Diplomatarium Danicum Template

En grænseflade til produktion af TEI P5-opmærkede dokumenter

Thomas Hansen, Det Danske Sprogog Litteraturselskab <th@dsl.dk>

Copyright © 2010 Det Danske Sprog- og Litteraturselskab, Thomas Hansen

Indholdsfortegnelse

Indledende bemærkninger	
1.1 Grundværdier	
1.2 Tegnsætning	2
2 Sådan udfyldes templaten	2
2.1 editorInitials	
2.2 textId	. 3
2.3 revision	. 3
2.4 textCreationTimeEarliest	4
2.5 textCreationTimeLatest	4
2.6 textCreationTimeCertainty	. 4
2.7 textCreationPlace	4
2.8 textCreationPlaceCertainty	. 4
2.9 summaryText	5
2.10 witness	. 5
2.11 samplingMethod	
2.12 textLanguage	10
2.13 text og p	11
2.14 translation og p	13

1 Indledende bemærkninger

Diplomatarium Danicum Template, DDTemplate, er et XML Schema Document¹ designet til produktion af dokumenter til Diplomatarium Danicums tekstbank i Det Danske Sprog- og Litteraturselskabs, DSL, regi. Et *dokument* defineres her som en kombination af metadata, en tekst på et middelalderligt sprog og en oversættelse til moderne dansk.

DDTemplate afspejler i vidt omfang elementer, som også er defineret af *Text Encoding Initiative*, nærmere bestemt i retningslinjerne for standarden TEI P5. Begrundelserne for overhovedet at udfærdige en template har været flere. Først og fremmest har ønsket om bedre kontrol over datainput end muligt med TEI P5 været udtalt, men også med henblik på det daglige arbejde med oprettelse af dokumenter har DDTemplates enkle og flade struktur vist sig lettere at betjene og dermed arbejdsmæssigt mere effektiv end TEI's forholdsvis dybe træstrukturer.

Når et instansdokument af DDTemplate er udfyldt, transformeres dokumentet vha. et XSLT-stylesheet til TEI P5; den faste struktur gør administration af det samlede materiale, såvel DDTemplate-instanser som TEI-dokumenter, lettere.

I øvrigt bemærkes følgende generelle forhold om brugen af DDTemplate:

1.1 Grundværdier

I modsætning til TEI tillader DDTemplate hverken at udelade uanvendte elementer eller at lade elementer være uden indhold. Såfremt information mangler i et felt, skal det således fremgå, om

¹Den nyeste version af dokumentet kan nedlades fra dd-drb-5_rk/xsd/DDTemplate.xsd

informationen mangler, fordi den ikke eksisterer, eller fordi der endnu ikke er taget stilling til, om den eksisterer. Hvor det kan godtgøres, at oplysninger ikke eksisterer, fx fordi et dokument er udstedt uden segl, indsættes for tekstværdier *empty*, i talfelter 0 og i datofelter 1000. Er der derimod endnu ikke taget stilling til feltets endelige indhold indsættes i tekstfelter værdien *nil* og i tal- og datofelter 999999999.

Når et felt udfyldes med værdien *empty*, 0 eller 1000 siger vi, at feltet "slukkes", fordi oplysningerne enten ikke eksisterer, eller er irrelevante og derfor kan udelades.

Tabel 1. Oversigt over defaultværdier i DDTemplate-felter

Værdi	Datatype	Betydning
empty	Tekststreng	Informationen eksisterer ikke
0	Talværdi	Informationen eksisterer ikke
1000	Datoværdi	Informationen eksisterer ikke
nil	Tekststreng	Informationen endnu ubestemt
99999999	Tal- og datoværdier	Informationen endnu ubestemt

Der er tre slags værdisæt til element- og attributværdier:

- 1. lukket sæt der vælges mellem et begrænset antal værdier
- 2. åbent sæt nogler værdier er prædefineret, men andre kan tilføjes
- 3. intensionalt redaktøren bestemmer selv hvad der skal stå

1.2 Tegnsætning

Der sættes generelt ikke slutpunktum i felterne. Kun i elementerne text og translation sætter redaktøren tegn.

2 Sådan udfyldes templaten

2.1 editorInitials

Lukket sæt. I editorInitials-elementet identificeres den redaktør, der medvirker til udfærdigelsen af dokumentet. Elementet tillader følgende seks værdier:

- 1. alk Anders Leegaard Knudsen
- 2. jon Jonathan Adams
- 3. mh Markus Hedemann
- 4. smb Sebastian Møller Bak
- 5. th Thomas Hansen
- 6. vw Vibeke Winge

Elementet kan gentages efter behov. Er der mere end en redaktør, oprettes endnu et element efter det første.

Eksempel 1. Mere end en redaktør

```
<ddTemplate>
    <editorInitials>mh</editorInitials>
    <editorInitials>vw</editorInitials>
    <textId>...
```

2.2 textld

Værdien af elementet textId er et tal med i alt 11 cifre fordelt efter mønstret: *yyyymmddxxx*. Tallet er opbygget, så det afspejler dateringen af dokumentet:

- 1. yyyy årstal mellem 1410-49, begge inklusive
- 2. mm måned mellem 01-12, begge inklusive
- 3. dd dag mellem 01-31, begge inklusive
- 4. xxx et løbenummer mellem 001-999, begge inklusive

Eksempel 2. Dokument 001 fra 1420. 27. november

```
...<editorInitials>vw</editorInitials>
<textId>14201127001</textId>
<revision>...
```



Husk: Er der andre dokumenter med samme dato?

De sidste tre cifre i talværdien for textId-elementet *skal* være unik. Undersøg derfor, om der allerede eksisterer et dokument fra samme dato med slutcifrene *001* eller højere. Der er plads til 999 dokumenter fra samme dag.

2.3 revision

Revision et lukket sæt, som underordner fire elementer til beskrivelse af dokumentets stade i udfærdigelsesprocessen:

- established Metadata, tekst og oversættelse er etableret
- proofFirst Første korrektur er læst
- proofSecond Anden korrektur er læst
- proofThird Tredje korrektur er læst

2.3.1 established, proofFirst, proofSecond, proofThird

Hver af de fire elementer established, proofFirst, proofSecond og proofThird indeholder de to attributter who og when.

2.3.1.1 who

Attributtet who er et lukket sæt med følgende seks muligheder:

- 1. #alk
- 2. #jon
- 3. #mh
- 4. #smb
- 5. #th
- 6. #vw

Initialerne henviser til en i elementet editorInitials anført redaktør.

2.3.1.2 when

Attributtet when indeholder en dato efter mønstret *yyyy-mm-dd*. Datoen angiver, hvornår et dokument er oprettet og/eller korrekturlæst.

Et dokument, der er oprettet og korrekturlæst første gang, kan være udfyldt som i Eksempel 3.

Eksempel 3. revision: Dokument oprettet og korrekturlæst første gang

2.4 textCreationTimeEarliest

Elementet textCreationTimeEarliest indeholder en tekststreng efter mønstret yyyy-mm-dd. Værdien angiver et terminus post quem for udstedelsen af dokumentet.

2.5 textCreationTimeLatest

Elementet textCreationTimeLatest indeholder en tekststreng efter mønstret: yyyy-mm-dd. Værdien angiver et terminus ante quem for udstedelsen af dokumentet. Hvis udstedelsestidspunktet kan angives nøjagtigt, udfyldes feltet med samme værdi som textCreationTimeEarliest, jf. Eksempel 4.

Eksempel 4. Datering af dokument fra 1420. 27. november

```
...</revision>
<textCreationTimeEarliest>1420-11-27</textCreationTimeEarliest>
<textCreationTimeLatest>1420-11-27</textCreationTimeLatest>
<textCreationTimeCertainty>...
```

2.6 textCreationTimeCertainty

Elementet textCreationTimeCertainty er et lukket sæt med fire mulige værdier:

- 1. empty Informationen eksisterer ikke
- 2. high Udstedelsestidspunktet fremgår af brevet
- 3. low Udstedelsestidspunktet fremgår ikke af brevet
- 4. nil Der er ikke taget stilling til udstedelsestidspunktet

2.7 textCreationPlace

Elementet textCreationPlace udfyldes efter behov med et stednavn i moderne form knyttet til udstedelsesstedet, jf. Eksempel 5.

2.8 textCreationPlaceCertainty

Elementet textCreationCertainty er et lukket sæt med fire værdier:

- 1. empty Informationen eksisterer ikke
- 2. high Udstedelsesstedet fremgår af brevet
- 3. low Udstedelsesstedet fremgår ikke af brevet
- 4. nil Der er ikke taget stilling til udstedelsesstedet

Eksempel 5. Dokument udstedt i Roskilde

```
...</textCreationTimeCertainty>
<textCreationPlace>Roskilde</textCreationPlace>
<textCreationPlaceCertainty>high</textCreationPlaceCertainty>
<summaryText>...
```

2.9 summaryText

Elementet summaryText udfyldes efter behov med en regest, der dækker hele tekstens indhold. Nøgleord som *Hansereces*, *vidimerer* og *tingsvidne* bør fremgå.

Eksempel 6. Regest i summaryText med nøgleordet "vidimerer"

```
...</textCreationPlaceCertainty>
<summaryText>Sjællands landsting vidimerer Anders Mortensen
    Peps brev af 1383. 19. april, hvori han sælger sit gods Bjergegård
    samt gods i Karlslunde og Strøby til Eskil Bosen Falk</summaryText>
<witness>
```

2.10 witness

Elementet witness er et åbent sæt, som samler information om et tekstvidne. Elementet, der selv kan gentages alt efter, hvor mange tekstvidner, der bruges i tekstkonstitueringen, underordner som udgangspunkt 16 elementer:

- 1. witnessSigil
- 2. archivePlaceName
- 3. archiveName
- 4. inventoryNumber
- 5. manuscriptName
- 6. manuscriptMaterial
- 7. manuscriptHeight
- 8. manuscriptWidth
- 9. manuscriptPlica

10.conditionDescription

11.layoutDescription

12.handDescription

13.additionsToText

14.sea1

15.witnessHistory

16.bibliographicEntry

2.10.1 witnessSigil

Elementet witnessSigil er et åbent sæt, som har til formål at tildele et sigil til et tekstvidne. Under *A*, *B*, *C*, osv. anføres originale breve, under *Aa*, *Ab*, *Aa1*, *Ab1*, osv. afskrifter. Elementets værdier forstås sådan:

- A er et originalt tekstvidne
- Aa er en afskrift af A
- Aa1 afskrift af Aa
- Aala afskrift af Aal
- Aalal afskrift af Aala

2.10.2 archivePlaceName

Elementet archivePlaceName udfyldes efter behov, fx <archivePlaceName>**Lund**</archivePlaceName>. Bemærk, at traditionen for at spare spalteplads ved at undlade oplysningen om arkivets beliggenhed *ikke* bør overføres til DDTemplate.

2.10.3 archiveName

Elementet udfyldes efter behov, fx <archiveName>Rigsarkivet</archiveName>.

2.10.4 inventoryNumber

Elementet udfyldes efter behov, fx <inventoryNumber>NKR c-2732; tidl. Tune Herred; Reg. 86 A, 5</inventoryNumber>.

Eksempel 7. Identifikation af et originalt manuskript fra Rigsarkivet

2.10.5 manuscriptName

Elementet manuscriptName udfyldes efter behov. Det bruges til kilder som *Langebeks Diplomatarium* og *Lundebogen*, fx <manuscriptName>**Langebeks Diplomatarium**</manuscriptName>. Ved angivelse af et manuskriptnavn, skal redaktionen tage stilling til, om det skal ledsages af en side- eller foliehenvisning. Ofte vil dokumentet ikke have anden betegnelse end den, der bringes i inventoryNumber-elementet, jf. 2.10.4. I så fald slukkes manuscriptName-elementet ved at udfylde med *empty*.

2.10.6 manuscriptMaterial

Elementet manuscriptMaterial er et lukket sæt med fem værdier:

- 1. empty Manuskriptets materiale er irrelevant
- 2. mixed Manuskriptet materialel er både papir og pergament
- 3. *nil* Der er ikke taget stilling til manuskriptets materiale
- 4. paper Manuskriptet er af papir
- 5. parch (parchment) Manuskriptet er af pergament

2.10.7 manuscriptHeight, -Width og -Plica

Elementerne manuscriptHeight, manuscriptWidth og manuscriptPlica udfyldes efter behov med et decimaltal. Da decimaltallet tilhører datatypen xs:decimal defineret af XML Schema, *skal* skilletegnet være '.' (punktum), jf. Eksempel 8.

Eksempel 8. Håndskrift på pergament med mål

```
...</manuscriptName>
<manuscriptMaterial>parch</manuscriptMaterial>
<manuscriptHeight>17.2</manuscriptHeight>
<manuscriptWidth>24.1</manuscriptWidth>
<manuscriptPlica>0.6</manuscriptPlica>
<conditionDescription>...
```

2.10.8 conditionDescription

Elementet udfyldes efter behov, fx <conditionDescription>Brevet er medtaget af fugt</conditionDescription>.

2.10.9 layoutDescription

Elementet udfyldes efter behov, fx <layoutDescription>Brevet er opstillet i to kolonner</layoutDescription>.

2.10.10 handDescription

Elementet handDescription udfyldes efter behov. Hvis skriverens navn er kendt, som i dele af Langebeks Diplomatarium, angives dette i handDescription, fx <handDescription>Afskrift af Jakob Bircherod 1736</handDescription>. Kombineret med (reference) ref-elementet kan der tilknyttes andre dokumenter, der kunne været skrevet af samme skriver, jf Eksempel 9.

Eksempel 9. Flere dokumenter af samme skriver?

Bemærk, at ref-elementerne *skal* indeholde værdien af referencedokumentets textId.

2.10.11 additionsToText

Elementet additionsToText udfyldes efter behov. Påskriftens ordlyd bringes i elementet q (quote), og forkortelser opløses vha. elementet ex (expansion), som vist i Eksempel 10.

Eksempel 10. Gengivelse af en bagsidepåskrift

2.10.12 seal

Elementet seal er et lukket sæt med fire elementer:

- 1. sealNumber
- 2. sealStatus
- 3. sealDescription
- 4. sealReferenceWork



Når dokumentet ikke har segl

Når et dokument aldrig har været forsynet med segl eller bomærker, skal redaktøren "slukke" for feltet. Dette gøres ved at tildele seal-elementets underelementer følgende værdier, jf. Afsnit 1.1:

2.10.12.1 sealnumber

Elementet sealNumber indeholder en talværdi.

2.10.12.2 sealStatus

Elementet sealStatus er et lukket sæt med følgende fire værdier:

- 1. empty Dokumentet viser ingen spor efter segl eller bomærker
- 2. missing Seglet mangler; dette gælder også ved seglsnit eller en resterende perg.rem.
- 3. nil Der er ikke taget stilling til eventuelle segl eller bomærker
- 4. pendant Seglet er bevaret

2.10.12.3 sealDescription

Elementet sealDescription udfyldes efter behov. Er dokumentets segl bevaret, angives sigillantens navn, seglets befæstelse og materiale, som i Eksempel 11 nedenfor.

2.10.12.4 sealReferenceWork

Elementet sealNumber indeholder en reference, typisk til sigilværkerne Adelige sigiller (AS), Danske kongelige sigiller (DKS), Danske gejstlige sigiller (DGS), Danske adelige sigiller (DAS) eller Danske Købstæders Segl.

Eksempel 11. Beskrivelse af bevarede og tabte segl

2.10.13 witnessHistory

Elementet witnessHistory opregner begivenheder i tekstvidnets historie, typisk proveniensforhold, såsom optegnelser i senmiddelalderlige gods- og klosterregistraturer, jf. Eksempel 12.

Eksempel 12. Original, som har befundet sig på Vallø Slot

```
...</sealReferenceWork>
<witnessHistory>Brevet er registreret i registratur over brevene
    på Vallø (1541), trykt i Thiset, Adel. Brevkister 137</witnessHistory>
<bibliographicEntry>...
```

2.10.14 bibliographicEntry

Elementet bibliographicEntry udfyldes efter behov med oplysninger om trykte udgaver, udtog, referater eller registraturer af teksten. De enkelte poster adskilles med ';'.

Eksempel 13. Trykhenvisninger

2.11 samplingMethod

Elementet samplingMethod er et lukket sæt med tre elementer:

- 1. textCompleteness
- 2. sourceSigil
- 3. samplingNote

2.11.1 textCompleteness

Elementet textCompleteness er et lukket sæt med fire muligheder:

- 1. empty Det er irrelevant, om teksten optages i sin helhed eller i udtog
- 2. *excerpt* Tekstvidnet optages i udtog, i hvilket tilfælde en bemærkning om udeladelser bør følge i samplingNote-elementet, jf. nedenfor

- 3. nil Der er ikke taget stilling til om teksten optages i udtog eller i sin helhed
- 4. version Teksten optages in extenso

2.11.2 sourceSigil

Elementet sourceSigil udfyldes efter behov og henviser til det håndskrift som er tekstens hovedkonstituent.

2.11.3 samplingNote

Elementet, som udfyldes efter behov, forklarer, hvilke passager redaktøren har udeladt. Oftest gengives dokumentets ordlyd faktisk *in extenso*, og elementet kan i så fald slukkes vha. **empty**, som vist i Eksempel 14.

Eksempel 14. Beskrivelse af fuldstændig tekstgengivelse med slukket samplingNote

Hvis dele af en tekst er trykt i tidligere bind af Diplomatarium Danicum, eller teksten omhandler forhold uden relevans for Danmark, kan redaktøren udelade disse. I vidisser er det således kutyme at udelade det vidimerede brev, hvis dette allerede er udgivet, mens omfattende udelandske kilder som Hanserecesser og lignende også kan bringes i udtog, jf. Eksempel 15 og 16.

Eksempel 15. Udeladelse af vidimeret brev

Eksempel 16. Hansereces i udtog

Når teksten bringes i udtog, bør dette ikke kun fremgå af samplingNote, men også de pågældende steder i tekst og oversættelse, jf. Afsnit 2.13.4.2 Redaktionelle udeladelser.

2.12 textLanguage

Elementet textLanguage er et åbent sæt med foreløbig følgende værdier:

- 1. gda Tekstsproget er gammeldansk
- 2. gmh (German Middle High) Tekstsproget er middelhøjtysk

- 3. gml (German Middle Low) Tekstsproget er middelnedertysk
- 4. la Tekstsproget er latin
- 5. xno Tekstsproget er normannerfransk

2.13 text og p

Elementet text er et lukket sæt, ordnet udelukkende vha. p-elementer. Elementet p (paragraph) er hentet fra TEI P5 og indeholder en blanding af tekst og elementer. På tekstniveau er følgende elementer anvendelige:

- 1. app (critical apparatus) kritisk apparat
- 2. cit (citation) citathenvisning
- 3. damage beskadiget men læselig tekst
- 4. ex (expansion) opløsning af forkortelse
- 5. gap udeladelse i teksten
- 6. hi (highlighted) typografiske afvigelse
- 7. ref (reference) henvisning
- 8. supplied tilføjelse

2.13.1 app Kritisk apparat

Elementet app gentages efter behov. Det indeholder to elementer:

- 1. lem (lemma) teksten, der skal kommenteres
- 2. rdg (reading) læsning og kommentar

2.13.1.1 lem

Tekstfelt med følgende tilladte elementer: damage, ex, gap, hi, supplied

2.13.1.2 rdg

Tekstfelt, som gentages efter behov, med følgende tilladte elementer: damage, ex, gap, hi, supplied Elementet *skal* indeholde et af de to attributter:

- 1. resp (responsible) henvisning til redaktør
- 2. wit (witness) henvisning til tekstvidne

Eksempel 17. Kritisk apparat med henvisning til manuskript Aa1a1

Eksempel 18. Håndtering af dittografi i app

```
...voce ista <app>
    <lem>verba</lem>
    <rdg wit="#A"><q>verba verba</q>
    A</rdg></app>
    que sequuntur...
```

2.13.2 cit

Elementet, som getages efter behov, underordner to elementer:

- 1. quote citatet
- 2. bibl bibliografisk henvisning

2.13.2.1 quote og bibl

Tekstfelter med følgende tilladte elementer: damage, ex, gap, hi, supplied

Eksempel 19. citathenvisning

2.13.3 damage

Er et dokument beskadiget, men teksten dog læselig, anvendes elementet damage, fx Katerina relicta Nicolai Tag<damage>esen</damage>.

2.13.4 ex

Opløsning af abbreviaturer annoteres vha. ex, fx Obligo ... Ingwaro Anderss<ex>øn</ex> et suis heredibus

2.13.5 gap

Elementet, som gentages efter behov, har to formål: 1. Markering af lakuner i et manuskript. 2. Markering af redaktørens udeladelse i selve teksten.

2.13.5.1 Lakuner

Hvis et manuskript er beskadiget og teksten ikke kan læses markeres en sådan lakune med et tomt gapelement med tilføjelse af attributterne reason udfyldt med værdien illegible. Almindeligvis vil elementet i denne forbindelse altid være underordnet lem og ledsaget af en kommentar om lakunens omfang, jf. Eksempel 20.

Eksempel 20. Lakune i håndskrift

```
...Niels Mikilsøn Barfoeth, Jesse <app><lem><gap reason="illegible"/></lem></rdy wit="#Aa">5-6 bogstaver ulæselige, Aa</rdy></app>wed
...
```

2.13.5.2 Redaktionelle udeladelser

Når redaktøren udelader dele af en tekst, skal dette ikke blot fremgå af samplingNote-elementet, jf. Afsnit 2.11.3, men må også fremgå varsles i tekst og oversættelse. Hertil bruges gap med et underordnet (description) desc-element. Bemærk at gap også her tilføjer attributtet reason, men her udfyldt med værdien sampling.

Eksempel 21. Ekscerpering i fortløbende tekst (14130405001)

2.13.6 hi

Elementet, som gentages efter behov, gengiver typografiske afvigelser som over- og understillet tekst. Det rummer attributtet rend, der er et åbent sæt med foreløbig to værdier:

- 1. sup (supralinear) overstillet tekst
- 2. sub (sublinear) understillet tekst

Kombinationen hi rend="sup" erisær hyppig i datoangivelser på latin, fx Anno Domini m<hi rend="sup">o</hi>cd<hi rend="sup">o</hi> tredecimo sabbato ...

2.13.7 ref

Elementet, som gentages efter behov, indsætter en henvisning til anden tektudgave, fx De uthschrift der ingeslothen copien <ref> = Hanserec. VI nr. 229</ref>....

2.13.8 supplied

Elementet, som gentages efter behov, angiver tekst, som redaktøren tilføjer i tilfælde af åbenlyse fejl, fx (14190801001) ...dominus Pet<supplied>r</supplied>us Naule rector ecclesie Er der tale om et længere supplement til teksten, anbefales det at lade supplied kommentere i det kritiske apparat.

2.14 translation og p

Elementettranslation indeholder p-elementer. Når p er underordnet translation kan elementet foruden tekst indeholde følgende elementer: hi, se 2.13.6, note, se 2.14.1, ref, se 2.13.7

2.14.1 note

Elementet note, som kan gentages efter behov, indeholder tekst og tillader følgende elementer:

```
em (emphasis)
hi (highlighted), jf. 2.13.6
Det indsættes således: ... vådt og tørt, rørligt og urørligt,<note>Dvs.
løsøre og fast ejendom</note> slet intet undtaget ...
```