

The Classical Text Services specification: version 1.1

Neel Smith, 2005

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/1.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

Contents

| | |
|---|----|
| | |
| Versions..... | 1 |
| Requirements for client-server communication..... | 1 |
| Common parameters..... | 3 |
| The CTS requests..... | 5 |
| The CTS DTDs..... | 19 |

This document defines version *1.1* of the Classical Text Services implementation specification (CTS). The specification defines interaction between a client and server providing identification of texts and retrieval of fragments texts using notions of "work" and "citation" traditional in classical studies. For an overview of the Classical Text Services, see <http://chs75.harvard.edu/projects/diginc/techpub/cts>.

Versions

Release versions are identified by two dot-separated integers, with the first being most significant. E.g., "2.0" is greater than "1.10" but "1.10" is greater than "1.2".

Experimental, pre-release versions may also be available from the CHS version control archive or directly from a CHS website. An identifier for a pre-release version is composed of a regular release-version number followed by a string indicating its test status (e.g., "1.2.beta" or "1.2.release-candidate"). Optionally, the identifier for a test version may have a further positive integer indicating its sequence in that test series (e.g., "1.2.beta.2").

A version number is a required attribute on the `TextInventory` element returned by a `GetCapabilities` request (defined below). As of version 1.1.release-candidate, client applications may optionally indicate a version number in a `version` parameter with any request. Use of a `version` parameter in client requests may be required in future versions of this protocol.

Requirements for client-server communication

CTS uses the HyperText Transfer Protocol [IETF RFC 2616] for communication between client and server. An instance of Classical Text Services, referred to hereafter as a CTS "base URL" is an HTTP Uniform Resource Locator (URL). This URL is not required to return any content, although it is strongly recommended that a request for a CTS URL return basic information

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/1.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

identifying the CTS in either text-only or XHTML format.

The CTS base URL serves as a URL prefix for CTS requests: CTS requests are formed by adding request parameters directly to the base request URL. More formally, the concatenation of the base request URL with URL parameters must produce a string forming a valid URL according to the requirements of the URL specification [IETF RFC 2396] and the HTTP Common Gateway Interface standard (CGI). Each parameter's name and value pair must be separated from any following parameters by an ampersand ("&", Unicode x0026); name and value for each individual parameter must be separated by an equals sign ("=", Unicode x003D). . All CTS requests include a URL parameter named `request`. The value of this parameter must be the name of a request defined in version 1.1. For all requests other than `GetCapabilities`, one or more additional parameters are required as defined in this specification. If any of the characters "/", "&", or "=" appear in the value of a parameter, they must therefore be encoded as defined in IETF RFC 2616. In version 1.1, CTS requires use of the GET method for all requests.

Examples of `GetCapabilities` requests formed from valid CTS base URLs:
`http://myhost/mycts?` is a valid base URL because the request string
`http://myhost/mycts?request=GetCapabilities` is a valid URL.
`http://myhost/mycts?configuration=default&` is also a valid base URL because
the request
`http://myhost/mycts?configuration=default&request=GetCapabilities`
is a valid URL.

In addition, CTS reserves a special meaning for two characters. When the period (".", Unicode x002E) appears in `ref` or `range` parameters, it separates hierarchical components of a reference. If a period instead represents a character in the reference value, it must be URL

escaped. Similarly, in the `range` parameter, the hyphen ("-"), Unicode x002D) separates two `ref` components of a range. If a hyphen instead represents a character in the range value, it must be URL escaped. See the following section on parameters.

Although the specification allows periods and hyphens to appear as literal values for `ref` and `range` parameters, editors of TEI documents for use in a CTS instance are strongly discouraged from including periods or hyphens in the literal values of individual components of a hierarchical reference scheme.

Parameter names and values are both case sensitive. Order of parameters in a request is not significant.

Common parameters

The following parameters are allowed or required in one or more CTS requests to refer to parts of a TextInventory's hierarchy of works. Using required and allowed parameters, a CTS client can refer to a work or group of works at any level of the TextInventory hierarchy. It is an error if the value of one of these parameters does not appear in the corresponding section of the TextInventory as summarized here:

- | | | |
|-------------------------|---|---|
| <code>textgroup</code> | • | Identifier of a TextGroup; corresponds to the <code>projid</code> attribute of a <code>textgroup</code> element in the TextInventory, and must be unique within the scope of the TextInventory. |
| <code>collection</code> | • | Identifier of a Collection; corresponds to the <code>id</code> attribute of a <code>collection</code> element in the TextInventory and must be unique within the scope of the TextInventory. |
| <code>work</code> | • | Identifier of a Work; corresponds to the <code>projid</code> attribute in the TextInventory and must be unique within the scope of the |

- containing TextGroup.
- edition • Identifier of an Edition; corresponds to the `projid` attribute on an `editioncomments` element in the TextInventory and must be unique among all edition and translation identifiers within the scope of the containing Work.
 - translation • Identifier of a Translation; corresponds to the `projid` attribute of a `translationcomments` element in the TextInventory and must be unique among all edition and translation identifiers within the scope of the containing Work.
 - exemplar • Identifier of an Exemplar; corresponds to the `projid` attribute of an `exemplar` element in the TextInventory and must be unique within the scope of the containing Edition or Translation.

The following parameters are allowed or required in one or more CTS requests to refer to portions of an on-line text in the citation scheme described in the TextInventory for that work.

- ref • A reference is a concatenation of zero or more strings representing components in a hierarchical citation scheme. Each component except the last is joined to following components by a period (".", Unicode x002E). It is an error if the components of a reference cannot be derived from the return value of a `GetValidReff` request by examining the `n` attributes of successively contained elements in the citation

scheme beginning with the root element of the citation scheme. See below on the `GetValidReff` request. A reference composed of zero elements (the empty string, "") is defined to mean a reference to an entire `Work`.

range

•

A range is a concatenation of exactly two valid references (as described above) joined by a hyphen ("-", Unicode x002D) . The individual references must contain the same number of reference components: it is an error if the references do not refer to the same depth of the citation hierarchy. It is an error if the first reference does not precede the second reference in the document order of the requested work.

The CTS requests

Classical Text Services requires implementation of six requests. These requests are named `GetCapabilities`, `GetWorks`, `GetValidReff`, `GetPassage`, `GetTEIHeader` and `DownloadText`, and include required and allowed parameters as defined in the following sections.

The reply to a CTS request is always a well-formed XML document with a root element having the same name as the CTS request. The reply document always contains exactly two elements, the first of which is named `request`. The format of the `request` element is unspecified, and may be used for any purpose the implementation wishes. As the name of the element is intended to imply, it is strongly recommended that implementing software use this element to report back parameters passed in with the request. Developers may also use this

element for other purposes such as embedding debugging information in a reply that complies with the CTS specification.

If the syntax and contents of request parameters do not fully comply with the specifications in this document, the second element of the reply is named `error`; as of version 1.1 of the CTS specification, the format and content of error reporting are left up to the implementing software. Future versions of the protocol may define structures for error reporting.

If the syntax and contents of request parameters fully comply with the specifications in this document, then the name and format of the second element vary according to the request, and are defined in the following sections. As specified below, four of the requests (`GetCapabilities`, `GetWorks`, `GetValidReff`, and `DownloadText`) require a second element that can be validated against DTDs documented or included as appendices to this specification; the `GetPassage` and `GetTEIHeader` requests return a second element containing either a well-formed fragment of a valid TEI document, or a well-formed aggregate of fragments of a valid TEI document.

GetCapabilities

Purpose: the `GetCapabilities` request retrieves the `TextInventory` of a CTS instance. The `TextInventory` defines the corpus of texts known to the server and any optional services it supports. In version 1.1, no optional services are defined, but drafts of future protocol versions have already proposed optional services, and optional services are scheduled to be part of future versions.

No parameters other than the `request` parameter are required, but as with all CTS requests, a request may optionally include a `version` parameter indicating the version of the protocol preferred by the client. As of version 1.1, servers are not required to recognize or alter

their reply in response to a `version` parameter.

| <i>Parameter</i> | <i>Required/optional</i> | <i>Description</i> |
|----------------------|--------------------------|---|
| <code>version</code> | optional | Client may indicate preferred version of CTS protocol |

Response: the second element validates against `TextInventory.dtd` (see appendix).

GetWorks

Purpose: The `GetWorks` request is a convenience method for rapid discovery of subsets of a `TextInventory`. If a `TextGroup` is requested, the `works` element of the response will contain a literal fragment of the server's `TextInventory` rooted at the `textgroup` element with `projid` attribute equal to the requested `TextGroup`. If a `Collection` is requested, the `works` element of the response will contain one or more `textgroup` elements, each containing one or more `work` elements. Only `work` elements containing a `memberof` element with `collection` attribute equal to the requested `Collection` will be included. It is an error if the requested `TextGroup` or `Collection` does not exist in the server's `TextInventory`.

| <i>Request</i> | <i>Required/optional</i> | <i>Description</i> |
|-------------------------|---|---|
| <code>textgroup</code> | One and only one of <code>textgroup</code> or <code>collection</code> attribute of the desired is required. | The parameter is the <code>projid</code> <code>textgroup</code> element. It is an error if this value is not present in the <code>TextInventory</code> of the CTS instance. |
| <code>collection</code> | One and only one of | The parameter is the <code>id</code> |

| | | |
|---------|--|---|
| | textgroup or collection attribute of the desired | |
| | is required. | collection. It is an error if this value is not present in the TextInventory of the CTS instance. |
| version | optional | Client may indicate preferred version of CTS protocol |

Response: the second element validates against `GetWorks.dtd` (see appendix).

GetValidReff

Purpose: the `citation` element of a `TextInventory` (exposed either through the `GetCapabilities` or `GetWorks` request) defines the syntax and semantic labels for a citation scheme; the `GetValidReff` request complements this by identifying all valid values for an on-line version of a requested work up to a specified level of the citation hierarchy.

| <i>Request</i> | <i>Required/optional</i> | <i>Description</i> |
|-----------------------|---------------------------------|--|
| level | required | The parameter is an integer value specifying the depth of the citation hierarchy to retrieve. The highest level of the citation hierarchy is defined to be a level of 0. |
| textgroup | required | The parameter is the <code>projid</code> attribute of the <code>textgroup</code> |

| | | |
|-------------|---|--|
| | | element containing the desired work. It is an error if this value is not present in the TextInventory of the CTS instance. |
| work | required | The parameter is the <code>projid</code> attribute of the desired work. It is an error if this value is not present in the TextInventory of the CTS instance. |
| edition | One and only one of <code>edition</code> or <code>translation</code> is required. | The parameter is the <code>projid</code> of the desired Edition. It is an error if the requested Edition contains neither an <code>online</code> element nor an <code>exemplar</code> element containing an <code>online</code> element. |
| translation | One and only one of <code>edition</code> or <code>translation</code> is required. | The parameter is the <code>projid</code> of the desired Translation. It is an error if the requested Translation contains neither an <code>online</code> element nor an <code>exemplar</code> element containing |

| | | |
|----------|----------|---|
| | | an online element. |
| exemplar | optional | The parameter is the <code>projid</code> of a specific Exemplar. It is valid to include an <code>exemplar</code> parameter only if one and only one of <code>translation</code> or <code>edition</code> parameters is also supplied. It is an error if the requested Exemplar does not exist within the requested Edition or Translation element of the server's TextInventory. |
| version | optional | Client may indicate preferred version of CTS protocol |

Response: validates against the `GetValidReff.dtd`

GetPassage

Purpose: retrieve a fragment of a TEI-conformant document by canonical reference at any granularity allowed by the citation scheme, at any notional level of a Work known to a server's TextInventory. It is an error if no Edition, Translation or Exemplar of a work contains an online element, but the `GetPassage` request does not require specific identification of an Edition, Translation or Exemplar.

| <i>Request</i> | <i>Required/optional</i> | <i>Description</i> |
|----------------|--------------------------|--------------------|
|----------------|--------------------------|--------------------|

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/1.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

| | | |
|--------------------------|---|---|
| <code>ref</code> | One and only one of <code>ref</code> or <code>range</code> is required. | A reference to a single passage in a hierarchical reference system. It is an error if the number of components in a Reference exceeds the number of citation elements in the applicable citation scheme. See discussion in section 3, "Common parameters," for further requirements on format and content of a Reference. |
| <code>range</code> | One and only one of <code>ref</code> or <code>range</code> is required. | A reference to a section of text spanning two or more elements as defined by a hierarchical reference system. See discussion in section 3, "Common parameters," for further requirements on format and content of a Range. |
| <code>contextsize</code> | The optional <code>contextsize</code> parameter is valid only in conjunction with a <code>ref</code> parameter. | The parameter specifies the number of units of context to include before and after the section of text identified by the |

| | | |
|------------------------|----------|---|
| | | <p><code>ref</code> parameter, at the same level of the citation hierarchy. If <code>contextsize</code> is defined, the CTS reply will include all preceding and following sections of the document at the same level of the citation hierarchy up to <code>contextsize</code> sections. It is <i>not</i> an error if the document contains fewer than <code>contextsize</code> preceding or following sections. It is an error if the parameter is not a positive integer.</p> |
| <code>textgroup</code> | required | <p>The parameter is the <code>projid</code> attribute of the <code>textgroup</code> element containing the desired work. It is an error if this value is not present in the TextInventory of the CTS instance.</p> |
| <code>work</code> | required | <p>The parameter is the <code>projid</code> attribute of the desired work. It</p> |

is an error if this value is not present in the TextInventory of the CTS instance.

edition Optionally, one and only one of The parameter is the `projid` edition or translation of the desired Edition. It is an error if this value is not present in the TextInventory of the CTS instance. In the absence of either an `edition` or `translation` parameter, choice of a default is left to implementing program. It is an error if the implementation's default (either Edition or Translation) is not present in the TextInventory of the CTS instance. It is an error if either the requested Edition, or in its absence the default Edition or Translation, contains neither an `online` element nor an `exemplar` element containing an `online` element.

| | | |
|-------------|---|---|
| translation | Optionally, one and only one of The parameter is the projid edition or translation of the desired Translation. It is may be included. | an error if this value is not present in the TextInventory of the CTS instance. In the absence of either an edition or translation parameter, choice of a default is left to implementing program. It is an error if the implementation's default (either Edition or Translation) is not present in the TextInventory of the CTS instance. It is an error if either the requested Translation, or in its absence the default Edition or Translation, contains neither an online element nor an exemplar element containing an online element. |
| exemplar | optional | The parameter is the projid of a specific Exemplar. It is valid to include an exemplar |

| | | |
|---------|----------|--|
| | | parameter only if one and only one of translation or edition parameters is also supplied. It is an error if the requested Exemplar does not exist within the requested Edition or Translation element of the server's TextInventory. Client may indicate preferred version of CTS protocol |
| version | optional | |

Response: the second element is named TEIFrag. It contains the well-formed fragment of the requested document referred to by either the ref or range parameter. The lang attribute on the TEIFrag element indicates the base language of the document.

GetTEIHeader

Purpose: retrieve metadata about a specific on-line Edition, Translation or Exemplar.

| <i>Request</i> | <i>Required/optional</i> | <i>Description</i> |
|----------------|--------------------------|---|
| textgroup | required | The parameter is the projid attribute of the textgroup element containing the desired work. It is an error if this value is not present in the TextInventory of the CTS |

| | | |
|-------------|--|--|
| | | instance. |
| work | required | The parameter is the <code>projid</code> attribute of the desired work. It is an error if this value is not present in the <code>TextInventory</code> of the CTS instance. |
| edition | One and only one of <code>edition</code> or <code>translation</code> must be included. | The parameter is the <code>projid</code> of the desired Edition. It is an error if this value is not present in the <code>TextInventory</code> of the CTS instance. |
| translation | One and only one of <code>edition</code> or <code>translation</code> must be included. | The parameter is the <code>projid</code> of the desired Translation. It is an error if this value is not present in the <code>TextInventory</code> of the CTS instance. |
| exemplar | optional | The parameter is the <code>projid</code> of a specific Exemplar. It is valid to include an <code>exemplar</code> parameter only if one and only one of <code>translation</code> or <code>edition</code> parameters is also supplied. It is an error if the |

| | | |
|---------|----------|--|
| | | requested Exemplar does not exist within the requested Edition or Translation element of the server's TextInventory. |
| version | optional | Client may indicate preferred version of CTS protocol |

Reponse: the second element, named `teiHeader`, may contain any content that would be valid as the `teiHeader` of a document according to the fourth public release of the Text Encoding Initiative's Guidelines (Sperberg-McQueen, C.M.. and Burnard, L., eds., 2002, *TEI P4: Guidelines for Electronic Text Encoding and Interchange*).

DownloadText

Purpose: the `DownloadText` request is a convenience method for easy retrieval of complete TEI-conformant documents. It is an error if the requested document does not contain an `online` element in the server's TextInventory.

| <i>Request</i> | <i>Required/optional</i> | <i>Description</i> |
|------------------------|---------------------------------|---|
| <code>textgroup</code> | required | The parameter is the <code>projid</code> attribute of the <code>textgroup</code> element containing the desired work. It is an error if this value is not present in the TextInventory of the CTS instance. |

| | | |
|--------------------------|--|--|
| <code>work</code> | <code>required</code> | The parameter is the <code>projid</code> attribute of the desired work. It is an error if this value is not present in the <code>TextInventory</code> of the CTS instance. |
| <code>edition</code> | One and only one of <code>edition</code> or <code>translation</code> must be included. | The parameter is the <code>projid</code> of the desired Edition. It is an error if this value is not present in the <code>TextInventory</code> of the CTS instance. |
| <code>translation</code> | One and only one of <code>edition</code> or <code>translation</code> must be included. | The parameter is the <code>projid</code> of the desired Translation. It is an error if this value is not present in the <code>TextInventory</code> of the CTS instance. |
| <code>exemplar</code> | <code>optional</code> | The parameter is the <code>projid</code> of a specific Exemplar. It is valid to include an <code>exemplar</code> parameter only if one and only one of <code>translation</code> or <code>edition</code> parameters is also supplied. It is an error if the requested Exemplar does not |

| | | |
|---------|----------|--------------------------------|
| | | exist within the requested |
| | | Edition or Translation element |
| | | of the server's TextInventory. |
| version | optional | Client may indicate preferred |
| | | version of CTS protocol |

Response: a complete document that validates against the DTDs in the fourth public release of the Text Encoding Initiative's Guidelines (Sperberg-McQueen, C.M. and Burnard, L., eds., 2002, *TEI P4: Guidelines for Electronic Text Encoding and Interchange*).

The CTS DTDs

The appended DTDs referred to previously in this specification are part of the Classical Text Services specification.

TextInventory.dtd

PUBLIC identifier: "-//Center for Hellenic Studies//DTD TextInventory
1.1//EN"

insert DTD here

textgroup.dtd

GetWorks.dtd

PUBLIC identifier: "-//Center for Hellenic Studies//DTD GetWorks
1.1//EN"

insert DTD here

Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/1.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

GetWorks.dtd

PUBLIC identifier: "-//Center for Hellenic Studies//DTD GetValidRef
1.1//EN"

insert DTD here