

Norsk Arkivråds Jubileumsseminar: Dokumenter og andre spor etter 22. juli

Gisle Hannemyr
Institutt for informatikk
27. oktober, 2011

22. Juli 2011. En historisk hendelse

- Etter avtale med Statsarkivaren i Trondheim vil Interkommunalt arkiv for Møre og Romsdal IKS ta vare på materiale frå vårt fylke. Dette omfattar alt utanom blomar og lys.
- Kondolanseprotokollane, andre skriftlege helsingar både på papir og elektronisk, bilde, teikningar og gjenstandar bør takast vare på.
- Det same gjeld materiale som har kome inn til kommunane, statlege institusjonar, kyrkjene, organisasjonar eller private, meiner Interkommunalt arkiv for Møre og Romsdal IKS.

Du ser: - Vi så slektningers skriftlige fortvilelse



Her samles folks hilsener etter 22. juli

Minnene etter Utøya ligger nå hos Riksarkivet. Bamses, kaniner, smykker og tegninger med varme hilsener ligger tett i tett.

Dokumenter og spor etter 22. juli

- NB: Jeg er ingen ekspert på arkiver eller arkivering.
- Jeg antar at grunnen til at jeg er invitert hit for å snakke i dag, er fordi dere *tror* jeg vet noe om elektroniske medier, og kanskje særlig om interpersonale medier (av noen også kalt "sosiale medier") på nettet.

Nettmedier



- Tradisjonelle medier finnes også i ulike inkarnasjoner på nettet (nettaviser som Aftenposten.no, Digi.no, ebok-utgivelser tilgjengelig gjennom bl.a. Bokskya.no).
- Nettmedier som er underlagt tradisjonelle medier, men som utnytter nettets egenskaper mht. engasjement og interaktivitet (VGdebatt.no, NRKbeta.no).
- Nettmedier som sjangermessig ligner på de omtalt over, men, men som har private utgivere (blogger, spesialnettsteder for ulike interesser).
- Interpersonale medier. Publiseringsflater stil til rådighet for enkeltpersoner av utgivere som kun bidrar med tilrettelegging, ikke innhold. (Underskog, Twitter, Facebook, Google+, Yahoo).

Nettmedier: Nye offentligheter

- Tradisjonell offentlighet: Aviser, tidsskrifter, bøker, radio og fjernsyn.
 - Arkiveres delvis av institusjonene selv.
 - Delvis av private aktører (*Retriever*).
 - Nasjonalbiblioteket (*pliktavlevering*).
- John Keane: *Structural transformations of the public sphere* (1995)

Helge Rønning (etter Keane)

- Den tradisjonelle offentligheten er i ferd med å oppløses:
 - Makro: De tradisjonelle redaktørstyrte mediene (også digitale utgaver)
 - Mezzo: Spesialnettsteder (*Foto.no*) og diverse blogger.
 - Mikro: Interpersonale medier (Facebook, Twitter, Google+, Yahoo).

Arkivering av disse nye offentlighetene

- Makro- og mezzo-offentlighetene er forholdsvis stabile og materiale fra disse kan samles inn ved hjelp av søkemotorteknologi.
- Å skille mellom innhold som tilhører "static web" og "deep web" kan by på tekniske utfordringer.
- Lovhjemmel for automatisk innsamling av materiale mangler (sannsynligvis).

Arkivering av disse nye offentlighetene

- Mikro-offentlighetene (Underskog.no, Twitter, Facebook, Google+, Yahoo, med flere) er ustabile, og det meste av materialet er ikke offentlig tilgjengelig, men krever en interpersonalt relasjon ("venn", "krets") for tilgang.
- Uoversikterlig (selv deltar jeg i minst 23 ulike interpersonale nettsted).
- Flyktige. Innhold forsvinner og endrer seg dynamisk.

Arkivlova

- **§ 1. *Føremål***
- Føremålet med denne lova er å tryggja arkiv som har monaleg kulturelt eller forskingsmessig verdi eller som inneheld rettsleg eller viktig forvaltningsmessig dokumentasjon, slik at desse kan verta tekne vare på og gjorde tilgjengelege for ettertida.

Private arkiv (§ 13)

- Riksarkivaren skal
 - a. halda oversyn over verneverdige privatarkiv.
 - b. føra register over privatarkiv som er tekne vare på av offentlege og private oppbevaringsinstitusjonar.
 - Riksarkivaren kan registrera visse privatarkiv som særskilt verneverdige. Melding om slik registrering skal sendast arkiveigaren.

Pliktavleveringslova

- **§ 1. *Føremål***
- Føremålet med denne lova er å tryggja avleveringa av dokument med allment tilgjengeleg informasjon til nasjonale samlingar, slik at desse vitnemåla om norsk kultur og samfunnsliv kan verta bevarte og gjorde tilgjengelege som kjeldemateriale for forskning og dokumentasjon.

§ 4. *Kva som skal avleverast*

- Dokument som er gjort tilgjengeleg for allmenta skal avleverast slik:
 - Dokument av papir eller papirliknande medium, mikroformer og fotografi i sju eksemplar.
 - Lydfesting, film, videogram, **edb-dokument** og kombinasjonar av desse dokumenttypane **i to eksemplar.**
 - Opptak av kringkastingsprogram i eitt eksemplar.

Tekniske utfordringer ved innsamling av materiale fra nettmedier

- Hva er en arkiveringsverdig enhet?
- Blanding av sjangre og formater (HTML, XML, flash, pdf, frames, inline).
- Hyperlenkede dokumenter kan spenne over mange domener.
- Dynamisk generert innhold.
- Webservere kan levere helt ulike versjoner avhengig av hvem som spør.

Nasjonalbibliotekets Paradigma

- Paradigma-prosjektet gikk i regi av Nasjonalbiblioteket 2001-2004.
- Målet var å finne gode rutiner for plikt-avlevering av digitalt materiale, bl.a. «høsting» av internettsider for langtidslagring.

Paradigma

- Søkerobot 'NEDLIB-harvester' var programmert for å samle inn materiale som befant seg på URLer fra .no (dvs. tekstfiler, bildefiler, lydfile osv.) med unntak av de URLene som ble etterfulgt av parametre (dvs. "deep web")
- Tok ikke hensyn til ROBOTS.TXT.
- Innsamlingen ble omkring 2003 utvidet slik at det også omfatter norske dokumenter som ligger i domener utenfor .no, f.eks. .com, .org, .net.
- Etter prøveprosjektet fortsatte dette som en ordinær del av Nasjonalbibliotekets oppgave.

Nasjonalbiblioteks tilnærming: Problemer

- Foregikk mye i "splendid isolation" – ingen særlig dialog med verken utgivere eller personer med kunnskaper om teknologien bak nettpublisering.
- Forsto ikke hvordan nettstedet behandler søkeroboter (for eksempel at en del nettsteder bruker en "robotfelle" for å nekte "uoppdragne" roboter tilgang).
- Forsto ikke at det meste av innholdet på moderne nettsteder er dynamisk generert ("deep web"), men at dette både kan og ikke kan reflekteres i URL (dvs. hvorvidt "URLene som er etterfulgt av parametre").
 - Når det gjelder å oppfylle det som er Nasjonalbibliotekets oppgave: Å sørge for at "desse vitnemåla om norsk kultur og samfunnsliv kan verta bevarte og gjorde tilgjengelege som kjeldemateriale for forskning og dokumentasjon." – så frykter jeg at oppgaven har blitt ivaretatt dårlig. Dvs. at det materialet som er innsamlet så langt både representerer en forholdsvis liten del av det som dette hadde vært relevant å samle inn, og denne delmengden er lite representativ for det som har blitt publisert i dette tidsrommet.

Hva bør gjøres?

1. Arkivarer og bibliotekarer bør bevisstgjøre seg i forhold til den langt mer fragmenterte offentlighet som vi har fått i dag.
2. Både Arkivlova og Pliktarkiveringslova bør revideres for å gi lovhjemmel til innsamling av materiale fra nettet.
3. Det bør opprettes en dialog med utgivere på alle tre nivåer (macro, mezzo og mikro), med sikte på å forankre innsamling og arkivering i samtykke fra utgivere.
4. (Automatisk) innsamling av elektronisk materiale bør skje i henhold til det som regnes som gode normer, og god skikk på nettet.