**New Version of Mapping - 11-5-17 - Chronologies**

|  |  |
| --- | --- |
| **TEI** | **schema** |
| div0 @id | schema:Dataset {@id/@type for json+ld} |
|  | schema:author <<http://viaf.org/viaf/44300868>> |
|  | schema:inLanguage “fr” |
| ->head->date @value | schema:temporalCoverage [schema:DateTime] |
| ->div1->p->name  ->div1->note->name | schema:mentions [schema:Person] |
| ->div1->p->title  ->div1->note->title | schema:mentions [schema:CreativeWork]\*\* |
| ->div1->(listBibl)->bibl | schema:citation [schema:CreativeWork]\*\* |
| ->div1->bibl->title@type=”es”, @type=”re”,  or @level=“a” | schema:headline [schema:Text] |
| ->div1->bibl->title [all other cases] | schema:name [schema:Text] |
| ->div1->(p)->bibl->author | schema:author [schema:Person] |
| ***\*\*If the bibl contains a second title then*** |  |
| ***---If @level=”j”*** | schema:isPartOf [schema:PublicationIssue] |
|  | schema:dateCreated [schema:DateTime] |
| ->div1->bibl->title [2nd title node] | schema:name [schema:Text] |
|  | schema:publisher [schema:Organization or schema:Person] |
|  | schema:datePublished [schema:DateTime] |
| ->bibl->biblScope@type=”vol”  # biblScope might not exist on Chronologies. | schema:volumeNumber [schema:Text or schema:Integer] |
| ->bibl->biblScope@type=”issue” | schema:issueNumber [schema:Text or schema:Integer] |
| ->bibl->biblScope@type=”pages” | schema:pageStart [schema:Integer or schema:Text] |
| ->bibl->biblScope@type=”pages” | schema:pageEnd [schema:Integer or schema:Text] |
| ***---All other cases of 2nd title node*** | schema:isPartOf [schema:PublicationVolume] |
|  | schema:publisher [schema:Organization or schema:Person] |
|  | schema:datePublished [schema:DateTime] |
|  | schema:volumeNumber [schema:Text or schema:Integer] |
|  | schema:issueNumber [schema:Text or schema:Integer] |
|  | schema:pageStart [schema:Integer or schema:Text] |
|  | schema:pageEnd [schema:Integer or schema:Text] |

\*\* Map bibl to more specific schema:CreativeWork type when title has @type or @level values according to the table below.

**New Version of Mapping - 11-5-17 - Bibliographies**

|  |  |
| --- | --- |
| **TEI** | **schema** |
| TEI@id=”id” | schema:Dataset {@id/@type for json+ld} |
| TEI->teiHeader->fileDesc->titleStmt->title | schema:name [schema:Text] |
| TEI->teiHeader->fileDesc->titleStmt->respStmt->name | schema:author <<http://viaf.org/viaf/44300868>> |
| TEI->teiHeader->fileDesc->editionStmt->edition->date->@when | schema:dateCreated [schema:DateTime] |
| TEI->text->body->div1->head->date->@when | schema:temportalCoverage [schema:DateTime] |
| TEI->text->body->div2->p->title  TEI->text->body->div2->note->title | schema:mentions [schema:CreativeWork]\*\* |
| TEI->text->body->div2->p->name  TEI->text->body->div2->note->name | schema:mentions [schema:Person] |
| TEI->text->body->div2->bibl  TEI->text->body->div2->p->bibl  TEI->text->body->div2->note->bibl | schema:citation [schema:CreativeWork]\*\* |
| ->div2->bibl->title@type=”es”, @type=”re”,  or @level=“a” | schema:headline [schema:Text] |
| ->div2->bibl->title [all other cases] | schema:name [schema:Text] |
| ->div2->bibl->author->name | schema:author [schema:Person] |
| ***\*\*If the bibl contains a second title then*** |  |
| ***---If @level=”j”*** | schema:isPartOf [schema:PublicationIssue] |
| ***---If @level=”m”*** | schema: locationCreated [schema:Place] |
|  | schema:dateCreated [schema:DateTime] |
| ->div2->bibl->title [2nd title node] | schema:name [schema:Text] |
| ->bibl->publisher | schema:publisher [schema:Organization or schema:Person] |
| ->div2->date->@when | schema:datePublished [schema:DateTime] |
| ->bibl->biblScope@type=”vol”  # biblScope might not exist on Chronologies. | schema:volumeNumber [schema:Text or schema:Integer] |
| ->bibl->biblScope@type=”issue” | schema:issueNumber [schema:Text or schema:Integer] |
| ->bibl->biblScope@type=”pages” | schema:pageStart [schema:Integer or schema:Text] |
| ->bibl->biblScope@type=”pages” | schema:pageEnd [schema:Integer or schema:Text] |
| ***---All other cases of 2nd title node*** | schema:isPartOf [schema:PublicationVolume] |
| ->bibl->publisher | schema:publisher [schema:Organization or schema:Person] |
|  | schema:datePublished [schema:DateTime] |
| ->bibl->biblScope@type=”vol” | schema:volumeNumber [schema:Text or schema:Integer] |
| ->bibl->biblScope@type=”issue” | schema:issueNumber [schema:Text or schema:Integer] |
| ->bibl->biblScope@type=”pages” | schema:pageStart [schema:Integer or schema:Text] |
| ->bibl->biblScope@type=”pages” | schema:pageEnd [schema:Integer or schema:Text] |

\*\* Map bibl to more specific schema:CreativeWork type when title has @type or @level values according to the table below.

**Title Level/Type mappings to sub-types of schema:CreativeWork**

|  |  |
| --- | --- |
| title@level="a" | schema:Article |
| title@level="j" | schema:Periodical |
| title@level="s" | schema:Series |
| title@level="m" | schema:Book |
| title@type="es" (prose nonfiction) | schema:CreativeWork |
| title@type="ag" (graphic arts) | schema:VisualArtwork |
| title@type="re" (prose fiction) | schema:CreativeWork |
| title@type="th" (theater) | schema:CreativeWork; schema:sameAs scp:StageWork |
| title@type="sc" (sculpture) | schema:Sculpture |
| title@type="sp" (variety show) | schema:CreativeWork; schema:sameAs scp:StageWork |
| title@type="ds" (dance) | schema:CreativeWork; schema:sameAs scp:StageWork |
| title@type="po" (poetry) | schema:CreativeWork |
| title@type="mu" (music) | schema:MusicComposition |
| title@type="op" (opera) | schema:CreativeWork; schema:sameAs scp:StageWork |

**Worked Example in turtle/txt**

<<http://example.org/c20090>> a schema:Dataset ,

   schema:author <<http://viaf.org/viaf/44300868>> ,

   schema:temportalCoverage “1899” ,

   schema:citation <\_:0001> ,

   schema:citation <\_:0002> ,

   schema:citation <\_:0003> .

<\_:0001> a schema:Article ,

    schema:headline “La Peste de Vienne et le danger que peut faire courir à l'Europe la peste du Turkestan.” ,

    schema:isPartOf <\_:0001j> .

<\_:0001j> a schema:PublicationIssue ,

   schema:name “Comptes-rendus des séances de l'Académie des sciences morales et politiques” ,

   schema:issueNumber “59e année” ,

   schema:pageStart 4 ,

   schema:pageEnd 4 .

<\_:0002> a schema:Article ,

   schema:headline “La conférence sanitaire internationale de Venise de 1897” ,

   schema:creationDate “1897” ,

   schema:isPartOf <\_:0002j> .

<\_:0002j> a schema:PublicationIssue ,

   schema:name “Revue d'Hygiène” ,

   schema:issueNumber “XIX” ,

   schema:pageStart 7 ,

   schema:pageEnd 7 .

<\_:0003> a schema:Article ,

   schema:headline “La défense de l'Europe contre la peste.” ,

   schema:creationDate “1897” ,

   schema:isPartOf <\_:0003j> .

<\_:0003j> a schema:PublicationIssue ,

   schema:name “Comptes-rendus des séances de l'Académie des sciences morales et politiques” .

   schema:issueNumber “57e année” ,

   schema:pageStart 4 ,

   schema:pageEnd 4 .

The summary of TEI records in ‘KP TEI Files Combined’ folder is shown below and on <https://uofi.box.com/s/au7tv4rchjzyen75glvdcwcjt43sa2iw>, as well.

|  |  |  |
| --- | --- | --- |
| **Tag** | **Location tree** | **Count** |
| bibl | /TEIList/TEI/text/body/div1/div2/**listBibl/bibl** | 12276 |
| /TEIList/TEI/text/body/div1/div2/**bibl** | 1436 |
| /TEIList/TEI/text/body/div1/div2/**p/bibl** | 184 |
| /TEIList/TEI/text/body/div1/div2/**note/bibl** | 11 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/note/bibl | 10 |
| /TEIList/TEI/text/body/div1/div2/p/note/bibl | 3 |
| /TEIList/TEI/text/body/div1/listBibl/bibl | 2 |
| /TEIList/TEI/text/body/div1/div2/bibl/note/bibl | 1 |
| author | /TEIList/TEI/text/body/div1/div2/listBibl/**bibl/author** | 1468 |
| /TEIList/TEI/text/body/div1/div2/**bibl/author** | 1323 |
| /TEIList/TEI/text/body/div1/div2/p/**bibl/author** | 23 |
| /TEIList/TEI/text/body/div1/div2/note/**bibl/author** | 4 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/note/**bibl/author** | 1 |
| name | /TEIList/TEI/text/body/div1/div2/**p/name** | 25770 |
| ~~/TEIList/TEI/teiHeader/fileDesc/titleStmt/respStmt/name~~ | ~~5312~~ |
| /TEIList/TEI/text/body/div1/div2/listBibl/**bibl/name** | 3984 |
| /TEIList/TEI/text/body/div1/div2/bibl/**author/name** | 1670 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/**author/name** | 1503 |
| /TEIList/TEI/text/body/div1/div2/**p**/hi/**name** | 157 |
| /TEIList/TEI/text/body/div1/div2/**p/note/name** | 112 |
| /TEIList/TEI/text/body/div1/div2/p/title/name | 68 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/**note/name** | 59 |
| /TEIList/TEI/text/body/div1/div2/**bibl/name** | 40 |
| /TEIList/TEI/text/body/div1/div2/**note/name** | 29 |
| /TEIList/TEI/text/body/div1/div2/p/**bibl/name** | 26 |
| /TEIList/TEI/text/body/div1/div2/p/**bibl/author/name** | 25 |
| /TEIList/TEI/text/body/div1/div2/lg/l/name | 21 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/title/name | 19 |
| /TEIList/TEI/text/body/div1/div2/p/name/name | 11 |
| /TEIList/TEI/text/body/div1/div2/bibl/respStmt/name | 6 |
| /TEIList/TEI/text/body/div1/div2/note/bibl/author/name | 6 |
| /TEIList/TEI/text/body/div1/head/note/name | 6 |
| /TEIList/TEI/text/body/div1/div2/bibl/note/name | 5 |
| /TEIList/TEI/text/body/div1/div2/note/bibl/name | 4 |
| /TEIList/TEI/text/body/div1/p/name | 3 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/note/bibl/author/name | 2 |
| /TEIList/TEI/text/body/div1/div2/bibl/title/rs/name | 1 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/name/name | 1 |
| /TEIList/TEI/text/body/div1/div2/p/emph/name | 1 |
| /TEIList/TEI/text/body/div1/div2/p/note/bibl/name | 1 |
| title | ~~/TEIList/TEI/teiHeader/fileDesc/titleStmt/title~~ | ~~10044~~ |
| /TEIList/TEI/text/body/div1/div2/listBibl/**bibl/title** | 8183 |
| /TEIList/TEI/text/body/div1/div2/**p/title** | 6190 |
| /TEIList/TEI/text/body/div1/div2/**bibl/title** | 2514 |
| /TEIList/TEI/text/body/div1/div2/**p/bibl/title** | 171 |
| /TEIList/TEI/text/body/div1/div2/**note/title** | 162 |
| /TEIList/TEI/text/body/div1/div2/**p/note/title** | 132 |
| /TEIList/TEI/text/body/div1/div2/p/title/title | 30 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/note/title | 28 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/note/bibl/title | 17 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/title/title | 13 |
| /TEIList/TEI/text/body/div1/div2/listBibl/bibl/note/bibl/title/title | 10 |
| /TEIList/TEI/text/body/div1/div2/**note/bibl/title** | 9 |
| /TEIList/TEI/text/body/div1/div2/note/p/title | 7 |
| /TEIList/TEI/text/body/div1/p/title | 7 |
| /TEIList/TEI/text/body/div1/div2/bibl/note/title | 5 |
| /TEIList/TEI/text/body/div1/div2/p/hi/title | 5 |
| /TEIList/TEI/text/body/div1/div2/lg/l/title | 3 |
| /TEIList/TEI/text/body/div1/div2/p/title/rs/title | 3 |
| /TEIList/TEI/text/body/div1/div2/bibl/note/bibl/title | 2 |
| /TEIList/TEI/text/body/div1/div2/bibl/note/bibl/title/title | 2 |
| /TEIList/TEI/text/body/div1/div2/p/bibl/note/title | 2 |
| /TEIList/TEI/text/body/div1/head/note/title | 2 |
| /TEIList/TEI/text/body/div1/listBibl/bibl/title | 2 |
| /TEIList/TEI/text/body/div1/div2/bibl/author/name/title | 1 |
| /TEIList/TEI/text/body/div1/div2/bibl/biblScope/title | 1 |
| /TEIList/TEI/text/body/div1/div2/bibl/title/rs/title | 1 |
| /TEIList/TEI/text/body/div1/div2/p/name/title | 1 |
| /TEIList/TEI/text/body/div1/div2/p/note/bibl/title | 1 |
| /TEIList/TEI/text/body/div1/div2/p/note/rs/title | 1 |
| /TEIList/TEI/text/body/div1/div2/p/rs/title | 1 |
| date | /TEIList/TEI/text/body/div1/head/date | 9503 |
| /TEIList/TEI/text/body/div1/div2/bibl/date | 1369 |
| ~~/TEIList/TEI/teiHeader/fileDesc/editionStmt/edition/date~~ | ~~1328~~ |
| biblScope | /TEIList/TEI/text/body/div1/div2/bibl/biblScope | 1425 |
| pubPlace | /TEIList/TEI/text/body/div1/div2/bibl/pubPlace | 117 |
| publisher | ~~/TEIList/TEI/teiHeader/fileDesc/publicationStmt/publisher~~ | ~~1328~~ |
| /TEIList/TEI/text/body/div1/div2/bibl/publisher | 158 |

**Feb. 20, 2017**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

<div0 id=”card no.”> -**schema:CreativeWork**

<head>

<date> **schema:dateCreated**

<div1> and <p> **schema:Text**

<bibl> **schema:Article (When title level=”a”)**

**<title> schema:name**

**<author><name>schema:author ->schema:Person**

**SHOULD CAPTURE schema:datePublished, schema:pageEnd, schema:pageStart IF THE VALUE IS AVAIABLE.**

<bibl> **schema:Series (When title-level=”s”)**

<bibl> **schema:Book (When title-level=”m”)**

**Further discussion:**

1. BUT schema:book and schema:Series do not support the pagination.  we may use schema:citation for everything, then use only <title> to appropriate types of the schema:CreativeWork?

2. How to deal with <ListBibl>

3. How to work with Title type and Title level.

1. Difference between level=p and level=s?
2. Value of title types?

TEI is fundamentally different from other metadata standards. TEI is an encoding schema for text analysis. It is designed for preserving structure of the text and encode specific parts included in the text while others are created for/used for encode the 'metadata,' I.e., it describes 'a thing' rather that 'a part' of the text. And this makes the mapping of TET document to schema.org semantics.

Our approaches of mapping the K-P collection's metadata are -

1. Identify entities that can be represented with schema.org semantics, such as personal names, place names, titles of the article/journal/book, events
2. Identify text structures that can be represented with schema.org semantics, such as <p> with <schema:text> and <bibl> with <schama:citation>.

Feb. 22, 2017

For Chronology ones –

1. Using the date included in TEI header, treat it as a schema:event?
   1. Then it is difficult to identify a correct specific Types of the Event
   2. Event can be connected to the card in bibliography by using <schema:workFeatured>?
2. Use each <bibl> as a <schema:creativeWork>?
3. Names included in <P> and Names included in <bibl>
   1. Names in <P> - meed roles – how? For example?

<div0 id="c20120">

<head>

<date value="18990100">1899 janvier</date>

</head>

<div1>

<p>Révélations de M. Quesnay de Beaurepaire [Jules ].</p>

<p>"À la suite des révélations de

<name key="quesna1">M. Quesnay de Beaurepaire</name> en janvier 1899, le gouvernement se décida, non sans peine, à préscrire une enquête sur les agissements de la Chambre criminelle. Cette enquête... se termina par un rapport où ces trois magistrats [<name key="mazeau1">Mazeau</name>, <name key="darest1">Dareste</name>, <name key="voisin2">Voisin</name>] déclaraient que les membres de la Chambre criminelle n'avaient plus le 'calme et la liberté morale indispensable pour faire l'office de juges'."</p>

* + 1. socialEvent (if we know!) ->using [**attendee**](https://schema.org/attendee)?
    2. PublicationEvent ->
    3. Otherwise Just use <Event>?

<schema:Event>

<schema:startDate> --from header <date>

<schema:description> -- text from the <P>

<schema:attendee> - names un <P>

<schema:recordedIn> - from <bibl>

<schema:CreativeWork>

<schema:name>

<schema:author>

<

* 1. Names in <bibl>

1. y

For Bibliography ones –

1. Use the schema:creativeWork

Should look at patterns (Melina) – 2 weeks?!

1. If there is a clear patterns to distinguish publication event
2. Caroline can provide Melina sample records (Chronology ones about 20 records)
3. Melina can record what kinds of event she can find (https://schema.org/Event):
   1. BusinessEvent
   2. ChildrensEvent
   3. ComedyEvent
   4. DanceEvent
   5. DeliveryEvent
   6. EducationEvent
   7. ExhibitionEvent
   8. Festival
   9. FoodEvent
   10. LiteraryEvent
   11. MusicEvent
   12. PublicationEvent
   13. SaleEvent
   14. ScreeningEvent
   15. SocialEvent
   16. SportsEvent
   17. TheaterEvent
   18. VisualArtsEvent
4. Jacob and Jooho will create a mapping table
5. Next step will be analyzing Bibliography cards – Caroline and Melina
6. \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

May 3, 2016.

1. Name database includes fictitious character names - Caroline will look at all names in the database entries to identify these names
   1. Caroline will also add additional information??
   2. These may have entries in VIAF and Wikipedia
2. Name:
   1. When used with <author>
   2. When used with ‘to’(à)
   3. Plain names included in the texts
   4. Unpublished conversation
   5. Add one more column for any linked data source URIs
   6. How to present linked data sources -additional contextual information (How about title of article, books and journals??)
3. Title
   1. By title:level - Journals, books, series, article:
      1. Janina will generate unique titles
      2. GA can find links for digitized journal Melina or someone else?

B. how to work with title:type?

1. Named entity database is in external server?
   1. When the display is built - how can we fetch the external information, search, display, facet, etc.
2. Jacob will see how to represent linked data uris in HTML/TEI
3. MJ will contact Alex/Ryan for Alex’s schedule for summer
4. Peter will work with JAcob for a mock-up of Proust collection for next meeting?
5. For oclc webinar:
   1. Motley: metadata mapping and mock-up page
   2. Proust collection: TEI to RDFa mapping in HTML page, and a mock-up

**Summer Schedule:**

* Tim: May 13-21, June 11-18
* Caroline: May 31-June 15
* Jacob: May 27-May 31 (or June 2), June 19-23
* Janina: May 24- 26
* MJ: May 19- June 7
* Alex: June/July

MJ - April 8, 2016

1. All **card** **entry <div0>** is represented under <schema:CreativeWork>
2. Associated **@id** is mapped to <schema:recordedAt>
3. Associated **<head><date>** should be mapped to <schema:dateCreated>
   1. (Question should be whether to use the <**date value**="18990102"> or string value.
   2. It would be consistent is we can use <date value="18990102"> - first four digits to yyyy, next two digitis to mm, and last two to dd.
4. All **div1** should map to <schema:text>
   1. ***Should we preserve <p>, <bibl> structure??***
5. All **name** is used for a person who is mentioned in the card, without any specific role. -
   1. All name key is mapped to **foaf: Agent->schema:Person**
   2. Person allows firstName and lastName
6. All title type will be mapped to appropriately according to its value -
   1. es=prose nonfiction
   2. ag=graphic arts (typeOf=”schema:VisualArt”)
   3. re= prose fiction
   4. th=theater (schema:StageWork (or our own subclass thereof))
   5. sc=sculpture
   6. sp=variety show
   7. ds=dance
   8. po= poetry
   9. mu=music
   10. op=opera
7. All title level will be mapped to appropriately according to its value -
   1. a=article (schema:Article)
   2. j=periodical (schema:Periodical)
   3. m=monograph (schema:Book) [what about schema:PeriodicalVolume?]
   4. s=serial - maybe series (schema:Periodical) [but might be schema:Book]

**Meeting April 25th -**

1. When there is a title without type or level?
2. For title level: we can simply reconcile title with any linked data sources?
3. Is there any master list of title - Janina will generate the list
4. There are some Abb. titles. - can we make it as a full title by magic? For example ‘Cor’ can be changed into --- collections of letters?? Caroline can provide a list?
5. All TEI structural elements can be used in the RDFa as well.
6. All title can be used for periodicals, publicationVolume, publicationIssue, or other detailed semantics.
7. Mapping for title=type will be lossy - for example, there are sculpture and graphic arts, variety show and dance, opera - not all have matching semantics
8. Jacob will create and share sample mappings with the group
9. What kind of manual or automatic clean up should be done?
   1. Can generate all titles without level or type
   2. Can find author without name tag
   3. Identify certain consistent patterns of string values (volumes, year, pages…)

|  |  |
| --- | --- |
| TEI | schema |
| Div 0->head->date @value | n/a it could be used for standard encoding later |
| Div 0->head->date | dateCreated |
| div 0->div 1-> bibl | text |
| title |  |
| @type= in Title |  |
| p |  |
| @Level= in Title |  |
| When @Level=a | CreativeAWork->article |
| author |  |
| name |  |
| key= |  |

The following email in the TEI archives might be relevant: https://listserv.brown.edu/archives/cgi-bin/wa?A2=ind1604&L=TEI-L&P=R10334&1=TEI-L&9=A&J=on&d=No+Match%3BMatch%3BMatches&z=4

Color Key:

Done

Skipped

Done But Have Questions About