

Swepub

Formatspecifikation

Stefan Andersson,
Digital publicering,
Uppsala universitetsbibliotek

Version 1.0
Juni 2008

ERSATT

Inledning

Denna formatspecifikation anger hur metadataposter i lärosätenas lokala publikationsdatabaser skall göras tillgängliga för de nationella tjänster som utvecklas i projektet *Samlad ingång till och redovisning av svensk vetenskaplig publicering*¹ (nu med namnet SwePub).

Specifikationen utgår från de rekommendationer till *Nationellt format för publiceringsdatabaser (lokala register över akademisk publicering)*² som godkändes av SUHF och publicerades i augusti 2005.

Först listas de metadataelement som är obligatoriska. I de obligatoriska elementen kan även finnas underelement som inte behöver tas med men som rekommenderas. Dessa listas för varje element under rubriken *Tillägg*. De icke obligatoriska element som sedan listas behöver inte ingå i metadataposten men rekommenderas om de är tillämpliga.

Några tillägg och förtydliganden av formatet har gjorts i förhållande till de ursprungliga rekommendationerna:

Översikt över ändringar från SVEP

Upphovsman

- Roller anges som standard med koder från *marcrelator*³.
- Valfritt antal hierarkiska nivåer kan anges för ansvarig institution men högsta nivån (lärosäte) samt institution eller motsvarande är fortfarande obligatoriska.

Publikationstyp

- Nya publikationstyper:
 - artikel, forskningsöversikt (tidigare recension)
 - proceedings (redaktörskap)
 - samlingsverk (redaktörskap)

¹ <http://www.kb.se/om/projekt/openaccess/projekt/#Samlad>

² http://epc.ub.uu.se/files/publ1_2.pdf

³ <http://www.loc.gov/marc/sourcecode/relator/relatorsource.html>

- Innehållsmärkning görs till ett eget element som kan kombineras enligt tabellen på sid. 6-7.

Värdpublikation

- ISSN för tidskrifter som värdpublikation görs obligatoriskt.

Identifikatorer för elektroniska dokument

- Alla metadatatposter skall ha en länk till motsvarande post i de lokala databaserna.

Beskrivning av annan medietyp

- Beskrivningsstandarden skall vara enligt MODS⁴.

ISBN

- Görs obligatoriskt (om det finns).

Språk

- Görs obligatoriskt. Standarden skall vara iso639-2b⁵.

Översikt över ändringar efter remissomgång

Under våren 2008 har kravspecifikationen för SwePub, inklusive formatspecifikationen, skickats på remiss till samtliga medlemmar i SUHF samt ytterligare några aktörer. Utifrån remissvaren har följande ändringar och tillägg har gjorts:

Upphovsman

- Adressangivelse kopplad till respektive upphovsman är inte längre obligatoriskt eftersom de lärosäten som använder DSpace⁶ inte kan leverera denna information då det inte är möjligt att länka ihop olika metadatafält i det programmet metadatabeskrivningar.
- Texten i avsnittet har omarbetats.

⁴ <http://www.loc.gov/standards/mods/v3/mods-userguide-elements.html#typeofresource>

⁵ <http://www.loc.gov/marc/languages/>

⁶ <http://www.dspace.org/>

Publikationstyp

- Enbart koder skall anges i metadataposterna. Svenska och engelska benämningar i klartext på olika typer kommer att skapas i webbgränssnittet istället.
- Patent läggs till som publikationstyp.
- En publikation behöver inte ingå i en rapportserie för att definieras som rapport.
- Forskningsöversikt kan nu även vara av typen *referee*granskad.
- Konstnärlig forsknings- och utvecklingsrapport döps om till *Konstnärligt forsknings- och utvecklingsarbete* som istället görs till en typ som kan kombineras med alla övriga, exempelvis rapport eller doktorsavhandling. Exempel på detta ges i texten.
- Innehållstyperna *referee*granskad och *övrig vetenskaplig* grupperas som vetenskapliga med inriktning mot vetenskapssamhället (i relation till *övrigt populärvetenskapligt* med inriktning mot allmänheten).

Värdpublikation

- Uppgifter om volym, häfte och sidor är viktiga om OpenURL⁷-länkar till fulltexter skall kunna skapas på bästa möjliga sätt i systemet. De lärosäten som använder DSpace kan dock inte ange volym, häfte och sidor i separata fält. Dessa kan då istället nu anges i textform i ett gemensamt fält motsvarande det som finns i DSpace.

Identifikatorer för elektroniska dokument

- Uppgift om publikationen finns i elektronisk form i det lokala systemet läggs i elementet *physicalDescription* istället för *genre* för bättre transparens om posterna hämtas av andra tjänster.

Fysisk beskrivning

- Omformulering av texten och exempel på hur fysisk beskrivning kan användas för att beskriva textilier specifikt.

⁷http://www.niso.org/kst/reports/standards?step=2&gid=&project_key=d5320409c5160be4697dc046613f71b9a773cd9e

Metadataformat

Allmänt

- Metadaposter skall hämtas från de lokala publiceringsdatabaserna via OAI-PMH⁸.
- Det skall vara möjligt att hämta endast de poster som är publicerade genom det egna lärosätet, dvs. att upphovsmannen var knuten till lärosätet vid publiceringstillfället
- Metadataformatet skall vara MODS⁹, version 3.2 eller högre.
- Alla element (fält) kan upprepas valfritt antal gånger.
- Element utan innehåll skall inte medtagas i beskrivningen.
- Teckenkodningen skall vara UTF-8.
- Om språkattributet *lang* används skall språket anges enligt iso639-2b¹⁰.

Obligatoriska metadataelement

Upphovsman

Med upphovsman menas person eller institution (hit räknas även exempelvis kongress eller forskningsprojekt) som är ansvarig för det intellektuella eller konstnärliga innehållet i ett dokument.

Upphovsmannen kan vara antingen författare eller redaktör, rollen skall specificeras i metadata. Rollen skall anges med koder från *marcrelator*¹¹. Obligatoriska är *aut* (författare) resp. *edt* (redaktör/utgivare).

För personnamn anges både för- och efternamn. En rekommendation är att dessa anges i separerade fält. Om detta inte är möjligt måste de anges enligt modellen Efternamn, Förnamn i ett *namePart*-element utan attribut.

Författare anges i den ordning de står i publikationen.

⁸ <http://www.openarchives.org/OAI/openarchivesprotocol.html>

⁹ <http://www.loc.gov/standards/mods/>

¹⁰ <http://www.loc.gov/marc/languages/>

¹¹ <http://www.loc.gov/marc/sourcecode/relator/relatorsource.html>

Adressangivelser

Adress (dvs. lärosäte plus institution eller motsvarande) bör, om det är möjligt i det lokala systemet, anges för åtminstone lärosätets egna författare (och helst för samtliga författare). Detta gör det möjligt att identifiera vilka författare som hör till vilket lärosäte. Sökning på lärosäte/institution kommer dock i systemet att göras mot den strukturerade informationen i elementet för *Ansvarig organisation*.

Adressen skall anges enligt modellen *Lärosäte, Institution el. motsv.*

Uppgiften skall gälla den adress som är kopplad till själva publikationen (ej den adress som upphovsmannen för närvarande är kopplad till – om den inte är densamma som i publikationen).

Tillägg

- Relevanta årtal (födelsedatum) kan läggas till personlig upphovsman.
- Upphovsmansinformationen kan knytas till ett lokalt auktoritetsregister. I *authority*-attributet anges då lärosätets domän eller annan unik kod och i *xlink:href*-attributet anges en URI eller annat unikt ID till den lokala posten.

Exempel

```
<name type="personal" authority="uu" xlink:href="N2-1690">  
  <namePart type="given">Leif</namePart>  
  <namePart type="family">Svensson </namePart>  
  <namePart type="date">1963-</namePart>  
  <role>  
    <roleTerm type="code" authority="marcrelator">aut</roleTerm>  
  </role>  
  <affiliation>Uppsala universitet, Litteraturvetenskapliga institutionen</affiliation>  
</name>
```

Ansvarig organisation

Den organisatoriska enhet som upphovsmannen tillhör vid det egna lärosätet, och som är knuten till publiceringen, skall anges i ett separat obligatoriskt element. De olika hierarkiska nivåerna skall läggas i olika underelement från den högsta till den lägsta. Lärosätet är alltid den högsta nivån och måste alltid anges. Nivåer motsvarande fakultet eller institution skall också anges. Rollen för ansvarig organisation skall anges med koden *pbl* från *marcrelator*. Lärosätets domän eller annan unik kod anges i *authority*-attributet.

Detta element kommer att ligga till grund för sökning på publikationer från ett visst lärosäte eller institution. Om en publikation har upphovsman från flera olika institutioner eller enheter vid det egna lärosätet skall alla dessa enheter anges i separata element med lärosätet som högsta nivå.

Ansvarig organisation skall inte blandas ihop med utgivare/förlag som, för icke-seriella publikationer, anges i elementet *originInfo*. I de fall ansvarig

organisation och utgivare/förlag är samma enhet kan uppgiften anges i båda elementen om så önskas.

Organisatoriska förändringar hanteras i de lokala systemen.

Tillägg

- Om organisationsnamnen även finns i engelska versioner kan ett separat element läggas till för dessa. *Lang*-attributet är då obligatoriskt.
- Organisationsinformationen kan knytas till ett lokalt auktoritetsregister via *xlink:href*-attributet.

Exempel

```
<name type="corporate" lang="swe" authority="uu" xlink:href="uu-XX">
  <namePart>Uppsala universitet</namePart>
  <namePart>Humanistisk-samhällsvetenskapliga vetenskapsområdet</namePart>
  <namePart>Historisk-filosofiska fakulteten</namePart>
  <namePart>Litteraturvetenskapliga institutionen</namePart>
  <role>
    <roleTerm type="code" authority="marcrelator">pbl</roleTerm>
  </role>
</name>
```

Titel

Ord eller fras som namnger ett verk. Huvudtitel och undertitel läggs i separata underelement. Om detta inte är möjligt skall titeln anges enligt modellen Huvudtitel : undertitel i *title*-elementet.

Alternativa titlar anges i separat element med värdet *alternative* i *type*-attributet.

```
<titleInfo>
  <title>Twenty years to nowhere</title>
  <subTitle>does land use affect regeneration perspectives?</subTitle>
</titleInfo>

<titleInfo lang="eng" type="alternative ">
  <title>Variations in Russian Verbal Government 1800–1840</title>
</titleInfo>
```

Publikationstyp

Publikationen skall klassificeras enligt följande typologi:

<i>Typ</i>	<i>Kod</i>	<i>Ref. vet.</i>	<i>Övr. vet.</i>	<i>Övr. pop.</i>
artikel	art	X	X	X
artikel, forskningsöversikt	for	X	X	-
artikel, recension	rec	-	X	X
doktorsavhandling	dok	-	X	-
licentiatavhandling	lic	-	X	-
bok	bok	-	X	X
kapitel i bok eller rapport	kap	-	X	X
konferensbidrag	kon	X	X	X
patent	pat	-	-	-
proceedings (redaktörskap)	pro	-	X	X
konstnärligt forsknings- och utvecklingsarbete ¹²	kfa	-	-	-
rapport	rap	-	X	X
samlingsverk (redaktörskap)	sam	-	X	X
övrigt	ovr	-	X	X

¹² Konstnärligt forsknings- och utvecklingsarbete skall kombineras med någon av de andra publikationstyperna i beskrivningen, se exempel nedan.

Ovanstående tabell anger vilka möjliga värden som finns för innehållsmärkning (X = möjlig). Innehållstyp måste kombineras med publikationstyp och endast en av vardera skall ges till varje publikation.

Innehållsmärkning

Typ	Innehåll	Kod
refereegranskat	Riktat sig mot vetenskapssamhället	ref
övrigt vetenskapligt	Riktat sig mot vetenskapssamhället	vet
övrigt (populärvetenskap, debatt)	Riktat sig mot allmänheten	pop

Både publikationstyp och innehållsmärkning måste alltid anges enligt koder i ovanstående tabell och följande modell:

Exempel

```
<genre authority="svep" type="publicationType">art</genre>  
<genre authority="svep" type="contentType">ref</genre>
```

Definitioner av publikationstyper

Artikel

Avser en artikel i vetenskaplig eller annan tidskrift.

Recensioner avser recension av bok (book review) i tidskrift eller dagstidning

Forskningsöversikt avser review/survey i vetenskapliga tidskrifter.

Bok

Monografisk publikation bestående av en eller flera fysiska delar, avsett att utgöra en avslutad helhet. Omfattar även e-bok.

Kapitel i bok eller rapport

Självständig del av monografisk publikation. Ingår ofta i samlingsverk, ett antal fristående verk som utgivits tillsammans eventuellt under redaktionell styrning till ett visst tema.

Doktorsavhandling

Avser godkänd avhandling för doktorsexamen.

Licentiatavhandling

Avser godkänd avhandling för licentiatexamen.

Konferensbidrag

Refereegranskat konferensbidrag avser sådant bidrag som på förhand granskats av en eller flera referenter och som utges i officiell proceeding eller konferensmeddelande.

Övriga konferensbidrag avser bidrag som publiceras i konferensskrift men som ej genomgått förhandsgranskning.

Proceedings (redaktörskap)

Avser eget redaktörskap för officiell konferensproceeding eller konferensmeddelande. Egna enskilda bidrag redovisas under *Konferensbidrag*.

Rapport

Avser publikation t.ex. ingående i rapportserie.

Samlingsverk (redaktörskap)

Avser publikation där ett antal var för sig fristående verk utgivits tillsammans. Endast redaktören skall redovisas under denna kategori. De enskilda verken skall var för sig redovisas som *Kapitel* eller del av bok.

Konstnärligt forsknings- och utvecklingsarbete

Arbete från ett konstnärligt forsknings- och/eller utvecklingsprojekt. Denna typ skall kombineras med någon av de övriga publikationstyperna. Ett sådant arbete kan ta sig uttryck även i andra medietyper än text. Detta anges då under rubriken *Beskrivning av annan medietyp*.

Exempel

```
<genre authority="svep" type="publicationType">kfa</genre>  
<genre authority="svep" type="publicationType">rap</genre>  
<genre authority="svep" type="contentType">vet</genre>  
<typeOfResource>sound recording-musical</typeOfResource>
```

Ett konstnärligt arbete publicerat som musik i rapportform.

Övrigt

Avser publikationer där ingen annan av ovanstående publikationstyper är tillämplig, t.ex. poster (presentationer), evenemang, konserter, utställningar, teaterföreställningar eller textilier.

Definitioner av innehållsmärkning

Refereegranskat

Avser att publikationen genomgått granskning av oberoende referenter vars samlade bedömning avgör om publikationen uppfyller tillräckliga krav på vetenskaplig kvalitet för att accepteras för publicering. Gäller oftast artiklar men även konferensbidrag kan vara refereegranskade.

Övrigt vetenskapligt

Avser att publikationens innehåll är av vetenskaplig karaktär och riktar sig mot vetenskapssamhället men ej genomgått refereegranskning.

Övrigt (populärvetenskap, debatt)

Avser att publikationens innehåll är av populärvetenskaplig karaktär och riktar sig mot allmänheten. Även debattinlägg räknas hit.

Utgivningsuppgift

Utgivningsår anges alltid. Med utgivningsår menas det år ett arbete publiceras och att man härvidlag följer förlagens utgivningstid. T.ex. skall en artikels upphovsår anges till det år tidskriften publiceras och inte det år artikeln accepteras.

Exempel

```
<originInfo>
  <dateIssued>1994</dateIssued>
</originInfo>
```

Tillägg

För monografier bör även utgivningsort och förlag anges.

Exempel

```
<originInfo>
  <place>
    <placeTerm>Uppsala</placeTerm>
  </place>
  <publisher>Institutionen för cell- och molekylärbiologi</publisher>
  <dateIssued>1994</dateIssued>
</originInfo>
```

Ämnesområde

Angivande av ämnesområde enligt rekommendationer för *Ämneskategorier för vetenskapliga publikationer*¹³ från DP 1 inom SVEP. Dessa bygger på

¹³ Se: <http://www.ub.uu.se/epub/categories/>

Statistiska centralbyråns, SCB:s *Nationell förteckning över forskningsämnen*¹⁴ samt Vetenskapsrådet utvidgade klassificering¹⁵. SVEP-kategorierna skall anges enligt nedanstående modell (svensk och engelsk form).

Exempel

```
<subject lang="swe" authority="svep" xlink:href="112104">
  <topic>HUMANIORA och RELIGIONSVETENSKAP</topic>
  <topic>Estetiska ämnen</topic>
  <topic>Litteraturvetenskap</topic>
</subject>
<subject lang="eng" authority="svep" xlink:href="112104">
  <topic>HUMANITIES and RELIGION</topic>
  <topic>Aesthetic subjects</topic>
  <topic>Literature</topic>
</subject>
```

Om det inte är möjligt att medtaga hela hierarkin kan kategorierna anges på en nivå enligt modellen:

Exempel

```
<subject lang="swe" authority="svep" xlink:href="112104">
  <topic>Litteraturvetenskap</topic>
</subject>
<subject lang="eng" authority="svep" xlink:href="112104">
  <topic>Literature</topic>
</subject>
```

Tillägg

Fria ämnesord kan anges med eller utan *lang*-attribut. Detta element kan användas för att lägga till ämnesord som saknas bland SVEP-kategorierna.

Exempel

```
<subject lang="eng">
  <topic>Hermeneutics</topic>
</subject>
```

Värdpublikation

Med värdpublikation menas en seriell eller monografisk publikation varur endast en mindre, självständig del utgör underlag för en referens, t.ex. en tidskriftsartikel eller ett kapitel i en bok. Värdpublikationens titel och siduppgifter anges samt, i förekommande fall vid artiklar, volym, häfte och tidskriftens ISSN. Flera ISSN (print/electronic) kan anges om så önskas.

Rekommendation: Utskriven titel föredras framför förkortningar vid angivande av tidskriftstitel.

¹⁴ Se: http://www.scb.se/templates/Standard_24458.asp

¹⁵ Se: <http://www.snac.vr.se/classifications.htm>

Exempel

```
<relatedItem type="host">
  <titleInfo>
    <title>Journal of Applied Ecology</title>
  </titleInfo>
  <identifier type="issn">1057-9478</identifier>
  <part>
    <detail type="volume">
      <number>31</number>
    </detail>
    <detail type="issue">
      <number>3</number>
    </detail>
    <extent>
      <start>11</start>
      <end>23</end>
    </extent>
  </part>
</relatedItem>
```

De lärosäten som använder DSpace kan inte ange uppgifter om volym, häfte och sidor i separata fält. Dessa kan då istället anges i textform i ett gemensamt fält enligt nedanstående exempel:

Exempel

```
<part>
  <detail type="citation">
    <caption>Vol. 117, no. 6, p. 30-38</caption>
  </detail>
</part>
```

Identifikatorer för elektroniska dokument

Alla metadataposter skall ha en länk till motsvarande post i de lokala databaserna. Om lärosätet själv publicerar dokumentet elektroniskt skall adressen (URL) för publiceringen anges.

Rekommendation: Om möjligt anges även ett s.k. DOI-nummer (Digital Object Identifier)¹⁶ vilken gör att publikationen kan spåras även om den byter publiceringsadress.

Även post-ID:n från andra externa databaser kan anges: PubMed, ISI eller Scopus. Enbart numret anges, ej inledande prefix som doi: eller ISI:.

Tillåtna värden i *type*-attributet är uri, doi, hdl, pmid, isi eller scopus.

Exempel

```
<identifier type="uri">http://opus.uu.se/publication.xml?id=103993</identifier>
```

En URL till den lokala posten. Obligatorisk.

```
<identifier type="doi">10.1000/182</identifier>
```

¹⁶ Se: <http://www.doi.org/>

Ett DOI-nummer.

```
<identifier type="hdl">1721.1/5759</identifier>
```

Ett handle-nummer.

```
<identifier type="pmid">121668435</identifier>
```

Ett PubMed-nummer.

```
<identifier type="isi">000249875900007</identifier>
```

Ett ISI-nummer.

```
<identifier type="scopus">12345678900</identifier>
```

Ett Scopus-nummer.

URL:er till fulltexter eller andra tillhörande filer kan anges i elementet *location/url*. En förklarande text kan läggas i *displayLabel*-attributet.

```
<location>  
  <url displayLabel="Fulltext, fritt tillgänglig">http://...</url>  
</location>
```

Om fulltexten dessutom är publicerad i ett lokalt öppet arkiv skall följande element (för utsökning av sådana) tillfogas posten:

```
<physicalDescription>  
  <form authority="marcform">electronic</form>  
</physicalDescription>
```

ISBN

ISBN (internationellt standardnummer för böcker). Anges alltid om det finns tilldelat publikationen.

Exempel

```
<identifier type="isbn">91-554-6028-3</identifier>
```

Språk

Uppgift om dokumentets språk. Anges alltid enligt iso639-2b.

Exempel

```
<language>  
  <languageTerm type="code" authority="iso639-2b">eng</languageTerm>  
</language>
```

Ej obligatoriska metadataelement

Beskrivning av annan medietyp

Om andra medietyper än text tas med i det lokala systemet kan dessa anges enligt MODS¹⁷.

Exempel

```
<typeOfResource>still image</typeOfResource>
```

Abstract

Ett referat av dokumentets innehåll. Språkattribut kan läggas till vid behov.

Exempel

```
<abstract lang="eng"> ...Text...</abstract>
```

Anmärkningar

Eventuella anmärkningar rörande dokumentet. Språkattribut kan läggas till vid behov.

Exempel

```
<note lang="eng">...Text...</note>
```

Fysisk beskrivning

Antal sidor eller annan fysisk beskrivning av dokumentet.

Exempel

```
<physicalDescription>  
  <extent>30 s.</extent>  
</physicalDescription>
```

Exempel

```
<physicalDescription>  
  <extent>1 bordsduk : blå- och vitrutig ; 90 x 90 cm</extent>  
</physicalDescription>
```

Serie

Uppgift om dokumentet ingår i en serie. Serietitel, ISSN och del anges alltid i förekommande fall.

¹⁷ <http://www.loc.gov/standards/mods/v3/mods-userguide-elements.html#typeofresource>

Exempel

```
<relatedItem type="series">
  <titleInfo>
    <title>Comprehensive Summaries of Uppsala Dissertations from the Faculty of
      Science and Technology</title>
  </titleInfo>
  <identifier type="issn">1104-232X</identifier>
  <identifier type="issue number">1008</identifier>
</relatedItem>
```

Relaterade publikationer

Uppgift om relaterade dokument, exempelvis olika versioner/format eller recensioner. Anges enligt MODS i *relatedItem*-elementet. Alla ovanstående element kan ingå. Följande varianter finns:

Exempel

```
<relatedItem type="otherVersion">
  <!--Innehåll, annan upplaga med reviderat/ändrat innehåll-->
</relatedItem>

<relatedItem type="otherFormat">
  <!--Innehåll, samma innehåll, annat format (tryckt, elektroniskt)-->
</relatedItem>

<relatedItem type="isReferencedBy">
  <!--Innehåll, refereras av -->
</relatedItem>

<relatedItem type="constituent">
  <!--Innehåll, del av-->
</relatedItem>

<relatedItem>
  <!--Annan typ av relation-->
</relatedItem>
```