

**Deliverable 1.1.**

**Application profile for  
OpenURL Context Objects and SUSHI**

<i>Version History</i>		
7 July 2009	Peter Verhaar	First draft
9 September 2009	Peter Verhaar	Specification of the usage of OpenURL Context Objects completed.
25 October 2009	Peter Verhaar	Specification of SUSHI and SOAP added.
7 January 2010	Peter Verhaar	Section 6 on the use of SUSHI has been expanded.
8 January 2010	Peter Verhaar	Several changes have been made in section 6.

## 1. Introduction

As a result of the growing dissatisfaction with traditional metrics for the assessment of the impact of academic output, a large number of research projects have begun to explore alternative methods. Projects such as PEER, PIRUS and OA-Statistik specifically aim on the analysis of download statistics. The assumption is that these usage data can shed light on the degree in which open access publishing can contribute to an improved visibility of scientific publications. Evidently, usage statistics do not necessary provide information on the scientific impact or the quality of publications, since a download of a document does not automatically imply that the document is also read. To investigate impact, analyses of citation behaviour may perhaps be more reliable.<sup>1</sup> Citation analyses have the potential disadvantage that they focus on developments which take place over a fairly long period of time, whereas usage statistics can in theory be studied immediately after the publication is made available.

The SURE Workgroup (Statistics on the Usage of Repositories), which is largely funded by SURFFoundation, aims to build a suite of software applications that can collect usage data from individual Dutch repositories and that can send this information to a central database. The agent that collects the usage data will be referred to as the Log Aggregator. After the data have been analysed centrally, a number of services can be built on the basis of this information.

The basic components of this infrastructure are visualised in figure 1.

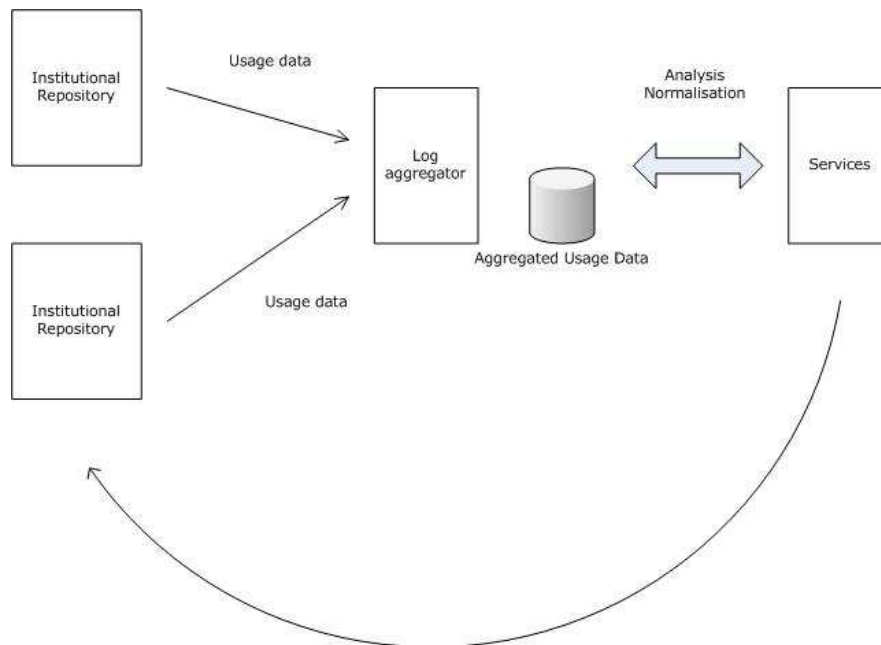


Figure 1.

<sup>1</sup> Recent studies such as MESUR have also demonstrated a clear correlation between citation impact and download frequency.

One of challenges in this report is to find an effective way of transferring data on usage events from local repositories to a central log aggregator. Various standards and protocols have recently been developed in this context, but so far the practical experiences with these have been rather limited. The SURE statistics workgroup will set up a pilot to demonstrate that a national infrastructure for the collection of usage data is feasible. Importantly, the project works with the assumption that data must be aggregated centrally to ensure that reliable and meaningful data can be produced.

This document provides the technical specifications for the software applications that will be developed in the course of this project. Firstly, an overview will be given of the data items that need to be recorded. Secondly, this document will provide specifications for the format in which information about usage events need to be recorded. Finally, this report will describe the protocol that will be used to transfer the data from local repositories to the central facility.

## 2. Aspects to be recorded

A usage event takes place when a **user** downloads a **document** which is managed in an **institutional repository**, or when a user view the **metadata** that is associated with this document, see also figure 1.

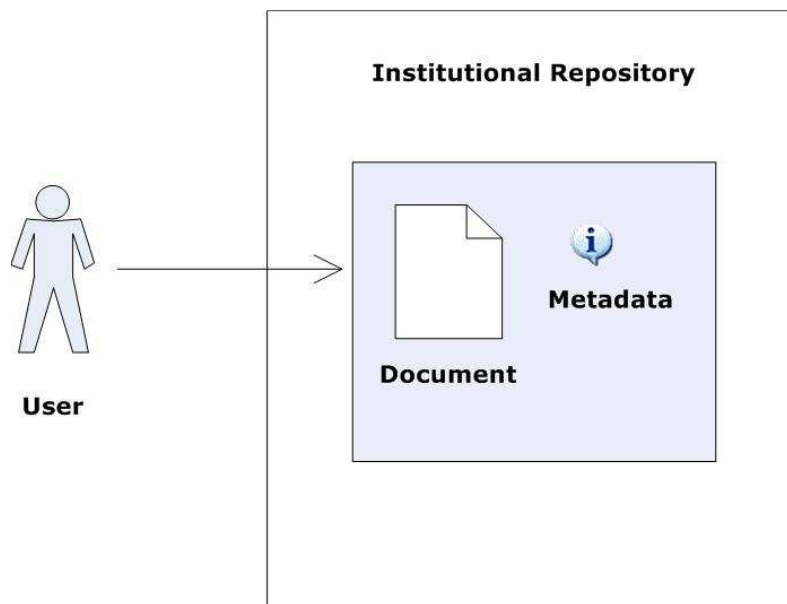


Figure 1.

The user may have arrived at this document through the mediation of a **referrer**. This is typically a search engine, such as Google or Yahoo.

The aim of the SURFShare SURE project is to describe the highlighted entities that are involved in the usage event as fully as possible. The

event took place in a certain country, at a certain time on a certain date, and from a machine which a certain IP-address. In the *JISC Usage Statistics Review*,<sup>2</sup> which was released in September 2008, it is stipulated that the following items of information must minimally be captured:

- Who (Identification of user/session)
- What (Item identification)
- Type of request performed (e.g. full-text, front-page, including failed/partially fulfilled requests)
- When (Date and time)
- Usage event ID

In line with these JISC recommendations, it has been decided in the SURF SURE project to provide the following data elements:

<i>Data element</i>	<i>Description</i>
IP-address of requestor	Providing the full IP-address is not permitted by international copyright laws. For this reason, the IP-address needs to be encrypted.
C-class Subnet	When the IP-address is encrypted, this will have the disadvantage that information on the geographic location, for instance, can no longer be derived. For this reason, the Class subnet, being the first three most significant bits from the IP-address must also be provided.
Geographic location	The country from which the request originated is also provided explicitly.
Persistent identifier of requested document	See also section 4.
URL of document	See also section 4.
Date and time of the request	
Request Type	It must be clear if a document was downloaded or if its metadata was viewed.
Host name	The institution that is responsible for the repository in which the requested document is stored.
Usage event ID	Unique number for a specific usage event.
Referrer URL	The URL which was received from the referring entity, if it was used. This URL often contains the search terms that were typed in by the user.
Referrer name	A classification of the referrer, based

<sup>2</sup> <http://www.jisc.ac.uk/whatwedo/programmes/digitalrepositories2007/usagestatisticsreview.aspx>

on a short list of search engines.

### 3. Source of information

All Dutch repositories make use of Apache server software for the maintenance of their repository websites. Each usage event that takes place generates an entry in the Apache logging files. These logging files will be used in the SURF SURE project as the primary source of information for usage statistics.

Figure 2 below contains a typical entry from an Apache log file.

```
13/Jul/2009:09:14:16 +0200] 193.173.52.133 TLSv1
RC4-MD5 openaccess.leidenuniv.nl "GET
/bitstream/1887/3674/1/360_138.pdf HTTP/1.1" 722168
-
"http://www.google.nl/search?hl=nl&q=beleidsregels+a
rtikel+4%3A84&meta=" "Mozilla/4.0 (compatible; MSIE
7.0; Windows NT 5.1; .NET CLR 1.1.4322; .NET CLR
2.0.50727)" 200 20352 15
```

Figure 2.

From figure 2, it can be seen that the aspects which were mentioned at the end of section 2 can normally be derived relatively easy from the log file.

### 4. Formatting guidelines

Different institutions may also have configured the logging facilities of their servers in different ways. Because of this, there may occasionally be variations in way in which, for instance, the time and the date are formatted. To avoid problems arising from variations in formatting, this section provides guidelines on the format in which the mandatory and optional data element need to be provided. A general principle, however, is that information should be passed along to the central database as 'pure' as possible, so that analysis can take place centrally and consistently.

#### IP-address of requestor

Description	The IP-address must be hashed using MD5 encryption.
Usage	Mandatory
Format	32 hexadecimal numbers.

Example `c06f0464f37249a0a9f848d4b823ef2a`

### **C-class Subnet**

Description The first three bytes of an IP-address, which are used to designate the network ID. It is similar to the IP-address, with the crucial difference that the final (most significant) byte, which designates the HOST ID is replaced with a '0'.

Usage Mandatory

Format Three decimal numbers separated by a dot, followed by a dot and a '0'.

Example `118.94.150.0`

### **Geographic location**

Description The country from which the request was

Usage Mandatory

Format A two-letter code in lower case, following the ISO 3166-1-alpha-2 code.<sup>3</sup>

Example `ne`

### **URL of document**

Description The document identifier provides a globally unique identification of the resource that is requested.

Usage Mandatory

Format The identifier must be given in the form of a URL. At the level of the service aggregator, this URL must be connected to the object's URN, so that the associated metadata can be obtained. The URN can be found by retrieving the URL of the object file in the DIDL file that is maintained by KNAW.

Example `/bitstream/1887/12100/1/Thesis.pdf`

### **OAI-PMH identifier**

Description The URN of the publication. This identifier must be

---

<sup>3</sup> [http://www.iso.org/iso/english\\_country\\_names\\_and\\_code\\_elements](http://www.iso.org/iso/english_country_names_and_code_elements)

the same as the identifier in the DIDL representation of the publication.

Usage	Optional
Format	URN
Example	<code>http://hdl.handle.net/1887/12100</code>

### **Request Type**

Description	The request type specifies whether an object file was downloaded or whether a metadata record was viewed.
Inclusion	Mandatory
Format	Two values are allowed: "objectFile" or "metadataView"
Example	<code>ObjectFile</code>

### **Host name**

Description	An identification of the repository that has recorded the usage event
Usage	Mandatory
Format	URL of the receiving institution's repository
Example	<code>www.openaccess.leidenuniv.nl</code>

### **Request Time**

Description	The exact time on which the usage event took place.
Usage	Mandatory
Format	The request time must be given in a format that that conforms to ISO8601. The YYYY-MM-DDTHH:MM:SSZ representation must be used. Note that this format may differ from the format that is provided in the Apache log file.
Example	<code>2009-07-29T08:15:46+01:00</code>

### **Referrer URL**

Description	The environment which has directed the user to the
-------------	--

requested object. This usually refers to the search engine which the client has used to find the object.

Usage	Optional
Format	URL
Example	<code>http://www.google.nl/search?hl=nl&amp;q=beleidsregels+artikel+4%3A84&amp;meta="</code>

### **Referrer Name**

Description	The referrer must be categorised on the basis of a limited list of known referrers.
Usage	Optional
Format	The following values are allowed: "google", "google scholar", "bing", "yahoo", "altavista"
Example	<code>google</code>

### **Usage Event ID**

Description	Unique identification of the usage event. This identification will be generated, and it can not be derived from the Apache log file.
Usage	Mandatory
Format	The identifier will be formed by combining the item, the date and a three-letter code for the institution. Next, this identifier will be encrypted using MD5, so that the identifier becomes a 32-digit number (hexadecimal notation).
Example	<code>b06c0444f37249a0a8f748d3b823ef2a</code>

## **5. Normalisation**

- The SURE Statistics project will attempt to restrict its focus to requests which have consciously been initiated by human users. For this reason, automated visits by internet robots must be filtered from the data as much as possible. The Log Aggregator must maintain a file which list the names of internet robots that individual repositories must use during the filtering of their results. The name of this file must indicate its version. The name of the first file that will be published will be robots-v1.xml. Repositories can use the version indication in the filename to check if they are working with the most recent list of internet robots.



- If a single user clicks repeatedly on the same document within the same 24 hours, this should be counted as a single request.
- One single publication may be split into a set different files. The impact of such variations in the organisation of complex objects must be nullified. The consultation of a part should count towards the statistic of the whole. It should make no difference if a publication consists of one pdf-files or of multiple pdf-files.

## 6. Data format

In compliance with the JISC Usage Statistics Review, individual usage events need to be serialized in XML using the syntax that is specified in the OpenURL Context Objects schema.<sup>4</sup> This section will describe a recommended practice for the use of this schema.

The root element of the XML-document must be <context-objects>. It must contain a reference to the official schema and declare two namespaces: `xmlns:ctx="info:ofi/fmt:xml:xsd:ctx"` and `xmlns:dcterms=http://dublincore.org/documents/2008/01/14/dcmi-terms/`.

Each usage event must be described in a separate <context-object> element, which must appear as a direct child of <context-objects>. Two attributes must be used:

- The time and date on which the usage event took place must be recorded in the *timestamp* attribute.
- The identification of the usage event must be captured in an attribute with the name *identifier*.

Within <contextobject>, a number of elements can be used which describe the context of the usage event. The names of these elements are as follows:

- <referent>: the object that was downloaded or viewed
- <requester> refers to the agent that has initiated the usage event
- <serviceType>: the type of service that was requested
- <referrer>: the system that has forwarded the reader to the downloaded object
- <resolver>: the institution that provides access to the requested item and which has received the usage event.

Information about these contextual entities can be given in four different ways. Firstly, they can be characterised using an <identifier>. Secondly, metadata can be included literally by wrapping these into the file using a <metadata-by-val> element. Thirdly, a reference to metadata stored elsewhere can be included by using <metadata-by-ref>. A fourth method is the use of the element <private-data>. In the SURE Statistics project,

---

<sup>4</sup> The XML Schema for XML Context Objects can be accessed at <http://www.openurl.info/registry/docs/info:ofi/fmt:xml:xsd:ctx>

only the first two methods shall be used. Listing 1 is an example of a full OpenURL Context Object document.

```
<?xml version="1.0" encoding="UTF-8"?>
<ctx:context-objects xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:dcterms="http://dublincore.org/documents/2008/01/14/dcmi-terms/"
  xsi:schemaLocation="info:ofi/fmt:xml:xsd:ctx
  http://www.openurl.info/registry/docs/info:ofi/fmt:xml:xsd:ctx"
  xmlns:ctx="info:ofi/fmt:xml:xsd:ctx">
  <ctx:context-object identifier="c06f0464f37249a0a9f848d4b823ef2a"
    timestamp="2009-07-13T09:14:16+02:00">

    <ctx:referent>
      <ctx:identifier>https://openaccess.leidenuniv.nl/dspace/
        handle/1887/584</ctx:identifier>
      <ctx:identifier>http://hdl.handle.net/
        handle/1887/584</ctx:identifier>
    </ctx:referent>

    <ctx:referring-entity>
      <ctx:identifier>http://www.google.nl/search?
        hl=nl&q=beleidsregels+artikel+4%3A84&meta="
      </ctx:identifier>
      <ctx:identifier>google</ctx:identifier>
    </ctx:referring-entity>

    <ctx:requester>
      <ctx:identifier>c06f0464f37249a0a9f84
        8d4b823ef2a</ctx:identifier>
      <ctx:identifier>118.94.150.0</ctx:identifier>
      <ctx:metadata-by-val>
        <ctx:format>
          http://dublincore.org/documents/dcmi-terms/
        </ctx:format>
        <ctx:metadata>
          <dcterms:spatial>nl</dcterms:spatial>
        </ctx:metadata>
      </ctx:metadata-by-val>
    </ctx:requester>

    <ctx:service-type>
      <ctx:metadata-by-val>
        <ctx:format>
          http://dublincore.org/documents/dcmi-terms/
        </ctx:format>
        <ctx:metadata>
          <dcterms:type>objectFile</dcterms:type>
        </ctx:metadata>
      </ctx:metadata-by-val>
    </ctx:service-type>

    <ctx:resolver>
      <ctx:identifier>openaccess.leidenuniv.nl</ctx:identifier>
    </ctx:resolver>

  </ctx:context-object>
</ctx:context-objects>
```

*Listing 1.*

- Under <referent>, the two identifiers for the requested document must both be given in separate <identifier> elements.

- Element <referring-entity> contains information on the referrer. The URL that was received from the referrer and the classification of the search engine, if it was used, must both be given in an <identifier> element.
- The <requester>, the agent who has requested the <referent> must be identified by providing the C-class Subnet, and the encrypted IP-address must both be given in separate <identifier>s. In addition, the name of the country where the request was initiated must be provided. The <metadata-by-val> element must be used for this purpose. The country must be given in <dcterms:spatial>. The dcterms namespace must be declared in the <format> element as well.
- The DC metadata term "type" is used to clarify whether the usage event involved a download of a object file or a metadata view.
- Finally, an <identifier> for the institution that provided access to the downloaded document must be given within <resolver>.

## **7. Transfer of data**

Information on usage events must be transferred to a central database. The SURE Statistics project has formulated as a requirement that the transfer of data must take place in a reliable manner. In the infrastructure to be built in this project, the log aggregator will bear the primary responsibility for obtaining the statistical data from individual repositories. Once every 24 hour, it will send a request to each repository for the usage events that have occurred on that particular day. Secondly, the log aggregator is also responsible for ensuring that automated requests by internet robots are filtered as much as possible.

Communication between log aggregator and repository must take place as is detailed in figure 3.

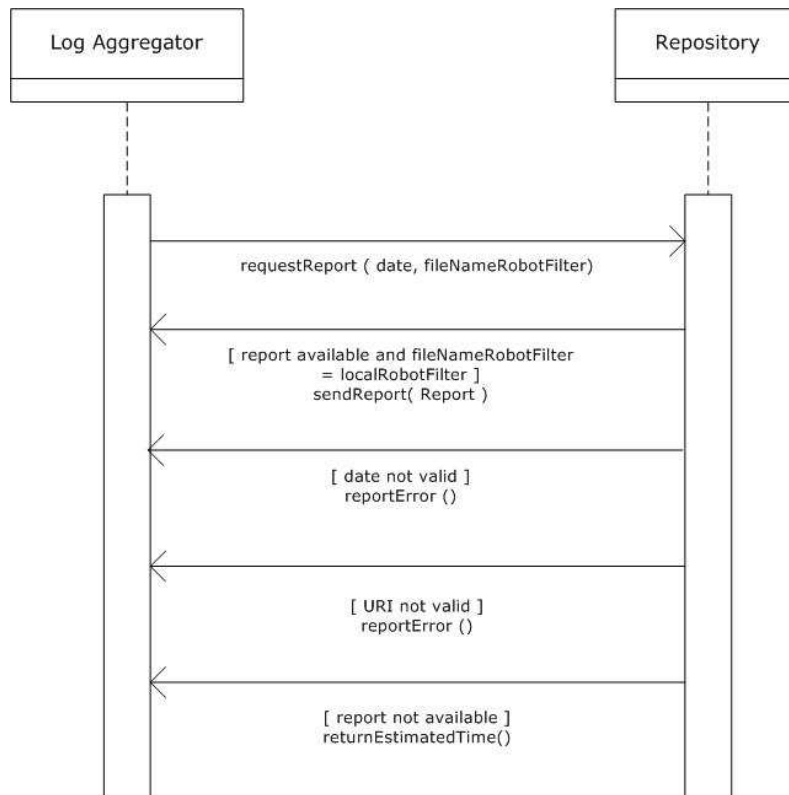


Figure 3.

The interaction commences when the log aggregator sends a request for a report about the daily usage of a certain repository. Two parameters must be sent as part of this request: (1) the date of the report and (2) the file name of the most recent robot filter. The filename that is mentioned in this request will be compared to the local filename. Four possible responses can be returned by the repository.

- If the filename that is mentioned in the request exactly matches the filename that is maintained locally, and if a report for the requested data is indeed available, this report will be returned immediately.
- In this protocol, only daily reports will be allowed. This was decided mainly to restrict the size of the data traffic between the servers. If a request is sent for a period that exceeds one day, an error message will be sent indicating that the date parameter is incorrect.
- If the URI of the robot filter file, for some reason, cannot be resolved to an actual file, an error message will be sent about this.
- If the parameters are correct, but if the report is not yet available, a message will be sent which provides an estimation of the time of arrival.

This protocol involves a conversation in two directions. To implement the transfer of data, the protocol that was developed by the *Standardised Usage Statistics Harvesting Initiative* (SUSHI) shall be used. SUSHI<sup>5</sup> was

<sup>5</sup> <http://www.niso.org/schemas/sushi/>

developed by NISO (National Information Standards Organization) in cooperation with COUNTER. SUSHI enables parties to harvest usage statistics. The protocol works with only two types of messages: requests and responses. The protocol was originally developed for the exchange of COUNTER reports, but fortunately, other types of reports can also be used as the payload of the transfer. The standard does stipulate, however, that the requirements for report naming are adhered to. SUSHI is based on SOAP. The services that can be offered by the Web Service are described in a WSDL document.<sup>6</sup>

In a number of other projects, OAI-PMH is used to synchronise the central database with local databases. In SURE, SUSHI was favoured over OAI-PMH because the latter technique only allows for a one-way-traffic. It was found that SUSHI allowed for a more reliable transfer of data.

In SUSHI version 1.0., the following information must be sent along with the request:

- Requestor ID
- Name of requestor
- E-mail of requestor
- CustomerReference ID (may be identical to the Requestor ID)
- Name of the report that is requested
- Version number of the report
- Start and end date of the report

This request will activate a special tool that can inspect the server logging and that can return the requested data. These data are transferred as OpenURL Context Object log entries, as part of a SUSHI response.

The response must repeat all the information from the request, and provide the requested report as XML payload

The usage data are subsequently stored in a central database. External parties can obtain information about the contents of this central database through specially developed web services. The log harvester must ultimately expose these data in the form of COUNTER-compliant reports.

Listing 2 is an example of a SUSHI request, sent from the log aggregator to a repository.

```
<?xml version="1.0" encoding="UTF-8"?>
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://schemas.xmlsoap.org/soap/envelope/
http://schemas.xmlsoap.org/soap/envelope/">
  <soap:Body>
    <ReportRequest
      xmlns:ctr="http://www.niso.org/schemas/sushi/counter"
      xsi:schemaLocation="http://www.niso.org/schemas/sushi/counter
```

---

<sup>6</sup> It can be accessed at [http://www.niso.org/schemas/sushi/counter\\_sushi2\\_5.wsdl](http://www.niso.org/schemas/sushi/counter_sushi2_5.wsdl).

```

    http://www.niso.org/schemas/sushi/counter_sushi3_0.xsd"
    xmlns="http://www.niso.org/schemas/sushi"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" >
    <Requestor>
      <ID>www.logaggregator.nl</ID>
      <Name>Log Aggregator</Name>
      <Email>logaggregator@surf.nl</Email>
    </Requestor>
    <CustomerReference>
      <ID>www.leiden.edu</ID>
      <Name>Leiden University</Name>
    </CustomerReference>
    <ReportDefinition Release="urn:robots-v1.xml" Name="Daily
Report v1">
      <Filters>
        <UsageDateRange>
          <Begin>2009-12-21</Begin>
          <End>2009-12-22</End>
        </UsageDateRange>
      </Filters>
    </ReportDefinition>
  </ReportRequest>
</soap:Body>
</soap:Envelope>

```

*Listing 2.*

Note that the intent of the SUSHI request above is to see all the usage events that have occurred on 21 December 2009. The SUSHI schema was originally developed for the exchange of COUNTER-compliant reports. In the documentation of the SUSHI XML schema, it is explained that COUNTER usage is only reported at the month level. In SURE, only daily reports can be provided. Therefore, it will be assumed that the implied time on the date that is mentioned is 0:00. The request in the example that is given thus involves all the usage events that have occurred in between 2009-12-21T00:00:00 and 2009-12-22T00:00:00.

As explained previously, the repository can respond in four different ways. If the parameters of the request are valid, and if the requested report is available, the OpenURL ContextObjects will be sent immediately. The Open URL Context Objects will be wrapped into element <Report>, as can be seen in listing 3.

```

<?xml version="1.0" encoding="UTF-8"?>
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://schemas.xmlsoap.org/soap/envelope/
http://schemas.xmlsoap.org/soap/envelope/">
  <soap:Body>
    <ReportResponse
      xmlns:ctr="http://www.niso.org/schemas/sushi/counter"
      xsi:schemaLocation="http://www.niso.org/schemas/sushi/counter
http://www.niso.org/schemas/sushi/counter_sushi3_0.xsd"
      xmlns="http://www.niso.org/schemas/sushi"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" >
      <Requestor>
        <ID>www.logaggregator.nl</ID>
        <Name>Log Aggregator</Name>
        <Email>logaggregator@surf.nl</Email>
      </Requestor>

```

```

<CustomerReference>
  <ID>www.leiden.edu</ID>
  <Name>Leiden University</Name>
</CustomerReference>
<ReportDefinition Release="urn:DRv1" Name="Daily Report v1">
  <Filters>
    <UsageDateRange>
      <Begin>2009-12-22</Begin>
      <End>2009-12-23</End>
    </UsageDateRange>
  </Filters>
</ReportDefinition>
<Report>
  <ctx:context-objects
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns:dcterms="http://dublincore.org/documents/2008/01/14/dcmi-terms/"
    xmlns:ctx="info:ofi/fmt:xml:xsd:ctx">
    <ctx:context-object timestamp="2009-11-
      09T05:56:18+01:00">
      ...
    </ctx:context-object>
  </ctx:context-objects>
</Report>
</ReportResponse>
</soap:Body>
</soap:Envelope>

```

*Listing 3.*

If the begin date and the end date in the request of the log aggregator form a period that exceeds one day, an error message must be sent. In the SUSHI schema, such messages may be sent in an <Exception> element. Three types of errors can be distinguished. Each error type is given its own number. An human-readable error message is provided under <Message>.

```

<?xml version="1.0" encoding="UTF-8"?>
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://schemas.xmlsoap.org/soap/envelope/
http://schemas.xmlsoap.org/soap/envelope/">
  <soap:Body>
    <ReportResponse
      xmlns:ctr="http://www.niso.org/schemas/sushi/counter"
      xsi:schemaLocation="http://www.niso.org/schemas/sushi/counter
http://www.niso.org/schemas/sushi/counter_sushi3_0.xsd"
      xmlns="http://www.niso.org/schemas/sushi"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" >
      <Requestor>
        <ID>www.logaggregator.nl</ID>
        <Name>Log Aggregator</Name>
        <Email>logaggregator@surf.nl</Email>
      </Requestor>
      <CustomerReference>
        <ID>www.leiden.edu</ID>
        <Name>Leiden University</Name>
      </CustomerReference>
      <ReportDefinition Release="urn:DRv1" Name="Daily Report v1">
        <Filters>
          <UsageDateRange>
            <Begin>2009-12-22</Begin>

```

```

        <End>2009-12-23</End>
      </UsageDateRange>
    </Filters>
  </ReportDefinition>
  <Exception>
    <Number>1</Number>
    <Message>The range of dates that was provided is not
      valid. Only daily reports are
      available.</Message>
  </Exception>
</ReportResponse>
</soap:Body>
</soap:Envelope>

```

*Listing 4.*

A second type of error may be caused by the fact that the file that is mentioned in the request can not be accessed. In this situation, the response will look as follows:

```

<?xml version="1.0" encoding="UTF-8"?>
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://schemas.xmlsoap.org/soap/envelope/
  http://schemas.xmlsoap.org/soap/envelope/">
  <soap:Body>
    <ReportResponse
      xmlns:ctr="http://www.niso.org/schemas/sushi/counter"
      xsi:schemaLocation="http://www.niso.org/schemas/sushi/counter
      http://www.niso.org/schemas/sushi/counter_sushi3_0.xsd"
      xmlns="http://www.niso.org/schemas/sushi"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" >
      <Requestor>
        <ID>www.logaggregator.nl</ID>
        <Name>Log Aggregator</Name>
        <Email>logaggregator@surf.nl</Email>
      </Requestor>
      <CustomerReference>
        <ID>www.leiden.edu</ID>
        <Name>Leiden University</Name>
      </CustomerReference>
      <ReportDefinition Release="urn:DRv1" Name="Daily Report v1">
        <Filters>
          <UsageDateRange>
            <Begin>2009-12-22</Begin>
            <End>2009-12-23</End>
          </UsageDateRange>
        </Filters>
      </ReportDefinition>
      <Exception>
        <Number>2</Number>
        <Message>The file describing the internet robots is not
          accessible.</Message>
      </Exception>
    </ReportResponse>
  </soap:Body>
</soap:Envelope>

```

*Listing 5.*

When the repository is in the course of producing the requested report, a response will be sent that is very similar to listing 6. The estimated time



of completion will be provided in the <Data> element. According to the documentation of the SUSHI XML schema, this element may be used for any other optional data.

```

<?xml version="1.0" encoding="UTF-8"?>
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://schemas.xmlsoap.org/soap/envelope/
  http://schemas.xmlsoap.org/soap/envelope/">
  <soap:Body>
    <ReportResponse
      xmlns:ctr="http://www.niso.org/schemas/sushi/counter"
      xsi:schemaLocation="http://www.niso.org/schemas/sushi/counter
      http://www.niso.org/schemas/sushi/counter_sushi3_0.xsd"
      xmlns="http://www.niso.org/schemas/sushi"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" >
      <Requestor>
        <ID>www.logaggregator.nl</ID>
        <Name>Log Aggregator</Name>
        <Email>logaggregator@surf.nl</Email>
      </Requestor>
      <CustomerReference>
        <ID>www.leiden.edu</ID>
        <Name>Leiden University</Name>
      </CustomerReference>
      <ReportDefinition Release="urn:DRv1" Name="Daily Report v1">
        <Filters>
          <UsageDateRange>
            <Begin>2009-12-22</Begin>
            <End>2009-12-23</End>
          </UsageDateRange>
        </Filters>
      </ReportDefinition>
      <Exception>
        <Number>3</Number>
        <Message>The report is not yet available. The estimated
          time of completion is
          provided under "Data".</Message>
        <Data>2010-01-08T12:13:00+01:00</Data>
      </Exception>
    </ReportResponse>
  </soap:Body>
</soap:Envelope>

```

*Listing 6.*

Error numbers and the corresponding Error messages are also provided in the table below.

<i>Error number</i>	<i>Error message</i>
1	The range of dates that was provided is not valid. Only daily reports are available.
2	The file describing the internet robots is not accessible
3	The report is not yet available. The estimated time of completion is provided under "Data"