EDIS 3 Web Service Guide

Data Extraction Support

Aug 13, 2010

Description
To support new Open Government initiatives, the Electronic Document Information System (EDIS) provides an interface to extract investigation, document and attachment metadata, as well as attachment Adobe Acrobat PDF files. EDIS uses a Representational State Transfer (RESTful) web service API to provide this data.

The base URL for all calls to the EDIS 3 Web Service (EWS) is:

https://edis.usitc.gov/data

Currently XML is the only format the EWS will return data in.

For complete explanations of the EDIS data elements used in the EWS, please refer to the EDIS 3 User Guide. It is available on the EDIS Help tab Data element definitions are also available through online field help on the EDIS Advanced Search parameters page.

Pagination
Certain methods will return up to 100 resources per request. These methods utilize a pageNumber parameter to control which page of resources will be returned. For example:

https://edis.usitc.gov/data/someMethod?pageNumber=2

This will return the 2nd page of results for “someMethod”. This would include the 101st resource through the 200th resource, if available. If the method called does not have any available resources in that range, the results returned will be empty.

Query Parameters

<table>
<thead>
<tr>
<th>Query Parameter</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>pageNumber</td>
<td>Controls which page of resources the method will return. If left unspecified, pageNumber will default to 1.</td>
</tr>
</tbody>
</table>

Errors
When an EWS request generates an error, the EWS will return a response with an 500 HTTP Status code. The response body will include additional information in the follow form:

<errors>
Methods
URIs listed below may include parameters embedded as URI path elements. For notation in this guide, parameters are identified with \{\}, for example:

https://edis.usitc.gov/data/someMethod/\{parameterEmbeddedInPath\}

Authentication
The EWS allows users to request API methods anonymously. However, due to the sensitivity of the available data, anonymous requests have access to a limited set of data. At the time of publishing this guide, anonymous access is currently limited in the following ways:

- Can only access metadata for public, limited and confidential documents
- Can only access metadata for public, limited and confidential attachments
- Cannot download any attachments

Please note that these limitations can change without notice.

If a EWS request includes authentication information, and the authentication in valid, the EWS request will be executed with the user’s full set of authorizations.

The EWS authentication is based on a valid EDIS 3 (https://edis.usitc.gov) user account and works in the following manner:

1. A user with a valid EDIS 3 user account makes a request to the Error! Reference source not found. method detailed below.
2. The user creates an authentication token using the secret key returned in step 1 above.
3. Include the authorization token created in step 2 above as the value for an HTTP Authorization header in an EWS request. If verified, that EWS request will be executed with the permissions authorized for the associated user account.

Secret Key
The following URIs return a secret key that can be used for authenticating subsequent EDIS WS requests. The secret key uses the submitters currently username and password in EDIS for validation. The user account must remain active in EDIS to be used with the EWS. The resulting secret key must then be hashed and passed back to the EWS as a authentication string.

**URI Structure**

<table>
<thead>
<tr>
<th>URI</th>
<th>Method</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>/secretKey/{username}</td>
<td>POST</td>
<td>Active EDIS login username.</td>
</tr>
</tbody>
</table>

**Post Parameters**
<table>
<thead>
<tr>
<th>Query Parameter</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>password</td>
<td>Password for user specified in URI.</td>
</tr>
</tbody>
</table>

**Pagination**

Does not support pagination.

**Request Example**

To obtain the secret key, the EDIS Username and corresponding password are posted with the secretKey command. In this example, the password is provided as a POST data parameter and the username is part on the URL.

If the password contains ‘&’ or ‘%’ these may need to be properly escaped within the parameter string.

```bash
... -d "password=JDoe1234" https://edis.usitc.gov/data/secretKey/JDoe
```

**Results Example**

```xml
<results>
  <secretKey>fbd45fca08df0d04dd3959769089ebbdd8957aff</secretKey>
</results>
```

**Authentication Token**

The authentication token is the value passed to the EWS as an HTTP Authorization header as part of an EWS request. It identifies the user issuing the request and includes the user’s secret key\(^1\).

The following is pseudo-code that shows how to create the authentication token:

```plaintext
RawAuthenticationToken =  username + ":" + secretKey
AuthenticationToken = "Basic" + " " + Base64(RawAuthenticationToken)
```

The AuthenticationToken is generated by Base64 encoding the RawAuthenticationToken.

**Authentication Token Value Example**

Given a username of “TestUser” and a secretKey of “981234972389473874837asdf” the Authentication Token would be created as follows:

```plaintext
RawAuthenticationToken =  “TestUser” + “:" + “981234972389473874837asdf”
AuthenticationToken = "Basic" + " " + Base64(RawAuthenticationToken)
```

The value of the RawAuthenticationToken should be:

---

\(^1\) All EWS requests are issued over HTTPS, limiting the chances that a secret key can be compromised. In the future the EWS may move to using a HMAC authentication scheme.
The value of the resulting AuthenticationToken should be:

"Basic VGVzdFVzZXI6OTgxMjM0OTcyMzg5NDcxODc0ODM3YXNkZg=="

**Authorization Header**

To authenticate an EWS request, the request must include a valid HTTP Authorization header. The Authorization header has the following form:

Authorization:AuthenticationToken

The Authorization Header for the Authentication Token created in “Authentication Token Value Example” above would be as follows:

Authorization:Basic VGVzdFVzZXI6OTgxMjM0OTcyMzg5NDcxODc0ODM3YXNkZg==

**Investigation**

The following URIs return a list of investigations.

**URI Structure**

<table>
<thead>
<tr>
<th>URI</th>
<th>Method</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>/investigation</td>
<td>GET</td>
<td></td>
</tr>
<tr>
<td>/investigation/{investigationNumber}</td>
<td>GET</td>
<td>Returns a list of investigations with a matching investigation number.</td>
</tr>
<tr>
<td>/investigation/{investigation number}/{investigationPhase}</td>
<td>GET</td>
<td>Returns a list of investigations with a matching investigation number and investigation phase.</td>
</tr>
</tbody>
</table>

**Optional Query Parameters**

<table>
<thead>
<tr>
<th>Query Parameter</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>investigationType</td>
<td>Entire string for investigationType must be specified, such as Industry%20and%20Economic%20Analysis</td>
</tr>
<tr>
<td>investigationStatus</td>
<td>Values are Preinstitution, Active, Inactive, Cancelled.</td>
</tr>
</tbody>
</table>

**Pagination**

Supports pagination.

**Request Example**

Example: Find investigations with an investigation number of 731-1103:

https://edis.usitc.gov/data/investigation/731-1103

Example: Find investigation 731-1103 Final phase:

https://edis.usitc.gov/data/investigation/731-1103/final

8/13/2010
Examples: Find all active investigations:

https://edis.usitc.gov/data/investigation?investigationstatus=Active

**Results Example**

```xml
<results>
  <investigations>
    <investigation>
      <investigationNumber>111-123</investigationNumber>
      <investigationPhase>Final</investigationPhase>
      <investigationStatus>Active</investigationStatus>
      <investigationTitle>Certain RESTful Web Services</investigationTitle>
      <investigationType>Import Injury</investigationType>
      <documentListUri>
        https://edis.usitc.gov/data/document?investigationNumber=111-123&investigationPhase=Final
      </documentListUri>
    </investigation>
    <investigation>...</investigation>
  </investigations>
</results>
```

<table>
<thead>
<tr>
<th>Result Element</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>investigationNumber</td>
<td>such as 337-406</td>
</tr>
<tr>
<td>investigationPhase</td>
<td>such as Violation, Preliminary, Final, Review2</td>
</tr>
<tr>
<td>investigationStatus</td>
<td>such as PreInstitution, Active, Inactive, Cancelled</td>
</tr>
<tr>
<td>investigationTitle</td>
<td>Investigation Title associated with the Investigation Number</td>
</tr>
<tr>
<td>investigationType</td>
<td>such as Import Injury, Sec 337, Industry and Economic Analysis, Byrd Amendment, Rule Making, or Tariff Affairs &amp; Trade Agreements.</td>
</tr>
<tr>
<td>documentListUri</td>
<td>provides the URI that returns document data for a given investigation</td>
</tr>
</tbody>
</table>

**Document**

The following URIs return a list of documents.

**URI Structure**

<table>
<thead>
<tr>
<th>URI</th>
<th>Method</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>/document</td>
<td>GET</td>
<td>Returns a single document with a matching document ID.</td>
</tr>
<tr>
<td>/document/{documentId}</td>
<td>GET</td>
<td></td>
</tr>
</tbody>
</table>

**Optional Query Parameters**

<table>
<thead>
<tr>
<th>Query Parameter</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>investigationNumber</td>
<td>Investigation number in EDIS format can be specified. Partial InvestigationNumbers may be specified.</td>
</tr>
<tr>
<td>investigationPhase</td>
<td>Specify any EDIS investigationPhase such as Review2</td>
</tr>
<tr>
<td>documentType</td>
<td>The complete documentType must be specified. Partial strings are not allowed.</td>
</tr>
<tr>
<td>firmOrg</td>
<td>Any words of the complete Firm Organization name may be provided. Each word is treated as an OR search.</td>
</tr>
<tr>
<td>securityLevel</td>
<td>Enter an EDIS security level such as Public to retrieve only Public metadata</td>
</tr>
</tbody>
</table>

**Pagination**
Supports pagination.

**Request Example**
Example: Find the second set of 100 document resources in investigation 731-1103 Final:


Example: Find all documents that are of document type ‘Questionnaire – Purchaser’:


Example: Find documents in 731-1103 with document type of ‘Questionnaire – Purchaser’:


**Results Example**

```xml
<results>
  <documents>
    <document>
      <id>111112</id>
      <documentType>Other</documentType>
      <documentTitle>RESTful Pre-Trial Statement</documentTitle>
      <securityLevel>Confidential</securityLevel>
      <investigationNumber>111-123</investigationNumber>
      <investigationPhase>Final</investigationPhase>
      <investigationStatus>Active</investigationStatus>
      <investigationTitle>Certain RESTful Web Services</investigationTitle>
      <investigationType>Import Injury</investigationType>
      <firmOrganization>Ruggles and Ruggles</firmOrganization>
  </document>
</documents>
```

8/13/2010 6
Result Element | Notes
---|---
id | unique ID number for the document in EDIS

documentType | identifies the type of document, such as ‘Action Request’ or ‘Correspondence – USITC’

documentTitle | investigationNumber such as 337-406

securityLevel | identifies the security level of the document

investigationPhase | such as Violation, Preliminary, Final

investigationStatus | such as PreInstitution, Active, Inactive, Cancelled

investigationTitle | Investigation Title associated with the Investigation Number

investigationType | such as Import Injury, Sec 337, Industry and Economic Analysis, Byrd Amendment and Rule Making

firmOrganization | name of the Firm filing the document

filedBy | attorney filing the document

onBehalfOf | Company being represented on the filing; or the Office for internal USITC filings

documentDate | date the document was authored or signed

officialReceivedDate | date the document was provided to the USITC Dockets Office

attachmentListUri | provides the URI that returns attachment data for the given investigation

Attachment Metadata
The following URI returns a list of attachments.

**URI Structure**

<table>
<thead>
<tr>
<th>URI</th>
<th>Method</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>/attachment/{documentId}</td>
<td>GET</td>
<td>Returns attachment data for a document with a matching document ID.</td>
</tr>
</tbody>
</table>
Optional Query Parameters
None

Pagination
Does not support pagination.

Request Example
https://edis.usitc.gov/data/attachment/111112

Results Example
```
<results>
  <attachments>
    <attachment>
      <id>654321</id>
      <documentId>111112</documentId>
      <title>RESTful Q&A</title>
      <fileSize>8112</fileSize>
      <originalFileName>\mopey\prodimages\a7f\61BAF6</originalFileName>
      <pageCount/>
      <createDate>2005-01-24 00:00:00.0</createDate>
      <lastModifiedDate/>
      <downloadUri>https://edis.usitc.gov/data/download/111112/654321</downloadUri>
    </attachment>
  </attachments>
</results>
```

<table>
<thead>
<tr>
<th>Result Element</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>unique ID number for the attachment in EDIS</td>
</tr>
<tr>
<td>documentId</td>
<td>unique ID number for the document in EDIS that the attachment is associated with</td>
</tr>
<tr>
<td>title</td>
<td>Attachment title</td>
</tr>
<tr>
<td>fileSize</td>
<td>Size of the attachment’s PDF document in bytes</td>
</tr>
<tr>
<td>originalFileName</td>
<td>Filename as uploaded into EDIS.</td>
</tr>
<tr>
<td>pageCount</td>
<td>Pages in the attachment</td>
</tr>
<tr>
<td>createDate</td>
<td>Original upload date</td>
</tr>
<tr>
<td>lastModifiedDate</td>
<td>Last file edit date. This may indicate the time the file was OCR’d, not necessarily last textual edit of the content.</td>
</tr>
<tr>
<td>downloadUri</td>
<td>provides the URI that returns the attachment’s PDF document</td>
</tr>
</tbody>
</table>

Attachment Document
The following URI returns the PDF for the given attachment.

8/13/2010 8
### URI Structure

<table>
<thead>
<tr>
<th>URI</th>
<th>Method</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>/download/{documentId}/attachmentId</td>
<td>GET</td>
<td>Returns PDF for an attachment with a matching document ID and attachment ID.</td>
</tr>
</tbody>
</table>

### Optional Query Parameters

None

### Pagination

Does not support pagination.

### Request Example

The authorization string must be submitted for these requests. Example: https://edis.usitc.gov/data/download/400316/925354

### Results Example

NA. Results are a stream of “application/pdf” bytes.