

Dato: 7. maj 2007

Bilag:

Version: 1.2

Dokumentation af datamodel

Danmarks Arealinformation

Produktions- og distributionssystem

Indholdsfortegnelse

1	INDLEDNING	3
1.1	DANMARKS AREALINFORMATION	FEJL! BOGMÆRKE ER IKKE DEFINERET.
2	BAGGRUND FOR VALG AF DATAMODEL.....	4
3	KONSEKVENSER FORBUNDET MED VALG AF DATAMODEL	5
3.1	KONSEKVENSER I ET FAGLIGT PERSPEKTIV	5
3.2	KONSEKVENSER I ET SYSTEMPERSPEKTIV	6
4	DATAMODELBESKRIVELSE	7
4.1	GENERELLE DESIGNKRAV TIL DATAMODEL.....	7
4.2	DATAMODEL SOM BESKRIVER DATA	7
BILAG 1: DATAPRODUKTION - HARMONISERINGSPROCES		FEJL! BOGMÆRKE ER IKKE DEFINERET.

1 Indledning

Dette dokument redegør for overvejelser og valg relateret til opstilling af en fysisk datamodel for Danmarks Arealinformation for så vidt angår de fælles arealdata.

Overordnet set er formålet med den fysiske datamodel meget kort beskrevet at skabe en datamæssig ramme på databaseniveau, der kan være med til at sikre, at de fælles arealdata kan distribueres og ajourføres hensigtsmæssigt den 1. januar 2007.

Undervejs i dokumentet redegøres der for, hvorfor der er valgt "kun" at udarbejde en fysisk datamodel på databaseniveau og ikke et "fuldt" datamodelarbejde med afsæt i domæne og begrebsafklaring og først derefter fastlæggelse af logiske og fysiske datamodeller. Der redegøres desuden kort for, hvilke valg og begrænsninger, der har ligget i harmoniseringsarbejdet med de fælles arealdata.

1.1 Danmarks Arealinformation

Kort beskrevet har etableringen af Danmarks Arealinformation til formål at sikre, at de fælles arealdata er tilgængelige for de rette myndigheder og offentlighed den 1. januar 2007. Systemet har endvidere til formål at sikre, at de fælles arealdata, når de ajourføres af de "nye" myndigheder efter den 1. januar 2007, også er tilgængelige i en fælles løsning, så den fortsatte sagsbehandling kan ske på det kvalitative bedste grundlag. Tilgængeligheden af de fælles arealdata er en af forudsætningerne for, at sikker drift kan opretholdes den 1. januar 2007.

Danmarks Arealinformation skal bl.a. understøtte konkret sagsbehandling og øvrig opgaveløsning på miljøområdet ved at sikre adgang til de fælles arealdata og sikre adgang til at ajourføre de fælles arealdata. Det fælles system skal også sikre at myndigheder inden for andre områder end miljøområdet også har adgang til de fælles arealdata. Derudover skal Danmarks Arealinformation understøtte mål og tanker omkring digital forvaltning. Endelig skal systemet også sikre offentligheden adgang til de informationer, den har brug for eller forventer adgang til.

Danmarks Arealinformation kommer til at indeholde en række fælles data på plan-, natur- og miljøområdet. De respektive myndigheder bliver ansvarlige for at ajourføre data i Danmarks Arealinformation, så det altid er i Danmarks Arealinformation andre myndigheder og borgere i almindelighed ser og henter de fælles arealdata.

Danmarks Arealinformation vil blive opbygget, så det via en række services og brugerflader muliggør ajourføring og distribution af arealdata. Der er i princippet tale om et "dataleverandørsystem", der skal kunne tilbyde ajourførte arealdata til brug i de respektive myndigheders egne fagsystemer, herunder GIS-systemer og diverse forvaltnings-systemer.

De fælles arealdata vil udover, at de stilles til rådighed for myndigheder m.fl. også blive præsenteret i et fælles offentligt arealinformationssystem i lighed med det amterne og HUR har på deres hjemmesider i dag. Det vil her være muligt for borgere og andre, at se de forskellige arealreservationer i forhold til hinanden og foretage en række gængse forespørgsler på data.

Under et vil Det fælles offentlige Arealssystem blive en del af Danmarks Miljøportal både for så vidt angår ajourføring og distribution af arealdata og for så vidt angår den fælles præsentation af de fælles arealdata på Internettet.

Der er i arbejdet med at fastlægge IT-arkitekturen og brugergrænsefladerne for Danmarks Arealinformation arbejdet med to faser. En fase I, der er de databaser, services og brugergrænseflader, der skal være på plads den 1. januar 2007 og en fase II, der er de services og brugerflader, som skal udvikles umiddelbart i forlængelse af fase I, men som det tidsmæssigt ikke er muligt at nå.

Det er primært den meget begrænsede tid, der er til rådighed, der har gjort det nødvendigt at arbejde med en fase I og en fase II. Dertil kommer, at en stor del af de fælles arealdata helt naturligt er stærkt præget af den amtslige struktur og administrative inddeling, der er derfor også datamæssige aspekter i forhold til standardiseringsbehov, der bevirker, at det er hensigtsmæssigt at arbejde med en fase I og fase II.

2 Baggrund for valg af datamodel

I forbindelse med sikringen af de amtslige arealdata og etableringen af Danmarks Arealinformation har der vist sig et stort behov for at udarbejde en datamodel eller flere datamodeller for de arealdata, der skal indgå i Danmarks Arealinformation. Grundlæggende skyldes behovet at der ikke pt. findes datamodeller for de fælles arealdata.

Der har umiddelbart været to konsekvenser af den forskellighed som data har. For det første har det været nødvendigt, at forsøge at harmonisere de amtslige data til landsdækkende datasæt i den udstrækning dette har kunnet lade sig gøre. For det andet har det været nødvendigt at fastlægge en logisk datamodel for, hvordan de landsdækkende datasæt lagres i databaserne i Danmarks Arealinformation.

Overordnet set er formålet med den logiske datamodel meget kort beskrevet at skabe en datamæssig ramme på databaseniveau, der kan være med til at sikre, at:

- de fælles arealdata er tilgængelige i en så struktureret og ensartet form, som de sikrede amtslige data tillader
- de fælles arealdata kan vises i en så struktureret og ensartet form, som de sikrede amtslige data tillader
- de fælles arealdata kan distribueres i en så struktureret og ensartet form, som de sikrede amtslige data tillader.
- de fælles arealdata kan ajourføres i en standardiseret form af de ansvarlige myndigheder
- det er muligt at etablere en hensigtsmæssig brugerstyring på Danmarks Arealinformation
- det er muligt at forberede Danmarks Arealinformation på håndtering af historik

I en ideel verden, hvor der var mere tid og flere ledige ressourcer, burde der inden 1. januar 2007 have været lavet egentligt datamodelarbejde for de data, der kommer til at indgå i Danmarks Arealinformation. Et datamodelarbejde, der også burde have ophæng i Sektorstandardiseringsudvalget for miljødata. Dette har dog ikke været muligt og

udarbejdelsen af en logisk datamodel på databaseniveau er det nødvendige kompromis for at kunne have et fælles system klart den 1. januar 2007.

Det nødvendige valg betyder også, at der så snart tid og ressourcer tillader det skal gås videre med datamodellarbejdet og at der nedsættes arbejdsgrupper under sektorstandardiseringsudvalget for miljødata, f.eks. på naturområdet.

På et område er der dog allerede fuldt datamodellarbejde i gang under Sektorstandardiseringsudvalget. Det drejer sig om planområdet, hvor der er nedsat en arbejdsgruppe, der arbejder med at udvide PlanDK2 datamodellen til også at rumme de regionplanda-
ta, er fremover bliver en del af kommuneplanlægning.

3 Konsekvenser forbundet med valg af datamodel

I et følgende beskrives kort nogle af de konsekvenser af faglig og systemmæssig karakter, som valget af datamodel, herunder det begrænsede omfang af egentligt datamodellarbejde, fører med sig.

3.1 Konsekvenser i et fagligt perspektiv

De amtslige arealdata er ikke skabt ud fra en fælles datamodel, og som følge heraf har harmoniseringen måttet gennemføres i en balance mellem fordelene ved at skabe ensartede landsdækkende datasæt overfor ulemperne ved at ændre i datasættene for at tilpasse dem en harmoniseret datamodel.

De fælles arealdata er ikke komplette på tværs af amterne. Det vil sige, at ikke alle amter har foretaget en registrering af samtlige arealdata. Det har derfor været nødvendigt at fastlægge en vurdering af konsekvenserne ved en harmonisering af et temakort, der ikke indeholder data for et eller flere amter. Danmarks Arealinformation kan blandt andet anvendes af borgere til at klarlægge hvilke registreringer og udpegninger, der vedrører en given ejendom. Hvis der i den forbindelse forekommer ”huller” i datasæt på grund af manglende registreringer kan der være risiko for fejlfortolkninger. Løsningen har været at udfylde ”hullerne” således at restriktioner vises overalt i de amter, hvor der ikke foreligger en registrering. Problemstillingen kræver under alle omstændigheder en særlig forklaring af disse forhold for brugerne.

I mange af de fælles arealdata forekommer der geografiske overlap ved amtsgrænserne. Et amts registrering begrænser sig ikke præcist til amtets arealer, og herved forekommer der overlap, som ikke er sammenfaldende. Det kræver ressourcer og faglig indsigt, at vurdere hvilke objekter, der bør bevares, og hvilke der bør slettes. En ”afklipping” af det der ligger uden for amtet giver ikke nødvendigvis et fornuftigt resultat. I de fleste tilfælde må de pågældende overlap og eventuelle dobbeltregistreringer afklares af de nye myndigheder.

En anden problemstilling er forskellige registreringsmetoder. Amterne har eksempelvis anvendt forskellige geometriske objekttyper til registrering af samme data. Fredede områder er i visse tilfælde registreret som polygoner, mens andre amter har benyttet objekttyperne linjer og punkter. Når datasæt harmoniseres til samme fælles geometri-

ske objekttype kan der dels være objekter der mister information, dels objekter, der får tillagt information, som de ikke er tiltænkt.

Informationen som er tilknyttet de enkelte objekter er ligeledes forskellig fra amt til amt. Typisk er den så forskelligartet, at en harmonisering ikke har været mulig. Det kan være oplysninger om registrator, registreringsgrundlag, registrerings- og ajourføringstidspunkt, kobling til databaser mv. Enkelte af amternes temakort har indeholdt information om klassifikation, som i en vis udstrækning har været harmoniseret. Eksempelvis de beskyttede naturtypers klassificering i Eng, Hede, Mose, Overdrev, Strandeng og Sø. I de tilfælde er harmoniseringen også gennemført for attributdata.

For ikke at miste de amtsspecifikke informationer er der dannet historiske attributtabel-ler til alle temakort, som via nøglefelt kan joines på det harmoniserede landsdækkende datasæt. De nye myndigheder vil således på objektniveau kunne tilknytte amternes historiske attributdata til det gældende datasæt.

3.2 Konsekvenser i et systemperspektiv

Valget af en datamodel har selvfølgelig betydning for den anvendelse af data, der er mulig. Implementering af den valgte model giver således mulighed for adgang til de egenskaber som hvert objekt består af, samt via databasens funktionalitet til de søgninger (ændringer) der understøttes af data og databaseskema.

Danmarks Arealinformation er opbygget på en SQL-baseret objekt-relationsdatabase (Oracle 10g) som tilgås via applikationsservere. De udvalgte egenskaber i datamodellen er pga. de tidligere omtalte vilkår gjort relativ smal (ved at reducere objekttypens specielle egenskaber) og med vilje ikke mere kompleks end de anvendte søgemekanismer nødvendiggør.

Et egentligt datawarehouse (database for data fra flere forskellige kilder) med et antal dimensions-tabeller som indgang til søgning og en bred fakta-tabel er det ideelle design for en database med de specificerede anvendelser. Løsningen dækker specielt fakta-tabellen som af systemmæssige årsager er splittet op i areal- og plandata fakta-tabeller.

Data-(objekter) er desuden skåret i de nye kommune- og regiongrænser, således at brugere ved opdatering ikke kommer til at rette i dataobjekter udenfor deres egen jurisdiktion.

Søgninger er koncentreret om følgende datafelter:

- Geometri (punkt-, linje- og fladegeometri)
- Gyldighedsperiode (datointerval som forberedelse til historik)
- Administrativ inddeling (kommune, region)
- Tema- og objekttype (temakode, objekttypekode)
- Objekt-identifikation (db-nøgle og UUID) og
- nogle mindre vigtige egenskabsfelter

De meget anvendte koder har også et tekstfelt med klartekst, således at brugeren vælger et søgekriterium i klartekst, men udtrækket anvender den tilhørende kode.

Der er valgt et landsdækkende koordinatsystem (UTM zone 32 i det nye europæiske datum ETRS89) – som også gælder for områder på Østsjælland (øst for 12° øst for Greenwich) og Bornholm som koordinatsystemteknisk ligger i zone 33.

Historikken er valgt implementeret som en gyldighedsperiode (*gyldig-fra-dato/tid* -- *gyldig-til-dato/tid*), således at alle ændringer gemmes. På et givet tidspunkt vil man kunne udvikle en funktion som vil kunne anvende historiske objekter – indtil da vil man kun kunne få at se den gældende situation.

Der er af hensyn til en kortsigtet fremtidssikring sat tomme felter af til egenskaber, som endnu ikke er efterspurgt og dermed specificeret.

4 *Datamodelbeskrivelse*

4.1 **Generelle designkrav til datamodel**

Den konceptuelle datamodel af de aktuelle arealdata burde egentlig beskrives ud fra de formål, forvaltning eller forretning hvortil de er indsamlede. Men omstændighederne omkring udviklingen af Danmarks Arealinformation har betinget at de indsamlede data har de karakteristika: geometriske, attributter og interne sammenhænge som formålet helliger.

Den logiske datamodel designes ud fra de indgående datasæts skemaer udvidet med enkelte nye felter:

- database identifikation,
- indførelse af forberedt historik ved hjælp af gyldighedsinterval (gyldig-fra og -til)
- indførelse af unikke identifikatorer (intern database identifikation og UUID)
- temakodesystem (4 klasser: tema- / undertema- / objekt- / detailobjekt- niveauer)
- indførelse af statusfelt, datareference og offentlighedsflag
- reduktion af specielle egenskabsfelter som senere kan sammenkobles ved hjælp af UUID

Den fysiske datamodel afledes af den logiske, således at systemmæssige, performancemæssige og arkitekturmæssige hensyn styrer designet af denne.

Det er valgt at modellere objekter i distributionsmodellen med både kode og klartekst, således at performance kan optimeres. Det er ikke nødvendigvis tilfældet i produktionsbasen.

4.2 **Datamodel som beskriver data**

Følgende tabel anvendes til at udtrykke de oplysninger som skal være generelle for produktionssystem og distributionssystem for arealdata.

Tabellen angiver hvilke felter, deres navne, deres indhold, værdimængde og et eksempel som skal illustrere formålet.

Generel datamodel for arealdataproduktionssystem (1/2)

	Navn	Formål	Datatype	Værdiområde	Eksempel
Database identifikator	db_ident	Entydig database nøgle	9-cifret heltal	100000000-999999999	123456789
Dato for start af gyldighed	gyldig_fra	Start gyldighedsperiode	ISO dato/tid	2006-12-31 23:59:00 – 2999-12-31 23:59:00	2006-12-31 23:59:00
Dato for slut af gyldighed	gyldig_til	Slut gyldighedsperiode	ISO dato/tid	2006-12-31 23:59:00 – 2999-12-31 23:59:00 (NULL = stadig gyldig)	2007-04-30 18:25:00
Administrativ inddeling	adm_kode	Ansvarlig myndighed	1-4-cifret heltal	1-9999	157
Navn svarende til adm. enhed	adm_tekst	Navn på ansvarlig myndighed	tekststreng	3-39 tegn	Gentofte
Reference til objekttype	dateref	Tilknytning til objekttype	tekststreng	3-22 tegn	matrikelskel
Objekttype	objektkode	Kode for objekt	4-cifret heltal	1000-9999	2350
Objekt (klartekst)	objekttkst	Klar tekst for objekttype	tekststreng	3-79 tegn	Kirkebyggelinjer
Symbologi	signatur	Angivelse af aktuel signatur	4-cifret heltal	1000-9999	1234
Status	status	Gældende status for objekt	Tekststreng	ukendt, gældende, under forberedelse/ behandling/drøftelse/, besluttet - men ikke i kraft	under forberedelse
Statuskode	statuskode	Gældende status i kode	1- cifret heltal	0-9	1
Offentliggørelse	offentlig	Kode for synlighed for borger, alle myndighed., én myndighed	1-cifret heltal	1-3	2
Bemærkninger	bemærkning	Evt. bemærkninger i klartekst	tekststreng	0-79 tegn	ingen data

Generel datamodel for arealdataproduktionssystem (2/2)

	Navn	Formål	Datatype	Værdiområde	Eksempel
Temakode1	temakode1	Informationsbærende felt, som er specifikt for et vist tema	langt heltal	0-999999999	
Tematekst1	tematekst1	Informationsbærende felt, som er specifikt for et vist tema	tekststreng	0-79 tegn	
Temakode2	temakode2	Informationsbærende felt, som er specifikt for et vist tema	langt heltal	0-999999999	
Tematekst2	tematekst2	Informationsbærende felt, som er specifikt for et vist tema	tekststreng	0-79 tegn	
Unik identifikator	UUID	Unik nøgle til ekstern reference	128 bit	16 bytes	{E6AB20EA-67E4-4C11-A051-B50A084788A3}
Bruger identifikation	bruger_id	Brugernavn ved opdatering	tekststreng	0-128 tegn	viborg1752
Geometri	shape	Geometrisk beskrivelse af objekt	Shape (2D) Punkt, linje, flade	E: -370000-1777483 N: 5200000-7347483 UTM z32N-ETRS89	For et punkt: 543210,999 6123456,111
Areal	shape.area	Arealstørrelse i m ²	Flydende (Double precision)	0-999999999999,999	101239,243
Længde	shape.len	Længde i m	Flydende (Double precision)	0-999999999999,999	12901,592

Forklaring til datafelter i den generelle datamodel:

Feltnavn	Forklaring
Database identifikator	Entydigt ID for hvert objekt. Historiske objekter har den samme ID. Bruges til at referere objektet via ID.
Gyldig fra dato	Start for objektets gyldighed fra dato og tid på døgnet. Bruges til at bestemme objektets gældende og historiske instanser.
Gyldig til dato	Slut for objektets gyldighed til dato og tid på døgnet. Bruges til at bestemme objektets gældende og evt. historiske instanser.
Administrativkode	Kode for objektets administrative tilhørsforhold. Bruges til af svar-tidshensyn at udtrække objektet udenom de geometriske udtræks-funktioner.
Administrativnavn	Navn på administrationen i klar tekst. Bruges til brugergrænseflade-formål.
Objektet tilknyttet tematype	Objektet er ved sin oprettelse tilknyttet tematypen, fx matrikelskel-temaet. Bruges til at give oplysning om relativ beliggenhed i forhold til andre objekttyper.
Objekttype	Kode for objektets type. (se kodeliste)
Objekt	Klartekst for objektets type til brugergrænsefladeformål
Symbologi	Kode for hvilken type grafisk udtræk objektet skal have
Status	Kode for hvilken status objektet er registreret med: 0: ukendt 1: gældende 2: under forberedelse 3: under behandling 4: under drøftelse 5: besluttet, men ikke i kraft 6-9: <not in use> Bruges til brugergrænsefladeformål.
Statuskode	Se status. Bruges af systemet.
Offentliggørelse	Kode for synlighed for offentligheden, den ansvarlige myndighed eller alle andre myndigheder 1: synlig for alle (offentlig) 2: synlig for den ansvarlige myndighed 3: synlig for alle andre myndigheder, men ikke offentligheden
Bemærkninger	Felt til tekstmæssige oplysninger som kan være nødvendige for at kunne forstå sammenhænge og indhold.
Temakode1	Informations-bærende felt, specifikt for et vist tema. Kan fx bruges som fremmednøgle til andre baser mhp. kobling til disse (Oracle datatype: number)
Tematekst1	Informations-bærende felt, specifikt for et vist tema
Temakode2	Informations-bærende felt, specifikt for et vist tema. Kan fx bruges som fremmednøgle til andre baser mhp. kobling til disse.
Tematekst2	Informations-bærende felt, specifikt for et vist tema
Unik identifikator	Instansobjekt ID. Bruges til at referere til objektets unikke instans i tid og rum. Universally Unique Identifier
Bruger identifikation	Brugernavn fra det fælles offentlige brugerstyringssystem. Regi-streres automatisk ved opdatering via online eller offline editering.
Geometri	Geometri for objektets udstrækning – i punkt, linje eller flade. Bru-ges til at fastlægge den absolutte beliggenhed på jordoverfladen.
Areal	Objektets størrelse for så vidt angår det er et areal. Bruges til at filtrere udtræk, symbologi eller generaliseringsmåde.
Længde	Objektets størrelse for så vidt angår længden eller omkredsen af en polygon. (se Areal)

Nedenstående tabel viser på skitseniveau hvordan en tematiseret klasseinddeling kan tænkes implementeret. Den er opbygget som en hierarkisk struktur, hvor der kan skydes nye over-, og undertemaer ind i en vis udstrækning og slettes igen.

Temakodetabel/klasseinddeling (skitse for beskyttet natur)

Fra objektkode	Til objektkode	Objekttekst	Definition
3000	3999	naturbeskyttelse	Naturbeskyttelse er et overbegreb der omfatter data indenfor naturbeskyttelses-området
3000	3199	Beskyttede naturtyper	
3055	3059	Beskyttede naturtyper : Eng	Eng er et periodevis vådt område som kan være fersk eller kultur som type
3060	3064	Beskyttede naturtyper : Hede	Landareal med udtalt lyngbevoksning med isprængte trævækster på sur bund
3065	3069	Beskyttede naturtyper : Mose	Mose er et blandingsareal af en våd bund, sur plantevækst og evt. trævækst
3070	3074	Beskyttede naturtyper : Overdrev	Overdrev er et normalt tørt blandingsareal med græs isprængt trævækst
3075	3079	Beskyttede naturtyper : Strandeng	Strandeng er et af det salte hav periodevis overskyttet ved kysten lavtliggende (eng)område
3080	3090	Beskyttede naturtyper : Sø	Vandareal på land
3170	3179	<not in use>	
3180	3189	<not in use>	
3190	3199	<not in use>	