From File Interoperability to Service Interoperability :

The Distributed Text Services

Wooden conduit once used to supply water to Edo residents. Made from hardwood to make it watertight. Displayed at Tokyo Edo Museum
Source: https://web-japan.org/niponica/niponica15/en/feature/feature02.html

Thibault Clérice

Head of TNAH MA
École Nationale des Chartes
thibault.clerice@chartes.psl.eu
Twitter: @ponteineptique
Github: @ponteineptique

TEI Conference 2018, September 13th. #TEI2018 - https://github.com/distributed-text-services/presentations

The Team

- Bridget May Almas
- Hugh Cayless
- Vincent Jolivet
- Emmanuelle Morlock
- Jonathan Robie
- James Tauber
- Jeffrey C Witt
- Pietro Liuzzo
- Matteo Romanello

Plan

1. History of the Project
2. Current options for Digital Edition APIs
3. A new standard? The Distributed Text Services (DTS)
   a. Requirements
   b. Specifications
   c. Demo
4. What's next?
   a. Invitation!
   b. Current tools and perspectives

History of the project

- Simple statement: no community-wide accepted standard for sharing text over HTTP queries.
- IIIF is a success!
- December 2015: meeting at Tufts with European and American colleagues
- Decision taken: we need to build something

Current API Options

- Three Options:
  - Custom APIs
  - OAI-PMH with links to some TEI
  - Canonical Text Services Protocol and Identifiers

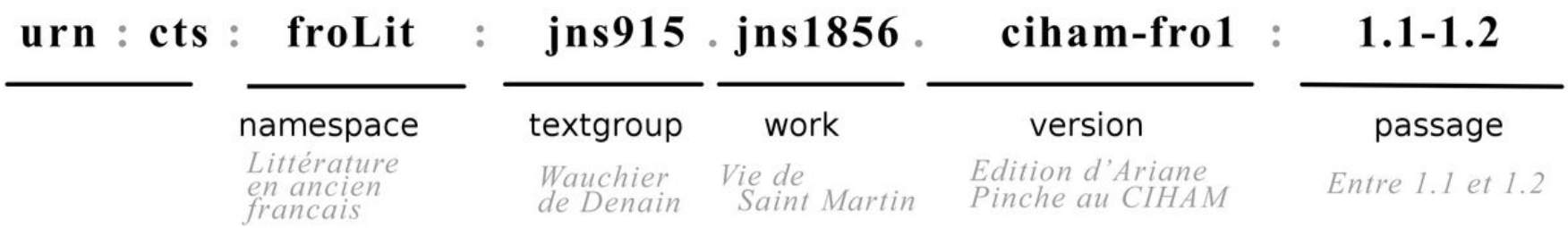
Current API Options (1): Custom APIs

Pros +++	Cons ---
• Fits the project perfectly • Quick to design	• A client needs to adapt to every single custom API • Data exchange is complicated • Interoperability ~ 0 • Documentation needs to be really clear

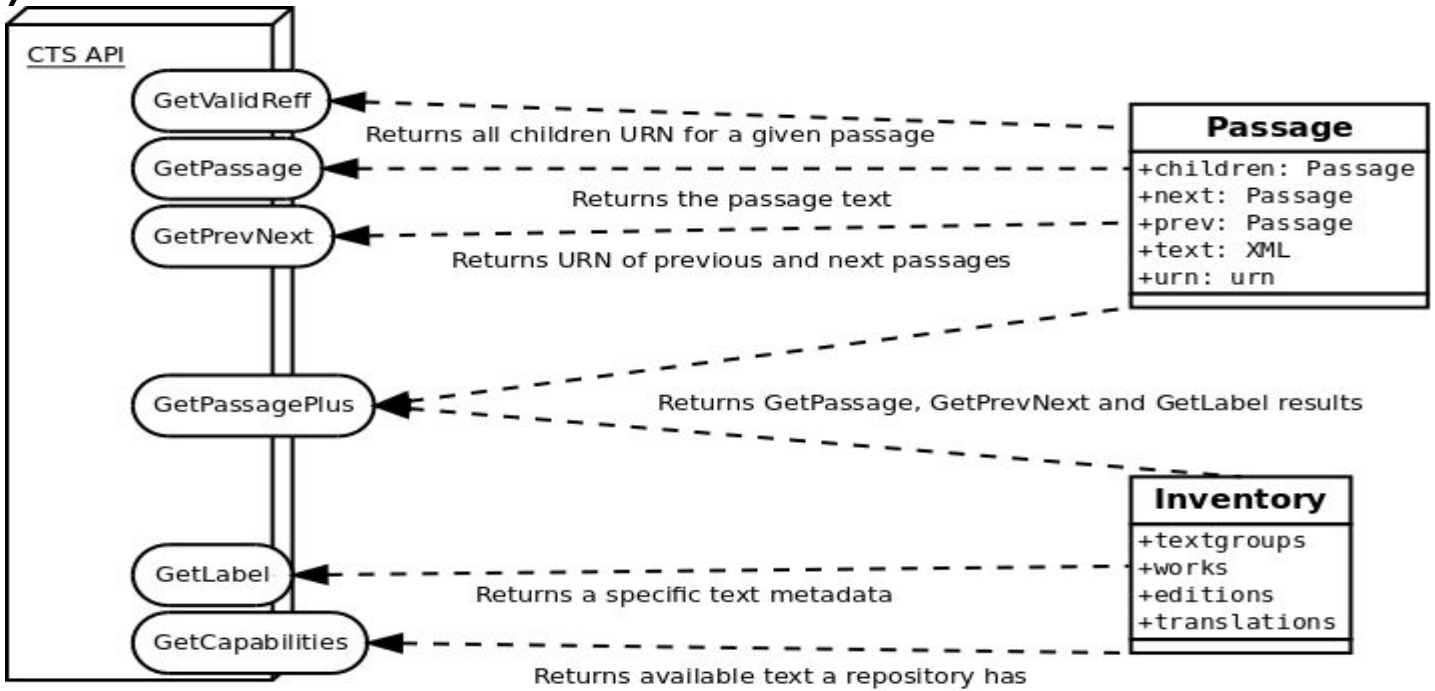
Current API Options (2): OAI-PMH

Pros +++	Cons ---
• OAI-PMH can be harvested by many services and reference data portals like Isidore in France: it's really common! • Cataloging with nice limited set of Metadata!	• The text can only really be a download link • Dublin Core "only" • Not exactly designed for this purpose (more general)

Current API Options (3): Canonical Text Services (CTS)



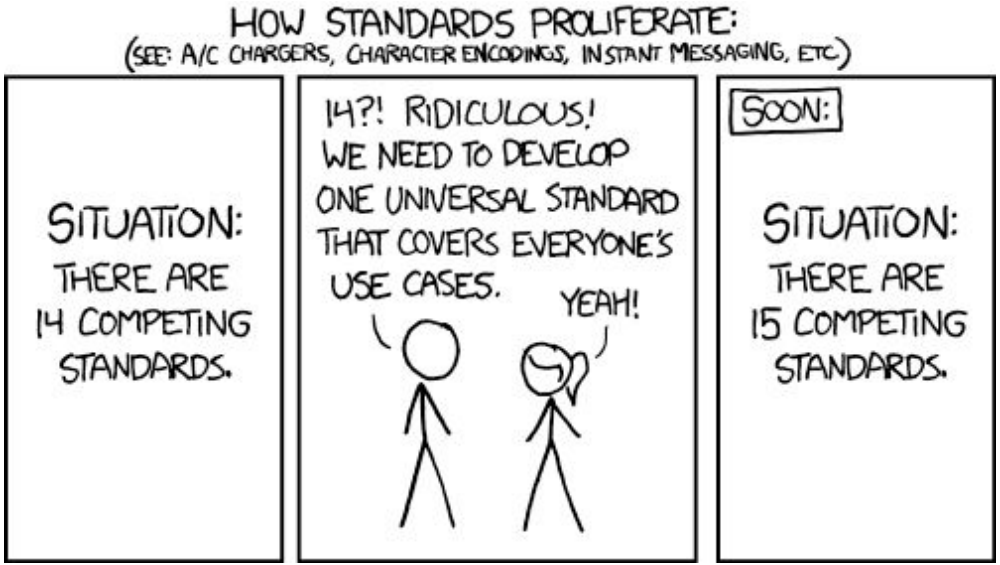
Current API Options (4): Canonical Text Services (CTS)



Current API Options (5): Canonical Text Services (CTS)

Pros +++	Cons ---
• Catalog is simple to read and understand • FRBR is enforced! • Retrieve passages and full text! • Retrieve list of passage identifiers!	• Some specific form of FRBR is enforced. • Tightly coupled to text identifier syntax • No pagination of catalogs. • Redundant response format. • Strong ontological requirements for a canonical citation scheme • Project Specific and as such not open.

A new standard: the Distributed Text Services



There was not 14 standards, so we agreed that one would be good.

Distributed Text Services: Requirements

1. The Distributed Text Services API should be able to provide a **catalogue of its collections.**
2. The DTS API should **not enforce any hierarchy of how a text collection is formed.**
3. The DTS API should support **any kind of valid URIs as identifiers** for texts and collections.
4. The DTS API should support **both partial and complete text retrieval.**
5. The DTS API should **reuse accepted standards** wherever possible.
6. The DTS API should be **flexible**, but specify **technical requirements** that can be **trusted** and **enforced**.
7. The DTS API should offer a solution compatible with the modern technological landscape using standard approaches to describing and navigating Web Service requests and responses.

Distributed Text Services: Endpoints

- *First Public Working Draft*
- 3 API Endpoints
  - Collection Endpoint : Catalogue
  - Navigation Endpoint : Index of passages
  - Document Endpoint : Text retrieval
- Optional endpoints
- 3 available draft implementations
  - Capitains Tool Suite
    - http://dev.chartes.psl.eu/api/nautilus/dts
    - http://texts.alpheios.net/api/dts
  - Project Specific
    - http://betamasaheft.eu/api/dts

http://w3id.org/dts

Distributed Text Services: Entry Point

Example URI : http://dev.chartes.psl.eu/api/nautilus/dts

```
{
  "@context": "/dts/api/contexts/EntryPoint.jsonld",
  "@id": "/dts/api/",
  "@type": "EntryPoint",
  "collections": "/dts/api/collections",
  "documents": "/dts/api/documents",
  "navigation" : "/dts/api/navigation"
}
```

Distributed Text Services: Collection Endpoint

Name	Description	Accepted value
id	Identifier for a collection or document	Any
page	Page	Integers
nav	Direction of the navigation	Children (default), parents

Distributed Text Services: Collection Endpoint


```
{
    "@context": { "@vocab": "https://www.w3.org/ns/hydra/core#",
      "dc": "http://purl.org/dc/terms/", "dts": "https://w3id.org/dts/api#"
    },
    "@id": "general", "@type": "Collection",
    "totalItems": 2, "title": "Collection Générale de l'École Nationale des Chartes",
    "dts:dublincore": {
        "dc:publisher": ["École Nationale des Chartes", "https://viaf.org/viaf/167874585"],
        "dc:title": [{"@lang": "fr", "@value": "Collection Générale de l'École Nationale des Chartes"}]
    },
    "member": [
        {
             "@id" : "lasciva_roma", "title" : "Lasciva Roma",
             "description": "Collection of primary sources of interest in the studies of Ancient World's sexuality",
             "@type" : "Collection", "totalItems" : 1
        },
        {
             "@id" : "lettres_de_poilus", "title" : "Correspondance des poilus",
             "description": "Collection de lettres de poilus entre 1917 et 1918",
             "@type" : "Collection", "totalItems" : 10000
        }
]}
```

Distributed Text Services: Navigation Endpoint

Name	Description
id	Identifier for a document
ref	Page
start	Start of a passage range
end	End of a passage range
groupBy	Ask the server to group references into range
page	Current page
max	Maximum number of references

```
{
    "@context": {
        "@vocab": "https://www.w3.org/ns/hydra/core#",
"dts": "https://w3id.org/dts/api#"
    },

"@id":"/api/dts/navigation/?id=urn:cts:greekLit:tlg0012.tlg001.o
pp-grc",
    "dts:citeDepth" : 2, "dts:level": 1,
    "member": [
      {"ref": "1"},
      {"ref": "2"},
      {"ref": "3"}
    ],
    "dts:passage":
"/dts/api/document/?id=urn:cts:greekLit:tlg0012.tlg001.opp-grc{&
ref}{&start}{&end}"
}
```

Distributed Text Services: Document Endpoint

Name	Description
id	Identifier for a document
ref	Page
start	Start of a passage range
end	End of a passage range

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
   <teiHeader>
      <fileDesc>
         <titleStmt>
            <title>bgu.11.2029</title>
         </titleStmt>
         <publicationStmt>
            <authority>Duke Collaboratory for Classics Computing
             (DC3)
            </authority>
            <idno type="filename">bgu.11.2029</idno>
         </publicationStmt>
      </fileDesc>
   </teiHeader>
   <dts:fragment xmlns:dts="https://w3id.org/dts/api#">
      <lb n="1"/><expan>τετελ<ex>ώνηται</ex></expan>
      <expan>δι<ex>ὰ</ex></expan> <expan>πύλ<ex>ης</ex></expan> Διονυσιάδος
   </dts:fragment>
</TEI>
```

Distributed Text Services: Demos !

1. Jupyter Notebook Demo on the three implementations
    a. https://github.com/distributed-text-services/implementation-demos/blob/master/Beta%20masaheft.ipynb
    b. https://github.com/distributed-text-services/implementation-demos/blob/master/Alpheios.net.ipynb
    c. https://github.com/distributed-text-services/implementation-demos/blob/master/Ecole%20des%20Chartes.ipynb
    d. https://github.com/distributed-text-services/implementation-demos
2. Julien Pilla's Browsing Interface
    a. http://dev.chartes.psl.eu/dts-demo/

What's next (1)

- Join us!
  - https://github.com/distributed-text-services/specifications/issues are open!
  - We review your implementation
- It's a First Draft, we need you to make it better!
- Ask questions, propose fixes to the documentation!

What's next (2)

- Work on server implementations with the Capitains.org suite
  - Decoupled from CTS identifiers to offer more choices
- Work on a client for the Capitains Suite to interact with an API
- SCTA implementation
- Fixes to current implementations!
- More?

Thanks!

Don’t forget the link: https://w3id.org/dts

-> This is the one to share because it’s a permanent one!

Thanks again to the team that met specifically really often in the last academic year to finish this work.

Thanks to Pelagios for funding a 2018 June workshop where we were able to have big discussions about Citation Scheme and other important ontological questions!