

Dnr 14-757/09
2010-04-07

Regeringskansliet
Kulturdepartementet
103 33 STOCKHOLM

Redovisning av uppdraget om underlag till nationell strategi för digitalisering, elektronisk tillgång och digitalt bevarande Ku2009/2152/KT

Institutet för språk och folkminnen insänder härmed redovisning av uppdraget om underlag till nationell strategi för digitalisering, elektronisk tillgång och digitalt bevarande.

Beslut i detta ärende har fattats av generaldirektör Ingrid Johansson Lind efter föredragning av avdelningschef Ola Wennstedt.



Ingrid Johansson Lind
Generaldirektör

Kopia:
ABM-centrum

Sändlista inom myndigheten:
Arkiv-, avdelnings- och enhetschefer
SACO-föreningen, ST-Kultur

Redovisning av uppdraget om underlag till en nationell strategi för digitalisering, elektronisk tillgång och digitalt bevarande, Ku2009/2152/KT

Institutet för språk och folkminnen har till uppgift att bedriva språkvård och på vetenskaplig grund öka, levandegöra och sprida kunskaper om språk, dialekter, folkminnen, namn och immateriella kulturarv i Sverige. Institutet har närmare hundra anställda och finns på fem orter i Sverige: Dialekt- och ortnamnsarkivet i Lund (DAL), Dialekt-, ortnamns- och folkminnesarkivet i Göteborg (DAG), Dialekt-, ortnamns- och folkminnesarkivet i Umeå (DAUM) samt Språkrådet i Stockholm. Huvudkontoret ligger i Uppsala med ledning och stödfunktioner (Avdelningen för lednings- och verksamhetsstöd (VLS) samt Avdelningen för digitalt arkiv och lokal service (DLS)). Där finns även arkiv, bibliotek samt tre forskningsavdelningar, Namnavdelningen (NA), Dialektavdelningen (DA) och Folkminnesavdelningen (FA). I samlingarna finns stora mängder arkivmaterial, främst omfattande dialekter, ortnamn och folkminnen.

I texten nedan besvaras de frågor som presenteras i uppdraget, Ku2009/2152/KT, utifrån Institutet för språk och folkminnens perspektiv. Referenser till Institutet för språk och folkminnen sker i texten oftast under beteckningen *institutet*.

Del I – Beskriva institutets digitala samlingar och deras användning samt redogöra för:

1. Hur stor andel är digitaliserad respektive ursprungligen i digitalt format. Specificera vilka delar av samlingarna eller vilka materialkategorier som är digitaliserade och i vilken utsträckning.

De samlingar som finns hos institutet har tillkommit i en rad olika sammanhang – från slutet av 1800-talet och fram till i dag. Institutets föregångare har bestått av ideella föreningar, av institut som varit kopplade till universitet och högskolor och av självständiga statliga myndigheter som ombildats under olika beteckningar. Arkivmaterialet är insamlat av såväl privatpersoner som statliga tjänstemän. Bland materialet finns även hela donerade samlingar. I myndighetens arkiv finns således såväl allmänna som enskilda handlingar

Vid Institutet för språk och folkminnen finns arkivmaterial av två skilda slag. Institutet är en statlig myndighet och ska därför samla allt diariefört material i *ämbetsarkivet*. Därutöver har insamling och dokumentation av språk och folkminnen pågått i över hundra år och detta material har kontinuerligt arkiverats i *samlingarna*. Beträffande digitalisering av institutets arkivmaterial har *samlingarna* prioriterats.

Dokumentation och bevarande av olika former av immateriella kulturarv har pågått sedan början av 1900-talet. På så sätt har en stor mängd material samlats vid institutets avdelningar. Genom åren har olika tekniker kommit till användning vid insamling och för bevarande – från anteckningar på papper, skrivmaskin och fotografier, via inspelningar på vaxrulle, tråd, magnetband, video, till i modern tid digital dokumentation. Vid digitalisering grundas valet av teknik på digitaliseringsunderlagets form. Samlingarna kan därför delas in i två kategorier utifrån databärare: *inspelningar* och *text/bild*.

För att få en överblick över samlingarna genomfördes under andra halvan av 2008 en inventering av materialet vid avdelningarna, som redovisas i tabellerna nedan.

För *inspelningar* gäller att uppgifterna för kassetter, rullband och lackskivor i tabellen nedan representerar referensvärden som har räknats fram genom en uppskattad procentuell fördelning baserad på resultaten från inventeringen.

Inspelningar

Databärare	Totalt antal inspelningar	Totalt antal timmar inspelade	Totalt antal timmar överförda till digitalt medium	Procentuell andel digitaliserad
Kassettband	9 000	8 000	1200	15 %
Rullband	20 000	12 300	1500	12 %
Lackskivor	15 000	4 600	3 600	78 %
Cylindrar (vaxrullar)	80	4	4	100 %
Video	150	150	0	0 % 0
S:a	c:a 44 200	c:a 25 000	c:a 6 300	26 %

Materialet i kategorin *text/bild* har vid arkiven traditionellt registrerats i antal ”blad” vilket därför blivit det kvantitativa mått som använts vid inventeringen. I digitaliseringssammanhang vore antal ”sidor” ett mer relevant mått eftersom om ett ”blad” har text på två sidor krävs dubbelt så stora resurser för skanning av bladets hela innehåll.

Beträffande antalet accessioner kan konstateras att principerna för registrering av *text/bild*-samlingarna skiljer sig mellan de olika arkiven och avdelningarna. Ett värde som beskriver det totala antalet registrerade poster har därför inte kunnat räknas fram. Antalet accessioner är dock inte relevant vid beskrivning av samlingarnas omfattning eftersom en accession kan bestå av allt från 1 sida upp till flera tusen sidor. Genom åren har stora mängder material kommit att kopieras och skickas mellan arkiven. I tabellen nedan har strävan varit att redovisa originalmaterial. Kopior är därför inte medräknade i de angivna värdena, vilket förklarar de något högre värdena som återges i institutets årsredovisningar, där det totala antalet blad redovisas.

Text/bild

Databärare	Totalt antal accessioner	Totalt antal blad	Totalt antal sidor överförda till digitalt medium (antalet färdiga filer)	Procentuell andel digitaliserad
A3-5, folio	ospec.	2 963 000	56 000	2 %
sedesblad	ospec.	17 017 000	4 418 000	26 %
foton, positiv/negativ, färg/svartvitt	ospec.	18 700	5 800	31 %
tidningsurklipp	ospec.	29 500	0	0 %
S:a	88 000	20 028 200	4 479 800	22 %

2. I vilken utsträckning register och kataloger har digitaliserats samt i vilken utsträckning de är kompatibla med register och kataloger hos andra institutioner.

Verksamheten vid Institutet för språk och folkminnen sträcker sig långt tillbaka i tiden, ända till början av 1900-talet. Avdelningarna på de fem orter där institutet finns representerat har samtliga byggts upp och organiserats som självständiga, fristående arkiv och avdelningar, och verksamheten har därför kommit att formeras delvis enligt skilda principer. Ingen gemensam uppsättning metadata har stått som grund för registrering, och register och kataloger har därför delvis kommit att fokusera på skilda särdrag i samlingarna. Under andra halvan av 00-talet har dock ett initiativ tagits för att slå samman avdelningars och arkivs skilda register. En gemensam uppsättning metadata har arbetats fram och arbetet pågår nu med att överföra registerdata till det nya gemensamma registret, som byggs upp i programmet FileMaker. Det gemensamma registret kommer enligt planen att stå färdigt under 2010 och publicering på Internet kommer att påbörjas under 2011

Uppsättningen metadata i det nya registret är dock inte baserad på någon nationell eller internationell standard (t.ex. Dublin Core eller Encoded Archival Description (EAD)) varför kompatibiliteten med register och kataloger hos andra institutioner inte är omedelbar. Emellertid finns alltid en möjlighet att anpassa metadatauppsättningar. Genom valet av standardprogrammet FileMaker för registret finns goda möjligheter för export och anpassning av data till andra format.

3. Vilka delar av samlingarna som består av material som är, eller kan förväntas vara, upphovsrättsligt skyddat och där upphovsrätten tillkommer annan än institutionen.

Samlingarna inom Institutet för språk och folkminnen har tillkommit under en lång period och stora delar har därmed också tillkommit under en tid när frågor om upphovsrätt ännu inte var aktuella. Någon övergripande bedömning om upphovsrätt kan inte göras utan varje källa måste undersökas och bedömas individuellt utifrån bl.a. lagen om upphovsrätt (URL).

Sedan en tid pågår ett arbete vid institutet att utforma en policy för etik- och upphovsrättsfrågor. Där konstateras det att:

”De levnadsberättelser, sagor, sägner, sång, musik eller annat material som bygger på individuella erfarenheter eller individuellt framförande omfattas också av lagen om upphovsrätt (URL) och är skyddat till utgången av det sjuttionde året efter upphovsmannens dödsår¹ Enligt denna rätt får upphovsmannen inte kränkas genom förändringar i verket och skall alltid namnges vid publicering. Denna rätt kan tala tvärt emot etiska ställningstaganden och måste beaktas i samband med ställningstaganden när det gäller skydd av personuppgifter.

Lagen om upphovsrätt återverkar på våra möjligheter att sprida material.

Tillgängliggörande on-line kan exempelvis bara ske med rättshavarens tillstånd. URL inverkar också på hur allmänheten kan få använda vårt material.”

Institutets anställda ska vidare ”lyfta fram och vara medvetna om de upphovsrättsliga frågorna, dels när vi samlar in material, dels inför varje form av publicering. Vi måste också informera våra användare om att material i våra samlingar kan vara upphovsrättsligt skyddat.

¹ För individuellt framförande av annans verk gäller skyddet i 50 år (detta kan t.ex. gälla individuella framföranden av musik som annan komponerat). När det gäller ett fotografi som inte har ”verkshöjd” gäller upphovsrätten 50 år efter framställningsåret.

Vi ska vara behjälpliga med kontakter inför ev. publicering, men när det gäller tillstånd av enskilda upphovsmän måste detta ansvar vila på var och en som vill publicera, inte på myndigheten.” Policyn innehåller också förslag till utformning av informationsmaterial som användare ska ta del av. ”Material som lämnas ut i alla former av kopior skall vara tydligt märkt – med institutets namn, samling och accessionsnummer. Om material läggs ut som digitala filer på nätet skall dessa också märkas med vattenstämpel.” Under 2010 kommer institutet att anlita en jurist för hjälp med dessa frågor.

Frågorna om upphovsrätt kan sägas inverka mekaniskt utifrån ett formellt perspektiv på hanteringen av materialet i samlingarna. En annan fråga att ta lika stor hänsyn till, särskilt vid tillgängliggörande, är *den etiska aspekten*. Samlingarnas innehåll beskriver på olika sätt människors, individers, liv, tankar och erfarenheter och måste därför behandlas med största varsamhet. Detta förhållande berör en stor del av materialet i samlingarna – såväl inspelningar som textmaterial.

En policy som rör den etiska aspekten är under utarbetande vid institutet.

4. Hur användarmönstren har utvecklats över tid vad gäller digitalt tillgängligt material (antal besök på hemsidor, antal visningar eller nedladdningar etc.).

En förändring kan märkas vid institutets avdelningar rörande användarmönstren – de fysiska besöken minskar, medan antalet besök vid webbplatserna ökar. Uppgifter hämtade ur institutets årsredovisningar för 2002–2009 beskriver förhållandet.

Antalet besök på myndighetens webbplatser

	2009	2008	2007	2006*	2005	2004	2003	2002
Totalt antal besök på institutets webbplatser	1 131 399	948 674	1 001 747	500 865	88 289	61 708	71 000	68 646
(Lexin)	311 856 150	342 067 707	287 644 774	221 000 000	-	-	-	-

*Språkrådet anslöts till institutet år 2006, liksom ordboksprojektet Lexin.

Antalet fysiska besökare vid institutets arkivavdelningar

	2009	2008	2007	2006	2005	2004	2003	2002
Totalt antal fysiska besökare vid institutet	2 259	2 372	2 114	2 897	3 062	3 234	3 443	3 575

5. Hur stora delar av det digitala materialet som är tillgängligt via Internet, respektive hur stor andel som endast finns tillgängligt för egen personal eller i samband med fysiska besök vid institutionen.

Avdelningarna i Umeå och Lund har sedan drygt tio år tillbaka publicerat en sökbar kopia av registren över avdelningarnas samlingar på Internet. Sökmöjligheterna har varit uppskattade av forskare och allmänhet och kommit till användning i relativt hög omfattning.

Benägenheten att på Internet publicera material ur samlingarna har begränsats bl.a. av frågorna om etik och upphovsrätt. Utifrån dessa faktorer kan materialet delas in på en skala utifrån hur känsligt det är att publiceras. Samlingar av ortnamn och dialektord kan publiceras tämligen fritt. Dialekt- och folkminnessamlingar som innehåller musik, bilder, levnadsminnen och andra berättelser är mer komplicerade, där hänsyn måste tas både till dess tillkomst – upphovsrätt – och dess innehåll – etik. Noter är ett exempel på skriftligt material där upphovsrätten är påtaglig. Den kategori som är mest komplicerad att hantera i dessa sammanhang är inspelningar. Där måste hänsyn tas såväl till vem som ansvarat och medverkat vid inspelningen, som till vem som spelats in samt även till vad som sägs under inspelningen.

Som en följd av dessa begränsningar har institutet hittills huvudsakligen kommit att offentligt publicera uppgifter om ortnamn och dialektord på Internet, men även, om än i mindre omfattning, inspelningar med dialekter och folkminnen.

Några publiceringsinsatser kan nämnas. Under de sista åren av 1990-talet drevs vid institutets namnavdelning det s.k. SODA-projektet (Svensk ortnamnsdatabas). Målet för projektet var att låta skanna in och tillgängliggöra på Internet samtliga ortnamnsbelägg i samlingarna, registrerade på sedesblad (16:o). Arkivkorten skannades in och den information som finns på korten indexerades. Projektet resulterade i att c:a 3,7 miljoner (3 747 600) ortnamnsblad skannades in, och materialet är nu åtkomligt för allmänheten från institutets webbplats.

Vid DAL drivs ett digitaliseringsprojekt där c:a 600 000 kort med uppgifter om ortnamn respektive dialektord har skannats in och publicerats genom ett MySQL-skript för allmänheten på Internet.

DAUM presenterar en mindre dialektordsamling (c:a 6000 ord), med sökfunktion, på sin webbplats.

På institutets webbplats kan man lyssna på dialekter och folkminnen i en sammanställning från Sveriges landskap, som är under arbete. CD:n Dialekter i Dalarna, med 30 dialektprov från alla delar av Dalarna, som publicerades år 2004, är också tillgänglig på webbplatsen.

På webbplatsen finns även ett material publicerat som behandlar dialekterna i Sörmland och Värmland.

NA presenterar Ordbok över Sveriges medeltida personnamn som nätpublikation på webbplatsen, liksom skriften Nordiskt runnamnslexikon.

Vid DA pågår ett arbete med en presentation av svenska dialekter med ljudexempel på en klickbar karta på webbplatsen.

Texter, foton och inspelat material finns alltså ännu inte tillgängliggjorda på Internet i någon större omfattning, förutom ett begränsat antal demonstrationsinspelningar av dialekter. Det material som digitaliserats och som beskrivs i tabellerna ovan på sid. 1–2 är tillgängligt för institutets personal och kan även distribueras till forskare och allmänhet efter förfrågan. Försiktighet iakttas vid distributionen av digitalt material eftersom detta format är enkelt att kopiera vidare. Av den policy omkring etik och upphovsrätt som arbetas fram inom institutet framgår att allt material som lämnas ut i kopior ska märkas och inte får spridas vidare.

Genom att vissa register publicerats på Internet har samlingar varit åtkomliga för forskare och allmänhet, inte omedelbart för nedladdning, men väl efter kontakt med respektive arkivavdelning.

”Tillgängligt via Internet” innebär inte nödvändigtvis enbart att materialet ska vara publicerat för nedladdning på Internet, utan även att Internet kan användas som distributionskanal för material t.ex. som bilagor till e-post, nedladdning genom HTML eller för distribution via FTP. Förutsättningen för distribution över Internet är dock att materialet föreligger i digitalt format. Beträktat utifrån denna aspekt är allt digitaliserat material vid institutet, redovisat ovan på sid. 2, ”tillgängligt via Internet”.

<p>6. I vilken utsträckning det digitala materialet är tillgängligt via externa system, t.ex. Europeana och K-samsök, men även kommersiella leverantörer som Google.</p>

Boksamlingarna vid avdelningarna i Uppsala och Umeå är registrerade i Libris, den gemensamma nationella samkatalogen för universitet och högskolor.

Även den samling samisk jojk, som finns vid institutet, och som i viss utsträckning är digitaliserad, är till största delen registrerad i Libris i samverkan mellan DAUM och Åjtte.

DAUMs förteckning över musikrelaterat material finns registrerad i The LICHEN Project: The LInguistic and Cultural Heritage Electronic Network, som är ett internationellt tvärvetenskapligt initiativ rörande språk och kulturer i det nordliga cirkumpolära området. Inom projektet uppmärksammas särskilt språken Meänkieli och Sverigefinska i Sverige, Kvänska i Norge, Karelska varieteter och Vepsiska i Ryssland, språken i Võro och Seto i Estland samt liviska i Lettland. Initiativtagare är forskare vid universiteten i Uleåborg och Glasgow.

DAUM och DAG har även kontakter med DISMARC, DIScovering Music ARChives, som är ett europeiskt initiativ som tillhandahåller ett presentationsformat för musiksamlingar, oberoende av metadatastruktur. Användare får överblick över innehållet i musiksamlingar i Europa. DISMARC ingår i Europeana, EU:s webbportal för presentation av kulturarvet.

European Cultural Heritage Online (ECHO) – an Open Access Infrastructure to bring Essential Cultural Heritage Online. Under 2000-talet tog DAUM kontakt med de svenska representanterna för ECHO vid Inst. för lingvistik i Lund och förde resonemang om medverkan i projektet.

Del II – Beskriva myndighetens eller institutionens arbete och resultatet av arbetet med digitalisering, elektronisk tillgång och digitalt bevarande samt redogöra för:

7. Om det finns en digitaliseringsplan eller motsvarande, i vilken mån planens mål uppfyllts samt vilka principer som gäller för att välja ut vad som ska digitaliseras och för mottagandet av digitalt material.

Inom arbetet i de långsiktiga projekten, som drivs vid institutet, t.ex. Svenskt dialektlexikon, Ordbok över Sveriges medeltida personnamn och Sveriges ortnamn, sker registrering med hjälp av dator och det arbetet kan på så sätt sägas ingå i digitaliseringsverksamheten. Utgångspunkten i resonemangen nedan är dock att termen *digitalisering* innebär verksamhet där huvudsyftet är att konvertera arkivmaterial från magnetiska och mekaniska ljudbärare samt från papper, i analogt format, till digitalt format.

Digitalisering av materialet i samlingarna sker dels organiserat i projektform och dels som löpande verksamhet.

Som några exempel på organiserade projekt kan nämnas att vid Avdelningen för digitalt arkiv och lokal service i Uppsala har digitalisering av lackskivor prioriterats sedan slutet av 1990-talet. Den prioriteringen är en följd av att lackskivorna håller på att förstöras av tiden; de krackelerar och lacket släpper från skivornas kärna. Prioriteringen har givit resultat på så sätt att år 2008 hade ca 80 % av den på lackskivor lagrade informationen överförs till digitalt format.

Vid DAUM har överföring av information på rullband prioriterats. Där har också fokuserats på överföring av information lagrad på papper. Skanning av text och bild har prioriterats alltsedan slutet av 1990-talet.

Vid DAL drivs ett projekt, som startade under 00-talet, rörande digitalisering, indexering och tillgängliggörande på Internet av ortnamns- och dialektordssamlingarna, Skånsk ortnamnsdatabas och Syd- och västsvensk dialektdatabas.

Vid DAG har arbetet med överföring av analoga data till digitalt format tagit fart under senare delen av 00-talet. Inspelat material har digitaliserats, bland annat stora samlingar kassetband, och text/bild skannas. Ett speciellt projekt digitaliserar och registrerar folkmusikinspelningar, främst rullband och kassetband, från Västra Götalandsregionen.

Digitalisering förekommer också i samband med beställningar av material från forskare och allmänhet. I samband med denna form av tillgängliggörande digitaliseras de efterfrågade källorna även i ett format för långsiktigt bevarande.

En fastställd digitaliseringsplan har inte upprättats bland annat p.g.a. att digitaliseringsverksamheten är ett uppdrag som tillkommit utöver institutets ordinarie uppdrag men som har måst ske inom befintliga resursramar. Detta har i många fall medfört att digitalisering har kunnat ske när tillfälle givits, ofta genom samarbeten med Arbetsförmedlingen, Försäkringskassan, kommunförvaltningar och privata rehabiliteringsinstitutioner, som bidragit med både utrustning och arbetskraft. Stor hjälp har också erhållits i digitaliseringsarbetet

genom statliga arbetsmarknadsåtgärder som Sesamprojektet och Access samt förväntas i nu aktuella Lyft.

Den långsiktiga digitaliseringsplan som styr arbetet vid Institutet för språk och folkminnen kan sammanfattas i några punkter:

1. Arkivverksamheten byter tillstånd och materialet konverteras därigenom från analogt format till digitalt format. Allt material i samlingarna vid institutet och dess avdelningar ska därför genomgå digitalisering.
2. Materialet är så omfattande att konvertering inom institutets ordinarie ramanslag kommer att löpa under många år, alternativt bli mycket kostsamt vid anlitande av extern entreprenör.
3. Digitalisering sker systematiskt organiserat i projektform och som löpande verksamhet.
4. Digitalisering av inspelningar, särskilt lackskivor, prioriteras före digitalisering av texter och foton
5. En för institutet gemensam registerkatalog över samlingarna är under upprättande och beräknas kunna publiceras på Internet under 2011.
6. Särskild uppmärksamhet riktas mot de nationella minoriteterna, teckenspråk samt mot att skapa användarvänliga lösningar för människor med särskilda behov.

Principen för urval av material som ska digitaliseras kan sägas vara styrd av behov och efterfrågan på så sätt att material som med tiden riskerar att förstöras, samt material som efterfrågas av forskare och allmänhet, prioriteras för digitalisering.

Mottagande av digitalt material sker i relativt liten omfattning vid institutet och de praktiska frågorna löses när de aktualiseras. Inkommet material som "fötts digitalt" registreras som digitalt original, och genomgår sedan samma process för lagring och tillgängliggörande, som det från analoga original digitaliserade materialet.

8. I vilken utsträckning avgifter tas ut för tillgången till digitalt material, vilka bemyndiganden som finns för ett eventuellt avgiftsuttag samt vilka principer i övrigt som gäller för avgiftsuttaget.

I instruktionen för Institutet för språk och folkminnen konstateras att "11 § När myndigheten utför de uppgifter som anges i 2 § 1 och 2 får avgifter tas ut inom ramen för full kostnadstäckning. Avgifternas storlek bestäms av myndigheten, utom i de fall som avses i 15 § avgiftsförordningen (1992:191). Myndigheten får disponera avgiftsinkomsterna." SFS 2009:743.

Institutet har efter samråd med ESV fastställt en avgiftsmodell, 2008-12-16, där uttaget av avgifter regleras. Avgifter tas ut endast för beställda digitala kopior. Det material som finns tillgängligt på Internet kan däremot användas kostnadsfritt.

9. Vilka typer av upphovsrättsligt skyddat material som gjorts tillgängligt digitalt och hur rättigheter i dessa fall inhämtats från rättighetshavarna samt kostnaderna för att publicera materialet digitalt.

Material ur samlingarna har publicerats digitalt i begränsad omfattning, och det är huvudsakligen material där upphovsrätten ligger hos institutet som gjorts tillgängligt digitalt. Det mesta av det tillgängliggjorda materialet härrör således ur institutets samlingar och frågan om rättigheter har därför behandlats internt. I den mån rättigheter inhämtats utifrån har detta vanligen skett genom muntliga överenskommelser, eller genom formella dokument. Inga kostnader har dock hittills uppkommit för institutet för digital publicering av upphovsrättskyddat material.

10. Vilka kommunikativa strategier som finns för att nå ut med digitalt material till olika grupper samt vilka insatser som har gjorts för tillgänglighet i den digitala miljön för grupper med särskilda behov.

Institutets ambition är att kunna tillgängliggöra allt material som inte är klassificerat som känsligt eller belagt med nyttjandebegränsningar. Det material som hittills lagts ut är framförallt databaser över ortnamn och dialektord. Det myndighetsgemensamma accessionsregistret kommer att möjliggöra goda sökmöjligheter via nätet. Vidare är också arbetet med översyn av etik- och upphovsrätsfrågor en del av en kommunikativ strategi. De tekniska resurserna vid institutet har inte möjliggjort en vidare anpassning av presentationsformen för materialet.

Institutet har en stor samling material på de nationella minoritetsspråken samiska och meänkieli, som utgör en viktig källa för revitaliseringsarbete. Arbetet med att digitalisera minoritetsspråkssamlingarna bör prioriteras för att underlätta för språkvårdande insatser.

Institutet har ett språkvårdande ansvar för fyra av de fem nationella minoritetsspråken – finska, meänkieli, romani och jiddisch (ansvaret för de samiska språken ligger hos Sametinget). Därutöver ska institutet ”samla kunskap om och följa användningen av även andra språk i Sverige” SFS 2009:743 § 2:2. Det är därför viktigt att understryka att från institutet tas hänsyn till människors olika språk och språkliga behov vid tillgängliggörandet av digitalt innehåll och annan information på Internet. Ett projekt pågår för närvarande vid Språkrådet, ”Språk och Internet”, om flerspråklig tillgänglighet på Internet. En rapport publiceras inom kort. Det finns en vägledning för flerspråklig information utlagd på Språkrådets webbplats: <http://www.språkradet.se/6485>.

Institutet har också ett ansvar för det svenska teckenspråket. Ett digitalt arkiv för teckenspråk är därför mycket angeläget för språkvårdsarbetet samt för forskare, studenter och andra intresserade. För talade och skrivna språk finns det gott om språkliga källor att hämta kunskaper ifrån och forska i. För teckenspråket är möjligheterna mycket begränsade trots att det finns en hel del material i dagsläget. Problemet är att materialet är spritt på olika platser och att det finns upphovsrättsliga och etiska frågor som måste lösas. Språkrådet arbetar för att bilda ett samlat digitalt arkiv för teckenspråkligt material. En inventering av inspelningar av tecknat språk har påbörjats under 2009 och målet är att ställa samman, digitalisera och långsiktigt bevara samt tillgängliggöra material på teckenspråk.

De insatser som hittills gjorts för att öka tillgängligheten i den digitala miljön för grupper med särskilda behov är att genom talsyntes erbjuda synsvaga tillgång till webbplatsernas innehåll. En totalrevidering av institutets webbplatser är under genomförande där målet är att de ska uppfylla alla krav på tillgänglighet. En kartläggning av tillgångar och behov är genomförd och arbetet med att anpassa sidorna sker under 2010.

11. Hur det långsiktiga bevarandet av digitaliserat respektive ursprungsdigitalt material säkerställs.

Det ”långsiktiga bevarandet av digitaliserat [...] material” innefattar ett flertal resurskrävande moment. Det första man associerar till är den diskryta som måste till för att lagra de stora mängder data som genereras, för institutets del uppskattas behovet uppgå till 60–70 terabyte (TB). Datamängden måste av säkerhetsskäl också lagras i (minst) en backup-kopia (=totalt 140 TB), och lagringsdiskarna måste säkras mot virus och intrångsförsök, något som kräver en hög teknisk kompetens. Vidare måste en nätverksförbindelse med hög överföringskapacitet finnas från varje plats där digitalisering sker, till lagringsorten. Datamängden kommer slutligen på sikt att återkommande få genomgå konverteringar för att följa förändringar av standarder i den tekniska utvecklingen.

Institutet saknar resurser att på egen hand bygga upp en organisation för långtidslagring av digitala data. Därför måste det långsiktiga bevarandet av digitaliserat material ske genom upphandling av lagringstjänster från externa leverantörer. För närvarande köper institutet tjänsten från Avdelningen för IT och inköp vid Uppsala universitet. Kostnaderna, som redan nu når betydande nivåer, kommer att öka i takt med att mängden lagrade digitala data ökar genom det fortsatta digitaliseringsarbetet.

Möjligen kunde en central infrastruktur för långtidslagring, gemensam för de statliga kulturminnesinstitutionerna, byggas upp eller upphandlas och administreras centralt. Kostnaderna kunde då begränsas för de medverkande genom de samordningsvinster som skulle uppnås.

12. Vilka samarbeten med andra offentligt finansierade institutioner som har genomförts eller initierats inom områdena digitalisering, elektronisk tillgång och digitalt bevarande.

För transport av digitaliserade data till långtidslagring krävs en väl utbyggd nätverksförbindelse. Institutet samverkar på detta område med universitet och högskolor på så sätt att institutet som s.k. extern organisation är anslutet till SUNET (Swedish University Computer Network), med medgivande om anslutning från universiteten vid samtliga orter där institutet finns representerat med avdelningar.

Vid institutets avdelningar har genom åren löpande kontakter skett med arbetsförmedlingen och kommunerna vid respektive orter för tillhandahållande av praktikplatser och arbetsuppgifter i samband med arbetsmarknadsåtgärder. Dessa arbetsuppgifter har i många fall handlat om digitalisering av material ur samlingarna, framförallt genom skanning, registrering i databaser och inskrivning av text i dator.

Som exempel på sådana samarbeten kan nämnas Umeå kommuns initiativ Scan-IT där långtidsarbetslösa ungdomar och ungdomar med särskilda behov erbjuds tidsbegränsad anställning för skanningsarbete. DAUM anlätade Scan-IT under 00-talet för inskanning av 10 000 sidor text ur samlingarna.

Ett liknande samarbete sker för närvarande med Uppsala kommun, som inrättar arbetsplatser för upp till 10 deltidsarbeten, där institutet tillhandahåller arbetsuppgifter och arbetsmaterial. Arbetsuppgifterna omfattar indexering och registrering i databas av information från skannade bilder av registerkort över dialektord ur samlingarna.

DAUM har även anlitat Forskningsarkivet vid Umeå universitet för inskanning av delar av Nils Tibergs estlandssvenska dialektordssamling samt 1883 års jordebok över Västerbotten.

Diskussioner omkring upprättande av ett teckenspråksarkiv pågår vid institutet. Avdelningen Språkrådet har kontakter med Stockholms universitet, Institutionen för lingvistik, avdelningen för teckenspråk, om insamlat material.

DAL har samverkat med Landsarkivet kring Demografisk Databas Södra Sverige (DDSS) i fråga om normalisering av ortnamn. DAL har tillhandahållit ett digitalt register över skånska bebyggelsenamn och får i gengäld av DDSS nya namn som påträffas vid inmatning av uppgifter ur kyrkoböckerna.

Språkrådet bedriver ett samarbete med övriga nordiska länders offentliga språkvårdsorganisationer om att skapa en gemensam taxonomi och terminologi för uppmärkning av språkfrågor och tillhörande svar. Planen är att sedan genom en gemensam sökportal göra det möjligt att söka i alla nordiska databaser samtidigt.

13. Vilka samarbeten med privata aktörer, såsom företag, folkrörelser och andra ideella organisationer, som har genomförts eller initierats inom områdena digitalisering, elektronisk tillgång och digitalt bevarande.

Under slutet av 1990- och början av 2000-talet pågick ett arbete omkring utveckling av metoder för digitalisering av inspelat material lagrat på grammofonskivor. För utvecklingsarbetet anlätade dåvarande Fonogramenheten vid institutet i Uppsala konsultfirman Swedish Sound Reference AB (SSRAB).

Årligen arrangeras en folkmusikarkivkonferens där såväl institutioner som ideella organisationer och privatpersoner medverkar, verksamma inom området för bevarande av musiksamlingar. Ett återkommande tema vid konferenserna är digital registrering, digitalisering och tillgängliggörande av musikmaterial.

Vid DAG pågår ett projekt för att digitalisera inspelningar och bygga upp ett arkiv för folklig sång och musik från Västra Götalandsregionen. "Hitta folkmusiken.nu" är ett samarbete med spelmansförbunden i Västsverige, Västarvet (förvaltningen för natur- och kulturarv inom Västra Götalandsregionen), Vara folkhögskola, Musik i Väst, ABF och Vuxenskolan. I samarbete med Arbetsförmedlingen bekostas en externt finansierad projektanställning.

Samverkan har initierats med Sveriges Dövas Riksförbund som har ett stort videoarkiv för tiden 1978-2000.

Del III – Respektive myndighet eller institution ska dessutom, utifrån de strategiska utvecklingsområden som beskrivits ovan, bedöma vilka särskilda utmaningar och möjligheter som den digitala tekniken innebär inom det egna verksamhetsområde

14. Utifrån de tre angivna strategiska områden, vilka särskilda utmaningar (U) och möjligheter (M) innebär den digitala tekniken inom det egna verksamhetsområdet?

I effektivare hantering av digital information (U / M)

II användarvänliga lösningar (U / M)

III metoder att hantera upphovsrättsfrågor i den digitala miljön (U / M)

I. En effektivare hantering av digital information

I:1. Samutnyttja investeringar för digitalisering och lagring.

Institutets samlingar kan indelas i två grupper utifrån metod och teknik i digitaliseringsprocessen *pappersburen information (text/bild)* och *inspelningar*. I stället för att bygga upp en egen resurskrävande infrastruktur av teknik och personal för storskalig konvertering av pappersburen information kan samutnyttjande ske t.ex. med Riksarkivets avdelning Mediakonverteringscentrum (MKC) i Fränsta, Ånge kommun. Där kan alla typer av historiska handlingar migreras till digitala bilder. Verksamheten är den största i sitt slag i Europa och vid MKC skannas c:a 100 000 bilder per dygn. <http://www.riksarkivet.se/default.aspx?id=2229&refid=1135>

Ett liknande resonemang kan föras för inspelat material, som kan omhändertas och storskaligt migreras t.ex. av Sveriges Radio Förvaltnings AB, SRF Ånge. I mars 2009 hade där 70 000 band digitaliserats åt SVT, UR och Statens ljud- och bildarkiv. <http://www.srf.se/Page.aspx?id=320>

I:2. Ta fram och sprida gemensamma standarder

En modell för upprättande och drift av digitala arkiv samt tekniska standarder för digitalt arkivmaterial presenteras i Riksarkivets föreskrifter och allmänna råd om elektroniska handlingar (upptagningar för automatiserad behandling), RA-FS 2009:1 och Riksarkivets föreskrifter och allmänna råd om tekniska krav för elektroniska handlingar (upptagningar för automatiserad behandling), RA-FS 2009:2.

Vid institutet har tekniska standarder utvecklats och fastställts för digitalisering av skriftligt material och av inspelningar.

Skriftligt material – ingen skillnad görs mellan text och bilder – skannas in i ≥ 300 dpi med referensvärdet >3000 pixlar för x- eller y-axeln. Filerna sparas i 8 bitars RGB-färg, i TIFF för långtidsbevarande och komprimeras vid behov till 72 dpi JPEG- eller PDF-format för publicering på Internet och för distribution.

Inspelningar sparas, oftast i mono, i 16–32 bitar 44,1 kHz, wave för långtidsbevarande och komprimeras till mp3 för presentationer på Internet och för distribution.

I:3 Ta fram modeller för digital långtidslagring

En grupp av representanter för myndigheter inom kulturarvssektorn kunde tilldelas uppdraget att utarbeta en modell för digital långtidslagring för statliga myndigheter.

Flera initiativ är på andra håll under utarbetande. T.ex. kunde ett samverkansprojekt liknande REDA upprättas för det statliga området. ”REDA står för *REgionalt Digitalt Arkiv* och är ett samverkansprojekt mellan Regionarkivet i Skåne län, Landsarkivet i Lund, Lunds universitetsarkiv och Lunds kommun. Projektets syfte är att utveckla ett digitalt arkiv avsett för långtidsbevarande av elektroniska handlingar. Utvecklingsarbetet har skett från grunden inom REDA men tar sin utgångspunkt i den internationella standarden OAIIS.” <http://www.redaprojektet.se/>.

I:4 Upphandla licenser gemensamt

För upphandling av programvarulicenser och mjukvarusystem utnyttjas leverantörer med statliga ramavtal, www.avropa.se. Inom institutet sker sedan ett par år tillbaka gemensam upphandling av programvarulicenser till avdelningarna på alla fem orterna i landet.

Om ytterligare samordningsvinster kan tänkas uppnås som överträffar de statliga ramavtalen, genom gemensam upphandling inom kulturarvssektorn, skulle någon framträdande part med övergripande ansvar inom sektorn kunna fungera som central samordnare för upphandlingen.

Upphandling av andra licenser, t.ex. för publiceringsrättigheter, sker i mycket liten omfattning inom institutets verksamhet.

II. Användarvänliga lösningar.

II:1 Samordning av kataloger och portaler

Samordning av kataloger och portaler i initiativ som det nationella K-samsök och det sameuropeiska Europeana ger goda möjligheter för användarna att få en överblick över det ämnesområde som efterfrågas (M). Utmaningen ligger bl.a. i de tekniska lösningarna och hur metadata ska struktureras för att samordningen ska bli möjlig (U).

II:2. Tillgängliggörande via standardiserade gränssnitt

Det viktiga för användarna är att komma åt informationen och få svar på sina frågor. Genom standardiserade gränssnitt underlättas för användarna att nå den efterfrågade informationen (M). De samlingar som ska presenteras i de gemensamma gränssnitten är ofta av olika karaktär och skilda aspekter har uppmärksamats vid fastställandet av beskrivande metadata. Arbetet med att utveckla standardiserade gränssnitt där varje material kan presenteras på ett uttömligt och informativt sätt och där varje ägare ges möjlighet att uttrycka sin samlings särart kan därför ses som en utmaning (U).

Flera exempel kan hämtas från Internet där initiativ tagits till standardiserade gränssnitt. Riksantikvarieämbetets samsöktjänst Kringla (www.kringla.nu) är ett strålande exempel på hur information från skilda organisationer kan samordnas under en portal.

II:3. Öppna för ökad interaktivitet med användarna

Institutet välkomnar en ökad interaktivitet med användarna. Den kunde ta sig många uttryck i verksamheten, t.ex. presentation från institutet av dialektord och ortnamn, och möjlighet för användarna att rapportera in belägg för dialektord och ortnamn ur sina egna erfarenhetsvärldar. Vidare kan t.ex. insamling av människors berättelser och svar på frågelistor ske online, jfr t.ex. WiGo, Västerbottens museums wiki-satsning, där allmänheten bereds möjlighet att skapa sitt lokala kulturarv på Internet genom att skriva in texter, som publiceras på muséets webbplats genom en koppling till Google maps. Ett

annat exempel är Folklivsarkivet i Lund, som publicerar frågelistor för nedladdning och besvarande. (M). Utmaningen ligger i att skapa resurser för utveckling och drift samt utarbetande av tekniska lösningar (U).

Verka för att göra material fritt tillgängligt, t.ex. Lexin. Då kan det vidareförädlas på olika sätt t.ex. genom wikiteknik eller i kommersiella tillämpningar. Frågor om upphovsrätt och licensiering kring detta måste dock först lösas.

II:4. Öka tillgängligheten för grupper med särskilda behov.

En stor grupp nya användare kan nås genom att kataloger och webbpresentationer anpassas, i samråd med representanter för grupper med särskilda behov (M). Utmaningen ligger i att anpassa presentationsmodeller utifrån skilda gruppers behov (U).

III. Metoder att hantera upphovsrätsfrågor i den digitala miljön

III:1. Insikt och kompetens om den gällande upphovsrättsliga lagstiftningen

Frågorna om upphovsrätt är komplexa och svårhanterade och en grundläggande kännedom och kunskap om den bakomliggande juridiken är nödvändig i den löpande verksamheten (U). För större beslut, som t.ex. inför publicering på Internet måste expertkompetens finnas att tillgå. En informations- och utbildningsfunktion med relevant juridisk kompetens, centralt placerad i den statliga arkivorganisationen, vore en tillgång för arbetet med det digitala arkivet (M).

III:2. Aktiva kontakter och nätverksbyggande med företrädare för berörda rättighetsgrupper

I den dagliga verksamheten är tillfällena till kontakter med rättighetsgrupperna inte så ofta återkommande. En central informationsfunktion att vända sig till med frågor, där relevant och uppdaterad kompetens finns tillgänglig skulle därför underlätta (M).

III:3. Principer för avgiftsuttag

I avgiftsförordningen, SFS 1992:191, anges riktlinjer och principer för avgiftsuttag. På basis av dessa har institutet utarbetat en Avgiftsmodell för tillämpning i den löpande verksamheten, som omnämns ovan. Institutet ställer vidare i verksamhetsbudgeten upp intäktsposter för försäljning av kopior av materialet i samlingarna samt försäljning av utgivna böcker. I vissa fall tas även ut en avgift för publicering av material ur samlingarna. Inga avgifter tas dock ut för nyttjande av det material och de tjänster som nu tillgängliggörs på institutets webbplats. Önskvärt är att kunna fortsätta att tillhandahålla det nationella kulturarvet fritt tillgängligt, utan avgiftsbeläggning, för forskare och allmänhet. Institutet har därför inte för avsikt att avgiftsbelägga användandet av digitalt tillgängliggjort material på Internet.

Möjligheterna att öka intresset för kulturarvet genom gemensamma och samordnade presentationer på Internet ökar om informationen är fritt tillgänglig (M).

Ska avgiftsbeläggning av material och tjänster tillämpas för sådant som är presenterat i en gemensam samordnad portal är det viktigt att det sker utifrån gemensamma principer (U). Det bör dock framhållas att den fortsatta uppbyggnaden av institutets *samlingar* i stor utsträckning sker genom frivilliga bidrag från forskare och allmänhet. Formaliseras avgiftsbeläggningen av uttag ur samlingarna striktare kan man vänta sig att även de som bidrar med material till samlingarna kommer att vilja ha ersättning i högre utsträckning än vad som nu är fallet. En sådan utveckling vore olycklig för verksamheten.

Ange även: Vilka samarbeten med andra offentligt finansierade institutioner som bör prioriteras inom områdena digitalisering, elektronisk tillgång och digitalt bevarande,

För mindre institut och myndigheter med begränsade resurser är kontakterna med de stora organisationerna inom kulturarvssektorn, Riksarkivet, Riksantikvarieämbetet, Kungliga biblioteket av stor betydelse för framstegen inom områdena digitalisering, elektronisk tillgång och digitalt bevarande.

Dessa organisationer kan dela med sig av tankar och erfarenheter inom området i utbildningar, seminarier och konferenser (M). ABM-centrum kan stå som koordinatör i uppdrag och kontakter mellan aktörerna.

På folkminnessidan finns ett etablerat samarbete mellan institutets avdelningar och Nordiska museet samt Folklivsarkivet i Lund. Våra institutioner har en lång, gemensam tradition när det gäller dokumentation av och forskning om olika former av immateriella kulturarv, som bör vidareutvecklas också när det gäller dessa frågor. I våra samlingar, där de äldsta etablerades i slutet av 1800-talet, finns likartat material och gemensamma utmaningar.

Myndigheten Handisam är en viktig part för kontakter med nya grupper användare (M).

Ett förslag finns hos Vetenskapsrådet om finansiering för en nationell språkbank med resurser för språkvetenskaplig och språkteknologisk forskning. Förslaget har anknytning till det europeiska infrastrukturprojektet Clarin, som med språkteknologiska metoder vill skapa nya möjligheter för humanistisk forskning att hantera språkliga data. Samtidigt vill det också skapa möjligheter för språkteknologisk utveckling, som t.ex. kan bidra till bättre tillgänglighet till digitalt innehåll, t.ex. textuppläsning, taligenkänning och teckenspråksgenerering. Behovet av en sådan nationell språkbank måste framhållas, liksom behovet av verktyg för att analysera, märka upp och söka i språkliga data, samt standarder för det. Andra som uppmärksammar frågan på olika sätt i sina svar är Handisam, Hörselskadades riksförbund och Kungliga Biblioteket. (U/M)

Samarbete bör också prioriteras med SVT och Utbildningsradion som båda har en hel del teckenspråklig produktion i sina arkiv, med Riksteaterns Tyst Teater som har videoinspelningar av sina teaterproduktioner på teckenspråk, Stockholms universitet, Institutionen för lingvistik, Avdelningen för teckenspråk, som för närvarande håller på att bygga upp en teckenspråkskorpus. (M)

samt: Vilka samarbeten med privata aktörer, såsom företag, folkrörelser och andra ideella organisationer, som är möjliga framöver inom områdena digitalisering, elektronisk tillgång och digitalt bevarande.

Organisationen Handikappförbunden med ett 40-tal medlemsförbund för att nå nya grupper användare (M).

Föreningar och organisationen som representerar de nationella minoritetsspråken samt teckenspråk (M).

Arkiv, muséer, studieförbund och hembygdsföreningar (M).