

Digitizing Charters with XML – TEI-Proposal

Charters, deeds, wills, writs etc. are major sources for the history of the middle ages and early modern times, for political, legal, cultural or linguistic research. There have been many attempts to transfer medieval charters to the digital world. Since the 1990s scholars in the field of diplomatics try to build digital corpora. In April this year scholars representing the majority of attempts using XML for this purpose met in Munich to talk about the possibility of bringing these projects together. They – i.e. we - built a working group called “Charters Encoding Initiative” (<http://www.cei.lmu.de>). One result of our work is to propose the following extensions to the TEI as we regard the TEI the most promising standard in digitising texts in the humanities.

Suggestions for Changes:

Going through the existing guidelines we found some point we would like to suggest some changes:

transcription of primary sources

In the chapter on the transcription of primary sources we would be happy to have an element `<textwitness>` in the `<witnesslist>` instead of `<witness>` as witness can also be a person notified in a charter who witnesses the legal content.

As the text from charters is often known by the original not a copy as the majority of literary texts it would be a good thing to mark the the type of textual tradition in `<traditioForm>`.

Suggestions for extensions

Manuscript description

In the chapter on manuscript-description we miss a proper structure to mark up archival material. We looked at the EAD for this purpose but found rather broad terms as the EAD has to deal with a great variety of material and archival structures. Our proposal would look more like this:

- `<Archive>` to name the archive by a full name or by some common abbreviation
- `<archivalFond>` to name the section in the archive. An `<archivalFond>` might be part of itself comparable to the `<c>` (Component) element in the EAD, but we see the `<archivalFond>` principally as a collection of single documents/files.
- `<refNumber>` The single item inside such fond is identified by a reference number. This could be the single charter, a chartulary containing the charter, a file containing other material alongside the charter.
- `<scope>` information on the location inside the item, mostly folios.

Types

type-attribute

`<text type="charter">`

It is crucial for an exchange between the different approaches to the material that the individual charter can be defined as a basic unit. A dedicated element `<document>` seemed necessary but we considered it best to stay near to the TEI guidelines and take each charter as

single <text>. That text has to be identified as a text of specific genre. Thus we propose to introduce the type-attribute to the <text> element.

<listBibl type="facsimilia|prints|regesta|studies">

Similarly we would like to suggest the extension of the <listBibl> element with a type-attribute, as the mentioned types of lists occur in scholarly editions of charters as distinct parts often without any explicit indication of their scope.

The single charter

The single charter will be divided in to two sections: one for the metadata and another for the text of the individual charter. <chDesc> (charterDescriptoin) and <tenor>.

In existing corpora of charters it is usual to identify each document by a number that could be encoded in an <idno>-Element which is now restricted to ISBN.

<chDesc>

```
<abstract>
  <issuer>, <issuePlace>, <issueDate>, <recipient>
<diplomaticAnalysis>
  <class>, <auth>, <nota>, <facsimilia>, <prints>,
  <regesta>, <studies>
```

taken from TEI: <physDesc>, <witList> with the suggested changes.

Since the charter is identified using issuer, recipient, place and date of issue, markups should be available for these elements: <issuer>, <issuePlace>, <issueDate>, <recipient>. These identifications could be part of the summary and could be encoded with the element <abstract> taken over from the EAD.

The diplomatic comment <diplomaticAnalysis> is a completely new area of diplomatic metadata. It consists of references to illustrations, other editions, and studies on the charter. It also includes a classification of the document and an attempt at putting it in its historical context. Therefore there are elements <class> for an classification, and the <listBibl> with the following values for a type-attribute: facsimilia, prints, regesta, studies.

Texts become legally relevant charters through different means of authentication. That is why the type of authentication used is an important part of the metadata and should be summarised as a group in the element <auth>entication. This element could contain some specialized elements describing the means of authentication, like <seal>, <notariusSign>, <notariusSub>.

Charters are produced by a bureaucracy and are kept by a bureaucracy. After the writing down of the actual text both bureaucracies add a large number of remarks to the charter. Iussio, tax annotations, registration notes, archiving marks of different times could all be subsumed under a class of texts outside the tenor and described in a <nota> element.

<tenor>

```
<protocol>
<context>
<eschatocol>
<invocatio>, <intitulatio>, <inscriptio>
<arenga>, <narratio>, <publicatio>, <petitio>, <dispositio>,
<sanctio>, <corroboratio>, <testes>
<subscriptiones>, <datatio>, <apprecatio>
```

A specific characteristic of medieval and early modern charters, however, is their linguistic structure. In the chanceries and in legal practice more or less firmly established wordings and set phrases for certain parts of the text emerged. These formalised elements of documents and the three main sections "protocol", "context", "eschatocol" as well as the different parts ranging from "invocatio", "intitulatio", "arenga", over "dispositio", "corroboratio" to "subscriptiones", "datatio", and "apprecatio" on the **one** hand reflect contemporary attitudes and mindsets as regards legal and representation issues and, on the other hand, are tools of diplomatic criticism. This textual structure should be encoded as well as individual formulae that were taken on from collections or became firmly established in legal practice.

While the classical charter form could be mapped using dedicated elements, the individual formulae and clauses such as the pertinents-clause or the non-obstante-clause should be put together in one class. In this regard it must be checked if the TEI elements <cl>, <phr>, or <seg> will suffice and which of them should be used.

In the following you find the long descriptions of the elements our working group suggest for the <tenor>. The function of most of them is best described by the Vocabulaire international de la diplomatie (ed. Maria Milagros Cárcel Ortí, 1997) of the Commission Internationale de Diplomatie in the Comité International des Sciences Historiques.

Elements anywhere in the <tenor>

<insert>

marks text passages containing text of a document inserted in another document (e.g. vidimus, transsumptum, "Vorurkunde").

<elongata>

marks parts of the text in the script the CID describes as follows:

*Les lettres allongées (lat.: litterae elongatae), caractérisées par leur élongation par rapport aux autres lettres de l'acte par le resserrement général de l'écriture, sont utilisées pour certains éléments du texte (*invocation, *subscription, première ligne, *souscriptions ou *signa, date ...) dans certaines chancelleries (chancellerie pontificale et carolingienne notamment). Elles peuvent aussi être simplement décoratives dans certains actes de nature solennelle.*

<formula>

any word sequence referring to a specific content. The type of the formula should be given in the TYPE-attribute using the semantics of the VID.

<nota>

contains the extra sigillum notes on the document:

Example

<nota type="iussio" resp="issuer">per mandatum domini nostri imperatoris in consilio</nota>

<nota type="" resp="adressee">...</nota>

Comment

The vocabulary in the type-Attribute should be taken out of the thesaurus provided by the Vocabulaire internationale de la diplomatie or the Lexicon of archival terminology.

Possible attributes

Date: gives the date when the text of an element is written.

id

lang

place: indicates where the text of the element is written contains, like 'left side under the plice',

'verso' ...

resp: (responsibility) signifies the responsible person for the text in the element, e.g. editor, transcriber, archivist of the 16th century, chancellor ...

type

<protocol>

marks the initial formal part of a document described by the CID as follows:

Le text s'insère dans un cadre formel initial (le protocole) et final (l'eschatocole), dont les éléments ne sont pas nécessairement formulés en fonction de l'acte en question, mais qui

répondent aux règles en usage pour un même type de document, compte tenu de son contenu juridique et de sa nature diplomatique.

<invocatio>

marks the part of the text the CID describes as follows:

*L'invocation verbale ou simplement invocation (lat. invocatio) est la formule de dévotion par laquelle s'ouvre le protocole des actes pour que le contenu en soit placé sous la protection divine et éventuellement (ou secondairement) sous le patronage d'un saint, le tout pouvant s'achever par 'Amen'. Elle peut s'accompagner d'une corix ou d'un *invocation figurée.*

<intitulatio>

marks the part of the text the CID describes as follows:

La suscription (lat. intitulatio, superscriptio, persona salutans) est l'élément du protocole qui fait connaître le nom de l'auteur de l'acte écrit et sa titulature. Elle peut prendre un aspect personnel et commencer par le pronom Ego, Nos.

<inscriptio>

marks the part of a document containing the adresse, describe by the CID as follows:

L'adresse (Lat. inscription, persona salutata) est l'élément du protocole par lequel sont indiqués le nom et éventuellement les titres et qualités de la personne (ou des personnes) à qui l'acte est adressé, qu'il soit comme bénéficiaire ou comme exécuteur ou pour simple information.

<context>

marks the part of the document the CID describes as follows:

Le texte proprement dit est la partie de la teneur se rapportant directement à l'acte juridique qui s'y trouve mis par écrit: justification de cet acte, exposé des circonstances qui l'ont provoqué, dispositions, clauses destinées à en préciser la portée et à en assurer l'exécution.

<arenga>

marks the part of the text defined by the CID as follows:

Le préambule (lat. arenga, exordium, proemium) est la partie du texte par laquelle celui-ci est justifié de façon générale par des considérations juridiques, religieuses, morales ou simplement de convenance (telle que 'quoniam labilis est memoria').

<publicatio>

marks the part of a document the CID describes as follows:

La notification (lat.: notificatio, publicatio, promulgatio) est une formule par laquelle ce qui suit est porté à la connaissance.

<narratio>

marks the part of the text the CID describes as follows:

*L'exposé (lat. narratio) est la partie du texte par laquelle sont expliquées les circonstances du *commandement de l'acte, ses raisons, éventuellement les antécédents de l'affaire (*requête, *intervention de tiers, production de pièces, enquêtes, litiges, arguments des parties en cause, etc.)*

<petitio>

marks a part of the document expressing the petition (cfr. the CID: *La trace de la requête peut apparaître dans le texte même de l'acte qui en découle sous la forme ad instanciam, ad petitioem N. (ou une formule analogue).*)

<dispositio>

marks the part of the text described by the CID as follows:

Le dispositif (lat.: dispositio) est la partie fondamentale du texte, par laquelle l'auteur manifeste sa volonté et fait naître l'acte juridique ou en reconnaît l'existence, en déterminant sa nature, sa portée, ses modalités, - éventuellement l'origine de propriété. Le verbe employé dans le dispositif oriente sur la nature juridique de l'acte ('concedimus' = 'concessio', 'vendidimus' = 'venditio' etc.

<sanctio>

marks the part of a document the CID describes as follows:

Les clauses pénales (lat.: sanctio negativa) sont des clauses destinées à assurer l'exécution de l'acte en prévoyant, contre ceux qui ne l'exécuteraient pas ou qui empêcheraient son exécution, des sanctions temporelles et en les menaçant de châtiments spirituels, ou de l'une autre de même valeur.

<corroboratio>

marks the part of the document the CID describes as follows:

*Par la clause de corroboration, parfois assortie d'un ordre de mise par écrit ou d'expédition, (qui peut aussi en être distinct: "hanc chartam scribere jussimus ...") sont annoncés les signes de validatio (*souscriptio: "manupropria subterfirmavimus"; *seings ou signatures: "manibus idoneorum testium signari ..."; sceaux: "de anulo nostro subter signari jussimus, "sigillum nostrum duximus apponendum ..."), en précisant qu'ils sont apposés pour donner validité à l'acte - Soit à titre de preuve de l'existence de celui-ci (corroboration testimoniale ou probatoire: "In cuius rei testimonium ..." - Soit pour lui conférer valeur perpétuelle (corroboration perpétuelle: "Quod ut ratum et stabile in perpetuum permaneat"*

<testes>

the part of the text naming the witnesses, as described by the CID as follows: *La liste de témoins est l'élément de l'acte qui contient les noms des témoins de l'acte écrit ou de l'acte juridique, ou des deux, lesquels constatent l'existence de cet acte et y donnent leur assentiment.*

<eschatocol>

marks the part of the document described by the CID as the final part of some stereotypical clauses:

Le text s'insère dans un cadre formel initial (le protocole) et final (l'eschatocole), dont les éléments ne sont pas nécessairement formulés en fonction de l'acte en question, mais qui répondent aux règles en usage pour un même type de document, compte tenu de son contenu juridique et de sa nature diplomatique.

<datatio>

marks the part of the text the CID describes as follows:

*La formule de date est, dans le texte d'un acte, l'indication qui permet d'en connaître la date. Elle comporte souvent une *date de temps et une *date de lieu. Elle peut être placée en tête de l'acte (date initiale) ou en fin (date finale) ou bien se trouver partagée entre le début et la fin de l'acte. Par exemple, dans les *privileges pontificaux du IXe au XIIe siècle, on trouve une 'petite date' (data brevis) qui comporte une indication sommaire (mois et indiction) dans la formule 'scriptum per ... N.', et une 'grande date' (data longa) dans la formule finale de *reconnaissance au pied de la pièce ('datum per manum ... N.') (Noter que dans les suppliques pontificales le terme de 'parva data' s'applique à la date du dépôt de la supplique). Introduite par les mots 'Datum, Actum, Factum, Scriptum, Confectum' etc. (parfois au féminin et, plus rarement, au masculin) ou par leurs équivalents en langue vulgaire, elle peut se développer en mentionnant le nom du notaire ou du scribe qui a rédigé ou écrit l'acte ('Scriptum per manum ...') ou encore de l'officier de chancellerie qui entend donner ainsi une indication sur la responsabilité qu'il a prise dans l'expédition de l'acte ('Datum per manum ...')*

<apprecatio>

marks the part of the text defined by the CID as follows:

L'appreciation (lat.: apprecatio) est une formule propitiatoire par laquelle, en pendant avec l'incantation initiale, l'acte s'achève. Elle prend souvent la forme 'In Dei nomine, feliciter. Amen'

<notariusSign> and <notariusSub>

Marks the place of the notarial sign and the notarial subscription.

<subscriptio>

contains the text of any kind of subscription (autograph, non-autograph, notarial subscription including the signs as monograms etc.