**Transkriptionsregler for Velux projektet**

Første udkast (27.10.2017)

1. Generelt:

* Vi transskriberer i ren text format (dvs .txt i hvilken som helst teksteditor eller .docx i Word). Hvis man arbejder i Word kan der dog ikke bruges formateringsfunktionerne, dvs. kursivering, farver, …
* Brug en MUFI-kompatibel font, fx Andron Scriptor Web eller Junicode (kan downloades her: <http://folk.uib.no/hnooh/mufi/fonts/index.html>).
* Vi transskriberer kun på faksimilært niveau.
* En linie i diplomet bliver gengivet med en linie i transskriptionen. Efter hver linie sættes dobbelt linieskift (2 x return).
* Interpunktuationstegn transskriberes som de forekommer i diplomet (fx ”.”, ”,” og ”/”), men et punktum der er skrevet lidt højere (en so-kaldt *punctus elevatus*) gengives som almindeligt punktum.
* Ord bliver separeret af mellemrum. Hvis noget er skrevet som to ord, der burde være ét, bruges ”=” (fx ”Hellig=ånd”, ”konge=magten”, ). Hvis noget er skrevet sammen, men burde behandles som to ord, indsættes ”==” (fx ”e==foruden”, ”før==sagt”, ”at==i”).
* Noter med kommentarer, fx når man har spørgsmål eller har lagt mærke til noget særligt, skrives indenfor <note>…</note>, fx ”<note>BS: Need to double check this!</note>

1. Bogstavsformer mm.:

* Vi skelner mellem langt s/rundt s, almindeligt r/r-rotunda og bevarer ligaturer, der deler en minim eller downstroke eller får en anderledes bogstavform (fx ligaturer af a+v, a+e, a+a, s+t, c+t, p+p, f+f, k+langt s, h+langt s), + bevarer vi rundt d????
* Vi markerer n og m, der har en højre descender. Der skelnes dog ikke mellem forskellige typer, men bruges hovedformen ”right descender” (dvs.  (&nrdes;) og  (&mrdes;)) (SKAL AFPRØVES OG ER MÅSKE IKKE RELEVANT)
* Der skelnes mellem store bogstaver, små bogstaver, small capitals og enlarged letters (fx ”E” vs. ”e” vs. ”ᴇ” (&escap;) vs. ”” (&eenl;).
* Unciale bogstaver opmærkes, fx  (&Eunc;),  (&eunc;) og  (&Munc;).
* Initialer bliver som udgangspunkt behandlet som store bogstaver.
* Diakritiske tegn skal transskriberes på alle vokaler + y (inkl. ligaturer). Dvs. der skelnes mellem ”ı” (&inodot;), ”i”, ”í”, ”ï” og ”” (&idblac;) mellem ”y”, ”ẏ” (&ydot;), ”ý” (&yacute;), ”ÿ” (&yuml;), ”” (&ydblac;) o.s.v. Skelnes mellem kodes streg over u og u uden streg?
* Det er lige meget om man vil sætte MUFI tegn ind med det sammen, eller om man hellere vil bruge entiteter for særtegn. MUFI character listen findes her: <http://folk.uib.no/hnooh/mufi/specs/MUFI-Alphabetic-3-0.pdf>

1. Abbreviaturer:

* Forkortelsestegn transskriberes som forkortelsestegn, dvs. ikke opløst, og i deres `standard form’ (se fil ”Forkortelser.pdf”). Hvilket tegn bruger vi for et rigtigt ampersand (og/et)? Og hvordan opmærkes d med en lille krølle nedenfor for at indikere afsluttende e?
* So vidt muligt skrives basistegnet, dvs. bogstaven på skrivelinien, først og så forkortelsestegnet bagefter (også hvis de er skrevet sammen i diplomet), dvs. ”a”+”&bar;” og ikke ”&amacr;”, ”h”+”&bar;” og ikke ”&hstroke;”. De eneste undtagelser er ”&pflour;” and ”&slongflour;”, fordi de pågældende forkortelsestegn p.t. ikke eksisterer som separate tegn (se også filen ”Forkortelser.pdf”).
* Forkortelsestegn der principielt også kunne fortolkes som almindelige tegn, fx punktum, markeres når de forekommer som forkortelsestegn med hjælp af parenteser rundt omkring (fx ”A(.)D(.)” for A*nno* D*omino*).
* Hvis et tegn kunne stå for både forkortelsestegn og sig selv, angives det med særtegnes ”·” (&middot;) (fx ”NN skal betale 10 m(.)s·” for ”NN skal betale 10 m*ark* s*ølv*.”)

1. Ændringer og problematiske læsninger:

* Tilføjelser (både af skriveren selv og senere tilføjelser) sættes indenfor <add>…</add>. Med hjælp af ”hand” og ”place” markører i den første del af mark-up’en kan det tillægges flere informationer om tilføjelsen, fx skriveren har først glemt a’et i ordet datter og senere tilføjet det ovenfor:

”d<add hand=”scribe” place=”supralinear”>a</add>tter, eller en senere tilføjelse i nedre margen: ”<add hand=”later” place=”margin-bot” >Denne gård er nu eget af mig, Jens Andersen</add>. Flere oplysninger kan findes i kapitel 9.2.1 af Menota håndbogen v.3 (<http://www.menota.org/HB3_ch9.xml>).

* Udslettede bogstaver eller ord markæres på lignende måde med <del>…</del>, som også kan specificeres med ”hand” og ”rend”. Et ord, der blev streget over af skriveren selv, for eksempel, markeres således: <del hand=”scribe” rend=”overstrike”>ord</del>. (Se også kapitel 9.2.2 i Menota håndbogen v.3 (<http://www.menota.org/HB3_ch9.xml>).
* Beskadiget tekst, som er fuldstændig forsvundet og ikke kan rekonstrueres, markæres med <gap>...</gap>. Indenfor <gap> sættes lige så mange 0-tal som der vurderes mangler, fx ”<gap>00000</gap>”.
* Suppleret tekst sættes indenfor <supplied>…</supplied>. Der kan angives en grund med hjælp af ”reason”, der kan være ”illegible” eller ”omitted”, fx

”Mar<supplied reason=”illegible”>ia</supplied>. Alternativt kan også skrives ”Mar[ia]” (hvis det er ulæseligt) eller ”Mar⟨ia⟩” (hvis bogstaverne aldrig blev skrevet).

Hvad gør vi i tilfælde af et ord, der er skrevet over 2 linier (og der er ikke indikeret bindestreg eller lignende)?

* Svært læselig tekst, bogstaver som kun delvis er synlige, eller noget som ikke kan bestemmes med sikkerhed sættes indenfor <unclear>…</unclear>, fx ”F<unclear>ø</unclear>rste”. Alternativt kan her anvendes tuborgklammer: ”F{ø}rste”.
* If there is too much text, i.e. a scribe repeated a word or parts of a word withput noticing and deleting it, the word(s) are palced inside of <surplus>, e.g. ”ett offentligt <surplus>be<surplus> beseglet bref”