

Elektronisk udgivelse af Henrik Pontoppidans tre store romaner

Kodebog

Karsten Kynde

Fileerne er kodet i TEI P5, jf. *P5: Guidelines for Electronic Text Encoding and Interchange*, <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/>

Tekstkorpus består af flg. filer:

dfland\1udg\m.xml
 dfland\1udg\df.xml
 dfland\1udg\dd.xml
 dfland\2udg\m.xml
 dfland\2udg\df.xml
 dfland\3udg\dfland.xml
 dfland\4udg\dfland.xml
 dfland\5udg\dfland.xml
 lykkeper\1udg\lp1.xml
 lykkeper\1udg\lp2.xml
 lykkeper\1udg\lp3.xml
 lykkeper\1udg\lp4.xml
 lykkeper\1udg\lp5.xml
 lykkeper\1udg\lp6.xml
 lykkeper\1udg\lp7.xml
 lykkeper\1udg\lp8.xml
 lykkeper\2udg\lp.xml
 lykkeper\4udg\lp.xml
 ddr\1udg\ddr1.xml
 ddr\1udg\ddr2.xml
 ddr\1udg\ddr3.xml
 ddr\1udg\ddr4.xml
 ddr\1udg\ddr5.xml
 ddr\1udg4\ddr1.xml
 ddr\2udg\ddr.xml
 ddr\3udg\ddr.xml
 ddr\4udg\ddr.xml

DET FORJÆTTEDE LAND, skrevet således med kapitæler, forkortet *dfland*, omfatter i sin førsteudgave tre selvstændige bind: *Muld*, forkortet *m*, *Det forjættede Land*, forkortet *dfl* og *Dommens Dag*, forkortet *dd*. Heraf findes *m* og *dfl* tillige i en andenudgave. Disse tre er slået sammen i ét bind i de følgende tredje-, fjerde- og femteudgaver.

Lykke-Per (*lykkeper*) udkom i førsteudgaven i 8 bind (*lp1* - *lp8*). Disse otte er i andenudgaven slået sammen i tre og fjerdeudgaven i to bind.

De Dødes Rige (*ddr*) udkom i førsteudgaven i fem bind (*ddr1* - *ddr5*), heraf nåede det første bind (*ddr1*) at udkomme i en let revideret 4.oplag. I anden-, tredje og fjerdeudgaven blev de fem bind slået sammen til to.

Dertil findes til *dfland* to tillæg, nemlig føljetonerne »Klørknægt« (*til1*) og »Ungdom« (*til2*) som kan betragtes som forarbejder til romanen og endelig efterskrift (*ps*) og noter (*noter*), jf. kommentarbindet til Gyldendals udgivelse 1997.

dfland\0til\til1.xml
 dfland\0til\til2.xml
 dfland\red\ps.xml
 dfland\red\noter.xml

Et <TEI> dokument består af en <teiHeader> samt en <text>.

<teiHeader>

Tei header er formuleret efter flg. eksempel:

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>Lykke-Per i det fremmede</title>
      <author>Henrik Pontoppidan</author>
    </titleStmt>
    <publicationStmt>
      <publisher>Det Danske Sprog- og Litteraturselskab</publisher>
      <authority>
        <name xml:id="KK">Karsten Kynde</name>
      </authority>
      <availability>
        <p>http://etxt.dk/pontoppidan/xml/lykkeper/ludg/lp4.xml</p>
      </availability>
    </publicationStmt>
    <sourceDesc>
      <bibl>
        <author>Henrik Pontoppidan</author>
        <title>Lykke-Per i det fremmede</title>
        <edition>1</edition>
        <publisher>Det nordiske Forlag</publisher>
        <publisher>Bogforlaget Ernst Bojesen</publisher>
        <pubPlace>København</pubPlace>
        <date>1899</date>
      </bibl>
      <listWit>
        <witness facs="http://www.kb.dk/e-mat/dod/130014824279.pdf">
          Det Kongelige Biblioteks eksemplar: 57,-237-8°</witness>
        </listWit>
      </sourceDesc>
    </fileDesc>
    <encodingDesc>
      <editorialDecl>
        <quotation marks="none"/>
      </editorialDecl>
      <tagsDecl>
        <rendition xml:id="beg" selector="q" scope="before">>></rendition>
        <rendition xml:id="end" selector="q" scope="after"><<</rendition>
      </tagsDecl>
    </encodingDesc>
    <revisionDesc>
      <change status="HTML" when="2002">(C) 2002 by K.Kynde og DSL</change>
      <change status="xml">Konverteret til TEI P5 med html2tei.xsl,
        KK 2015-12-11</change>
      <change who="#KK" when="2018-02-24"/>
      <change who="#KK" when="2019-03-29"/>
      <change who="#KK" when="2019-09-02">Autokorrektur mod KBs pdf</change>
    </revisionDesc>
  </teiHeader>
```

<quotation marks="none"/> betegner at der ikke er sat anførselstegn i brødteksten; i stedet er

disse anført som <q>elementer, se nedenfor. Samtidig er disses udseende i grundteksten anført i <tagsDecl>.

En ny <change> med angivelse af @who og @when bør ubetinget tilføjes ved hver større ændring teksten eller serie af ændringer af teksten, gerne med en specificering af hvad der er sket.

<text>

Et <text>element repræsenterer et bind, fx *Muld* eller etbindsudgaven af *Det forjættede Land*, 3.udg.

Hvor et værk alene pga. omfanget er delt i flere sammenhængende bind, fx *Lykke-Per*, 2.udg. i tre bind eller samme værks 4.udg. i to bind, er bindene samlet i ét dokument hvis <text> udgøres af en gruppe på tre <text> elementer:

```
<text>
  <group>
    <text n="1"> ... </text>
    <text n="2"> ... </text>
    <text n="3"> ... </text>
  </group>
</text>
```

@n angiver bindnummer.

Førsteudgavernes separate bind er omvendt fordelt på hver sin fil, og her angiver @next det følgende og @prev det foregående bind, fx:

```
<text n="2" prev="lp1.xml" next="lp3.xml">
```

<text>elementet består videre af elementerne <front>, <body> og <back>

<front>

'Front matter' består typisk af titelblad og evt. forord eller andet, fx

```
<front>
  <titlePage facs="fax/lykkeper1.iv.jpg">
    <docAuthor>HENRIK PONTOPPIDAN</docAuthor>
    <figure type="horisontalRuler" rend="short"/>
    <docTitle>
      <titlePart>LYKKE-PER</titlePart>
      <titlePart type="sub">I DET FREMMEDE</titlePart>
    </docTitle>
    <docImprint>
      <graphic url="img/nordugle.gif" />
      <pubPlace>KØBENHAVN</pubPlace>
      <publisher>DET NORDISKE FORLAG</publisher>
      <publisher rend="petit">
        BOGFORLAGET
      </publisher>
      <hi rend="spaced">ERNST BOJESEN</hi>
      <date>1899</date>
    </docImprint>
  </titlePage>
  <div>
    <head n="H3" rend="center">FORHISTORIEN</head>
    <p rend="petit">Per Sidenius, en ung Ingeniør paa et Par og tyve
      Aar, har udarbejdet et kæmpemæssigt Kanalprojekt, [...]</p>
  </div>
</front>
```

<back>

'Back matter' udgøres af diverse efterskrifter, evt. forskræp for flg. bøger, fx

```

<back>
  <pb ed="1" n="139" />
  <p rend="center">FORTÆLLINGER</p>
  <p rend="center">om</p>
  <p rend="center">LYKKE-PER</p>
  <p rend="center">udkommer i syv (højst otte) tvangfri Hefter</p>
  <figure type="horizontalRuler" rend="short"/>
  <p>Udkommen er:</p>
  <table>
    <row>
      <cell>I. </cell><cell>Hans Ungdom . . . . .
</cell><cell>1 Hefte</cell>
    </row>
    <row>
      <cell>II.</cell><cell>Lykke-Per finder
Skatten . . . . .</cell><cell>1 -</cell>
    </row>
    <row>
      <cell>III.</cell><cell>Hans
Kærlighed . . . . .</cell><cell>1 -</cell>
    </row>
    <row>
      <cell>IV. </cell><cell>Lykke-Per i det fremmede . . .</cell><cell>1
-</cell>
    </row>
  </table>
</back>

```

<body>

Body udgør bindets 'brødtekst', som følger:

<div>

Teksten underindeles på forskellig måde med <div> elementer:

```

type="part"
type="book"
type="superChapter"
type="chapter"

```

OBS Bind mærkes ikke som <div type="volume">, men som <text>, se ovenfor.

dflands oprindelige tre bind optræder fra tredjueudgaven som dele ("part"), hver bestående af fem bøger ("book") opdelt i et antal kapitler ("chapter").

lykkeper er fra andenudgaven kun opdelt i kapitler, forløbende nummereret gennem bindene.

ddr udkom oprindelig i fem bind, der fra andenudgaven bliver til otte tildels sammenfaldende *bøger*. De fem oprindelige bind er opdelt i enheder, hver udsmykket med en nummereret vignet, og har mangel af bedres havelse fået betegnelsen "superChapter". Disse er yderligere inddelt i unummererede kapitler. Fra andenudgaven og frem nummereres kapitlerne fortløbende inden for hver bog.

fx (1. udg)

```
<div type="superChapter" n="1">
```

```

    <pb ed="1" n="1" />
    <head><graphic
url="img/ddr1.i.png"><desc>I</desc></graphic></head>
    <div type="chapter" n="1" xml:id="skt1.01"
corresp="../../ddr/skakt.xml#skt1.01">
        <pb ed="1" n="3" />

```

fx (4. udg)

```

<div type="book" n="1">
    <pb ed="4" n="5" />
    <head>
        FØRSTE BOG
        <lb/>
        TORBEN OG JYTTE
    </head>
    <div type="chapter" n="1" xml:id="skt1.01"
corresp="../../ddr/skakt.xml#skt1.01">
        <head>I</head>

```

@n angiver enhedens nummer i arabertal, uanset forfatterens angivelse. Om @xml:id og @corresp, se senere.

Et kapitels overskrift eller nummer angives i et indledende <head>. Dette består inogle af udgaverne af en vignet, fx

```

<div type="book" n="1">
<head n="H3" rend="center"><graphic url="img/mibog.gif"><desc>FØRSTE
BOG</desc></graphic></head>

```

Førsteudgavens vignet indeholder udover billede teksten "FØRSTE BOG", her gengivet i <desc>.

<p>

<p> benyttes til afgrænsning af linjebrud (alinea), evt. en replik, fx

```

<p>Den Gamle stirrede.</p>
<p><q>Er De Præst? ...</q></p>

```

Evt. indrykning angives ikke. @rend kan angive

```

rend="petit"
rend="center"

```

<lg>, <l> og <lb/>

Lyrikciteringer markeres med *line group*, og *line*, fx

„Fred hviler over Land og By
ej Verden larmer mer — —.“

```

<lg type="BLOCKQUOTE">
    <l><q rendition="#beg"/>Fred hviler over Land og By </l>
    <l>ej Verden larmer mer — —.<q rendition="#end"/></l>
</lg>

```

Samme teknik benyttes hvor forf. har fremhævet fx et brev på lignende måde, dvs. i en særlig blok:

```

<lg>

```

```
<opener>Kære Fru Gehejmeraadinde!</opener>
<l>Mon De veed, hvad Alverden taler om, at Deres Datter er bleven
  Maleren Karsten Froms Elskerinde? Jeg gratulerer!</l>
<closer>En Veninde.</closer>
</lg>
```

Dog ikke ved breve i almindelighed, hvor de indgår i teksten med uforandret margen og skriftstørrelse.

Andre betydningsbærende lineskift i fx overskrifter er markeret som *line beginning*, `<lb/>`

FØRSTE BOG TORBEN OG JYTTE

```
<head>
  FØRSTE BOG
  <lb/>
  TORBEN OG JYTTE
</head>
```

<ab>

Anonymous block benyttes hvor kapitlerne er underopdelt af blanke linjer eller asterisk, der findes tre typer:

```
type="lineBreak"
type="asterisk"
type="head"
```

fx

```
<ab type="lineBreak" n="H5"/>

<ab type="asterisk" n="H4" rend="center">* &#xA0; &#xA0; &#xA0; *<lb
/>*</ab>

<ab type="head" n="H5" rend="right">17de Oktb.</ab>
```

` ` angiver et *non breaking space*. (`@n` angiver hvilken header der var anvendt i en ældre HTML-udgave af disse tekster og er irrelevant for det følgende).

Den sidste type *"head"* benyttes hvor der findes en overskrift på et sted hvor TEI ikke tillader et `<head>`

<pb/>

Placeringen af originaltrykkenes sideskift er bevaret i teksten, markeret med *page beginning*, fx

```
Man <pb ed="1" n="4"/> sendte ham til Herlufsholm

Men trods disse glørværdige Forbilleder brød han selv afgjort med
Slægts<pb ed="1" n="5"/>traditionen
```

`@n` angiver sidetallet og `@ed` udgavenummeret. `ed="0"` angiver den trykte udgave af DET FORJÆTTEDE LANDS 1. udgave, som udgivet af Gyldendal 1997 sammen med den elektroniske tekst.

OBS i *dfland* blev sideskiftene i HTML-udgaven rykket til efter ordet og er ikke rettet siden.

Tekstfremhævning, `<emph>`, `<foreign>`, `<mentioned>`, `<name>` etc.

Tekst som i originaltrykket er fremhævet på forskellig vis er kodet med et TEI-element efter bedste skøn og @rend efter flg. tabel

rend="spaced"	spatieret, spærrede typer
rend="italic"	kursiv
rend="bold"	fed
rend="smallCaps"	kapitæler
rend="init"	initialbogstav
rend="initWord"	initialord (kapitæler)
rend="petit"	mindre skriftgrad
rend="high"	løftet

Eksempler

Kirkens e g n e Rammer

Kirkens `<emph rend="spaced">egne</emph>` Rammer

De er en attraet Dame,

`<emph rend="italic">De</emph>` er en attraet Dame,

Altsaa — nu gentager du mine Ord — *Jeg* —«

»Jeg«, gentog Per.

Altsaa — nu gentager du mine Ord — `<mentioned`

`rend="italic">Jeg</mentioned>` —`</q></p>`

`<p><q>Jeg</q>, gentog Per.</p>`

den berømte *Stærkodder II*,

den berømte `<name rend="italic">Stærkodder II,</name>`

Per amor di Dio!

`<foreign rend="italic">Per amor di Dio!</foreign>`

Forf.

`<hi rend="bold">Forf.</hi>`

H. JOHANSEN, Sagfører

`<name rend="smallCaps">H. Johansen,</name>` Sagfører.

Klokken 12¹⁵

Klokken 12`<hi rend="super">15</hi>`

DER var nu gaaet tre Aar siden den Septemberdag, da den unge Godsejer Torben Dihmer vendte hjem

`<hi rend="initWord"><hi rend="init">D</hi>er</hi>` var nu gaaet tre Aar siden den Septemberdag, da den unge Godsejer Torben Dihmer vendte hjem

Korrektion af tekstfejl, `<choice>`, `<corr>`, `<sic>`

Tekstrettelser er mærket som fx

Han gik i hvert

`<choice>`

`<sic>Fad</sic>`

`<corr>Fald</corr>`

`</choice>` videre.

Evt. er er anført andre udgaver som støtte for rettelsen, særligt i sådanne tilfælde hvor andre rettelser kunne komme på tale, fx i *Lykkeper*, 4. udg.:

```
Men havde han paa denne Maade faaet Børnene til at le, forsømte han
aldrig tilsidst at give sin Fortælling Advarslens Form og
<choice>
  <sic>formanede</sic>
  <corr resp="1udg 2udg">formanende</corr>
</choice> at vende deres Sind og Tanker mod Livets Alvor og Pligtens
Bud.
```

5.udg. har her rettet til "formanede dem". Også i tilfælde hvor man kunne acceptere læsningen som en ortografisk variant, fx i anvendelsen af fugebogstaver, her i 2. udg.:

```
mellem Ølkasserne,
<choice>
  <sic>Brænvinsankrene</sic>
  <corr resp="1udg 4udg">Brændevinsankrene</corr>
</choice> og Sukkerfustagerne.
```

<q>

Anførselstegn er mærket med *quote*, som vist flere steder ovenfor, fx

```
»Er De Præst? ...«
```

```
<q>Er De Præst? ...</q>
```

Der anvendes typisk ikke samme tegn i forskellige udgaver, hvilket er grunden til at tegnet ikke er indsat direkte. Det frigør en nyudgivelse til at anvende hvilken typografi man måtte ønske, og ved sammenligninger af udgaver tages den slags 'accentalier' ikke i betragtning. Hvis man er interesseret i grundtekstens repræsentation kan det ses i TEI-headeren (se ovenfor), hvor fx 1.udg. af *ddr* har

```
<rendition xml:id="beg" selector="q" scope="before">»</rendition>
<rendition xml:id="end" selector="q" scope="after">«</rendition>
```

4. udg. af samme værk, der til anførselstegn anvender '99-66' har samme tekst, men i headeren

```
<rendition xml:id="beg" selector="q" scope="before">„</rendition>
<rendition xml:id="end" selector="q" scope="after">“</rendition>
```

Pga. xml's struktur medfører denne teknik lidt besvær hvor anførselstegn undtagelsesvist står uparrede, fx i lyrik, se ovenfor <lg>, eller i de ganske hyppige tilfælde hvor et anførselstegn er anbragt omvendt eller helt udeladt, og der derfor er behov for en tekstrettelse. I disse tilfælde anvendes <q/> som milepæl med @rendition til angivelse om der er tale om et begyndelses- eller slutanførselstegn, fx

```
<p><q rendition="#beg"/>Vil du da styrte dig selv i Ulykke, Barn! ...
Sig dog, at det ikke er
<choice>
  <sic>sandt!<q rendition="#beg"/></sic>
  <corr>sandt!<q rendition="#end"/></corr>
</choice></p>
```

I citater der strækker sig over flere afsnit, er der også ofte i grundteksten anvendt en konvention om at gentage begyndelsesanførselstegnet i alle afsnit og først sætte slutanførselstegn i det sidste.

Tankestreger

Hertil er anvendt unicode `—` *em-dash* —. Tankestreger må ikke forveksles med bindestreger i sammensatte ord, som i grundteksten gengives noget kortere.

Spatiering

Som typografisk fremhævelse

Gengives med `@rend="spaced"` og ikke bogstaveligt med mellemrum i ordet, se ovenfor.

Ved punktueringstegn

Punktum, komma, semikolon, udråbs- og spørgsmålstegn sættes efter ordene uden mellemrum, men efterfulgt af mellemrum. Tankestreger omgives med mellemrum efter skøn. Således altid når det er brugt som indskudsmarkør, fx

```
der – læg vel mærke hertil – af
```

og til indbyrdes adskillelse når de optræder i par. Derimod ikke foran afsluttende anførelsestegn, fx

```
kære Ven – –</q>
```

og ikke når de undtagelsesvist er anbragt inde i ord

```
Men–n
```

og det klart fremgår af grundteksten at tegnene skal stå klos.

Forkortelser

Der sættes dog ikke mellemrum efter punktum midt i forkortelser som f.Eks. d.v.s. o.s.v.

Ellipser

Der sættes altid mellemrum omkring ellipser: tre eller flere sammenhængende punktummer (hvis antal respekteres), fx

```
det kender en anden jo ingen Ting til .... Men ellers
```

Dette gælder også hvor foregående ord slutter med et punktueringstegn

```
Hun er en striks Kommandant! ...
```

Dette punktueringstegn kan specielt være punktum, hvilket i så fald i grundteksten vil kunne skelnes fra ellipsen ved at være anbragt klos op ad det foregående ord. Typisk vil der da efter ellipsen fortsættes med stort, fx

```
Deres egen Mund. ... Jeg tvivler
```

Noter (punktkommentarer)

Opføres i et særskilt dokument, og henvises fra tekstdokumentet med en `<ref>`, fx

```
en ny Katekismus<ref type="note" target="../../red/noter.xml#n12.0" />
trykket krampagtigt beskyttende ind i Favnen.
```

I den fil anført som `@target` findes noten som flg.

```
<note xml:id="n12.0">
  <label>Katekismus</label>
```

```
<p><title rend="italic">Luthers Kathekismus med en kort
  Forklaring</title> af biskop Carl F. Balslev (1805-1895) blev
  fra 1849 brugt i skolerne som lærebog i den kristne
  børnelærdom.</p>
</note>
```

Skakter

En skakt er en datastruktur som sammenkæder sammenlignelige passager i forskellige tekstfiler. I teksten fx

```
<div type="chapter" n="1" xml:id="skt19"
  corresp="../../lykkeper/skakt.xml#skt19">
```

@corresp udpeger skakten, der ser ud som følger (*TEI Guidelines*: kap. 16 Linking, Segmentation, and Alignment):

```
<linkGrp type="alignment" n="lp7">
  <link xml:id="skt19"
    target="1udg/lp7.xml#skt19 2udg/lp.xml#skt19 4udg/lp.xml#skt19 "/>
  ...
</linkGrp>
```

Sammenligningsprogrammet virker ved at @xml:id i skakten og i de sammenlignede filer er ens.

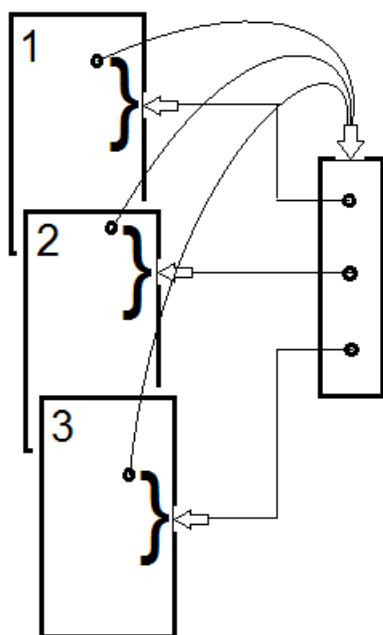
I dette projekt er den flyttede tekst medtaget i sammenligningen af den øvrige tekst (hvor den altså ikke matcher i sammenhængen), men dertil er der foretaget en særlig sammenligning af de flyttede dele. Disse anbringes i en særlig </linkGrp> og fil, navngivet efter @n. <link>'et (skakten) gives en særlig type="movedText" for dermed at kunne skelne den fra hovedteksten. @n beskriver den flyttede tekst og kan benyttes til vejledning i visningen.

Skakten for en flyttet tekst ser således ud:

```
<linkGrp type="alignment" n="flyt.12a">
  <link xml:id="skt12a" type="movedText" n="Pers julelov"
    target="1udg/lp4.xml#skt12a
      2udg/lp.xml#skt12a
      4udg/lp.xml#skt12a "/>
</linkGrp>
```

Der vil komme en særlig vejledning i kodningen af skakter.

Kodning af skakter i Henrik Pontoppidans tre store romaner



En skakt er en datastruktur som sammenkæder sammenlignelige passager i tekstfiler. De sammenlignelige dele af den enkelte tekst kaldes et *skaktmål*, som refererer til samme *skakt*. Selve skakten er en ekstern fil, hvorfra der refereres tilbage de enkelte mål, se figuren.

Ideen er at tilvejebringe en 'lodret' forbindelse. Som man kan se udpeges her de sammenlignelige punkter i de tre udgavers tekst, idet læseteksten opfattes som 'vandret'. Skakter har to formål: At lede læseren til at læse de andre versioner i deres sammenhæng og at kunne foretage en maskinel kollation.

Der anvendes dobbeltreferencer idet teksten refererer til skakten og skakten modsat til teksten. Det kan synes overflødigt, men har med de to formål at gøre. Hvis læseren skal ledes til de andre versioner vil det være praktisk at referencen starter i teksten. Det vil også være en naturlig fremgangsmåde for den kodende filolog. Herefter kan skakten konstrueres maskinelt, hvilket man må anbefale for at

minimere risikoen for fejl¹. Hvis vi imidlertid fokuserer på det andet og teknisk mest krævende formål, kan vi simpelthen bruge skakten som udgangspunkt for kollationeringen idet den jo udpeger i dette tilfælde tre sammenlignelige kapitler. Til kollationen af versionerne benyttes programmet *CollateX*, se <https://collatex.net>.

Kodningen af en skakt er i TEI foretaget efter reglerne i [TEI Guidelines 16.5](#), om "**alignment**", fx

lykkeper/skakt.xml:

```
<linkGrp type="alignment" n="lp1">
  <link xml:id="skt01"
        target="ludg/lp1.xml#skt01 2udg/lp.xml#skt01 4udg/lp.xml#skt01" />
  <link xml:id="skt02"
        target="ludg/lp1.xml#skt02 2udg/lp.xml#skt02 4udg/lp.xml#skt02" />
</linkGrp>
```

`<linkGrp>` bør have en `type="alignment"`, som anført i *Guidelines*. Hver `<linkGrp>` giver anledning til en kollationsfil efter behandling med *collateX* som tager navn efter `@n`, her *lp1.col.xml*.

Et `<link>` kan derudover have en `@type="movedText"` som angiver at dette tekststed ikke har samme position i alle versioner. De enkelte skaktmål angives som `@target` i en række adskilt med mellemrum, jf. *Guidelines*. Endelig kan tekststedet gives et navn, `@n`, som kan være nyttigt, specielt ved flyttede tekststeder.

I den enkelte version kodes en henvisning til et `<link>` i denne skaktstruktur med en reference fra det relevante element, fx

lykkeper/ludg/lp1.xml:

```
<div type="chapter" n="1" xml:id="skt01" corresp="../../lykkeper/skakt.xml#skt01">
<head n="H3" rend="center">FØRSTE KAPITEL</head>
<p>I en af de østjyske Smaakøbstæder levede der for omtrent en Menneskealder siden ... </p>
```

Der henvises til skaktstrukturen med `@corresp` og stedet gives samme `@xml:id` som skakten.

¹ Se særskilt vejledning [skaktgen.pdf](#)

I princippet behøver vi kun én skakt, i begyndelsen af teksten. Der er flere grunde til at det er utilstrækkeligt. Skakten benyttes i teksten til at vise læseren hen til de sammenlignelige steder i andre tekster, og det vil derfor være bekvemt om der altid er en skakt i nærheden af det sted man læser. Derfor er det godt med et rimeligt fintmasket net. Dertil kommer at skaktreferencerne i teksterne er udgangspunkter for en maskinel sammenligning. Det vil for det første sige at man kan bruge skakterne til at komme tilbage på sporet hvis en sammenligning er faret vild. For det andet fordi selve kollationsprocessen med *CollateX* kompliceres ganske voldsomt med længden af de sammenlignede tekster. Fx faldt behandlingstiden for 3. kapitel i *Lykkeper* fra 51 minutter til 3 minutter efter kapitlet var blevet opdelt i tre. Vel kan man sige at kollationen kun skal foretages én gang, men der er dog grænser for hvor lang tid det i praksis må tage.

På den anden side skal enhver skakt der etableres kontrolleres, dvs. den kodende filolog skal overbevise sig om at disse tekstafsnit virkelig dækker noget sammenligneligt, fx ved at studere en allerede forsøgt kollation.

Gennem alle sine omarbejdninger er Pontoppidan rimelig trofast overfor sine kapitler. Fx består hans føljeton *Ungdom* af et antal afsnit adskilt af blanke linjer. Disse afsnit svarer nøje til kapitlerne, som historien genfortælles i *Muld*, og som, bortset fra nogle få ombytninger, videreføres i 3. udg. Det er altså et godt udgangspunkt.

Dfland består i 3. udg. af 102 kapitler, nogle er slået sammen af flere kapitler fra 1. udg. hvilket resulterer i 117 skakter, kollationen med *CollateX* tager 35 minutter. *Lykkeper* består af 28 kapitler som igen er opdelt i afsnit, markeret med tre asterisker, ialt 83 *Ddr* består af 86 kapitler.

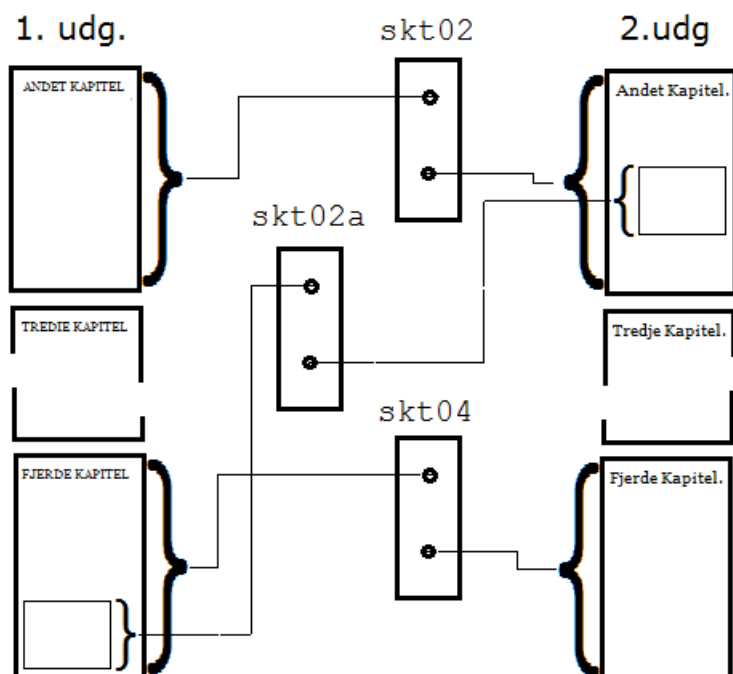
Lykkeper består i skrivende stund (2020-03-19) af 49 skakter med en kollationstid på 4 timer 35 minutter. *Ddr* er skaktet, men ikke kollationeret med *CollateX*.

Navngivningsstrategi

Eksempel, *lykkeper*: Der tages udgangspunkt i seneste udgaves kapitler og navngives "skt01", "skt02", ... "skt28". Hvor et kapitel i 1. udg. bestod af 2 kapitler, tilføjes 'decimal', fx "skt08.1", "skt08.2". Det samme hvor der kan opdeles yderligere efter asteriskseparerede afsnit.

Flyttede afsnit kan med fordel navngives med litra, fx "skt02a".

Flyttede afsnit



Nogle steder flytter Pontoppidan afsnit rundt, fx fortælles der første gang om Pers store projekt som en slags *cliffhanger* i 1. udg.s fjerde og sidste kapitel af det første bind, *Lykke-Per. Hans Ungdom*. Sidenhen er denne beskrivelse flyttet bagud til de samlede udgavers (2.-4. udg.) kapitel 2.

Det har været strategien i dette projekt at sådanne flyttede afsnit markeres som skaktmål indlejret i skaktmål, dvs. kapitel 2 og kapitel 4 udgør skaktmål i egen ret, sammenlignelige med 2 og 4 i de andre udgaver, men omtalte afsnit afgrænses og mærkes inde i disse, se ill.

Sammenligningsalgoritmen instrueres til at opfatte det flyttede afsnit som et hele, hvorved det forbliver repræsenteret i det omgivende afsnit i begge versioner, men udgår af sammenligningen. I stedet kollationeres det flyttede afsnit mod sin nye version, og den gengives særskilt.

Skakterne kodes:

```
<linkGrp type="alignment" n="lp1">
  ...
  <link xml:id="skt02"
        target="1udg/lp1.xml#skt02 2udg/lp.xml#skt02 ..."/>
  ...
  <link xml:id="skt04"
        target="1udg/lp1.xml#skt04 2udg/lp.xml#skt04 ..."/>
</linkGrp>

<linkGrp type="alignment" n="flyt.02a">
  <link xml:id="skt02a" type="movedText" n="Projektet beskrives"
        target="1udg/lp1.xml#skt02a 2udg/lp.xml#skt02a ..."/>
</linkGrp>
```

Skakten tilhørende den flyttede tekst anbringes i sin egen `<linkGrp>` og `n="flyt.02a"` angiver at den anbringes i en særskilt fil. `<link>` har typen `"movedText"` og `@n` angiver en titel for det flyttede afsnit (eksemplet er her gengivet lettere forenklet).

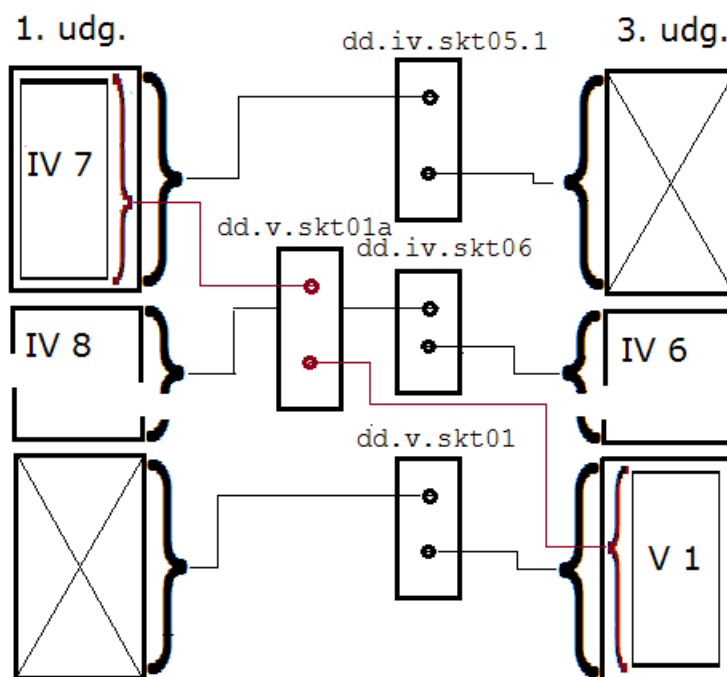
Skaktmålet kodes (i 1.udg.)

```
<p xml:id="skt02a" corresp="../../lykkeper/skakt.xml#skt02.1">Kort
  efter, at Eberhardt havde forladt ham, sad Per allerede bøjet over
  ...
  opblomstrende Nybygger-Byer derinde i Forbindelse med Havet.</p>
```

Den flyttede tekst er her afgrænset som et afsnit (`<p>`). Så heldige er vi ikke i 2.udg., hvor beskrivelsen er udvidet til to afsnit. De må altså omkranses af et element, en `<div>`:

```
<div xml:id="skt02a" corresp="../../lykkeper/skakt.xml#skt02.1">
  <p>Han havde endogsaa for længe siden udset sig Midlet, der ... </p>
  <p>I tre Aar, lige siden han fik lært sig at opsætte en Kanalprofil
  ...
  Nybygger-Byer derinde i Forbindelse med Havet paa begge Sider.</p>
</div>
<div>
  <p>...
  ...
</div>
```

Eftersom man i TEI ikke må efterfølge en `</div>` med en `<p>` (en i anden sammenhæng helt fornuftig regel), er vi nu nødsaget til at anbringe resten af kapitlets afsnit i endnu en (overflødig) `<div>`. Sådanne findes altså og kan kendes på at de hverken har `@type` eller andre attributter (et 'neutralt' element findes ikke i TEI, *anonymous block*, `<ab>` kunne lyde som en god kandidat, men passer ikke syntaktisk ind alle steder, fx ikke her). Som antydet ovenfor hænder det også



at hele kapitler er byttet rundt. Fx i *dfland*, 's sidste del, *Dommens Dag*, beskrives afslutningen af det store vennemøde på Sandinge Højskole i et kapitel. Dette findes i 1. udg. som 4. bogs 7. kapitel. Det bytter plads med 8. kapitel og bliver samtidig rykket til den næste, 5. bogs 1. kapitel i 3. udg. Vi opretholder politikken med at al teksten skal være repræsenteret på rette plads i hver version, hvilket betyder at stedet der svarer til det flyttede kapitel, bliver tom i den modsvarende version (det kræver lige at vi bestemmer os for hvilket kapitel der flyttes når to kapitler bytter plads. Vi beslutter os for at det er kapitlet der bevæges *frem* i teksten der betragtes som flyttet).

De overkrydsede skaktmål er 'tomme', svarende til de to <milestone>s i kodningen:

1. udg:

```
<div type="chapter" n="7" xml:id="dd.iv.skt05.1"
  corresp="../../dfland/skakt.xml#dd.iv.skt05.1">
  <div xml:id="dd.v.skt01a"
    corresp="../../dfland/skakt.xml#dd.v.skt01a"><!-- flyttet -->
    <head n="H4" rend="center">VII.</head>
    <p>Indtrykket af denne uhyggelige Scene havde næppe ...
  </div>
</div> <!--chapter-->

<div type="chapter" n="8" xml:id="dd.iv.skt06"
  corresp="../../dfland/skakt.xml#dd.iv.skt06">
  <head n="H4" rend="center">VIII.</head>
  <p>Hjemme hos Fru Betty havde der imidlertid ogsaa været stor Uro.
  ...
</div> <!--chapter-->

<milestone xml:id="dd.v.skt01" unit="chapter"
  corresp="../../dfland/skakt.xml#dd.v.skt01"/>
```

3. udg:

```
<milestone xml:id="dd.iv.skt05.1" unit="chapter"
  corresp="../../dfland/skakt.xml#dd.iv.skt05.1"/>

<div type="chapter" n="6" xml:id="dd.iv.skt06"
  corresp="../../dfland/skakt.xml#dd.iv.skt06">
  <head n="H4" rend="center">VI</head>
  <p>Hjemme hos Fru Betty havde der imidlertid ogsaa hersket stor Uro.
  ...
</div> <!--chapter-->

<div type="book" n="5">
  <head n="H3" rend="center">FEMTE BOG</head>
  <div type="chapter" n="1" xml:id="dd.v.skt01"
    corresp="../../dfland/skakt.xml#dd.v.skt01">
    <div xml:id="dd.v.skt01a"
      corresp="../../dfland/skakt.xml#dd.v.skt01a"><!-- flyttet -->
      <head n="H4" rend="center">I</head>
      <p>Stemningen paa Højskolens Formiddagsmøde den næste Dag var i
        hvert Fald i Begyndelsen endnu præget af den Uhygge, som ...
    </div>
  </div> <!--chapter-->
```

og skakten:

```
<linkGrp type="alignment" n="dfland3">
  <link xml:id="dd.iv.skt05.1"
    target="1udg/dd.xml#dd.iv.skt05.1 3udg/dfland.xml#dd.iv.skt05.1
    ... "/>
```

```

<link xml:id="dd.iv.skt06"
      target="1udg/dd.xml#dd.iv.skt06 3udg/dfland.xml#dd.iv.skt06 ..."/>
<link xml:id="dd.v.skt01"
      target="1udg/dd.xml#dd.v.skt01 3udg/dfland.xml#dd.v.skt01 ... "/>
...
</linkGrp>

<linkGrp type="alignment" n="flyt.dd.v.01a">
  <link xml:id="dd.v.skt01a" type="movedText" n="Vennemødets slutning"
        target="1udg/dd.xml#dd.v.skt01a 3udg/dfland.xml#dd.v.skt01a ..."/>
</linkGrp>

```

En komplikation opstår der hvor der ikke er en pæn kapitelstruktur at opmærke. Her må man i sidste ende ty til det TEI kalder 'milepæle', fx i dette tilfælde hvor der er tale om flyttet tekst:

```

<p>
  Siden ... havde han en eneste Gang genset sit Hjem.
  <milestone unit="movedText" xml:id="skt12a" spanTo="#skt12a.end"
            corresp=".../lykkeper/skakt.xml#skt12a"/>
  Den let bevægede Stemning, der til hans egen Overraskelse ...
</p>
<p>
  ... straks efter Nytaar vendte han tilbage til København.
</p>
<milestone unit="movedText" xml:id="skt12a.end" />

```

<milestone> skal efter TEI have en obligat @unit, her er "movedText" anvendt. Milepæle er ganske omstændelige, vanskelige at få anbragt i teksten og det er også vanskeligt at få en ordentlig visning af de involverede skakter. Det anbefales derfor at undgå denne metode, hvilket erfaringsmæssigt let kan gøres i det foreliggende materiale. Fx kan man i ovenstående eksempel overveje om ikke den indledende frase kan medtages i den flyttede tekst:

```

<div xml:id="skt12a" spanTo="#skt12a.end"
      corresp=".../lykkeper/skakt.xml#skt12a">
  <p>
    Siden ... havde han en eneste Gang genset sit Hjem. Den let bevægede
    Stemning, der til hans egen Overraskelse ...
  </p>
  <p>
    ... straks efter Nytaar vendte han tilbage til København.
  </p>
</div>

```

Yderligt interesserede henvises til "Maskinel kollation af et positivintegralt multivariantapparat eller Sammenlignende billeddannelse i ubevogtede øjeblikke" i *From Text to Artefact. Studies in Honour of Anne Mette Hansen*, <http://kismet.press/portfolio/from-text-to-artefact/#undefined248>.