

# danMARC2 tegnrepertoiret

Version 5. februar 2010

Trykfejl rettet 3. september 2010

Note vedrørende indkodning af diakritiske tegn i unicode og danMARC2 indsat 31. oktober 2011

## Introduktion

Notatet beskriver danMARC2 tegnrepertoiret, herunder relation til det gamle fællestegnsæt og til unicode.

De bibliografiske baser Alba og Basis benyttede fællestegnsættet. Det var en variant af LC-tegnsettet<sup>1</sup>, som på daværende tidspunkt var det bedste bud på et standardtegnsett til biblioteksbrug.

DanMARC2-tegnsettet blev skabt i forbindelse med etableringen af DanBib. Det skete i 1993 ved sammenlægning af fælleskatalogerne Alba og Basis. Ved den lejlighed blev det besluttet at benytte 8 bit tegnsættet ISO 8859-1 (Latin 1)<sup>2</sup> med mulighed for at indlejre tegn fra multioktettegnsettet unicode, som netop var publiceret i sin første version<sup>3</sup>. Indlejringen skete ved hjælp af en særlig @-notation, som blev opfundet til lejligheden.

## Oversigt over tegnrepertoiret

DanMARC2 anvender ISO 8859-1 (Latin1), men tegnrepertoiret omfatter alle tegn i ISO 10646 Basic Multilingual Plane (BMP), da disse kan indlejres ved brug af @-notation. Tegnene i række 00 svarer til ISO 8859-1.

Eksempel:

Unicodetegnet U+0393 (greek capital letter gamma) indlejres som tekststrengen @0393.

Tegnrepertoiret indeholder tre tegn med særlig betydning:

- @ bruges som skiftetegn; skal det bruges i sin normale betydning noteres @@
- ¤ bruges som alfabetiseringstegn; skal det bruges i sin normale betydning noteres @¤ (bør ikke forekomme)
- \* kan benyttes som delfeltskilletegn i linjeforamt; hvis det forekommer i posten skal det altid noteres @\*

Tegnrepertoiret rummer tillige såkaldte @U koder, som anvendes for tegn fra fællestegnsættet, som oprindeligt ikke kunne findes i unicode. De fleste af disse tegn er efterfølgende fundet i unicode.

Endelig rummer det Aa og aa (gammelt dansk å) med en helt speciel indkodning, nemlig @Å og @å.

**Tabel 1. indlejring af unicode i danMARC2**

Unicode	danMARC2	Kommentar
U+0020 til U+00FF	20 til FF	Række 0 i BMP svarer til ISO 8859-1 (Latin-1). Tre tegn behandles specielt, fordi de har særlig betydning i danMARC2-tegnsettet – se nedenfor.
U+002A	40 2A	@* - asterisk
U+0040	40 40	@@ - commercial at
U+00A4	40 A4	@¤ - currency sign
U+0100 – U+FFFF	@0100 - @FFFF	Tegnet repræsenteres i danMARC2 ved @ efterfulgt af 4 hexadecimale tegn som svarer til kodeværdien
Tegn udenfor BMP		Det er pt. ikke muligt at repræsentere tegn som ligger udenfor BMP.

<sup>1</sup> Library of Congress' tegnsæt – også også kaldet MARC8 eller ALA-tegnsettet

<sup>2</sup> ISO/IEC 8859-1 Information technology -- 8-bit single-byte coded graphic character sets -- Part 1: Latin alphabet No. 1

<sup>3</sup> ISO/IEC 10646-1 Information technology -- Universal Multiple-Octet Coded Character Set (UCS). Part 1: Architecture and Basic Multilingual Plane. Marts 1993.

## Særlige forhold knyttet til diakritiske tegn:

En række diakritiske tegn i fællestegnsættet har to repræsentationer i unicode – en selvstændig og en kombinatorisk udgave. Alle kombinatoriske værdier findes i række 03 i unicode. For hvert af de kombinatoriske tegn anfører unicode hvilke beslægtede tegn der forekommer.

Unicode angiver særlige regler for diakritiske tegn, herunder:

- Diakritiske tegn findes både som selvstændige tegn, som kombinationstegn og sammenbygget med andre tegn.
- Kun kombinationstegn kan kombineres med andre tegn, og de efterstilles i givet fald det tegn de kombineres med.
- Regler for brug af flere kombinationstegn samtidigt.

Praksis i danMARC2-tegnrepertoiret er som udgangspunkt at anvende de sammenbyggede tegn. Når dette ikke er muligt anvendes følgende praksis, som desværre afviger fra unicode-praksis:

- Der byttes om på den selvstændige og den kombinatoriske værdi for de tegn der findes i begge udgaver – bortset fra tilde
- Det diakritiske tegn placeres foran det tegn som det skal kombineres med - dette gælder alle tegn

Nedenstående tabel viser et udvalg af de mest anvendte diakritiske tegn. Det understreges at Unicode rummer flere kombinatoriske tegn end de her viste.

**Tabel 2. Praksis for anvendelse af diakritiske tegn**

Dansk betegnelse	Kode	Selvstændigt tegn	Kode	Kombinatorisk tegn	Dansk praksis
circumflex	U+005E	circumflex accent	U+0302	combining circumflex accent	ombyttes
understreg	U+005F	low line	U+0332	combining low line	ombyttes
Grave	U+0060	grave accent	U+0300	combining grave accent	ombyttes
Tilde	U+007E	tilde	U+0303	combining tilde	ombyttes ikke
umlaut	U+00A8	diaeresis	U+0308	combining diaeresis	ombyttes
macron	U+00AF	macron	U+0304	combining macron	ombyttes
Aigu	U+00B4	acute sign	U+0301	combining acute accent	ombyttes
Cedille	U+00B8	cedilla	U+0327	combining cedilla	ombyttes
Hacek	U+02C7	caron	U+030C	combining caron	ombyttes
Breve	U+02D8	breve	U+0306	combining breve	ombyttes
bolle over bogstav (overcirkel)	U+02DA	ring above	U+030A	combining ring above	ombyttes
højrekrog	U+02DB	ogonek	U+0328	combining ogonek	ombyttes

## Note vedrørende indkodning af diakritiske tegn i unicode og danMARC2:

Unicode tillader alle sekvenser af kodeværdier (code points), også selvom de ikke giver mening. Accenter, diakritter mm kaldes samlet for kombinationstegn (combining characters). Praksis er, at kodeværdierne for sådanne kombinationstegn placeres efter kodeværdien for det grundtegn, som de skal kombineres med. Hvis der er tale om en sekvens af flere kombinatoriske tegn, skal kodeværdierne som hovedregel placeres efter inside-out reglen. Hvis kombinationstegnene er stakket enten over eller under grundtegnet, indebærer denne regel, at kodeværdien for det inderste tegn noteres først, dernæst kodeværdien for det næstinderste tegn etc. afsluttende med kodeværdien det yderste tegn. Hvis man ønsker et a med circumflex og aigu (som begge befinder sig over bogstavet), bestemmes rækkefølgen således af, om det er circumflex'en eller aigu'en, der er øverst.

Grafik	Tegnsekvens	Kommentar
á	LATIN SMALL LETTER A (U+0061) + COMBINING CIRCUMFLEX (U+0302) + COMBINING ACUTE (U+0301)	aigu er øverst
â	LATIN SMALL LETTER A (U+0061) + COMBINING ACUTE (U+0301) + COMBINING CIRCUMFLEX (U+0302)	circumflex er øverst

DanMARC2 tegnrepertoiret rummer som bekendt alle unicodetegn, og sætter ligesom unicode ingen berænsning for i hvilken rækkefølge de forekommer. Hvor unicode sætter kodeværdierne for de kombinatoriske tegn efter kodeværdien for grundtegnet, anvender danMARC2 det spejlede princip, hvor kodeværdierne sættes foran kodeværdien for det grundtegn, som det skal kombineres med, som hovedregel således at det yderste kommer først og det inderste sidst. Konvertering mellem danMARC2 og unicode baserer sig på denne spejlvending.

Bemærk at mange kombinationstegn (eksempelvis i græsk, hebræisk og arabisk) har særlige typografiske traditioner, som indebærer at inside-out reglen ikke gælder. Se nærmere i unimarc-dokumentationen. Bemærk endvidere at samme grafiske præsentation i visse tilfælde kan opnås på flere forskellige måder i unicode:

Grafik	Tegnsekvens	Kommentar
á	LATIN SMALL LETTER A WITH CIRCUMFLEX AND ACUTE (U+1EA5)	én kodeværdi
á	LATIN SMALL LETTER A WITH CIRCUMFLEX (U+00E2) + COMBINING ACUTE (U+0301)	sekvens af to kodeværdier
á	LATIN SMALL LETTER A (U+0061) + COMBINING CIRCUMFLEX (U+0302) + COMBINING ACUTE (U+0301)	sekvens af tre kodeværdier

Denne eksistens af ækvivalente kodesekvenser i unicode har medført definitionen af fire normaliseringsformer - NFD, NFKD, NFC og NFKC - som kort fortalt indebærer, at man enten bruger flest mulige tegn eller færrest mulige tegn.

## Konvertering til unicode

Mens det hele tiden har været muligt at konvertere fra unicode til danMARC2, har det modsatte ikke været tilfældet. Men det er det nu – jævnfør tabel 3.

Uddybende kommentarer:

- Tegnene \*, @ og ꞥ har særlig betydning i danMARC2-tegnsettet. Disse tegn konverteres ikke til unicode.
- Gammelt dansk å og Å kan repræsenteres på to måder i danMARC2, henholdsvis @å, @Å og @A733, @A732.
- De fleste @U-koder har enten fundet en placering i unicode eller anvendes ikke i danMARC2.
- Brugen af typografisk op og ned har ikke fulgt en entydig praksis. Som ny praksis anvendes unicode superscript og subscript, som omfatter cifrene fra 0 til 9, venstre- og højreparentes, plus, minus, lighedstegn samt n.
- De tre skilletegn anvendes ikke i danMARC2, jævnfør tabel 4.
- Ifølge gældende praksis skelnes ikke mellem umlaut og trema, jævnfør tabel 4.

**Tabel 3. Mapping fra danMARC2-tegnrepertoiret til unicode og tilbage igen**

DM2	Unicode		Unicode	dM2
20 – FF	U+0020 - U+00FF	ISO 8859-1 (Latin-1) svarer række 0 i BMP	U+0020 - U+00FF	20 - FF
*		delfeltskilletegn i præsentationsformater		
@		skiftetegn i dM2		
ꞥ		alfabetiseringstegn i dM2		
@0100 - @FFFF	U+0100 - U+FFFF	standardmapping	U+0100 - U+FFFF	@0100 - @FFFF
@*	U+002A	asterisk	U+002A	@*
@@	U+0040	commercial at	U+0040	@@
@ꞥ	U+00A4	currency code – anvendes ikke i danske poster!	U+00A4	@ꞥ
@å	U+A733	gammelt dansk å - alternativ form @A733	U+A733	@A733
@Å	U+A732	gammelt dansk Å - alternativ form @A732	U+A732	@A732
@U5E		skiftetegn		
@UA7	U+02B9	slavisk blødt tegn	U+02B9	@02B9
@UAE	U+02BC	alif	U+02BC	@02BC
@UB0	U+02BB	ayn	U+02BB	@02BB
@UB7	U+02BA	slavisk hårdt tegn, tommer	U+02BA	@02BA
@UE0	U+0309	pseudo spørgsmålstegn	U+0309	@0309
@UEB	U+0361	ligatur - første halvdel	U+0361	@0361
@UEC		ligatur - anden halvdel		
@UF7	U+0326	venstrekrog	U+0326	@0326
@UF8	U+031C	omvendt cedille	U+031C	@031C
@UF9	U+032E	upadhmaniya	U+032E	@032E
@UFA	U+0360	dobbelt tilde - første halvdel	U+0360	@0360
@UFB		dobbelt tilde - anden halvdel		
@UD8	U+2014	lang streg (geviert-streg)	U+2014	@2014
@UD9		typografisk op (superscript) – overflødig		
@UDA		typografisk ned (subscript) – overflødig		
@UDD		postskilletegn – overflødig		
@UDE		feltskilletegn – overflødig		
@UDF		delfeltskilletegn – overflødig		
@UFC		trema – overflødig		

## **Konvertering fra fællestegnsættet til danMARC2**

Mapningen fremgår af tabel 4. Tabellen blev oprindeligt dannet for at kunne konvertere poster fra de Alba og Basis, en opgave der fandt sted under stærkt tidspres i 1993.

Da den hidtidige katalogiseringspraksis byggede på tegn fra fællestegnsættet, har tabellen også fungeret som et vigtigt grundlag for praksis for brug af danMARC2-tegnrepertoiret.

Ved inddatering til Basis anvendes en delmængde af danMARC2-tegnrepertoiret, som stort set svarer til det oprindelige fællestegnsæt. Det skal dog understreges, at det er betegnelser og kodeværdier fra unicode der er gældende i danMARC2-tegnrepertoiret.

### **Generelle kommentarer til tabel 4.**

Visse tegn i fællestegnsættet er mappet til flere værdier i unicode. Dette gælder eksempelvis diakritiske tegn, som findes både i en kombinatorisk og en ikke kombinatorisk udformning (ex. 5F, E1, E2, E3). Desuden har man i fællestegnsættet i flere tilfælde anvendt samme tegn i flere betydninger (ex. 9E, A3, B3)

En række tegn fra fællestegnsættet kunne i første omgang ikke lokaliseres i unicode. Som hovedregel blev disse tegn repræsenteret ved @U efterfulgt af to tegn som udpegede tegnets placering i fællestegnsættet.

I juli 1999 udgav LC en tabel som beskrev konvertering fra LC-tegnættet til unicode. Dette bevirkede at de fleste tegn fra fællestegnsættet nu kunne lokaliseres i unicode. LC's konverteringstabel er efterfølgende blevet revideret tre gange i 2004 og en enkelt gang i 2005. LC betragtes som autoriteten på dette område. Ændringerne slår derfor også igennem for danMARC2 tegnsættet.

**Tabel 4. Mapning fra fællestegnsættet til danMARC2**

Fællestegnsæt			UNICODE		danMARC2	
kode	tegn	betydning	kode	betydning	kode(r)	grafisk
20		blank	00 20	space	20	space
21	!	udråbstegn	00 21	exclamation mark	21	!
22	"	anførselstegn	00 22	quotation mark	22	"
23	#	nummertegn	00 23	number sign	23	#
24	\$	dollartegn	00 24	dollar sign	24	\$
25	%	procent	00 25	percent sign	25	%
26	&	et-tegn	00 26	ampersand	26	&
27	'	apostrof	00 27	apostrophe	27	'
28	(	rund venstre parentes	00 28	left parenthesis	28	(
29	)	rund højre parentes	00 29	right parenthesis	29	)
2A	*	asterisk	00 2A	asterisk	40 2A	@*
2B	+	plus tegn	00 2B	plus sign	2B	+
2C	,	komma	00 2C	comma	2C	,
2D	-	bindestreg, minus	00 2D	hyphen-minus	2D	-
2E	.	punktum, decimalpunkt	00 2E	full stop	2E	.
2F	/	skråstreg	00 2F	solidus	2F	/
30	0	ciffer 0	00 30	digit zero	30	0
31	1	ciffer 1	00 31	digit one	31	1
32	2	ciffer 2	00 32	digit two	32	2
33	3	ciffer 3	00 33	digit three	33	3
34	4	ciffer 4	00 34	digit four	34	4
35	5	ciffer 5	00 35	digit five	35	5
36	6	ciffer 6	00 36	digit six	36	6
37	7	ciffer 7	00 37	digit seven	37	7
38	8	ciffer 8	00 38	digit eight	38	8
39	9	ciffer 9	00 39	digit nine	39	9
3A	:	kolon, division	00 3A	colon	3A	:
3B	;	semikolon	00 3B	semicolon	3B	;
3C	<	mindre end tegn	00 3C	less-than sign	3C	<
3D	=	lighedstegn	00 3D	equals sign	3D	=
3E	>	større end tegn	00 3E	greater-than sign	3E	>
3F	?	spørgsmålstegn	00 3F	question mark	3F	?
40	ü	lille bogstav tysk y	00 FC	latin small letter u with diaeresis	FC	ü
41	A	stort bogstav a	00 41	latin capital letter a	41	A
42	B	stort bogstav b	00 42	latin capital letter b	42	B
43	C	stort bogstav c	00 43	latin capital letter c	43	C
44	D	stort bogstav d	00 44	latin capital letter d	44	D
45	E	stort bogstav e	00 45	latin capital letter e	45	E
46	F	stort bogstav f	00 46	latin capital letter f	46	F
47	G	stort bogstav g	00 47	latin capital letter g	47	G
48	H	stort bogstav h	00 48	latin capital letter h	48	H
49	I	stort bogstav i	00 49	latin capital letter i	49	I
4A	J	stort bogstav j	00 4A	latin capital letter j	4A	J
4B	K	stort bogstav k	00 4B	latin capital letter k	4B	K
4C	L	stort bogstav l	00 4C	latin capital letter l	4C	L
4D	M	stort bogstav m	00 4D	latin capital letter m	4D	M
4E	N	stort bogstav n	00 4E	latin capital letter n	4E	N
4F	O	stort bogstav o	00 4F	latin capital letter o	4F	O

Fællestegnsæt			UNICODE		danMARC2	
kode	tegn	betydning	kode	betydning	kode(r)	grafisk
50	P	stort bogstav p	00 50	latin capital letter p	50	P
51	Q	stort bogstav q	00 51	latin capital letter q	51	Q
52	R	stort bogstav r	00 52	latin capital letter r	52	R
53	S	stort bogstav s	00 53	latin capital letter s	53	S
54	T	stort bogstav t	00 54	latin capital letter t	54	T
55	U	stort bogstav u	00 55	latin capital letter u	55	U
56	V	stort bogstav v	00 56	latin capital letter v	56	V
57	W	stort bogstav w	00 57	latin capital letter w	57	W
58	X	stort bogstav x	00 58	latin capital letter x	58	X
59	Y	stort bogstav y	00 59	latin capital letter y	59	Y
5A	Z	stort bogstav z	00 5A	latin capital letter z	5A	Z
5B	Æ	stort bogstav æ	00 C6	latin capital letter ae	C6	Æ
5C	Ø	stort bogstav ø	00 D8	latin capital letter o with stroke	D8	Ø
5D	Å	stort bogstav å	00 C5	latin capital letter a with ring above	C5	Å
5E	↑	skiftetegn <sup>1</sup>				
5F	_	understreg	00 5F	low line	5F	_
5F	_	understreg	03 32	combining low line	40 30 33 33 32	@0332
60	ä	lille bogstav svensk æ	00 E4	latin small letter a with diaeresis	E4	ä
61	a	lille bogstav a	00 61	latin small letter a	61	a
62	b	lille bogstav b	00 62	latin small letter b	62	b
63	c	lille bogstav c	00 63	latin small letter c	63	c
64	d	lille bogstav d	00 64	latin small letter d	64	d
65	e	lille bogstav e	00 65	latin small letter e	65	e
66	f	lille bogstav f	00 66	latin small letter f	66	f
67	g	lille bogstav g	00 67	latin small letter g	67	g
68	h	lille bogstav h	00 68	latin small letter h	68	h
69	i	lille bogstav i	00 69	latin small letter i	69	i
6A	j	lille bogstav j	00 6A	latin small letter j	6A	j
6B	k	lille bogstav k	00 6B	latin small letter k	6B	k
6C	l	lille bogstav l	00 6C	latin small letter l	6C	l
6D	m	lille bogstav m	00 6D	latin small letter m	6D	m
6E	n	lille bogstav n	00 6E	latin small letter n	6E	n
6F	o	lille bogstav o	00 6F	latin small letter o	6F	o
70	p	lille bogstav p	00 70	latin small letter p	70	p
71	q	lille bogstav q	00 71	latin small letter q	71	q
72	r	lille bogstav r	00 72	latin small letter r	72	r
73	s	lille bogstav s	00 73	latin small letter s	73	s
74	t	lille bogstav t	00 74	latin small letter t	74	t
75	u	lille bogstav u	00 75	latin small letter u	75	u
76	v	lille bogstav v	00 76	latin small letter v	76	v
77	w	lille bogstav w	00 77	latin small letter w	77	w
78	x	lille bogstav x	00 78	latin small letter x	78	x
79	y	lille bogstav y	00 79	latin small letter y	79	y
7A	z	lille bogstav z	00 7A	latin small letter z	7A	z
7B	æ	lille bogstav æ	00 E6	latin small letter ae	E6	æ
7C	ø	lille bogstav ø	00 F8	latin small letter o with stroke	F8	ø
7D	å	lille bogstav å	00 E5	latin small letter a with ring above	E5	å
7E	ö	lille bogstav svensk ø	00 F6	latin small letter o with diaeresis	F6	ö
7F		benyttet ikke <sup>2</sup>				

Fællestegnsæt			UNICODE		danMARC2	
kode	tegn	betydning	kode	betydning	kode(r)	grafisk
80	Δ	stort græsk bogstav delta	03 94	greek capital letter delta	40 30 33 39 34	@0394
81	Θ	stort græsk bogstav theta	03 98	greek capital letter theta	40 30 33 39 38	@0398
82	Λ	stort græsk bogstav lambda	03 9B	greek capital letter lamda	40 30 33 39 42	@039B
83	Ξ	stort græsk bogstav xi	03 9E	greek capital letter xi	40 30 33 39 45	@039E
84	Π	stort græsk bogstav pi	03 A0	greek capital letter pi	40 30 33 41 30	@03A0
85	Σ	stort græsk bogstav sigma	03 A3	greek capital letter sigma	40 30 33 41 33	@03A3
86	Φ	stort græsk bogstav phi	03 A6	greek capital letter phi	40 30 33 41 36	@03A6
87	Ω	stort græsk bogstav omega	03 A9	greek capital letter omega	40 30 33 41 39	@03A9
88	α	lille græsk bogstav alfa	03 B1	greek small letter alpha	40 30 33 42 31	@03B1
89	β	lille græsk bogstav beta	03 B2	greek small letter beta	40 30 33 42 32	@03B2
8A	γ	lille græsk bogstav gamma	03 B3	greek small letter gamma	40 30 33 42 33	@03B3
8B	δ	lille græsk bogstav delta	03 B4	greek small letter delta	40 30 33 42 34	@03B4
8C	ε	lille græsk bogstav epsilon	03 B5	greek small letter epsilon	40 30 33 42 35	@03B5
8D	ζ	lille græsk bogstav zeta	03 B6	greek small letter zeta	40 30 33 42 36	@03B6
8E	η	lille græsk bogstav eta	03 B7	greek small letter eta	40 30 33 42 37	@03B7
8F	θ	lille græsk bogstav theta	03 B8	greek small letter theta	40 30 33 42 38	@03B8
90	ι	lille græsk bogstav iota	03 B9	greek small letter iota	40 30 33 42 39	@03B9
91	κ	lille græsk bogstav kappa	03 BA	greek small letter kappa	40 30 33 42 41	@03BA
92	λ	lille græsk bogstav lambda	03 BB	greek small letter lamda	40 30 33 42 42	@03BB
93	μ	lille græsk bogstav my	03 BC	greek small letter mu	40 30 33 42 43	@03BC
94	ν	lille græsk bogstav ny	03 BD	greek small letter nu	40 30 33 42 44	@03BD
95	ξ	lille græsk bogstav xi	03 BE	greek small letter xi	40 30 33 42 45	@03BE
96	ο	lille græsk bogstav omikron	03 BF	greek small letter omicron	40 30 33 42 46	@03BF
97	π	lille græsk bogstav pi	03 C0	greek small letter pi	40 30 33 43 30	@03C0
98	ρ	lille græsk bogstav rho	03 C1	greek small letter rho	40 30 33 43 31	@03C1
99	σ	lille græsk bogstav sigma-2	03 C3	greek small letter sigma	40 30 33 43 33	@03C3
9A	τ	lille græsk bogstav tau	03 C4	greek small letter tau	40 30 33 43 34	@03C4
9B	υ	lille græsk bogstav ypsilon	03 C5	greek small letter upsilon	40 30 33 43 35	@03C5
9C	φ	lille græsk bogstav phi	03 C6	greek small letter phi	40 30 33 43 36	@03C6
9D	χ	lille græsk bogstav chi	03 C7	greek small letter chi	40 30 33 43 37	@03C7
9E	Ψ	stort græsk bogstav psi <sup>3</sup>	03 A8	greek capital letter psi	40 30 33 41 38	@03A8
9E	ψ	lille græsk bogstav psi	03 C8	greek small letter psi	40 30 33 43 38	@03C8
9F	ω	lille græsk bogstav omega	03 C9	greek small letter omega	40 30 33 43 39	@03C9
A0	ς	lille græsk bogstav sigma-1	03 C2	greek small letter final sigma	40 30 33 43 32	@03C2
A1	Ł	stort polsk bogstav l	01 41	latin capital letter l with stroke	40 30 31 34 31	@0141
A2	Γ	Stort græsk bogstav gamma	03 93	greek capital letter gamma	40 30 33 39 33	@0393
A3	Ð	stort islandsk bogstav blødt d	00 D0	latin capital letter eth (icelandic)	D0	Ð
A3	Đ	stort serbokroatisk bogstav d	01 10	latin capital letter d with stroke	40 30 31 31 30	@0110
A4	Þ	stort islandsk bogstav thorn	00 DE	latin capital letter thorn (icelandic)	DE	Þ
A5		benyttes ikke <sup>4</sup>				
A6	Œ	stort fransk oe	01 52	latin capital ligature oe	40 30 31 35 32	@0152
A7	’	slavisk blødt tegn	02 B9	modifier letter prime	40 30 32 42 39	@02B9 <sup>5</sup>
A8	·	prik midt på linjen	00 B7	middle dot	B7	·
A9	b	musikalsk b-tegn	26 6D	musical flat sign	40 32 36 36 44	@266D
AA	®	registreret varemærke	00 AE	registered sign	AE	®
AB	±	plusminus tegn	00 B1	plus-minus sign	B1	±
AC	Ō	stort bogstav o med krog	01 A0	latin capital letter o with horn	40 30 31 41 30	@01A0
AD	Ū	stort bogstav u med krog	01 AF	latin capital letter u with horn	40 30 31 41 46	@01AF
AE	’	alif	02 BC	modifier letter apostrophe	40 30 32 42 43	@02BC <sup>6</sup>
AF		benyttes ikke <sup>7</sup>				



Fællestegnsæt		UNICODE		danMARC2	
kode	tegn betydning	kode	betydning	kode(r)	grafisk
B0	´ ayn	02 BB	modifier letter turned comma	40 30 32 42 42	@02BB <sup>8</sup>
B1	ł lille polsk bogstav l	01 42	latin small letter l with stroke	40 30 31 34 32	@0142
B2	» anførselstegn (start)	00 BB	right-pointing double angle quot. mark	BB	»
B3	đ lille serbokroatisk bogstav d	01 11	latin small letter d with stroke	40 30 31 31 31	@0111
B3	lille islandsk bogstav blødt d	01 11	latin small letter d with stroke	40 30 31 31 31	@0111
B4	þ lille islandsk bogstav thorn	00 FE	latin small letter thorn (icelandic)	FE	þ
B5	« anførselstegn (slut)	00 AB	left pointing double angle quotation mark	AB	«
B6	œ lille fransk oe	01 53	latin small ligature oe	40 30 31 35 33	@0153
B7	´ slavisk hårdt tegn, tommer	02 BA	modifier letter double prime	40 30 32 42 41	@02BA <sup>9</sup>
B8	ı lille tyrkisk bogstav i	01 31	latin small letter dotless i	40 30 31 33 31	@0131
B9	£ pundtegn	00 A3	pound sign	A3	£
BA	ð islandsk/færøsk eth	00 F0	latin small letter eth (icelandic)	F0	ð
BB	benyttes ikke <sup>10</sup>				
BC	σ lille o med krog	01 A1	latin small letter o with horn	40 30 31 41 31	@01A1
BD	υ lille u med krog	01 B0	latin small letter u with horn	40 30 31 42 30	@01B0
BE	§ paragraftegn	00 A7	section sign	A7	§
BF	° gradtegn	00 B0	degree sign	B0	°
C0	† kors/daggert	20 20	dagger	40 32 30 32 30	@2020
C1	‰ promille	20 30	per mille sign	40 32 30 33 30	@2030
C2	∫ integraltegn	22 2B	integral	40 32 32 32 42	@222B
C3	∞ uendelighedstegn	22 1E	infinity	40 32 32 31 45	@221E
C4	√ rodtegn	22 1A	square root	40 32 32 31 41	@221A
C5	÷ minus	00 F7	division sign	F7	÷
C5	÷ division	00 2D	hyphen-minus	2D	-
C6	≠ ulighedstegn	22 60	not equal to	40 32 32 36 30	@2260
C7	→ højrepil	21 92	rightwards arrow	40 32 31 39 32	@2192
C8	× gange, kryds	00 D7	multiplication sign	D7	×
C9	Aa stort dansk bogstav gammelt å	A7 32	latin capital letter AA	40 41 37 33 32	@A732
C9	Aa stort dansk bogstav gammelt å			40 C5	@Å
CA	Ü stort tysk bogstav y	00 DC	latin capital letter u with diaeresis	DC	Ü
CB	Ä stort svensk bogstav æ	00 C4	latin capital letter a with diaeresis	C4	Ä
CC	Ö stort svensk bogstav ø	00 D6	latin capital letter o with diaeresis	D6	Ö
CD	[ skarp venstreparentes	00 5B	left square bracket	5B	[
CE	aa lille dansk bogstav gammelt å	A7 33	latin small letter aa	40 41 37 33 33	@A733
CE	aa lille dansk bogstav gammelt å			40 E5	@å
CF	{ krøllet venstreparentes (tuborg)	00 7B	left curly bracket	7B	{
D0	] skarp højreparentes	00 5D	right square bracket	5D	]
D1	} krøllet højreparentes (tuborg)	00 7D	right curly bracket	7D	}
D2	ß tysk dobbelt s	00 DF	latin small letter sharp s (german)	DF	ß
D3	´ lille grønlandsk bogstav k	01 38	latin small letter kra (greenlandic)	40 30 31 33 38	@0138
D4	benyttes ikke <sup>11</sup>				
D5	benyttes ikke <sup>12</sup>				
D6	¿ spansk omvendt spørgsmålstegn	00 BF	inverted question mark	BF	¿
D7	¡ spansk omvendt udråbstegn	00 A1	inverted exclamation mark	A1	¡
D8	– lang streg (geviert-streg)	20 14	em dash	40 55 44 38	@2014 <sup>13</sup>
D9	↑ typografisk op (superscript) <sup>14</sup>				
DA	↓ typografisk ned (superscript) <sup>15</sup>				
DB	∖ fremhævelse	00 5C	reverse solidus	5C	∖
DC	¤ alfabetiseringstegn	00 A4	currency sign	A4	¤
DD	^ post skilletegn <sup>16</sup>				
DE	➤< felt skilletegn <sup>17</sup>				
DF	◆ delfelt skilletegn <sup>18</sup>				

Fællestegnsæt		UNICODE		danMARC2	
kode	tegn betydning	kode	betydning	kode(r)	grafisk
E0	´ pseudo spørgsmålstegn	03 09	combining hook above	40 30 33 30 39	@0309 <sup>19</sup>
E1	` accent grave	00 60	grave accent	60	`
E1	` accent grave	03 00	combining grave accent	40 30 33 30 30	@0300
E2	ˆ accent aigu	00 B4	acute accent	B4	ˆ
E2	ˆ accent aigu	03 01	combining acute accent	40 30 33 30 31	@0301
E3	˘ circumflex	00 5E	circumflex accent	5E	˘
E3	˘ circumflex	03 02	combining circumflex accent	40 30 33 30 32	@0302
E4	˜ tilde	00 7E	tilde	7E	˜
E4	˜ tilde	03 03	combining tilde	40 30 33 30 33	@0303
E5	- macron, længdetegn	00 AF	macron	AF	-
E5	- macron, længdetegn	03 04	combining macron	40 30 33 30 34	@0304
E6	_ breve, bue eller korttegn	02 D8	breve	40 30 32 44 38	@02D8
E6	_ breve, bue eller korttegn	03 06	combining breve	40 30 33 30 36	@0306
E7	· prik over bogstav	03 07	combining dot above	40 30 33 30 37	@0307
E8	¨ omlyd (påvirker sorteringen)	00 A8	diaeresis	A8	¨
E8	¨ omlyd (påvirker sorteringen)	03 08	combining diaeresis	40 30 33 30 38	@0308
E9	ˇ hacek	02 C7	caron	40 30 32 43 37	@02C7
E9	ˇ hacek	03 0C	combining caron	40 30 33 30 43	@030C
EA	◌̇ bolle over bogstav (påv. sort.)	03 0A	combining ring above	40 30 33 30 41	@030A
EA	◌̇ bolle over bogstav (påv. sort.)	02 DA	ring above	40 30 32 44 41	@02DA
EB	‘ ligatur - første halvdel	03 61	combining double inverted breve	40 30 33 36 31	@0361 <sup>20</sup>
EC	‘ ligatur - anden halvdel <sup>21</sup>				
ED	ᵹ højt diakritisk komma	03 15	combining comma above right	40 30 33 31 35	@0315
ED	ᵹ højt diakritisk komma	03 13	combining comma above	40 30 33 31 33	@0313
EE	ˆˆ dobbelt aigu (påvirker sort.)	02 DD	double acute accent	40 30 32 44 44	@02DD
EE	ˆˆ dobbelt aigu (påvirker sort.)	03 0B	combining double acute accent	40 30 33 30 42	@030B
EF	‘ candrabindu	03 10	combining candrabindu	40 30 33 31 30	@0310
F0	¸ cedille	00 B8	cedilla	B8	¸
F0	¸ cedille	03 27	combining cedilla	40 30 33 32 37	@0327
F1	ˆ højrekrog	02 DB	ogonek	40 30 32 44 42	@02DB
F1	ˆ højrekrog	03 28	combining ogonek	40 30 33 32 38	@0328
F2	˙ prik under bogstav	03 23	combining dot below	40 30 33 32 33	@0323 <sup>22</sup>
F3	ˆˆ dobbelt prik under bogstav	03 24	combining double dot below	40 30 33 32 34	@0324
F4	˙ bolle under bogstav	03 25	combining ring below	40 30 33 32 35	@0325
F5	‘ dobbelt understregning	03 33	combining double low line	40 30 33 33 33	@0333
F6	no-brake blank (EN-tegn)	00 A0	no-brake space	A0	no-brake space
F7	˘ venstrekrog	03 26	combining comma below	40 30 33 32 36	@0326 <sup>23</sup>
F8	‘ omvendt cedille	03 1C	combining left half ring below	40 30 33 31 43	@031C <sup>24</sup>
F9	˘ upadhmaniya	03 2E	combining breve below	40 30 33 32 45	@032E <sup>25</sup>
FA	‘ dobbelt tilde - første halvdel	03 60	combining double tilde	40 30 33 36 30	@0360 <sup>26</sup>
FB	‘ dobbelt tilde - anden halvdel <sup>27</sup>				
FC	¨ trema (påvirker ikke sortering)	03 08	combining diaeresis	40 30 33 30 38	@0308 <sup>28</sup>
FD	benyttes ikke <sup>29</sup>				
FE	benyttes ikke <sup>30</sup>				
FF	benyttes ikke <sup>31</sup>				

- 
- <sup>1</sup> Har tidligere figureret som @U5E. Tegnet benyttes ikke i danMARC2.
- <sup>2</sup> Har tidligere figureret som @U7F. Tegnet benyttes ikke i danMARC2.
- <sup>3</sup> Fællestegnsættet brugte samme værdi for stor og lille psi.
- <sup>4</sup> Har tidligere figureret som @UA5. Tegnet benyttes ikke i danMARC2.
- <sup>5</sup> I 1999 rettet fra @UA7 til MODIFIER LETTER PRIME @02B9.
- <sup>6</sup> I 1999 rettet fra @UAE til MODIFIER LETTER RIGHT HALF RING @02BE. I 2008 rettet til MODIFIER LETTER APOSTROPHE @02BC.
- <sup>7</sup> Har tidligere figureret som @UAF. Tegnet benyttes ikke i danMARC2.
- <sup>8</sup> I 1999 rettet fra @UB0 til MODIFIER LETTER LEFT HALF RING @02BF. I 2008 rettet til MODIFIER LETTER TURNED COMMA @02BB.
- <sup>9</sup> I 1999 rettet fra @UB7 til MODIFIER LETTER DOUBLE PRIME @02BA.
- <sup>10</sup> Har tidligere figureret som @UBB. Tegnet benyttes ikke i danMARC2.
- <sup>11</sup> Har tidligere figureret som @UD4. Tegnet benyttes ikke i danMARC2.
- <sup>12</sup> Har tidligere figureret som @UD5. Tegnet benyttes ikke i danMARC2.
- <sup>13</sup> I 2009 rettet fra @UD8 til EM DASH @2014
- <sup>14</sup> Værdien @UD9 kan forekomme i poster, men der er svingende praksis for brug af dette tegn. Brug i stedet unicode superscript-tegn.
- <sup>15</sup> Værdien @UDA kan forekomme i poster, men der er svingende praksis for brug af dette tegn. Brug i stedet af unicode subscript-tegn.
- <sup>16</sup> Værdien @UDD kan forekomme i ældre poster. Tegnet benyttes ikke i danMARC2.
- <sup>17</sup> Værdien @UDE kan forekomme i ældre poster. Tegnet benyttes ikke i danMARC2.
- <sup>18</sup> Værdien @UDF kan forekomme i ældre poster. Tegnet benyttes ikke i danMARC2.
- <sup>19</sup> I 1999 rettet fra @UE0 til COMBINING HOOK ABOVE @0309.
- <sup>20</sup> I 1999 rettet fra @UEB til COMBINING LIGATURE LEFT HALF @FE20. Fra 2008 ændres dette til COMBINING DOUBLE INVERTED BREVE @0361. Praksis er nu at dette tegn træder i stedet for de to halvdele som tidligere blev inddateret.
- <sup>21</sup> I 1999 rettet fra @UEC til COMBINING LIGATURE RIGHT HALF @FE21. Udgået fra 2008.
- <sup>22</sup> I 1999 rettet fra @0340 til COMBINING DOT BELOW @0323.
- <sup>23</sup> I 1999 rettet fra @UF7 til COMBINING COMMA BELOW @0326.
- <sup>24</sup> I 1999 rettet fra @UF8 til COMBINING LEFT HALF RING BELOW @031C.
- <sup>25</sup> I 1999 rettet fra @UF9 til COMBINING BREVE BELOW @032E.
- <sup>26</sup> I 1999 rettet fra @UFA til COMBINING DOUBLE TILDE LEFT HALF @FE22. Fra 2008 ændres dette til COMBINING DOUBLE TILDE @0360. Praksis er nu at dette tegn nu træder i stedet for de to halvdele som tidligere blev inddateret.
- <sup>27</sup> I 1999 rettet fra @UFB til COMBINING DOUBLE TILDE RIGHT HALF @FE23. Tegnet udgår fra tabellen i 2008.
- <sup>28</sup> Praksis ændret således at der ikke skelnes mellem trema og umlaut
- <sup>29</sup> Har tidligere figureret som @UFD. Tegnet benyttes ikke i danMARC2.
- <sup>30</sup> Har tidligere figureret som @UFE. Tegnet benyttes ikke i danMARC2.
- <sup>31</sup> Har tidligere figureret som @UFF. Tegnet benyttes ikke i danMARC2.