMHerrorcodeType"
use="required"/>
          </extension>
        </simpleContent>
    </complexType>

    <simpleType name="OAI-PMHerrorcodeType">
        <restriction base="string">
          <enumeration value="cannotDisseminateFormat"/>
          <enumeration value="idDoesNotExist"/>
          <enumeration value="badArgument"/>
          <enumeration value="badVerb"/>
          <enumeration value="noMetadataFormats"/>
          <enumeration value="noRecordsMatch"/>
          <enumeration value="badResumptionToken"/>
```

```
        <enumeration value="noSetHierarchy"/>
      </restriction>
   </simpleType>

   <!-- define OAI-PMH verb containers -->
   <!-- ============================== -->

   <complexType name="IdentifyType">
      <sequence>
         <element name="repositoryName" type="string"/>
         <element name="baseURL" type="anyURI"/>
         <element name="protocolVersion" type="oai:protocolVersionType"/>
         <element name="adminEmail" type="oai:emailType"
maxOccurs="unbounded"/>
         <element name="earliestDatestamp" type="oai:UTCdatetimeType"/>
         <element name="deletedRecord" type="oai:deletedRecordType"/>
         <element name="granularity" type="oai:granularityType"/>
         <element name="compression" type="string" minOccurs="0"
maxOccurs="unbounded"/>
         <element name="description" type="oai:descriptionType"
                  minOccurs="0" maxOccurs="unbounded"/>
      </sequence>
   </complexType>

   <complexType name="ListMetadataFormatsType">
      <sequence>
         <element name="metadataFormat" type="oai:metadataFormatType"
maxOccurs="unbounded"/>
      </sequence>
   </complexType>

   <complexType name="ListSetsType">
      <sequence>
         <element name="set" type="oai:setType" maxOccurs="unbounded"/>
         <element name="resumptionToken" type="oai:resumptionTokenType"
minOccurs="0"/>
      </sequence>
   </complexType>
```

```
  <complexType name="GetRecordType">
    <sequence>
      <element name="record" type="oai:recordType"/>
    </sequence>
  </complexType>

  <complexType name="ListRecordsType">
    <sequence>
      <element name="record" type="oai:recordType" maxOccurs="unbounded"/>
      <element name="resumptionToken" type="oai:resumptionTokenType"
minOccurs="0"/>
    </sequence>
  </complexType>

  <complexType name="ListIdentifiersType">
    <sequence>
      <element name="header" type="oai:headerType" maxOccurs="unbounded"/>
      <element name="resumptionToken" type="oai:resumptionTokenType"
minOccurs="0"/>
    </sequence>
  </complexType>

  <!-- define basic types used in replies to
         GetRecord, ListRecords, ListIdentifiers -->
  <!-- ======================================== -->

  <complexType name="recordType">
    <annotation>
      <documentation>A record has a header, a metadata part, and
        an optional about container</documentation>
    </annotation>
    <sequence>
      <element name="header" type="oai:headerType"/>
      <element name="metadata" type="oai:metadataType" minOccurs="0"/>
      <element name="about" type="oai:aboutType" minOccurs="0"
maxOccurs="unbounded"/>
    </sequence>
```

```
  </complexType>

  <complexType name="headerType">
    <annotation>
      <documentation>A header has a unique identifier, a datestamp,
        and setSpec(s) in case the item from which
        the record is disseminated belongs to set(s).
        the header can carry a deleted status indicating
        that the record is deleted.</documentation>
    </annotation>
    <sequence>
      <element name="identifier" type="oai:identifierType"/>
      <element name="datestamp" type="oai:UTCdatetimeType"/>
      <element name="setSpec" type="oai:setSpecType" minOccurs="0"
maxOccurs="unbounded"/>
    </sequence>
    <attribute name="status" type="oai:statusType" use="optional"/>
  </complexType>

  <simpleType name="identifierType">
    <restriction base="anyURI"/>
  </simpleType>

  <simpleType name="statusType">
    <restriction base="string">
      <enumeration value="deleted"/>
    </restriction>
  </simpleType>

  <complexType name="metadataType">
    <annotation>
      <documentation>Metadata must be expressed in XML that complies
        with another XML Schema (namespace=#other). Metadata must be
        explicitly qualified in the response.</documentation>
    </annotation>
    <sequence>
      <any namespace="##other" processContents="strict"/>
    </sequence>
```

```
  </complexType>

  <complexType name="aboutType">
    <annotation>
      <documentation>Data "about" the record must be expressed in XML
      that is compliant with an XML Schema defined by a
community.</documentation>
    </annotation>
    <sequence>
      <any namespace="##other" processContents="strict"/>
    </sequence>
  </complexType>

  <complexType name="resumptionTokenType">
    <annotation>
      <documentation>A resumptionToken may have 3 optional attributes
       and can be used in ListSets, ListIdentifiers, ListRecords
       responses.</documentation>
    </annotation>
    <simpleContent>
      <extension base="string">
        <attribute name="expirationDate" type="dateTime" use="optional"/>
        <attribute name="completeListSize" type="positiveInteger"
use="optional"/>
        <attribute name="cursor" type="nonNegativeInteger" use="optional"/>
      </extension>
    </simpleContent>
  </complexType>

  <complexType name="descriptionType">
    <annotation>
      <documentation>The descriptionType is used for the description
      element in Identify and for setDescription element in ListSets.
      Content must be compliant with an XML Schema defined by a
      community.</documentation>
    </annotation>
    <sequence>
      <any namespace="##other" processContents="strict"/>
```

```
    </sequence>
  </complexType>

  <simpleType name="UTCdatetimeType">
    <annotation>
      <documentation>Datestamps are to either day (type date)
      or to seconds granularity (type oai:UTCdateTimeZType)</documentation>
    </annotation>
    <union memberTypes="date oai:UTCdateTimeZType"/>
  </simpleType>

  <simpleType name="UTCdateTimeZType">
    <restriction base="dateTime">
      <pattern value=".*Z"/>
    </restriction>
  </simpleType>

  <!-- define types used for Identify verb only -->
  <!-- ======================================== -->

  <simpleType name="protocolVersionType">
    <restriction base="string">
      <enumeration value="2.0"/>
    </restriction>
  </simpleType>

  <simpleType name="emailType">
    <restriction base="string">
      <pattern value="\S+@(\S+\.)+\S+"/>
    </restriction>
  </simpleType>

  <simpleType name="deletedRecordType">
    <restriction base="string">
      <enumeration value="no"/>
      <enumeration value="persistent"/>
      <enumeration value="transient"/>
    </restriction>
```

```
</simpleType>

<simpleType name="granularityType">
   <restriction base="string">
      <enumeration value="YYYY-MM-DD"/>
      <enumeration value="YYYY-MM-DDThh:mm:ssZ"/>
   </restriction>
</simpleType>

<!-- define types used for ListMetadataFormats verb only -->
<!-- ================================================ -->

<complexType name="metadataFormatType">
   <sequence>
      <element name="metadataPrefix" type="oai:metadataPrefixType"/>
      <element name="schema" type="anyURI"/>
      <element name="metadataNamespace" type="anyURI"/>
   </sequence>
</complexType>

<simpleType name="metadataPrefixType">
   <restriction base="string">
      <pattern value="[A-Za-z0-9\-_\.!~\*'\(\)]+"/>
   </restriction>
</simpleType>

<!-- define types used for ListSets verb -->
<!-- ================================== -->

<complexType name="setType">
   <sequence>
      <element name="setSpec" type="oai:setSpecType"/>
      <element name="setName" type="string"/>
      <element name="setDescription" type="oai:descriptionType"
                  minOccurs="0" maxOccurs="unbounded"/>
   </sequence>
</complexType>
```

```
  <simpleType name="setSpecType">
    <restriction base="string">
      <pattern value="([A-Za-z0-9\-_\.!~\*'\(\)])+(:[A-Za-z0-9\-_\.!~\*'\(\)]+)*"/>
    </restriction>
  </simpleType>

</schema>
```

This Schema is available at http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd

4.3 UTCdatetime

Dates and times are uniformly encoded using ISO8601 and are expressed in UTC throughout the protocol. When time is included, the special UTC designator ("Z") must be used. UTC is implied for dates although no timezone designator is specified. For example, 1957-03-20T20:30:00Z is UTC 8:30:00 PM on March 20th 1957. UTCdatetime is used in both protocol requests and protocol replies, in the way described in the following sections.

4.3.1 UTCdatetime in Protocol Requests

Datestamps used as values of the optional arguments from and until in the ListIdentifiers and ListRecords requests are encoded using ISO8601 and are expressed in UTC. These arguments are used to specify datestamp-based selective harvesting. These arguments support the "Complete date" and the "Complete date plus hours, minutes and seconds" granularities defined in ISO8601. The legitimate formats are YYYY-MM-DD and YYYY-MM-DDThh:mm:ssZ. Both arguments must have the same granularity. All repositories must support YYYY-MM-DD. A repository that supports YYYY-MM-DDThh:mm:ssZ should indicate so in the Identify response. A request by a harvester with finer granularity than that supported by a repository must produce an error.

4.3.2 UTCdatetime in Protocol Responses

Datestamps appear in the headers of records that are returned in response to ListIdentifiers , GetRecord and ListRecords requests. These datestamps are encoded using ISO8601 and are expressed in UTC; they must be expressed in the finest granularity supported by the repository. The value of the datestamp must correspond to the rules for datestamp-based selective harvesting.
Each protocol response includes a responseDate element, which must be the time and date of the response in UTC. This is encoded using the "Complete

date plus hours, minutes, and seconds" variant of ISO8601. This format is YYYY-MM-DDThh:mm:ssZ.

A resumptionToken in a protocol reply may include an optional argument expirationDate, which is expressed in UTC. This is encoded using the "Complete date plus hours, minutes, and seconds" variant of ISO8601. This format is YYYY-MM-DDThh:mm:ssZ.

4.4 metadataPrefix and Metadata Schema

OAI-PMH supports the dissemination of records in multiple metadata formats from a repository. The ListMetadataFormats request returns the list of all metadata formatsavailable from a repository, each of which has the following properities:

(4) The metadataPrefix - a string to specify the metadata format in OAI-PMH requests issued to the repository. metadataPrefix consists of any valid URI unreserved characters. metadataPrefix arguments are used in ListRecords, ListIdentifiers, and GetRecord requests to retrieve records, or the headers of records that include metadata in the format specified by the metadataPrefix;

(5) The metadata schema URL - the URL of an XML schema to test validity of metadata expressed according to the format;

(6) The XML namespace URI that is a global identifier of the metadata format.

The metadata in each record returned by ListRecords and GetRecord must comply with the conventions of the XML namespace specification. This means that the root element of the metadata part must contain an xmlns attribute, the value of which is the XML namespace URI of the metadata format. The root element must also contain an xsi:schemaLocation attribute that has a value that includes the URL of the XML schema for validation of the metadata. This URL must match the URL of the metadata schema for the metadataPrefix included as an argument to the ListRecords or GetRecord request (the mapping from metadataPrefix to metadata schema is defined by the repository's response to the ListMetadataFormats request).

For purposes of interoperability, repositories must disseminate Dublin Core, without any qualification. Therefore, the protocol reserves the metadataPrefix `oai_dc', and the URL of a metadata schema for unqualified Dublin Core, which is http://www.openarchives.org/OAI/2.0/oai_dc.xsd. The corresponding XML namespace URI is http://www.openarchives.org/OAI/2.0/oai_dc/.
The metadataPrefix `all' is reserved for future use. Implementations should not use this metadataPrefix.

Communities should adopt guidelines for sharing of metadataPrefixes, metadata schema and XML namespace URI's of metadata formats. Such guidelines are outside of the scope of the OAI-PMH. The accompanying Implementation Guidelines document provides some sample XML Schema and instance documents for common metadata formats such as MARC and RFC 1807.

4.5 Flow Control

A number of OAI-PMH requests return a list of discrete entities: ListRecords returns a list of records, ListIdentifiers returns a list of headers, and ListSets returns a list of sets. Collectively these requests are called list requests. In some cases, these lists may be large and it may be practical to partition them among a series of requests and responses. This partitioning is accomplished as follows:

(3) A repository replies to a request with an incomplete list and a resumptionToken;

(4) In order to make the response a complete list, the harvester will need to issue one or more requests with resumptionTokens as arguments. The complete list then consists of the concatenation of the incomplete lists from the sequence of requests, known as a list request sequence.

Details of flow control and the resumptionToken are as follows:

(6) The only defined use of resumptionToken is as follows:

(g) a repository must include a resumptionToken element as part of each response that includes an incomplete list;

(h) in order to retrieve the next portion of the complete list, the next request must use the value of that resumptionToken element as the value of the resumptionToken argument of the request;

(i) the response containing the incomplete list that completes the list must include an empty resumptionToken element;

All other uses of resumptionToken by a harvester are illegal and must return an error.

(7) In all cases when a resumptionToken is issued, the incomplete list must consist of complete entities; e.g., all individual records returned in an incomplete record list from a ListRecords request must be intact.

(8) The format of the resumptionToken is not defined by the OAI-PMH and should be considered opaque by the harvester.

(9) The protocol does not define the semantics of incompleteness. Therefore, a harvester should not assume that the members in an incomplete list conform to some selection criteria (e.g., date ordering).

(10) Before including a resumptionToken in the URL of a subsequent request, a harvester must encode any special characters in it.

The following optional attributes may be included as part of the resumptionToken element along with the resumptionToken itself:

(4) expirationDate -- a UTCdatetime indicating when the resumptionToken ceases to be valid.

(5) completeListSize -- an integer indicating the cardinality of the complete list (i.e., the sum of the cardinalities of the incomplete lists). Because there may be changes in a repository during a list request sequence, as described under Idempotency of resumptionTokens, the value of completeListSize may be only an estimate of the actual cardinality of the complete list and may be revised during the list request sequence.

(6) cursor -- a count of the number of elements of the complete list thus far returned (i.e. cursor starts at 0).

The following example is a series of ListRecords requests where the complete list consists of 175 records and the repository only returns 100 records per response.

(5) The harvester issues a ListRecords request.

(6) The repository responds with an incomplete list of 100 records. The repository marks this list as incomplete by including in the response a non-empty resumptionToken element, with two attributes: a completeListSize of 175, and a cursor of 0.

(7) The harvester issues a subsequent ListRecords request that includes the resumptionToken that it received in the previous response.

(8) The repository responds with an incomplete list of 75 records. The repository marks this list as the final incomplete list by including in the response an empty resumptionToken element with two attributes: a completeListSize of 175, and a cursor of 100.

This flow control mechanism, in combination with HTTP transport layer facilities, provides some basic tools with which a repository can enforce an acceptable use policy for its harvesting interface. Communities implementing the OAI-PMH may need more extensive tools to enforce acceptable use policies for either the harvesting interface of their repositories or for the metadata harvested from those repositories. The enforcement of such additional policies is outside of the scope of the OAI-PMH.

4.5.1 Idempotency of resumptionTokens

Repositories that implement resumptionTokens must do so in a manner that allows harvesters to resume a sequence of requests for incomplete lists by re-issuing a list request with the most recent resumptionToken. The purpose of this is to allow harvesters to recover from network or other errors that would otherwise mean that the list request sequence would have to be started again. A re-issue of a list request with a resumptionToken occurs in two contexts:

(3) When there are no changes in the repository. There are no changes to the complete list returned by the list request sequence. In this case, the repository must return the same incomplete list when the most recent list request, i.e. the one with the most recent non-expired resumptionToken, is re-issued.

(4) When there are changes in the repository. There may be changes to the complete list returned by the list request sequence. These changes occur when the records disseminated in the list move in or out of the datestamp range of the request because of changes, modifications, or deletions in the repository. In this case, strict idempotency of the incomplete-list requests using resumptionToken values is not required. Instead, the incomplete list returned in response to a re-issued request must include all records with unchanged datestamps within the range of the initial list request. The incomplete list returned in response to a re-issued request may contain records with datestamps that either moved into or out of the range of the initial request. In cases where there are substantial changes to the repository, it may be appropriate for a repository to return a badResumptionToken error, signaling that the harvester should restart the list request sequence.

4.6 Error and Exception Conditions

In event of an error or exception condition, repositories must indicate OAI-PMH errors, distinguished from HTTP Status-Codes, by including one or more error elements in the response. While one error element is sufficient to indicate the presence of the error or exception condition, repositories should report all errors or exceptions that arise from processing the request. Each error element must have a code attribute that must be from the following table; each error element may also have a free text string value to provide information about the error that is useful to a human reader. These strings are not defined by the OAI-PMH.

Error Codes	Description	Applicable Verbs
badArgument	The request includes illegal arguments, is missing required arguments, includes a repeated argument, or values for arguments have an illegal syntax.	all verbs
badResumptionToken	The value of the resumptionToken argument is invalid or expired.	ListIdentifiers ListRecords ListSets
badVerb	Value of the verb argument is not a legal OAI-PMH verb, the verb argument is missing, or the verb argument is repeated.	N/A
cannotDisseminateFormat	The metadata format identified by the value given for the metadataPrefix argument is not supported by the item or by the repository.	GetRecord ListIdentifiers ListRecords
idDoesNotExist	The value of the identifier argument is unknown or illegal in this repository.	GetRecord ListMetadataFormats
noRecordsMatch	The combination of the values of the from, until, set and metadataPrefix arguments results in an empty list.	ListIdentifiers ListRecords
noMetadataFormats	There are no metadata formats available for the specified item.	ListMetadataFormats
noSetHierarchy	The repository does not support sets.	ListSets ListIdentifiers ListRecords

The following example demonstrates error handling in the case of an illegal verb argument. All request URLs shown from now on will be wrapped to make them more readable.

Request
http://arXiv.org/oai2?
verb=nastyVerb

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-05-01T09:18:29Z</responseDate>
  <request>http://arXiv.org/oai2</request>
  <error code="badVerb">Illegal OAI verb</error>
</OAI-PMH>
```

The following example demonstrates error handling in the case of a ListSets request to a repository that does not handle sets.

Request
http://arXiv.org/oai2?
verb=ListSets

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-05-01T09:18:29Z</responseDate>
  <request verb="ListSets">http://arXiv.org/oai2</request>
  <error code="noSetHierarchy">This repository does not
    support sets</error>
</OAI-PMH>
```

5.Protocol Requests and Responses

This section lists the requests, or verbs, defined in the OAI-PMH. The documentation for each request is organized as follows:

(6) A section title corresponding to the token used to specify the request as the required verb argument to an HTTP request.

(7) A brief summary of the meaning of the verb and notes on its usage.

(8) The list of additional arguments for the request. Arguments are of three types:

  (a) required, the argument must be included with the request (the verb argument is always required, as described in HTTP Request Format).

  (b) optional, the argument may be included with the request.

  (c) exclusive, the argument may be included with request, but must be the only argument (in addition to the verb argument).

(9) Error and exception conditions specific to the protocol request.

(10) One or more example requests and corresponding responses, with explanatory notes if appropriate.

An XML Schema defines the format of valid replies to all OAI-PMH requests.

5.1 GetRecord

Summary and Usage Notes

This verb is used to retrieve an individual metadata record from a repository. Required arguments specify the identifier of the item from which the record is requested and the format of the metadata that should be included in the record. Depending on the level at which a repository tracks deletions, a header with a "deleted" value for the status attribute may be returned, in case the metadata format specified by the metadataPrefix is no longer available from the repository or from the specified item.

Arguments

(3) identifier a required argument that specifies the unique identifier of the item in the repository from which the record must be disseminated.

(4) metadataPrefix a required argument that specifies the metadataPrefix of the format that should be included in the metadata part of the returned record . A record should only be returned if the format specified by the metadataPrefix can be disseminated from the item identified by the value of the identifier argument. The metadata formats supported by a repository and for a particular record can be retrieved using the ListMetadataFormats request.

Error and Exception Conditions

(4) badArgument - The request includes illegal arguments or is missing required arguments.

(5) cannotDisseminateFormat - The value of the metadataPrefix argument is not supported by the item identified by the value of the identifier argument.

(6) idDoesNotExist - The value of the identifier argument is unknown or illegal in this repository.

Examples

Request

Request a record in the Dublin Core metadata format [URL shown without encoding to be more readable].

http://arXiv.org/oai2?
verb=GetRecord&identifier=oai:arXiv.org:cs/0112017&metadataPrefix=oai_dc

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T08:55:46Z</responseDate>
  <request verb="GetRecord" identifier="oai:arXiv.org:cs/0112017"
```

```
                metadataPrefix="oai_dc">http://arXiv.org/oai2</request>
  <GetRecord>
   <record>
    <header>
      <identifier>oai:arXiv.org:cs/0112017</identifier>
      <datestamp>2001-12-14</datestamp>
      <setSpec>cs</setSpec>
      <setSpec>math</setSpec>
    </header>
    <metadata>
      <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>Using Structural Metadata to Localize Experience of
                  Digital Content</dc:title>
        <dc:creator>Dushay, Naomi</dc:creator>
        <dc:subject>Digital Libraries</dc:subject>
        <dc:description>With the increasing technical sophistication of
          both information consumers and providers, there is
          increasing demand for more meaningful experiences of digital
          information. We present a framework that separates digital
          object experience, or rendering, from digital object storage
          and manipulation, so the rendering can be tailored to
          particular communities of users.
        </dc:description>
        <dc:description>Comment: 23 pages including 2 appendices,
          8 figures</dc:description>
        <dc:date>2001-12-14</dc:date>
      </oai_dc:dc>
    </metadata>
   </record>
  </GetRecord>
</OAI-PMH>
```

Request

Request a record in the Dublin Core metadata format. The requested record, however, can not be returned because the identifier does not exist. Therefore, the response does not contain a record container. It does have an error element with a code attribute that has the value idDoesNotExist. [URL shown without encoding for better readability].

http://arXiv.org/oai2?
verb=GetRecord&identifier=oai:arXiv.org:quant-ph/02131001&metadataPrefix=oai_dc

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T08:55:46Z</responseDate>
  <request verb="GetRecord" identifier="oai:arXiv.org:quant-ph/0213001"
           metadataPrefix="oai_dc">http://arXiv.org/oai2</request>
  <error code="idDoesNotExist">No matching identifier in arXiv</error>
</OAI-PMH>
```

Request

Request a record in the oai_marc metadata format. However, the requested metadata format can not be disseminated for this identifier. Therefore, the response contains no record. It does contain an error element with a code attribute that has the value cannotDisseminateFormat. [URL shown without encoding for better readability].

http://arXiv.org/oai2?
verb=GetRecord&identifier=oai:arXiv.org:quant-ph/9901001&metadataPrefix=oai_marc

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T08:55:46Z</responseDate>
```

```
  <request verb="GetRecord" identifier="oai:arXiv.org:quant-ph/9901001"
           metadataPrefix="oai_marc">http://arXiv.org/oai1</request>
  <error code="cannotDisseminateFormat"/>
</OAI-PMH>
```

5.2 Identify

Summary and Usage Notes

This verb is used to retrieve information about a repository. Some of the information returned is required as part of the OAI-PMH. Repositories may also employ the Identify verb to return additional descriptive information.

Arguments

None

Error and Exception Conditions

(2) badArgument - The request includes illegal arguments.

Response Format

The response must include one instance of the following elements:

(7) repositoryName : a human readable name for the repository;

(8) baseURL : the base URL of the repository;

(9) protocolVersion : the version of the OAI-PMH supported by the repository;

(10) earliestDatestamp : a UTCdatetime that is the guaranteed lower limit of all datestamps recording changes, modifications, or deletions in the repository. A repository must not use datestamps lower than the one specified by the content of the earliestDatestamp element. earliestDatestamp must be expressed at the finest granularity supported by the repository.

(11) deletedRecord : the manner in which the repository supports the notion of deleted records. Legitimate values are no ; transient ; persistent with meanings defined in the section on deletion.

(12) granularity: the finest harvesting granularity supported by the repository. The legitimate values are YYYY-MM-DD and YYYY-MM-DDThh:mm:ssZ with meanings as defined in ISO8601.

The response must include one or more instances of the following element:

(2) adminEmail : the e-mail address of an administrator of the repository.

The response may include multiple instances of the following optional elements:

(3) compression : a compression encoding supported by the repository. The recommended values are those defined for the Content-Encoding header in Section 14.11 of RFC 2616 describing HTTP 1.1. A compression element should not be included for the identity encoding, which is implied.

(4) description : an extensible mechanism for communities to describe their repositories. For example, the description container could be used to include collection-level metadata in the response to the Identify request. Implementation Guidelines are available to give directions with this respect. Each description container must be accompanied by the URL of an XML schema describing the structure of the description container.

Examples

Request

http://memory.loc.gov/cgi-bin/oai?verb=Identify

Response

The below example of a response to the Identify request contains three description containers:

(4) The oai-identifier container complies to an XML Schema, which is available at http://www.openarchives.org/OAI/2.0/oai-identifier.xsd. This schema, provided in the accompanying Implementation Guidelines document, is used by repositories that choose to comply with a specific

format of unique identifiers for items. The format of that identifier is explained by means of comments in the oai-identifier.xsd XML Schema.

(5) The eprints container complies to an XML Schema, which is available at http://www.openarchives.org/OAI/1.1/eprints.xsd. This schema, provided in the accompanying Implementation Guidelines document, has been agreed upon by the OAI e-print community, and contains information specific to repositories in that community.

(6) The friends container complies to an XML Schema, which is available at http://www.openarchives.org/OAI/2.0/friends.xsd. This schema, provided in the accompanying Implementation Guidelines document, is used by repositories that want to point harvesters to other repositories, by listing their base URLs. Usage of the friends container is recommended; it may support harvesters in discovering the network-location of repositories.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T12:00:01Z</responseDate>
  <request verb="Identify">http://memory.loc.gov/cgi-bin/oai</request>
  <Identify>
    <repositoryName>Library of Congress Open Archive Initiative
                    Repository 1</repositoryName>
    <baseURL>http://memory.loc.gov/cgi-bin/oai</baseURL>
    <protocolVersion>2.0</protocolVersion>
    <adminEmail>somebody@loc.gov</adminEmail>
    <adminEmail>anybody@loc.gov</adminEmail>
    <earliestDatestamp>1990-02-01T12:00:00Z</earliestDatestamp>
    <deletedRecord>transient</deletedRecord>
    <granularity>YYYY-MM-DDThh:mm:ssZ</granularity>
    <compression>deflate</compression>
```

```
      <description>
        <oai-identifier
          xmlns="http://www.openarchives.org/OAI/2.0/oai-identifier"
          xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
          xsi:schemaLocation=
                "http://www.openarchives.org/OAI/2.0/oai-identifier
          http://www.openarchives.org/OAI/2.0/oai-identifier.xsd">
          <scheme>oai</scheme>
          <repositoryIdentifier>lcoa1.loc.gov</repositoryIdentifier>
          <delimiter>:</delimiter>

<sampleIdentifier>oai:lcoa1.loc.gov:loc.music/musdi.002</sampleIdentifier>
        </oai-identifier>
      </description>
      <description>
        <eprints
           xmlns="http://www.openarchives.org/OAI/1.1/eprints"
           xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
           xsi:schemaLocation="http://www.openarchives.org/OAI/1.1/eprints
           http://www.openarchives.org/OAI/1.1/eprints.xsd">
          <content>

<URL>http://memory.loc.gov/ammem/oamh/lcoa1_content.html</URL>
            <text>Selected collections from American Memory at the Library
                  of Congress</text>
          </content>
          <metadataPolicy/>
          <dataPolicy/>
        </eprints>
      </description>
      <description>
        <friends
           xmlns="http://www.openarchives.org/OAI/2.0/friends/"
           xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
           xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/friends/
          http://www.openarchives.org/OAI/2.0/friends.xsd">
         <baseURL>http://oai.east.org/foo/</baseURL>
         <baseURL>http://oai.hq.org/bar/</baseURL>
```

```
        <baseURL>http://oai.south.org/repo.cgi</baseURL>
      </friends>
    </description>
  </Identify>
</OAI-PMH>
```

5.3 ListIdentifiers

Summary and Usage Notes

This verb is an abbreviated form of ListRecords, retrieving only headers rather than records. Optional arguments permit selective harvesting of headers based on set membership and/or datestamp. Depending on the repository's support for deletions, a returned header may have a status attribute of "deleted" if a record matching the arguments specified in the request has been deleted.

Arguments

(6) from an optional argument with a UTCdatetime value, which specifies a lower bound for datestamp-based selective harvesting.

(7) until an optional argument with a UTCdatetime value, which specifies a upper bound for datestamp-based selective harvesting.

(8) metadataPrefix a required argument, which specifies that headers should be returned only if the metadata format matching the supplied metadataPrefix is available or, depending on the repository's support for deletions, has been deleted. The metadata formats supported by a repository and for a particular item can be retrieved using the ListMetadataFormats request.

(9) set an optional argument with a setSpec value , which specifies set criteria for selective harvesting.

(10) resumptionToken an exclusive argument with a value that is the flow control token returned by a previous ListIdentifiers request that issued an incomplete list.

Error and Exception Conditions

(6) badArgument - The request includes illegal arguments or is missing required arguments.

(7) badResumptionToken - The value of the resumptionToken argument is invalid or expired.

(8) cannotDisseminateFormat - The value of the metadataPrefix argument is not supported by the repository.

(9) noRecordsMatch- The combination of the values of the from, until, and set arguments results in an empty list.

(10) noSetHierarchy - The repository does not support sets.

Examples

Request

List the headers of records in the oldarXiv metadata format that are added, modified or deleted since January 15, 1998 in the set physics:hep. [URL shown without encoding for better readability].

http://an.oa.org/OAI-script?verb=ListIdentifiers&from=1998-01-15&metadataPrefix=oldArXiv&set=physics:hep

Response

A list of four headers is returned. One header has a deleted status, indicating that a record in the metadata format specified by the metadataPrefix is no longer available. In addition, a resumptionToken (non-empty, value xxx45abttyz) has been returned, indicating that the list of headers is incomplete and that one or more subsequent requests will need to be issued to retrieve a complete list. In the example, the resumptionToken comes with all of the 3 optional attributes: expirationDate indicates that the resumptionToken will become unusable after 11:20 PM UTC on June 1st 2002; completeListSize indicates that the complete list consists of 6 identifiers; the zero-value for cursor indicates that no headers have been returned previous to this reply.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-01T19:20:30Z</responseDate>
  <request verb="ListIdentifiers" from="1998-01-15"
              metadataPrefix="oldarXiv"
              set="physics:hep">http://an.oa.org/OAI-script</request>
  <ListIdentifiers>
   <header>
    <identifier>oai:arXiv.org:hep-th/9801001</identifier>
    <datestamp>1999-02-23</datestamp>
    <setSpec>physic:hep</setSpec>
   </header>
   <header>
    <identifier>oai:arXiv.org:hep-th/9801002</identifier>
    <datestamp>1999-03-20</datestamp>
    <setSpec>physic:hep</setSpec>
    <setSpec>physic:exp</setSpec>
   </header>
   <header>
    <identifier>oai:arXiv.org:hep-th/9801005</identifier>
    <datestamp>2000-01-18</datestamp>
    <setSpec>physic:hep</setSpec>
   </header>
   <header status="deleted">
    <identifier>oai:arXiv.org:hep-th/9801010</identifier>
    <datestamp>1999-02-23</datestamp>
    <setSpec>physic:hep</setSpec>
    <setSpec>math</setSpec>
   </header>
   <resumptionToken expirationDate="2002-06-01T23:20:00Z"
       completeListSize="6"
       cursor="0">xxx45abttyz</resumptionToken>
 </ListIdentifiers>
</OAI-PMH>
```

Request

Issue a subsequent request to the one issued above. The single resumptionToken argument has the value returned in the previous response. [URL shown without encoding for better readability].

http://an.oa.org/OAI-script?
verb=ListIdentifiers&resumptionToken=xxx45abttyz

Response

Two more headers are returned. The resumptionToken element at the end of the list has no value, indicating that the list is now complete. The value of the completeListSize attribute remains 6, while the value of the cursor attribute has changed to 4, indicating that a previous reply has (or previous replies have) already delivered 4 identifiers.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-01T19:30:00Z</responseDate>
  <request verb="ListIdentifiers"
           resumptionToken="xxx45abttyz">http://an.oa.org/OAI-script</request>
  <ListIdentifiers>
    <header>
      <identifier>oai:arXiv.org:hep-th/9801020</identifier>
      <datestamp>1999-02-23</datestamp>
      <setSpec>physic:hep</setSpec>
    </header>
    <header>
    <identifier>oai:arXiv.org:hep-th/9801060</identifier>
      <datestamp>1999-02-23</datestamp>
      <setSpec>physic:hep</setSpec>
    </header>
    <resumptionToken completeListSize="6" cursor="4"/>
  </ListIdentifiers>
</OAI-PMH>
```

Request

List the headers of olac-formatted records, added or modified on January 1, 2001 in the set Perseus:collection:PersInfo. There are no matches for this request, hence, the response contains an error tag and does not contain any header elements [URL shown without encoding for better readability].

http://www.perseus.tufts.edu/cgi-bin/pdataprov?verb=ListIdentifiers&metadataPrefix=olac&from=2001-01-01&until=2001-01-01&set=Perseus:collection:PersInfo

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T14:27:19Z</responseDate>
  <request verb="ListIdentifiers" metadataPrefix="olac"
    from="2001-01-01" until="2001-01-01"
    set="Perseus:collection:PersInfo">
    http://www.perseus.tufts.edu/cgi-bin/pdataprov</request>
  <error code="noRecordsMatch"/>
</OAI-PMH>
```

5.4 ListMetadataFormats

Summary and Usage Notes

This verb is used to retrieve the metadata formats available from a repository. An optional argument restricts the request to the formats available for a specific item.

Arguments

identifier an optional argument that specifies the unique identifier of the item for which available metadata formats are being requested. If this argument is omitted, then the response includes all metadata formats supported by this repository. Note that the fact that a metadata format is supported by a repository does not mean that it can be disseminated from all items in the repository.

Error and Exception Conditions

(4) badArgument - The request includes illegal arguments or is missing required arguments.

(5) idDoesNotExist - The value of the identifier argument is unknown or illegal in this repository.

(6) noMetadataFormats - There are no metadata formats available for the specified item.

Examples

Request

List the metadata formats that can be disseminated from the repository http://www.perseus.tufts.edu/cgi-bin/pdataprov for the item with unique identifier oai:perseus.tufts.edu:Perseus:text:1999.02.0119 [URL shown without encoding for better readability].

http://www.perseus.tufts.edu/cgi-bin/pdataprov?verb=ListMetadataFormats&identifier=oai:perseus.tufts.edu:Perseus:text:1999.02.0119

Response

The response shows that 3 metadata formats are supported for the given identifier: oai_dc, olac and perseus. For each of the formats, the location of an XML Schema describing the format, as well as the XML Namespace URI is given.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-02-08T14:27:19Z</responseDate>
  <request verb="ListMetadataFormats"
    identifier="oai:perseus.tufts.edu:Perseus:text:1999.02.0119">
    http://www.perseus.tufts.edu/cgi-bin/pdataprov</request>
  <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>olac</metadataPrefix>
        <schema>http://www.language-archives.org/OLAC/olac-0.2.xsd</schema>
        <metadataNamespace>http://www.language-archives.org/OLAC/0.2/
          </metadataNamespace>
    </metadataFormat>
    <metadataFormat>
        <metadataPrefix>perseus</metadataPrefix>
        <schema>http://www.perseus.tufts.edu/persmeta.xsd</schema>
        <metadataNamespace>http://www.perseus.tufts.edu/persmeta.dtd
          </metadataNamespace>
    </metadataFormat>
  </ListMetadataFormats>
</OAI-PMH>
```

Request

List the metadata formats that can be disseminated from the repository http://memory.loc.gov/cgi-bin/oai.

http://memory.loc.gov/cgi-bin/oai?verb=ListMetadataFormats

Response

The response shows that the repository supports two metadata formats: oai_dc, and oai_marc. For each of the formats, the location of an XML Schema describing the format is given. The support of these formats at the repository-level does not imply support of each format for each item of the repository.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-08T15:19:13Z</responseDate>
  <request verb="ListMetadataFormats">
              http://memory.loc.gov/cgi-bin/oai</request>
  <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_marc</metadataPrefix>
    <schema>http://www.openarchives.org/OAI/1.1/oai_marc.xsd</schema>
    <metadataNamespace>http://www.openarchives.org/OAI/1.1/oai_marc
      </metadataNamespace>
   </metadataFormat>
  </ListMetadataFormats>
</OAI-PMH>
```

Request

List the metadata formats that can be disseminated for the unique identifier oai:lcoa1.loc.gov:loc.rbc/rbpe.00000111 in the repository http://memory.loc.gov/cgi-bin/oai. The identifier, however, does not exist and therefore, the response contains an error element and no metadataFormat container. [URL shown without encoding for better readability].

http://memory.loc.gov/cgi-bin/oai?
verb=ListMetadataFormats&identifier=oai:lcoa1.loc.gov:loc.rbc/rbpe.00000111

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-08T15:19:13Z</responseDate>
  <request verb="ListMetadataFormats"
            identifier="oai:lcoa1.loc.gov:loc.rbc/rbpe.00000111">
            http://memory.loc.gov/cgi-bin/oai</request>
```

```
  <error code="idDoesNotExist">oai:lcoa1.loc.gov:loc.rbc/rbpe.00000111 has the
        structure of a valid LOC identifier, but it maps to no known
        item</error>
</OAI-PMH>
```

5.5 ListRecords

Summary and Usage Notes

This verb is used to harvest records from a repository. Optional arguments permit selective harvesting of records based on set membership and/or datestamp. Depending on the repository's support for deletions, a returned header may have a status attribute of "deleted" if a record matching the arguments specified in the request has been deleted. No metadata will be present for records with deleted status.

Arguments

(6) from an optional argument with a UTCdatetime value, which specifies a lower bound for datestamp-based selective harvesting.

(7) until an optional argument with a UTCdatetime value, which specifies a upper bound for datestamp-based selective harvesting.

(8) set an optional argument with a setSpec value , which specifies set criteria for selective harvesting.

(9) resumptionToken an exclusive argument with a value that is the flow control token returned by a previous ListRecords request that issued an incomplete list.

(10) metadataPrefix a required argument (unless the exclusive argument resumptionToken is used) that specifies the metadataPrefix of the format that should be included in the metadata part of the returned records. Records should be included only for items from which the metadata format matching the metadataPrefix can be disseminated. The metadata formats supported by a repository and for a particular item can be retrieved using the ListMetadataFormats request.

Error and Exception Conditions

(6) badArgument - The request includes illegal arguments or is missing required arguments.

(7) badResumptionToken - The value of the resumptionToken argument is invalid or expired.

(8) cannotDisseminateFormat - The value of the metadataPrefix argument is not supported by the repository.

(9) noRecordsMatch - The combination of the values of the from, until, set and metadataPrefix arguments results in an empty list.

(10) noSetHierarchy - The repository does not support sets.

Examples

Request

List the records expressed in oai_rfc1807 metadata format, that have been added or modified since January 15, 1998 in the hep subset of the physics set [URL shown without encoding for better readability].

http://an.oa.org/OAI-script?verb=ListRecords&from=1998-01-15&set=physics:hep&metadataPrefix=oai_rfc1807

Response

Two records are returned:

(3) The first record is expressed in the oai_rfc1807 metadata. This record also has an about part, and the item from which it was disseminated belongs to two sets (physics:hep and math).

(4) The second has a header with a status="deleted" attribute (and therefore no metadata part).

Note: The reply only includes records for those items from which metadata in oai_rfc1807 can be disseminated. No records are returned for those items that fit

the from, until, and set arguments but from which the specified format can not be disseminated.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-01T19:20:30Z</responseDate>
 <request verb="ListRecords" from="1998-01-15"
               set="physics:hep"
               metadataPrefix="oai_rfc1807">
               http://an.oa.org/OAI-script</request>
 <ListRecords>
  <record>
     <header>
       <identifier>oai:arXiv.org:hep-th/9901001</identifier>
       <datestamp>1999-12-25</datestamp>
       <setSpec>physics:hep</setSpec>
       <setSpec>math</setSpec>
     </header>
     <metadata>
      <rfc1807 xmlns=
           "http://info.internet.isi.edu:80/in-notes/rfc/files/rfc1807.txt"
       xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
       xsi:schemaLocation=
        "http://info.internet.isi.edu:80/in-notes/rfc/files/rfc1807.txt
         http://www.openarchives.org/OAI/1.1/rfc1807.xsd">
         <bib-version>v2</bib-version>
         <id>hep-th/9901001</id>
         <entry>January 1, 1999</entry>
         <title>Investigations of Radioactivity</title>
         <author>Ernest Rutherford</author>
         <date>March 30, 1999</date>
      </rfc1807>
     </metadata>
     <about>
       <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:publisher>Los Alamos arXiv</dc:publisher>
          <dc:rights>Metadata may be used without restrictions as long as
              the oai identifier remains attached to it.</dc:rights>
        </oai_dc:dc>
      </about>
    </record>
    <record>
      <header status="deleted">
        <identifier>oai:arXiv.org:hep-th/9901007</identifier>
        <datestamp>1999-12-21</datestamp>
      </header>
    </record>
  </ListRecords>
</OAI-PMH>
```

Request

Request records in the oai_dc metadata format, modified or added between 2:15pm and 2:20pm UTC on May 1st 2002. [URL shown without encoding for better readability].

http://www.perseus.tufts.edu/cgi-b:in/pdataprov?verb=ListRecords&from=2002-05-01T14:15:00Z&until=2002-05-01T14:20:00Z&metadataPrefix=oai_dc

Response

Two records are returned. The second one has a provenance container in its about element, giving an insight in its chain of provenance.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-01T19:20:30Z</responseDate>
  <request verb="ListRecords" from="2002-05-01T14:15:00Z"
```

```
            until="2002-05-01T14:20:00Z" metadataPrefix="oai_dc">
            http://www.perseus.tufts.edu/cgi-bin/pdataprov</request>
<ListRecords>
 <record>
   <header>
     <identifier>oai:perseus:Perseus:text:1999.02.0084</identifier>
     <datestamp>2002-05-01T14:16:12Z</datestamp>
   </header>
   <metadata>
     <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>Opera Minora</dc:title>
       <dc:creator>Cornelius Tacitus</dc:creator>
       <dc:type>text</dc:type>
       <dc:source>Opera Minora. Cornelius Tacitus. Henry Furneaux.
        Clarendon Press. Oxford. 1900.</dc:source>
       <dc:language>latin</dc:language>
       <dc:identifier>http://www.perseus.tufts.edu/cgi-bin/ptext?
         doc=Perseus:text:1999.02.0084</dc:identifier>
     </oai_dc:dc>
   </metadata>
 </record>
 <record>
   <header>
     <identifier>oai:perseus:Perseus:text:1999.02.0083</identifier>
     <datestamp>2002-05-01T14:20:55Z</datestamp>
   </header>
   <metadata>
     <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>Germany and its Tribes</dc:title>
            <dc:creator>Tacitus</dc:creator>
            <dc:type>text</dc:type>
            <dc:source>Complete Works of Tacitus. Tacitus. Alfred John Church.
              William Jackson Brodribb. Lisa Cerrato. edited for Perseus.
              New York: Random House, Inc. Random House, Inc. reprinted 1942.
                </dc:source>
            <dc:language>english</dc:language>
            <dc:identifier>http://www.perseus.tufts.edu/cgi-bin/ptext?
              doc=Perseus:text:1999.02.0083</dc:identifier>
          </oai_dc:dc>
          </metadata>
        <about>
          <provenance
            xmlns="http://www.openarchives.org/OAI/2.0/provenance"
            xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
            xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/provenance
            http://www.openarchives.org/OAI/2.0/provenance.xsd">
            <originDescription harvestDate="2002-01-01T11:10:01Z"
altered="true">
              <baseURL>http://some.oa.org</baseURL>
              <identifier>oai:r2.org:klik001</identifier>
              <datestamp>2001-01-01</datestamp>

<metadataNamespace>http://www.openarchives.org/OAI/2.0/oai_dc/</metadata
Namespace>
              </originDescription>
          </provenance>
        </about>
    </record>
  </ListRecords>
</OAI-PMH>
```

Request

Request records in the the oai_marc metadata format, modified or added between 2:00am and 3:00am UTC on June 1st 2002. The specified granularity is not supported by the repository and therefore, an error with code attribute of badArgument is returned. [URL shown without encoding for better readability].

http://memory.loc.gov/cgi-bin/oai?verb=ListRecords&from=2002-06-01T02:00:00Z&until=2002-06-01T03:00:00Z&metadataPrefix=oai_marc

Response

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-06-01T19:20:30Z</responseDate>
  <request verb="ListRecords" from="2002-06-01T02:00:00Z"
           until="2002-06-01T03:020:00Z"
           metadataPrefix="oai_marc">
           http://memory.loc.gov/cgi-bin/oai</request>
  <error code="badArgument"/>
</OAI-PMH>
```

5.6 ListSets

Summary and Usage Notes

This verb is used to retrieve the set structure of a repository, useful for selective harvesting.

Arguments

resumptionToken an exclusive argument with a value that is the flow control token returned by a previous ListSets request that issued an incomplete list.

Error and Exception Conditions

(4) badArgument - The request includes illegal arguments or is missing required arguments.

(5) badResumptionToken - The value of the resumptionToken argument is invalid or expired.

(6) noSetHierarchy - The repository does not support sets.

Examples

Request

http://an.oa.org/OAI-script?verb=ListSets

Response

The following response indicates a set hierarchy with two top level sets with respective setSpec music and video. The music set has two subsets, with setSpec music:(muzak) and music:(elec). The subsets identified by setSpec music:(elec), has a setDescription element which holds a Dublin Core container, used to describe its contents.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2002-08-11T07:21:33Z</responseDate>
 <request verb="ListSets">http://an.oa.org/OAI-script</request>
 <ListSets>
  <set>
    <setSpec>music</setSpec>
    <setName>Music collection</setName>
  </set>
  <set>
    <setSpec>music:(muzak)</setSpec>
    <setName>Muzak collection</setName>
  </set>
  <set>
    <setSpec>music:(elec)</setSpec>
    <setName>Electronic Music Collection</setName>
    <setDescription>
      <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:description>This set contains metadata describing
              electronic music recordings made during the 1950ies
              </dc:description>
      </oai_dc:dc>
```

```
    </setDescription>
  </set>
  <set>
    <setSpec>video</setSpec>
    <setName>Video Collection</setName>
  </set>
 </ListSets>
</OAI-PMH>
```

Request

http://purl.org/alcme/etdcat/servlet/OAIHandler?verb=ListSets

Response

The response shows that the repository does not have a set hierarchy.

```
<?xml version="1.0" encoding="UTF-8"?>
<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>2001-06-01T19:20:30Z</responseDate>
  <request verb="ListSets">
           http://purl.org/alcme/etdcat/servlet/OAIHandler</request>
  <error code="noSetHierarchy">This repository does not
         support sets</error>
</OAI-PMH>
```

6.Dublin Core

The following table shows the XML Schema for Dublin Core without qualification, which is associated with the reserved metadataPrefix oai_dc in the OAI-PMH. All examples in this document that include Dublin Core metadata, validate against this XML schema. Schema for other metadata formats are provided in the accompanying Implementation Guidelines document.

A XML schema for validating Unqualified Dublin Core metadata associated with the reserved oai_dc metadataPrefix

```
<schema targetNamespace="http://www.openarchives.org/OAI/2.0/oai_dc/"
        xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
        xmlns:dc="http://purl.org/dc/elements/1.1/"
        xmlns="http://www.w3.org/2001/XMLSchema"
```

```
            elementFormDefault="qualified" attributeFormDefault="unqualified">

<annotation>
   <documentation>
         XML Schema 2002-03-18 by Pete Johnston.
         Adjusted for usage in the OAI-PMH.
         Schema imports the Dublin Core elements from the DCMI schema for
unqualified Dublin Core.
         2002-12-19 updated to use simpledc20021212.xsd (instead of
simpledc20020312.xsd)
   </documentation>
</annotation>

<import namespace="http://purl.org/dc/elements/1.1/"

schemaLocation="http://dublincore.org/schemas/xmls/simpledc20021212.xsd"/>

<element name="dc" type="oai_dc:oai_dcType"/>

<complexType name="oai_dcType">
   <choice minOccurs="0" maxOccurs="unbounded">
      <element ref="dc:title"/>
      <element ref="dc:creator"/>
      <element ref="dc:subject"/>
      <element ref="dc:description"/>
      <element ref="dc:publisher"/>
      <element ref="dc:contributor"/>
      <element ref="dc:date"/>
      <element ref="dc:type"/>
      <element ref="dc:format"/>
      <element ref="dc:identifier"/>
      <element ref="dc:source"/>
      <element ref="dc:language"/>
      <element ref="dc:relation"/>
      <element ref="dc:coverage"/>
      <element ref="dc:rights"/>
   </choice>
</complexType>
```

```
</schema>
```

This Schema is available at http://www.openarchives.org/OAI/2.0/oai_dc.xsd

Examples

```
<?xml version="1.0" encoding="UTF-8"?>
<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 xml:lang="en">The Cornell Law Quarterly</dc:title>
   <dc:date>1915-1916</dc:date>
   <dc:identifier>http://heinonline.org/HeinOnline/show.pl?
         handle=hein.journals/clqv1%26id=1%26size=4</dc:identifier>
   <dc:rights>Available by Subscription.
                  See http://www.wshein.com</dc:rights>
</oai_dc:dc>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<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 xml:lang="en">Grassmann's space analysis</dc:title>
   <dc:creator>Hyde, E. W. (Edward Wyllys)</dc:creator>
   <dc:subject>LCSH:Ausdehnungslehre; LCCN QA205.H99</dc:subject>
   <dc:publisher>J. Wiley &amp;amp; Sons</dc:publisher>
   <dc:date>Created: 1906; Available: 1991</dc:date>
   <dc:type>text</dc:type>
   <dc:identifier>http://resolver.library.cornell.edu/math/1796949
        </dc:identifier>
   <dc:language>english</dc:language>
   <dc:rights xml:lang="en">Public Domain</dc:rights>
</oai_dc:dc>
```

參考資料

Open Archives Initiative Protocol for Metadata Harvesting （OAI-PMH）2.0, 2004.
< http://www.openarchives.org/OAI/openarchivesprotocol.html >

爭議事項

無

英中名詞對照表

英中名詞對照表

-A-

-B-

-C-

-D-

Datestamp *日戳*

-E-

-F-

-G-

-H-

Harvester *擷取器*

Header 標頭

-I-

-J-

-K-

-L-

-M-

-N-

-O-

-P-

-Q-

-R-

Record 紀錄

Request 請求

Repository *儲存庫*

-S-

Service Provider 服務提供者

Status-code *狀態碼*

-T-

-U-

Unique identifier 唯一識別符

-V-

-W-

-X-

-Y-

-Z-

中英名詞對照表（English-Chinese Index）

日戳	*Datestamp*
服務提供者	Service Provider
狀態碼	*Status-code*
紀錄	Record
唯一識別符	Unique identifier
採集者	*Harvester*
標頭	Header
請求	Request
儲存庫	*Repository*