

# Slots- og Kulturstyrelsen

---

## Udlægning til og ordning af registre i onlinekataloger

udarbejdet af en arbejdsgruppe nedsat af Bibliografisk Råd

November 2017

## Indhold

Forord .....	3
Indledning .....	5
1. Terminologi.....	5
2. Generelt om udlægning .....	6
3. Tegnenes indbyrdes ordning .....	6
4. Udlægning og normalisering af bogstaver .....	8
5. Udlægning og normalisering af tal.....	9
6. Udlægning og normalisering af symboler .....	10
7. Udlægning og normalisering af blanktegn (mellemrum) o.lign. ....	11
8. Udlægning og normalisering af skilletegn .....	12
9. Udlægning til sætningsregistre .....	12
10. Forkortelser.....	14

## Forord

I 1999 udkom vejledningen *Ordning af registre i onlinekataloger*, som omhandlede registeroplægning inklusive tegnsortering, normaliseringsregler vedrørende alfabetisering og ledopdelingsregler.

Den faktiske implementering i bibliotekerne har senere vist et behov for en enklere model. Samtidig har verden ændret sig. Bibliotekskatalogen udstilles mere og mere uden for bibliotekerne. En tilpasning til denne virkelighed vil gøre et samspil mellem biblioteker og verden omkring nemmere.

Som en konsekvens heraf er der i Bibliografisk Råd blevet taget en principbeslutning om at afskaffe alfabetiseringstegn (æ) og gammelt dansk aa, og som følge heraf revidere Ordning af registre i onlinekataloger.

Resultatet er den foreliggende *Udlægning til og ordning af registre i onlinekataloger*, som er udarbejdet af en arbejdsgruppe bestående af Leif Andresen, Hanne Hørl Hansen, Leif Makhholm og Camilla Riis Petersen, og som blev principielt godkendt af Bibliografisk Råd september 2012. Endelig beslutning blev udsat indtil der var truffet beslutning om håndtering af foranstillede artikler ved dansk implementering af RDA. Herefter har Leif Andresen, Hanne Hørl Hansen og Leif Makhholm foretaget en gennemskrivning og Bibliografisk Råd har september 2017 endeligt godkendt specifikationen.

Hovedprincippet i den nye specifikation er at alle tegn normaliseres maskinelt til ingen, et eller flere tegn, der udlægges til søgeregistre og som ved listning på skærm ordnes alfanumerisk tegn-for-tegn. Der er i en række tilfælde behov for flere udlægningsformer for at tilgodese brugerbehov. Dette håndteres i enkelte tilfælde som automatisk dublering (eksempelvis å som både å og aa) og dubleringen af former henholdsvis med og uden artikel i henhold til §9.1.1.-9.1.2. I andre tilfælde håndteres brugernes behov via supplerende inddatering og udlægning af et alternativt "skriftbillede" (eksempelvis således at en form med romertal suppleres af den tilsvarende form med arabertal eller udvalgte præfikser sammenskrives med navnet). I andre tilfælde igen sker relevant supplerende inddatering via inddatering som variantform (eksempelvis får titlen *Kjerlighed uden Strømper* supplerende inddateret variantformen *Kærlighed uden Strømper*).

Der er ingen afvigelser fra den rent alfanumeriske ordning tegn-for-tegn.

Der er endvidere i en række tilfælde sket forenklinger i forhold til udgaven fra 1999. Alle regler er formuleret beregnet for maskinel håndtering eller for håndtering med stort delfelt ved inddatering.

Specifikationen skal ses i sammenhæng med anvendelse af UNICODE som tegnsæt. Principielt har danMARC2 fuld UNICODE som referencetegnsæt, også selv om nogle systemer (DanBib, DDElibra m.v.) har anvendt den særlige indkodning til Latin1.

Alle UNICODE tegn er principielt dækket af denne specifikation, men der er i sagens natur ikke en specifikation tegn-for-tegn, men nogle generelle regler. Der er eksempler på hvordan ofte forekommende tegn uden for det danske alfabet og cifrene 0 - 9 normaliseres.

Det skal understreges at specifikationen alene omhandler udlægning og ordning. Hvorledes sætningsregistre præsenteres ligger uden for denne specifikation.

Det skal afslutningsvis understreges, at Bibliografisk Råd er opmærksom på at der kan være nogle anbefalinger, som det kan være vanskeligt at få implementeret i internationalt baserede systemer.

*Alfabetiseringsregler for danske biblioteker og bibliografier* kan fortsat betragtes som vejledende for ordning af bibliografiske lister i trykt form.

## Indledning

Specifikationen tager udgangspunkt i det danske tegnsæt, da den retter sig mod danske kataloger.

Hovedprincipperne i ordningen er:

- det danske alfabet inklusive æ ø å skal anvendes
- den fastsatte rækkefølge af tegn skal følges konsekvent
- alfabetiseringen skal baseres på skriftbilledet

Reglerne for numerisk ordning af tal er ophævet. Tal ordnes ciffer for ciffer. Der kan dog være problemer med implementeringen af meget specielle regler, som § 5.5. Se også § 5.6, der giver mulighed for at ordne særlige oplysninger numerisk.

## 1. Terminologi

I specifikationen anvendes følgende terminologi:

**Normalisering af enkelttegn:** Den værdi et tegn i ordnings- og søgesammenhæng skal opfattes som værende lig med hvad enten værdien er lig tegnet selv (fx a som a) eller et eller flere andre tegn (fx ä som æ, ß som ss og ½ som 1 2).

**Ordning:** Den indbyrdes sortering af de normaliserede enkelttegn.

**Udlægning til registre:** Den sekvens af tegn man ønsker repræsenteret i registret. Der kan være tale om en tegnsekvens ud fra en normalisering af de enkelte tegn i sekvensen, eller en tegnsekvens repræsenterende en semantisk værdi forskellig fra værdien af den normaliserede tegnsekvens .

**Supplerende udlægning til registre:** Når enkelttegn eller en tegnsekvens formodes at ville blive søgt på flere forskellige måder af brugerne, laves der supplerende udlægninger, så brugeren får et resultat uanset søgemåde.

**Dublerende udlægning til registre:** En automatisk udlægning af et tegn eller en tegnsekvens som to forskellige enkelttegn eller tegnsekvenser.

## 2. Generelt om udlægning

**2.1** Udgangspunktet for udlægning til registre er tegnsekvenser, som de fremgår af felterne i katalogposten (skriftbilledet). Den konkrete udlægning er derfor teksten som den står justeret efter reglerne i afsnit 4 – 10.

**2.2** Supplerende udlægning til registre kan ske ved hjælp af en af følgende metoder:

1. Ud fra tegnbestemte regler, der angiver at en dublerende udlægning indeholdende en supplerende normalisering bør finde sted – fx at udlægge å som både aa og å. Dublerende udlægning forudsættes løst maskinelt.
2. Ud fra supplerende inddatering.

## 3. Tegnenes indbyrdes ordning

**3.1** Ordning af tegn sker på den normaliserede form. Typografiske koder kan ikke normaliseres til tegn. I enkelte tilfælde dog til blanktegn, se § 5.5.

### 3.2 Tegnenes rækkefølge

Den grundlæggende rækkefølge er:

1. Blanktegn (mellemrum)
2. Symboler
3. Tal
4. Bogstaver

### 3.3 Blanktegn

Blanktegn ordnes foran alle andre tegn (»intet kommer før noget«). Reglen bevirker, at et kortere ord ordnes foran et længere ord, hvis begyndelse er identisk med det kortere.

### 3.4 Symboler

Følgende tegn regnes som symboler og ordnes i følgende rækkefølge (Jævnfør UNICODE, locale da\_DK):

#	nummertegn
\$	dollar
%	procent
&	et (og, ampersand)
*	asterisk
+	plus
<	mindre end
=	lig med

>	større end
@	commercial at
¢	cent
£	pund sterling
¥	yen
§	paragraf
©	copyright
¬	logisk ikke
®	registreret varemærke
°	grader
±	plus minus
·	gange
×	gange
÷	division/minus
€	euro
∞	uendelig

*Note: I de tilfælde, hvor et system anvender nogle af disse tegn som operatører, er det op til det pågældende system at håndtere brugen af det/de pågældende tegn som data.*

### 3.5 Tal

Ordningen baseres på cifrene 0-9 i denne rækkefølge:

0123456789

### 3.6 Bogstaver

Ordningen baseres på det danske alfabet, dvs. følgende små bogstaver i denne rækkefølge:

abcdefghijklmnopqrstuvxyzæøå

## 4. Udlægning og normalisering af bogstaver

**4.1** De små bogstaver udlægges som de er.

**4.2 Andre bogstaver og bogstavforbindelser** end tegn nævnt i § 3.6 normaliseres efter følgende principper:

**4.2.1** Store bogstaver normaliseres til de tilsvarende små bogstaver

**4.2.2** Umlautbogstaverne normaliseres til det tilsvarende bogstav i det danske alfabet.

*Eksempler:*

ä normaliseres til æ

ö og ungarsk ő normaliseres til ø

ü og ungarsk ú normaliseres til y (se også § 4.4.3)

**4.2.3** I øvrigt ses der bort fra accenter og andre diakritiske tegn.

*Eksempler:*

ë, é, è, ê og ē normaliseres til e

ç og č normaliseres til c

† normaliseres til l

**4.2.4** Ikke-danske bogstaver i det latinske alfabet normaliseres til "nærmeste" danske bogstaver.

*Eksempler:*

œ (sammenskrevet oe) normaliseres til oe

ð (islandsk ð) normaliseres til d

þ (islandsk thorn) normaliseres til th

ı (tyrkisk i uden prik) normaliseres til i

**4.3 Andre alfabeter.** Bogstaver fra andre alfabeter end det latinske *translittereres til* latinske bogstaver og *normaliseres* som disse. Valg af translitterationsstandard ligger uden for denne specifikation.

*Eksempler:*



Bogstav	Translittereret efter	Latinsk bogstav	Normaliseret til
Б	ISO 9	b	b
В	ISO 9	v	v
Ж	ISO 9	ž	z
Ю	ISO 9	Ů	u
Ю	DS 378	Ju	ju
Δ	DS/ISO 843	D	d
Θ	DS/ISO 843	TH	th
Ψ	DS/ISO 843	PS	ps

#### 4.4 Supplerende og dublerende normalisering

For at sikre at bibliotekssystemernes brugere kan få relevante svar på søgninger, er det muligt supplerende at normalisere bogstaver til alternative bogstaver i dublerende eller supplerende udlægning til søgeregistrene.

**4.4.1** Da aa og å i mange sammenhænge repræsenterer samme "bogstav", normaliseres de supplerende til den anden form, dvs. å får dublerende normalisering til aa og aa får dublerende normalisering til å.

*Note: I det omfang at "gammelt dansk å" tidligere indtastet som "@å" fortsat forefindes, udlægges dette tegn også både som aa og å.*

**4.4.2** De 2 danske bogstaver æ og ø kan normaliseres dublerende til tobogstavsforbindelserne ae og oe.

**4.4.3** For nogle ord kan de foranstående regler suppleres med andre supplerende udlægninger, hvis det i det enkelte tilfælde sikrer en søgbarhed, som kan forventes at afspejle opslag fra brugere.

*Eksempel i danMARC2:*

\*A Leus \*a Leüs

*Note: Ovenfor er givet et eksempel, hvor det diakritiske tegn "ü" kun har betydning for udtalen (navnet Leüs) og ikke for ordning af f.eks. franske ord og navne. I dette tilfælde kan stort delfelt bruges til at angive den supplerende udlægning. Herved opnås samme funktion som trema gav i det tidligere DanBib-tegnsæt.*

## 5. Udlægning og normalisering af tal

**5.1** Ved udlægning af tal til søgeregistre sker dette som hovedregel efter dokumentets form med de tilføjelser, som fremgår af efterfølgende.

### 5.2 Udlægning til sætningsregistre

Ved udlægning til sætningsregistre skal tal udlægges ciffer-for-ciffer.

### 5.3 Normalisering af romertal

For romertal inddateres supplerende en værdi for det tilsvarende arabertal.

*Eksempel i danMARC2:*

\*a Louis \*E 14 \*e XIV \*f konge af Frankrig

### 5.4 Brøker

Et brøk-tegn udlægges først med tælleren, dernæst med nævneren. Brøkstregen behandles som et blanktegn.

- 5.4.1** Er et helt tal sat foran brøken, ordnes talstørrelsen som om der var et blanktegn mellem det hele tegn og brøken.

*Eksempel:*

2½ ordnes som 2 1 2

- 5.5** Hævede og sænkede tal normaliseres til de tilsvarende cifre. Hvis det hævede eller sænkede tal er kombineret med et tal, indføres et blanktegn mellem tallet og det hævede eller sænkede tal. Dette gælder både ved typografiske tegn og ved specifikke UNICODE-tegn.

*Eksempel:*

10<sup>2</sup> inddateret som 10<sup>2</sup>, udlægges som 10 2

CO<sub>2</sub> inddateret som CO@UDA2, udlægges som CO2

10<sup>4</sup> inddateret som 10@UD94, udlægges som 10 4

- 5.6 Numerisk ordning af tal.** Hvis der ønskes en numerisk ordning af tal, kan der maskinelt udlægges foranstillede nuller i udpegede delfelter efter behov, f.eks. som stort delfelt ved serienumre.

- 5.7** For nogle tal og talkombinationer kan de foranstående regler suppleres med andre supplerende udlægningsregler, hvis det i det enkelte tilfælde sikrer en søgbarhed, som kan forventes at afspejle opslag fra brugere.

## 6. Udlægning og normalisering af symboler

- 6.1** Symbolerne nævnt i § 3.4 udlægges som de er.

- 6.2** Symboler, udover de i § 3.4 nævnte, normaliseres til # (nummertegn).

- 6.3** For nogle symboler kan de foranstående regler suppleres med andre supplerende udlægningsregler, hvis det i det enkelte tilfælde sikrer en søgbarhed, som kan forventes at afspejle opslag fra brugere, f.eks. ved en titel udelukkende bestående af symboler.

*Eksempler:*

Knud, Otto & Carmen Rosita  
*suppleres med opslag på varianten:* Knud, Otto og Carmen Rosita

§ 11 udbygningsplaner i henhold til den sociale styrelseslov  
*suppleres med opslag på varianten:* Paragraf 11 udbygningsplaner i henhold til den sociale styrelseslov

## **7. Udlægning og normalisering af blanktegn (mellemlinje) o.lign.**

- 7.1** Blanktegn eller et andet tegn normaliseret til blanktegn i starten af en tegnsekvens ignoreres.
- 7.2** Kombinationer af blanktegn og tegn, der *normaliseres til* blanktegn *normaliseres til* ét blanktegn.

*Eksempel:*

Post- og telegrafvæsen *normaliseres til* post og telegrafvæsen

## 8. Udlægning og normalisering af skilletegn

8.1 Skilletegn	Skriftbillede	Normaliseres som
Anførselstegn	"	Ingenting
Apostrof	'	Ingenting
Bindestreg	-	Blanktegn
Kolon	:	Ingenting
Komma	,	Ingenting
Parenteser	()[]{}	Ingenting
Punktum	.	Ingenting
Semikolon	;	Ingenting
Skråstreg	/	Blanktegn
Tankestreg	-	Blanktegn
Udråbstegn	!	Blanktegn
Underscore	_	Blanktegn

*Eksempler:*

*2 4/5 udlægges som 2 4 5*

*"Basse" udlægges som basse*

**8.2** Et skilletegn ignoreres, når det optræder som første tegn i en tegnsekvens (se § 7.1).

**8.3** Ved udlægning til særlige registre, f.eks. for klassifikationskoder, kan skilletegn have betydning for ordningen. For disse kan udlægning ske med anden udlægning af skilletegn end defineret i denne specifikation.

Note: Eksempler på felter, der udlægges i særlige registre, kan være 080 UDK (og 082 DDC)

## 9. Udlægning til sætningsregistre

### 9.1 Titelse registre og andre ordsammenstillinger

**9.1.1** Ved udlægning til sætningsregistre bestående af titler, skal den fulde titel udlægges inklusiv eventuel indledende artikel.

**9.1.2** Der bør foretages dublerende udlægning til sætningsregistre, hvis titlen indledes med ét af ordene på efterfølgende liste. Den dublerende udlægning sker uden det pågældende ord:

en, et, den, det, the, a, der, die, das

**9.1.3** Der kan foretages dublerende udlægning til sætningsregistre, hvis titlen indledes med andre ord, som på et anvendt sprog er en artikel.

**9.1.4** Reglen om dublerende udlægning i 9.1.2 gælder også selvom det pågældende ord ikke er en artikel, men et talord eller påpegende stedord, som f.eks. den, det, m.v.

## **9.2 Navneregistre**

**9.2.1** Særskilte præfikser og artikler og kombinationer heraf, udlægges som de fremgår af navnet.

**9.2.2** Ved efternavne, der begynder med et særskilt skrevet præfiks (en præposition, en artikel eller en kombination heraf), bør der foretages dublerende udlægning, der ser bort fra blanktegn, som adskiller præfikset fra resten af navnet eller adskiller præfiksets enkelte ord (fx »de la«). Det samme gør sig gældende for stednavne af samme art.

*Eksempel i danMARC2:*

700 00 \*A delaroché \*a De la Roche \*h Mazo

**9.2.3** Særlig regel: Præfikserne Mc og M<sup>c</sup> (hævet c) normaliseres til mc. Der foretages supplerende udlægning under mac.

*Eksempel i danMARC2:*

700 00 \*A macmanus \*a McManus

## **9.3 Generelt for sætningsregistre**

**9.3.1** Foranstående regler vedrører alene udlægning. For udlægning af tal i specielle tilfælde, henvises til § 5.6 og § 8.3. Præsentation ved listning af sætningsregistre ligger uden for denne specifikation.

**9.3.2** For nogle titler og navne kan de foranstående regler suppleres med andre supplerende udlægningsregler, hvis det i det enkelte tilfælde sikrer en søgbarhed, som kan forventes at afspejle opslag fra brugere.

## 10. Forkortelser

10.1 Udlægning af initialer og forkortelser bestemmes ved inddatering af blanktegn.

*Eksempler:*

*Titler:*

H.C Andersen som ung.  
H.C. *udlægges som hc*

*Navneformen:*

Andersen, H. C.  
H. C. *udlægges som h c*

*Navneformen:*

Hansen , H.-C.  
H.-C. *udlægges som h c*