

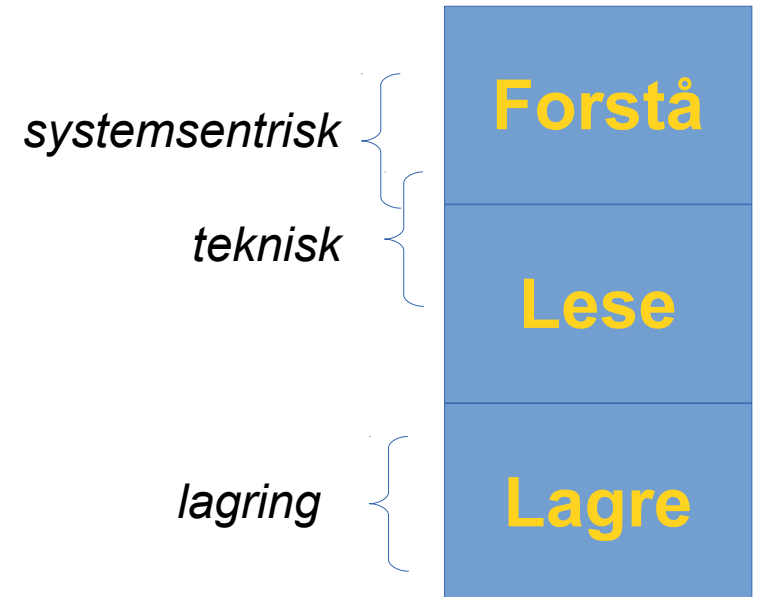
Noark 5 utvidelser og virksomhetspesifikke Metadata: En praktisk forklaring

Thomas Sødring
thomas.sodring@hioa.no

HiOA

Hvordan bevarer vi

- Dagens (elektronisk) bevaringstrategi er basert på en tenking som er enten
 - Teknisk
 - Lagre databasedump (og krysse fingre)
 - Syntaktisk
 - Tabelluttrekk i XML (feks SIARD)
 - Systemsentrisk / Semantisk
 - Noark 4/5 i XML
- Stykkevis med tanke på
 - Tid
 - System / Standard



Semantisk ??

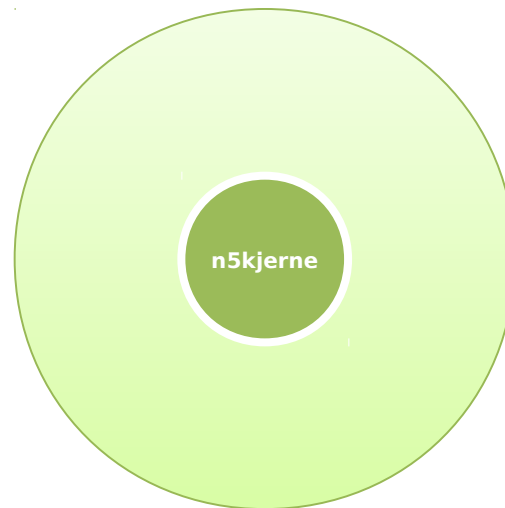
- Ordet Semantisk er i dag ofte brukt for semantisk web teknologi
 - RDF osv
- Med semantisk mener jeg at uttrekket har en uavhengig forståelse av innhold

Hvordan n5 var tenkt

Sak/arkivsystem



Fagsystem

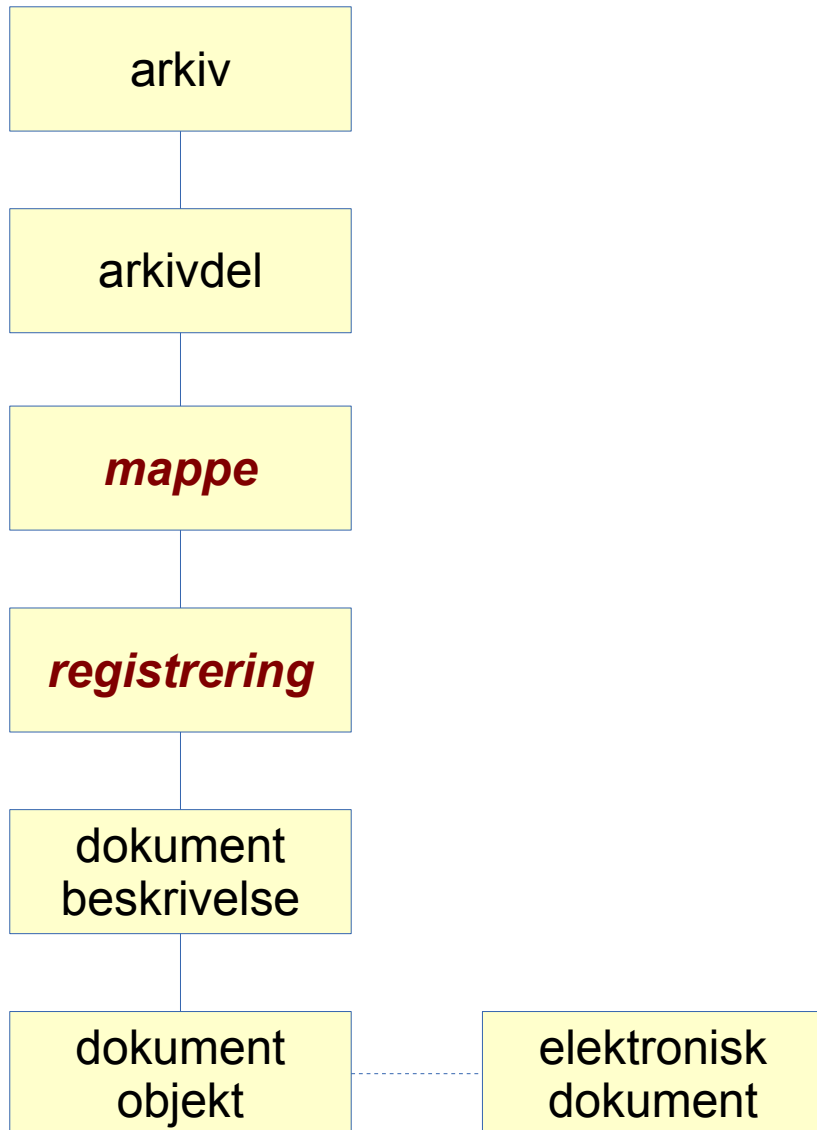


Fellesarkiv



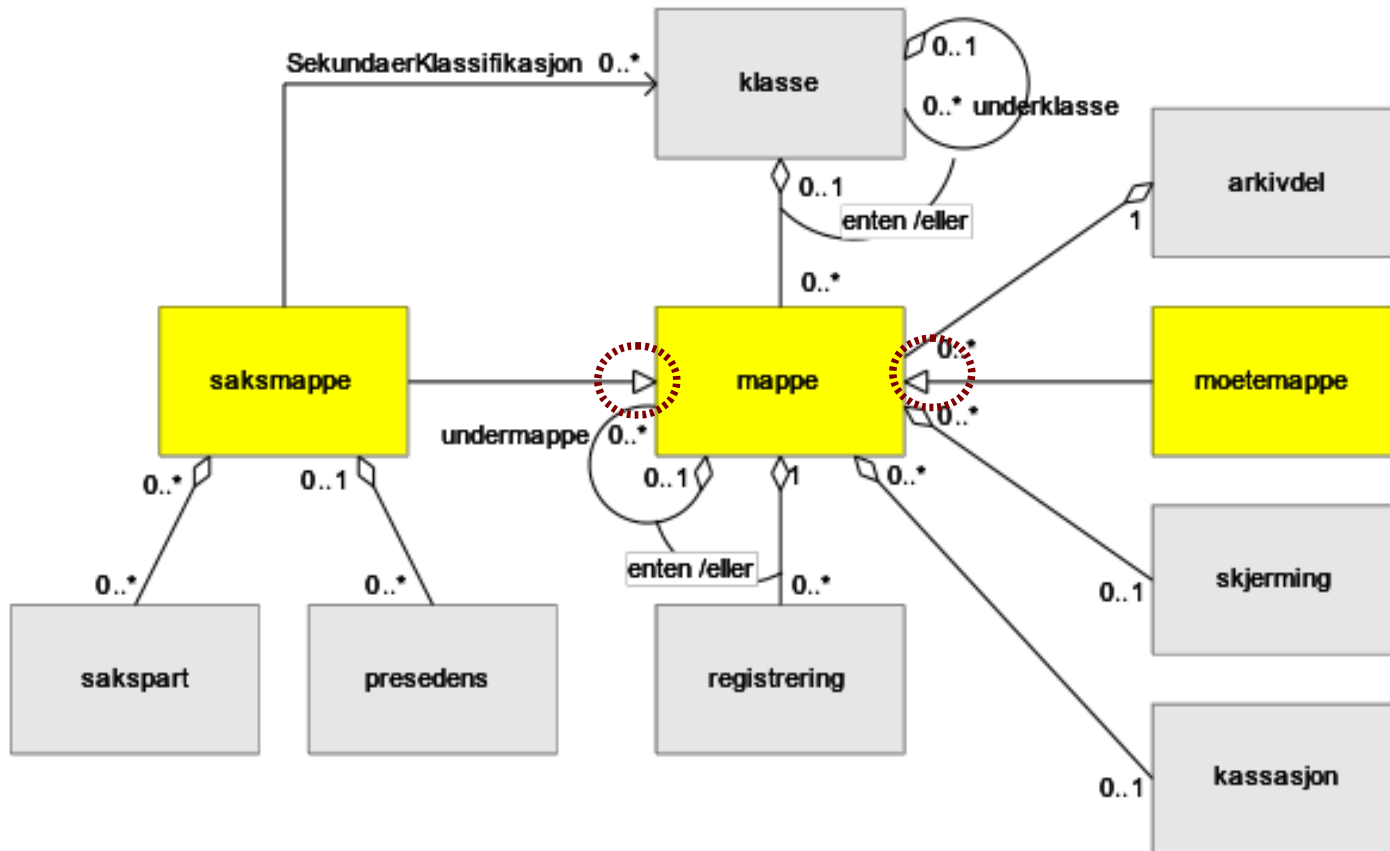
Men det måtte nok være nyutvikling, ikke eksisterende Noark og fagsystem

Plassering i n5 strukturen



```
<arkiv>
  <arkivdel>
    <mappe>
      <registrering>
        <dokumentbeskrivelse>
          <dokumentobjekt>
          </dokumentobjekt>
        </dokumentbeskrivelse>
      </registrering>
    </mappe>
  </arkivdel>
</arkiv>
```

Mappe og arv



Metadata for mappe

Nr.	Navn	Forek	Avl	Datatype
M001	systemID	1	A	Tekststreng
M003	mappeID	1	A	Tekststreng
M020	tittel	1	A	Tekststreng
M025	offentligTittel	0-1	A	Tekststreng
M021	beskrivelse	0-1	A	Tekststreng
M022	noekkelord	0-M	A	Tekststreng
M300	dokumentmedium	0-1	A	Tekststreng
M301	oppbevaringssted	0-M		Tekststreng
M600	opprettetDato	1	A	Dato og klokkeslett
M601	opprettetAv	1	A	Tekststreng
M602	avsluttetDato	1	A	Dato og klokkeslett
M603	avsluttetAv	1	A	Tekststreng
M208	referanseArkivdel	0-M	A	arkivdel.systemID
M711	virksomhetsspesifikke Metadata	0-1	A	Vilkårlig struktur

Metadata for saksmappe

Inkluderer all metadata fra Mappe

Nr.	Navn	Forek	Avl	Datatype
M011	saksaar	0-1	A	Heltall
M012	sakssekvensnummer	0-1	A	Heltall
M100	saksdato	1	A	Dato
M305	administrativEnhet	1	A	Tekststreng
M306	saksansvarlig	1	A	Tekststreng
M308	journalenhet	0-1	A	Tekststreng
M052	saksstatus	1	A	Tekststreng
M106	utlaantDato	0-1		Dato
M309	utlaantTil	0-1		Tekststreng
M209	referanseSekundaer Klassifikasjon	0-M	A	klasse.systemID

mappe - saksmappe

- Dette løses enkelt med to tabeller
 - mappe og saksmappe
 - Det er en én-til-én kobling mellom disse
 - Samme PK brukes på tvers av begge tabeller



- saksmappe må inneholde FK til mappe

mappe og saksmappe

mappe

systemID
mappID
tittel
offentligTittel
beskrivelse
noekkelord
dokumentmedium
oppbevaringssted
opprettetDato
opprettetAv
avsluttetDato
avsluttetAv
referanseArkivdel

saksmappe

saksaar
sakssekvensnummer
saksdato
administrativEnhet
saksansvarlig
journalenhet
saksstatus
utlaantDato
utlaantTil
referanseSekundaerKlassifikasjon
FK_MAPPE_SYSTEM_ID

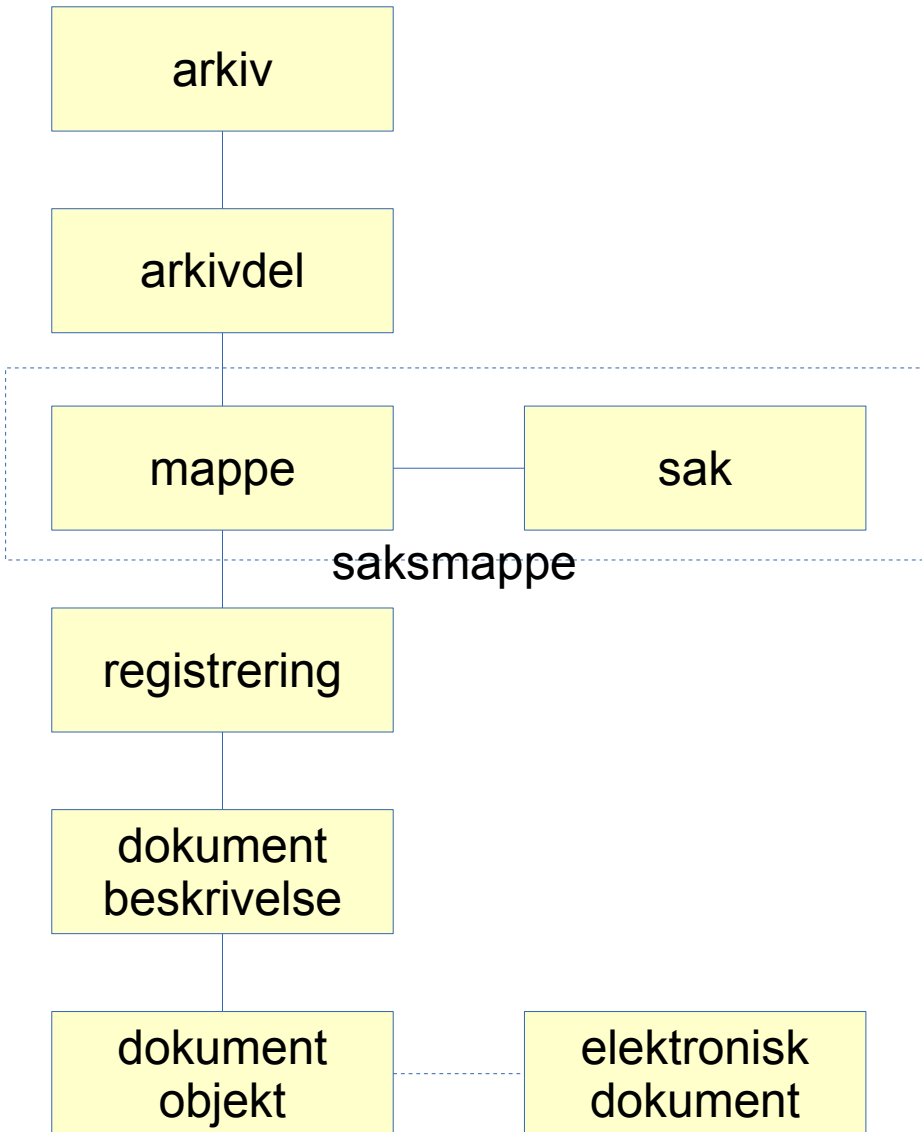


mappe og saksmappe

```
CREATE TABLE mappe (  
  
    systemID varchar(255),  
    mappelD varchar(255),  
    tittel varchar(255),  
    offentligTittel varchar(255),  
    beskrivelse varchar(255),  
    noekkelord varchar(255),  
    dokumentmedium varchar(255),  
    opprettetDato datetime,  
    opprettetAv varchar(255),  
    avsluttetDato datetime,  
    avsluttetAv varchar(255),  
    referanseArkivdel varchar(255),  
    PRIMARY KEY (systemID)  
  
) ENGINE=InnoDB;
```

```
CREATE TABLE saksmappe (  
    saksaar char(4),  
    sakssekvensnummer int,  
    saksdato date,  
    administrativEnhet varchar(255),  
    saksansvarlig varchar(255),  
    journalenhet varchar(255),  
    saksstatus varchar(255),  
    utlaantDato date  
    utlaantTil varchar(255),  
    systemID varchar(255),  
    FOREIGN KEY (systemID)  
        REFERENCES mappe (systemID),  
    PRIMARY KEY (systemID)  
  
) ENGINE=InnoDB;
```

saksmappe



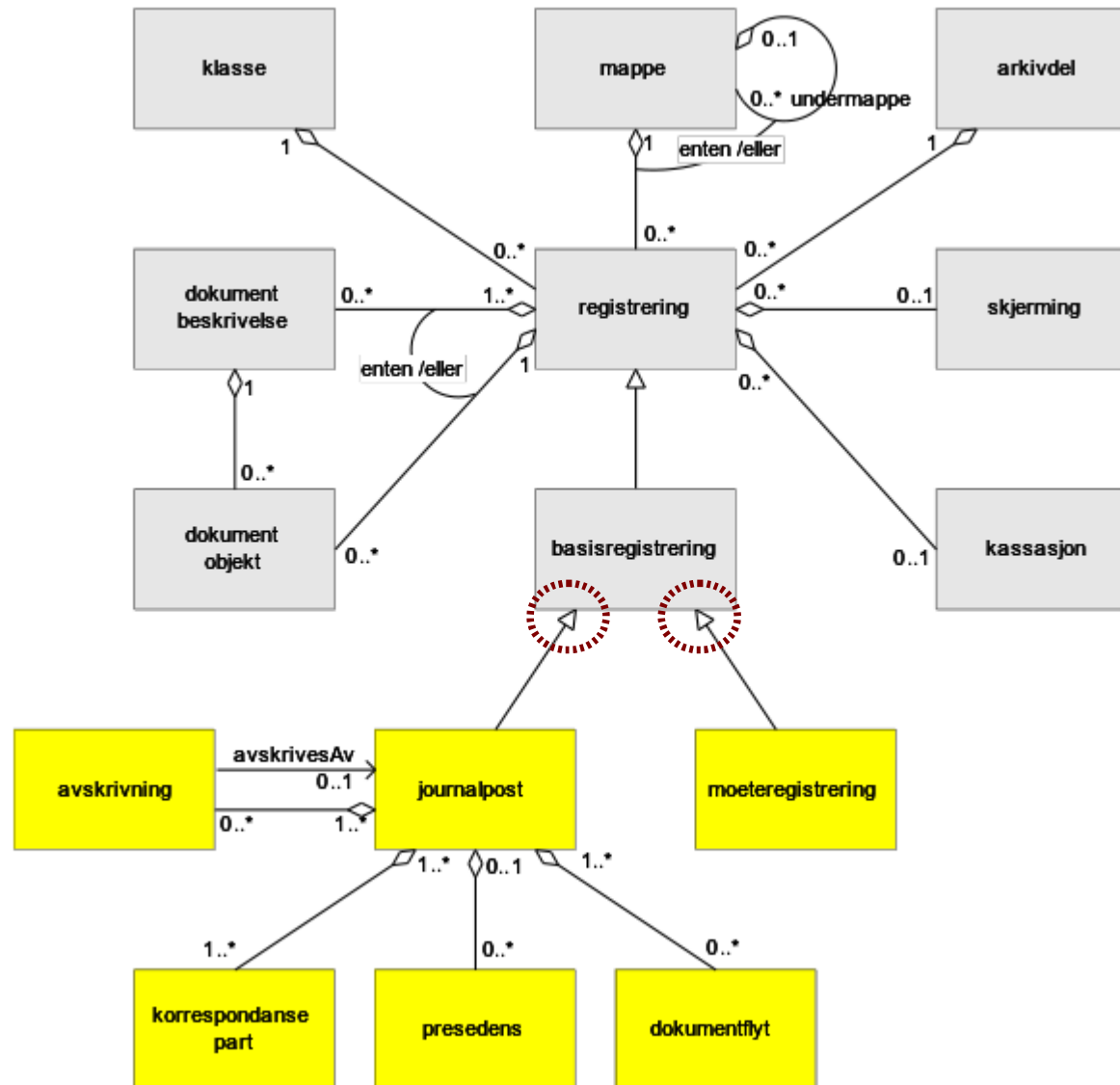
```
<arkiv>
  <arkivdel>
    <mappe xsi:type="saksmappe">
      <!-- mappe elementene -->
      <registrering>
        <dokumentbeskrivelse>
          <dokumentobjekt>
          </dokumentobjekt>
        </dokumentbeskrivelse>
      </registrering>
      <!-- saksmappe elementene -->
    </mappe>
  </arkivdel>
</arkiv>
```

Registrering

- Registrering bruker også arv
 - Registrering -> BasisRegistrering -> Journalpost
 - Registrering -> BasisRegistrering -> MøteRegistrering
- En *Registrering* inneholder all metadata som er nødvendig for å knytte registreringen til arkivstrukturen
 - Kan brukes for arkivering uten journalføring
 - Slike dokumenter skal
 - ikke kunne søkes fram etter innhold
 - ikke inngå i den ordinære identifikasjonen (nummereringen) av journalposter
 - Ikke komme på offentlig journal
 - Men skal inngå i uttrekk hvis de har bevaringsverdi*

* Viktig at det avtales med leverandør når det taes i bruk

registrering og arv



Metadata for registrering

Nr.	Navn	Forek	Avl	Datatype
M001	systemID	1	A	Tekststreng
M600	opprettetDato	1	A	Dato og klokkeslett
M601	opprettetAv	1	A	Tekststreng
M604	arkivertDato	1	A	Dato og klokkeslett
M605	arkivertAv	1	A	Tekststreng
M208	referanseArkivdel	0-M	A	arkivdel.systemID

Metadata for basisregistrering

Nr.	Navn	Forek	Avl	Datatype
M004	registreringsID	1	A	Tekststreng
M020	tittel	1	A	Tekststreng
M025	offentligTittel	0-1	A	Tekststreng
M021	beskrivelse	0-1	A	Tekststreng
M022	noekkelord	0-M	A	Tekststreng
M024	forfatter	0-M	A	Tekststreng
M300	dokumentmedium	0-1	A	Tekststreng
M301	oppbevaringssted	0-M		Tekststreng
M711	virksomhetsspesifikke Metadata	0-1	A	Vilkårlig struktur

registrering – basisregistrering

```
CREATE TABLE registrering (  
  
    systemID varchar(255),  
    mappelD varchar(255),  
    opprettetDato datetime,  
    opprettetAv varchar(255),  
    arkivertDato datetime,  
    arkivertAv varchar(255),  
    referanseArkivdel varchar(255),  
  
    PRIMARY KEY (systemID)  
  
) ENGINE=InnoDB;
```

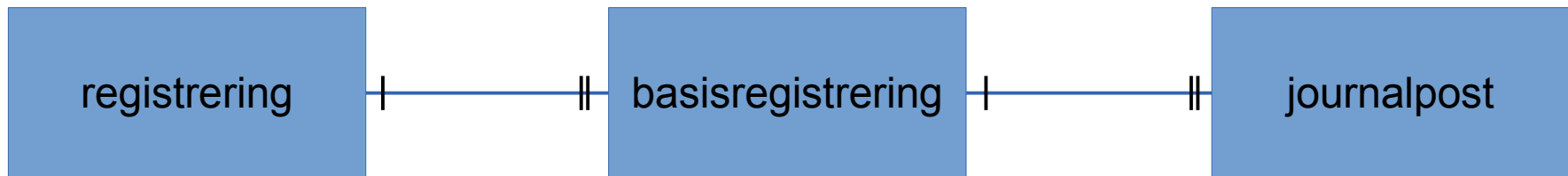
```
CREATE TABLE basisregistrering (  
  
    registreringsID varchar(255),  
    tittel varchar(255),  
    offentligTittel varchar(255),  
    beskrivelse varchar(255),  
    noekkelord varchar(255),  
    forfatter varchar(255),  
    dokumentmedium varchar(255),  
    systemID varchar(255),  
  
    FOREIGN KEY (systemID)  
        REFERENCES registrering (systemID),  
    PRIMARY KEY (systemID)  
  
) ENGINE=InnoDB;
```

Metadata for journalpost

Nr.	Navn	Forek	Avl	Datatype
M013	journalaar	0-1	A	Heltall
M014	journalsekvensnummer	0-1	A	Heltall
M015	journalpostnummer	0-1	A	Heltall
M082	journalposttype	1	A	Tekststreng
M053	journalstatus	1	A	Tekststreng
M101	journaldato	1	A	Dato
M103	dokumentetsDato	0-1	A	Dato
M104	mottattDato	0-1	A	Dato
M105	sendtDato	0-1	A	Dato
M109	forfallsdato	0-1		Dato
M110	offentlighetsvurdertDato	0-1		Dato
M304	antallVedlegg	0-1	A	Heltall
M106	utlaantDato	0-1		Dato
M309	utlaantTil	0-1		Tekststreng
M308	journalenhet	0-1	A	Tekststreng

registrering-basisregistrering-journalpost

- Dette løses enkelt med tre tabeller
 - registrering basisregistrering og journalpost
 - Det er en én-til-én kobling mellom disse
 - Samme PK brukes på tvers av alle tre tabeller



- basisregistrering må inneholde FK til registrering og journalpost må inneholde en FK til basisregistrering

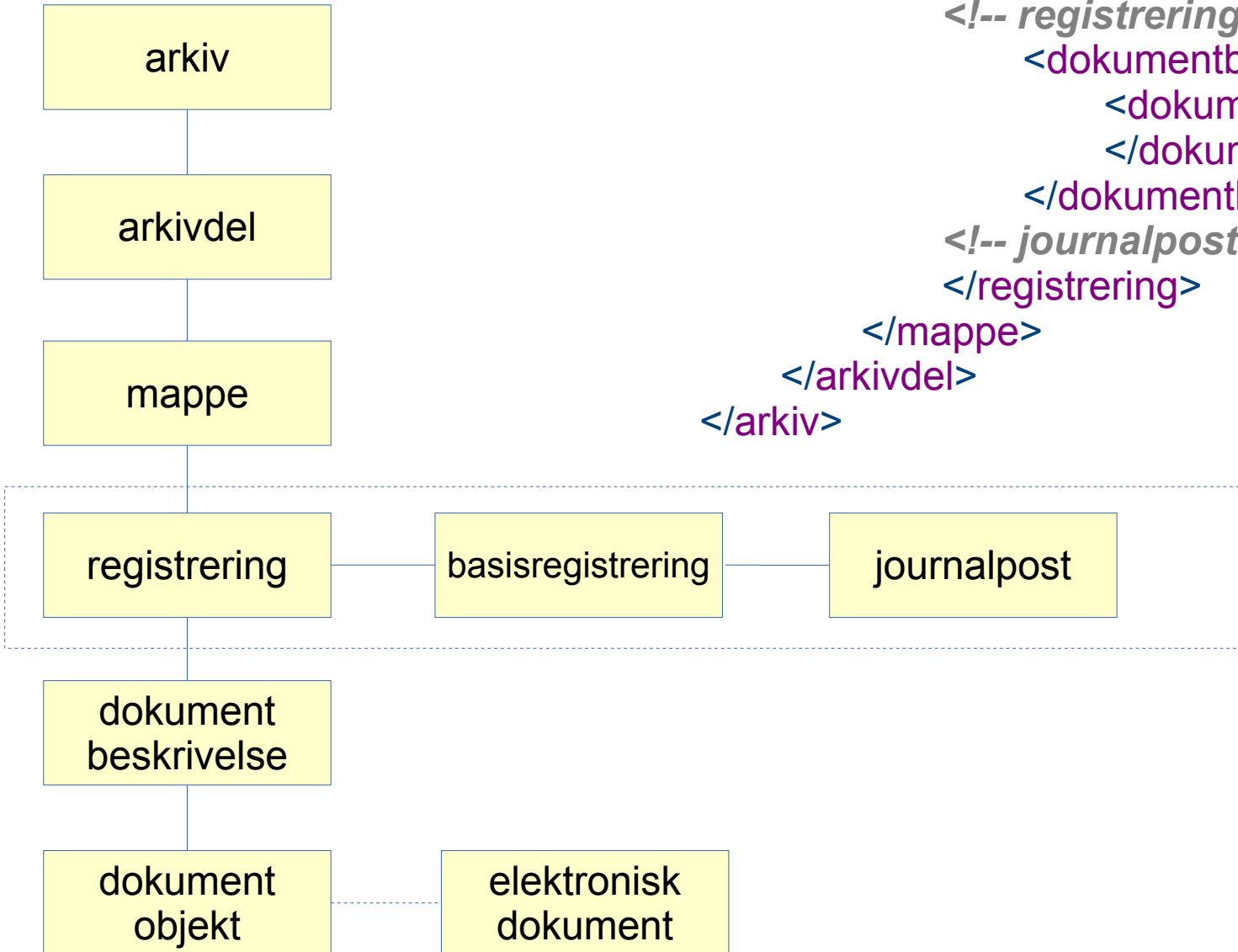
basisregistrering - journalpost

```
CREATE TABLE basisregistrering (  
  
    registreringsID varchar(255),  
    tittel varchar(255),  
    offentligTittel varchar(255),  
    beskrivelse varchar(255),  
    noekkelord varchar(255),  
    forfatter varchar(255),  
    dokumentmedium varchar(255),  
    systemID varchar(255),  
  
    FOREIGN KEY (systemID)  
        REFERENCES registrering (systemID),  
    PRIMARY KEY (systemID)  
  
) ENGINE=InnoDB;
```

```
CREATE TABLE journalpost (  
  
    journalaar char(4),  
    journalsekvensnummer char(4),  
    journalpostnummer int,  
    journalposttype varchar(255),  
    journalstatus varchar(255),  
    journaldato date,  
    dokumentetsDato date,  
    mottattDato date,  
    sendtDato date,  
    forfallsdato date,  
    offentlighetsvurdertDato date,  
    antallVedlegg int,  
    utlaantDato date  
    utlaantTil varchar(255),  
    journalenhet varchar(255),  
    systemID varchar(255),  
  
    FOREIGN KEY (systemID)  
        REFERENCES basisregistrering (systemID),  
    PRIMARY KEY (systemID)  
  
) ENGINE=InnoDB;
```

journalpost

```
<arkiv>  
  <arkivdel>  
    <mappe xsi:type="saksmappe">  
      <registrering xsi:type="journalpost">  
        <!-- registrering elementene -->  
        <dokumentbeskrivelse>  
          <dokumentobjekt>  
        </dokumentobjekt>  
        </dokumentbeskrivelse>  
        <!-- journalpost elementene -->  
      </registrering>  
    </mappe>  
  </arkivdel>  
</arkiv>
```



Det negative

- Dette er ikke en optimal løsningen
 - Når du setter inn innholdet i én saksmappe vil du bruke to INSERT kommandoer
 - Når du setter inn innholdet i én journalpost vil du bruke tre INSERT kommandoer
 - To SELECT kommandoer eller en JOIN for å hente ut en saksmappe
 - Tre SELECT kommandoer eller to JOIN for å hente ut en journalpost

Det positive

- Dette vil gjøre det langt enklere når man skal lage uttrekk fordi all metadata er samlet på et sted
- En database struktur kan lett fange sak/arkiv og flere fagsystem

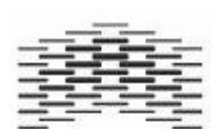


- Det er et nytt system bygget fra bunnen av som løser dette

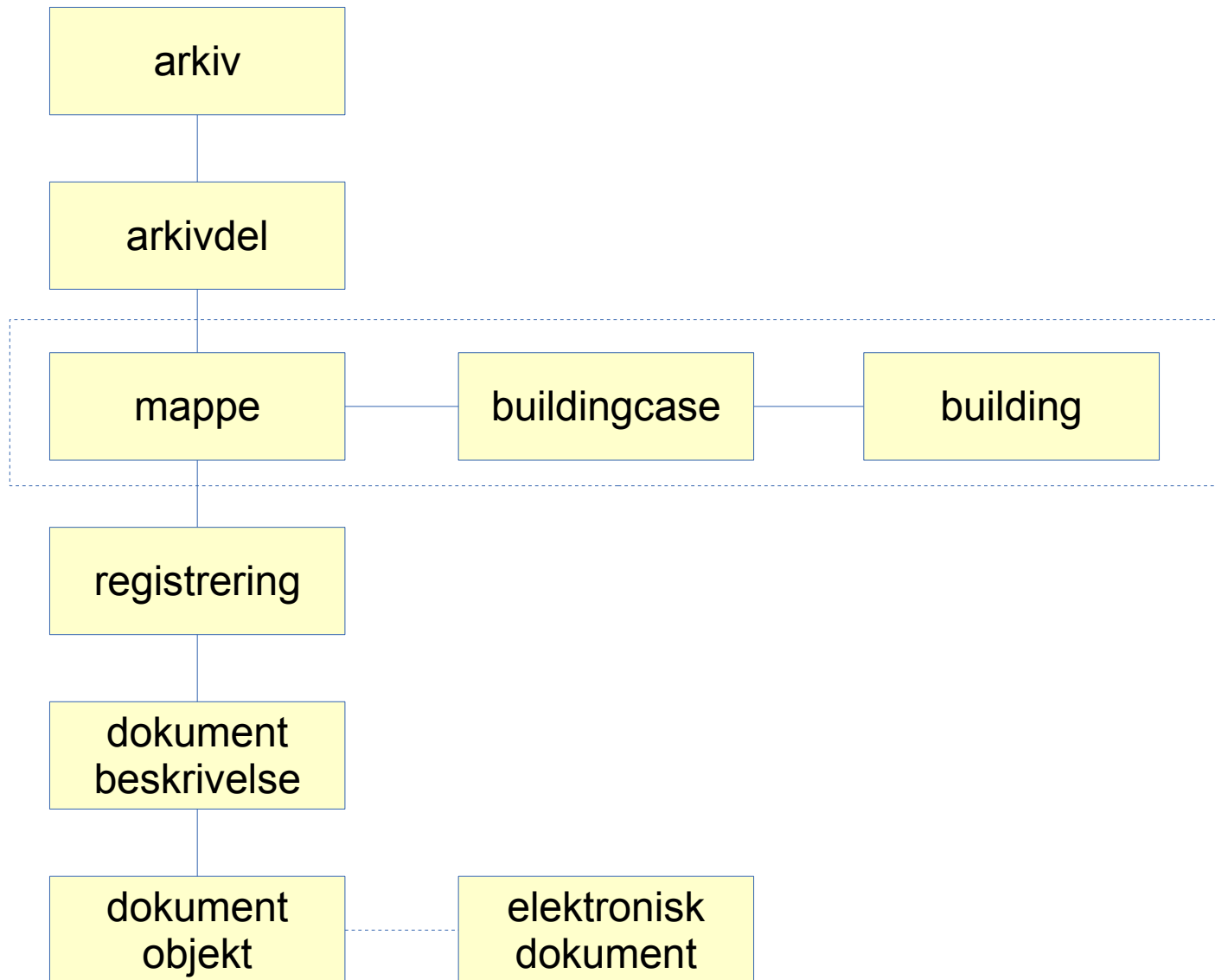
Virksomhetsspesifikke Metadata (VSM)

- Kan plasseres i mappe, basisregistrering og sakspart

```
<?xml version="1.0" encoding="UTF-8"?>
<arkiv xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="http://www.arkivverket.no/standarder/noark5/arkivstruktur"
  xmlns:n5mdk="http://www.arkivverket.no/standarder/noark5/metadatakatalog"
  xmlns:dots="http://www.dots.no/n5/extensions" |
  xsi:schemaLocation="http://www.arkivverket.no/standarder/noark5/arkivstruktur
  arkivstruktur.xsd http://www.dots.no/n5/extensions custom_metadata.xsd">
  <arkivdel>
    <mappe>
      <virksomhetsspesifikkeMetadata>
        <dots:buildingcase>
          <dots:boksProtokollnummer>108</dots:boksProtokollnummer>
          <dots:property>
            <dots:bruksnummer>102</dots:bruksnummer>
            <dots:gaardsnummer>313</dots:gaardsnummer>
            <dots:kommunennummer>1601</dots:kommunennummer>
            <dots:building>
              <dots:byggningsnummer>182481033</dots:byggningsnummer>
            </dots:building>
            <dots:building>
              <dots:byggningsnummer>21066117</dots:byggningsnummer>
            </dots:building>
            <dots:building>
              <dots:byggningsnummer>182481025</dots:byggningsnummer>
              <dots:address>
                <dots:adressekode>1032</dots:adressekode>
                <dots:gatenavn>Anders Haarstads veg</dots:gatenavn>
                <dots:gatenummerMedBokstav>3</dots:gatenummerMedBokstav>
              </dots:address>
            </dots:building>
          </dots:property>
        </dots:buildingcase>
      </virksomhetsspesifikkeMetadata>
    </mappe>
  </arkivdel>
</arkiv>
```



TK-eksemplet



Virksomhetsspesifikke Metadata (VSM)

- Skal mapper defineres som egne utvidelser av mappe eller som VSM?
- sosialmappe, byggesakmappe osv er foreløpig relegert til spesifisering som VSM
 - `<mappe xsi:type="byggesakmappe">` ikke
 - `<virksomhetsspesifikkeMetadata><byggesakmappe>`
- Er kommunal sektor blitt prioritert in Noark standarden?
 - Saksmappe / møtemappe brukes i statlig sektor

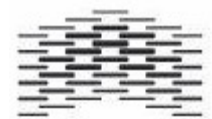
VSM kontra utvidelser

- Når skal det være VSM, når skal det være utvidelse
- De tydelige fagområdene bør ha utvidelser til mappe og registrering
- Litt mer utydelig kan bruke VSM
- Hvor legger vi listen med tanke på bevaring?
 - Lagre, lese, forstå?

Forstå

Lese

Lagre



Standardisering

- SIARD er i utgangspunktet en lagre-tilnærming
- Hvis du ikke har dokumentasjon av struktur ved siden av og har testet alle dokumentet for korrekthet og konvertert til arkivformat
 - Hvordan fungerer RAW INT datatypen fra Oracle i SIARD?
- Hvis vi først skal undersøke og forstå database strukturene da kan vi bevare semantisk
 - Da bør vi bruke en noark 5 struktur
 - Men den bør hete arkivstruktur