

Digitala arkiv för e-tidskrifter

Karolinska Institutet Universitetsbiblioteket, 2014-08-26

Johnny Carlsson
Darlene Duterrier
Helena Skyllberg



**Karolinska
Institutet**

Digitala arkiv för e-tidskrifter

Vi har fått i uppdrag att undersöka vilka digitala arkiv som finns tillgängliga för bevaring och lagring av elektroniska tidskrifter. Syftet med detta projekt är att komma fram till om KIB ska fortsätta vara medlem i ett av arkiven eller om vi ska byta till något annat arkiv, eller inte vara med i något arkiv alls. Även om det finns e-böcker i arkiven också, tar vi inte med dem i beräkningen eftersom de inte ingick i uppdraget.

Definitioner

Digitala arkiv – elektroniska arkiv som tillhandahåller bland annat vetenskapliga e-tidskrifter och e-böcker. Dessa bevaras i arkiv när upphovsrättsinnehavaren, dvs. förlaget, av olika skäl inte längre kan tillhandahålla accessen till dem. Vem som äger och underhåller arkiven kan variera, det kan vara antingen helt fristående företag eller universitet och högskolor. Syftet med dessa arkiv är att arkivera, tillhandahålla och bevara det elektroniska innehållet för kommande generationer.

Light archives – öppna arkiv som är åtkomliga för medlemmar (i vissa fall även för icke medlemmar), ungefär som i ett eget repositoryum.

Dark archives – stängda arkiv, inte åtkomliga för någon, inte ens för medlemmar. När en titel triggas i arkivet så blir den fritt tillgänglig för medlemmarna. För att en titel ska triggas krävs att förlaget inte längre har möjlighet att tillgängliggöra titeln av någon anledning.

Triggade titlar – triggade titlar innebär att de blir tillgängliga när någon av följande händelser inträffar: förlaget upphör, förlaget inte längre kan eller vill tillhandahålla en titel, arkivet försvinner från förlagets egen hemsida, samt om det uppstår ett plattformshaveri hos förlaget. Vem de triggade titlarna blir tillgängliga för beror på vilket arkiv de finns i.

Observera att nedanstående beskrivningar av de olika arkiven innehåller uppgifter och siffror som gäller vid skrivande stund.

Hathi Trust

Hathi Trust-samarbetet startades 2008 vid University of California och är ett arkiv med medlemmarnas egna digitaliseringar. Arkivet består av både fritt tillgängligt och upphovsrättsskyddat material, innehåller ungefär 5,8 milj. boktitlar samt drygt 290 000 seriella publikationer av olika slag.

Alla medlemmar (cirka 90 stycken) är amerikanska bibliotek. De betalar en avgift som inkluderar infrastruktur, lagring, backup, servrar och viss personal. Det upphovsrättskyddade kommer bara den egna institutionen till del. Hathi Trust är så kallat "light archive", ett öppet arkiv, där medlemmarna har tillgång till allt direkt.

LOCKSS

LOCKSS (Lots Of Copies Keep Stuff Safe) startade 1999 och är ett open source program utvecklat vid Stanford University. Det används av bibliotek i olika LOCKSS nätverk. För att bevara material i ett så kallat "LOCKSS arkiv" måste biblioteken uppfylla följande tre villkor: kontakta förlaget för att få rättigheter till arkiveringen (det är förlagsversionen som arkiveras), skapa ett kontrollerat "LOCKSS box" på egen server eller PC, samt att boxen ska sedan registreras inom ett LOCKSS nätverk. Ett LOCKSS arkiv ligger lokalt hos varje bibliotek och allt i arkivet är direkt åtkomligt för bibliotekens egna användare. I ett sådant arkiv kan biblioteket även spara annat lokalt material. I LOCKSS deltar ungefär 600 förlag (både tidskrifts- och e-boksförlag).

Liksom Hathi är LOCKSS ett så kallat "light archive", ett öppet arkiv.

CLOCKSS

Ett av de största nätverken är CLOCKSS (Controlled LOCKSS) är ett slutet arkiv som styrs och administreras gemensamt av medlemmarna. Att delta i CLOCKSS kräver medlemskap och aktivt deltagande. Ungefär 260 bibliotek och ca 180 förlag är medlemmar. Arkivet har nio triggade titlar fritt tillgängliga för alla, inte bara för medlemmar.

Det finns även ett så kallat Global LOCKSS Network (GLN). GLN är ett globalt öppet nätverk för LOCKSS arkiv. I detta nätverk medverkar ca 530 förlag (både tidskrifts- och e-boksförlag), men det framgår inte hur många bibliotek är med. CLOCKSS och GLN är båda så kallade "dark archives", det vill säga stängda arkiv med stort antal titlar som så länge de inte är triggade förblir otillgängliga för medlemmarna.

Efter att ha undersökt dessa arkiv har vi kommit fram till slutsatsen att välja bort Hathi, LOCKSS, CLOCKSS och GLN. Detta för att det Hathi har enbart amerikanska medlemmar och de andra kräver förutom medlemskap också egen arbetsinsats och egna servrar samt eventuellt annan utrustning.

Portico

Portico startade 2002 och drivs av ITHAKA som är en icke kommersiell organisation som stödjer universitet och högskolor inom bevarandefrågor. Porticos uppdrag är att långsiktigt bevara digitaliserat material. Drygt 900 bibliotek och 250 e-tidskriftsförlag är medlemmar från hela världen. Idag finns cirka 19 000 tidskriftstitlar och över 200 000 boktitlar arkiverade.

Medlemskapet baseras på en viss procent av bibliotekets totala inköpsbudget.

KIB är medlem i Portico sedan 2006 och är en av de så kallade "archive-founders". År 2014 betalade KIB ca 45 000 SEK i medlemsavgift.

Portico är också ett "dark archive", det vill säga ett stängt arkiv och i dagsläget

finns 11 stycken triggade titlar som är enbart tillgängliga för Porticos medlemmar.

Jämförelse mellan Porticos och KIBs tidskriftsbestånd

Vi har jämfört KIBs tidskriftsbestånd i SFX med tidskriftstitlarna i Porticos arkiv. KIB har ca 10 000 unika prenumerationstitlar (titlar vi betalat för) aktiverade i SFX. Vi valde att titta närmare på de 2 000 mest nedladdade titlarna. Att vi valde just 2000 titlar har att göra med att de utgör 20 % av vårt totala tidskriftsbestånd men står för drygt 93 % av den totala användningen. Den mest nedladdade titeln, Nature, är nedladdad drygt 76 000 gånger och den 2000:e har enbart 186 nedladdningar. KIBs totala användning (2013 års siffror) låg på 2,8 miljoner nedladdningar per år och dessa 2000 står för 2,6 miljoner. Jämförelsen visade att av dessa 2000 titlar finns 1747 stycken i Porticos arkiv. De 253 titlar som saknas i arkivet har laddats ned ca 558 000 gånger. Av KIBs fem mest använda titlar saknas tre i Portico: New England Journal of Medicine, Science och Journal of Biological Chemistry. De två andra, Nature och PNAS, finns med i Portico.

Ett stort tack till Henrik Åkerfelt för hjälpen med titeljämförelsen.

Rekommendation

Av de arkiv vi har tittat på har Portico flest bibliotek som medlemmar. De flesta stora förlagen är också medlemmar i Portico, även om LOCKSS/CLOCKSS samarbetar med många fler förlag totalt men där ingår också e-boksförlagen. Till skillnad från de andra arkiven kräver Portico inte någon egen arbetsinsats eller egen infrastruktur. Det finns en etablerad organisation som har kunskap och resurser för att upprätthålla ett arkiv anpassat för förändringar inom både åtkomst och lagring. De är dessutom förlagsoberoende.

Vår rekommendation är, att så länge KIB har ekonomisk uppbackning, vore det bra att fortsätta vara medlem i Portico. Förutom att biblioteket slipper en egen arbetsinsats är det en bra försäkring inför framtiden och inför eventuella problem när det gäller digital bevaring av e-tidskrifter.

Källor

<http://www.hathitrust.org>

<http://www.lockss.org>

<http://www.clockss.org>

<http://www.portico.org/>

<http://www.ithaka.org/>

Reilly, B. F. J., & Waltz, M. E. (2014). Trustworthy data repositories. In J. M. Ray (Ed.), *Research data management* (pp. 109-125). West Lafayette, IN: Purdue University Press.

Rosenthal, D. S. H. (2014). Architectural choices in LOCKSS networks. *Library Hi Tech*, 32(1), 2-10. doi: 10.1108/LHT-09-2013-0120

Seadle, M. (2010). Archiving in the networked world: LOCKSS and national hosting. *Library Hi Tech*, 28(4), 710-717. doi: 10.1108/07378831011096321

Seadle, M. (2011a). Archiving in the networked world: by the numbers. *Library Hi Tech*, 29(1), 189-197. doi: 10.1108/07378831111117001

Seadle, M. (2011b). Archiving in the networked world: open access journals. *Library Hi Tech*, 29(2), 394-404. doi: 10.1108/07378831111138251

What's the difference between LOCKSS and CLOCKSS? [Internet]. (2008). Retrieved Aug 26, 2014, from <http://www.hurstassociates.com/ppt/LOCKSSCLOCKSSChart20608.pdf>