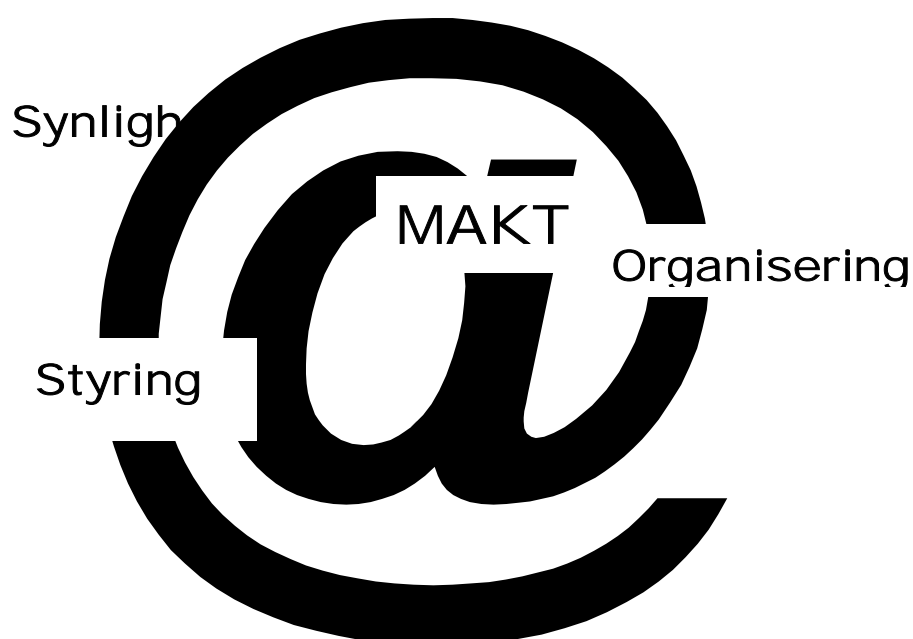

Universitetet i Oslo
Avdeling for forvaltningsinformatikk

DRI 1002 IKT og informasjonssøking – vår 2005

-
Tredje obligatoriske oppgave
-



-
Vi erklærer at alle i gruppen har deltatt aktivt i skrivingen av dokumentet og at besvarelsen er vårt eget produkt. Vi har ikke kopiert kildekode, tegninger, skjemaer eller andre deler av andres arbeid uten å kreditere kilde.

Aase Heien
Ida Martinussen
Sigrid Bæverfjord
Stig Hagelid Fjeldstad

Søkemotorer og rangering

Vi har valgt artikkelen: *How Search Engines Rank Web Pages*
<http://searchenginewatch.com/webmasters/article.php/2167961>

Beskrivelse og drøfting av artikkelen

Hvis vi sammenligner Internett med et bibliotek, (det er gode grunner for å hevde det motsatte, fordi et bibliotek inneholder litteratur som er valgt ut og anskaffet av noen) så er søkemotorene bibliotekarene¹. På denne måten kan man gå til bibliotekaren og oppgi et tema for å hjelp til å finne relevant litteratur. Her stanser likhetene mellom søkemotorer og bibliotekarer. Fordi mennesker kan tenke selv kan de stille spørsmål tilbake for å klargjøre og avgrense søket. Dette kan ikke søkemotorene. Her vil det være opp til brukeren å avgrense og eventuelt spørre på nytt i et mer avgrenset søk. Dersom man sier "hus" til en bibliotekar vil bibliotekaren stille oppfølgingsspørsmål for å avgrense temaet. Spør man en søkemotor om det samme får man listet alle kjente nettsider hvor dette ordet forekommer². Ingen bibliotekar vil gi deg alle bøkene som handler om hus. På denne måten har søkemotorene blitt svært opptatt av å tilby relevante treff for Internettbrukere som blir stadig mer utålmodige. Denne i utgangspunktet vanskelige øvelsen blir ikke lettere av at søkemotorene må ta vanskelige valg med tanke på bruk av sponsede treff og plassering av disse. Det er her søkemotorene må ta en del valg. Valgene man tar kan være avgjørende for butikken.

Artikkelen beskriver med bakgrunn i dette ulike teknikker forskjellige søkemotorer bruker for å indeksere og rangere treff. Den dokumenterer at det er ulikheter i teknologier og ulike løsninger som gjør at det samme søket gir ulike treff fra søkemotor til søkemotor. I tillegg omtaler den det menneskene bak avgjør; hvor stor plass betalt annonsering/plassering skal få og hvordan den skal plasseres i forhold til ikke betalte søk. Videre beskriver den hvordan vevsjefer tilpasser nettsider til søkemotorer og hva søkemotorer gjør for at denne optimaliseringen ikke går over grensen til manipulering av søkemotorenes søke- og indekseringsteknikker.

Vurdering av troverdighet

Artikkelen er lagt på Search Engine Watch (heretter SEW), et nettsted som kaller seg "the source for search engine marketing"(SEW 2005). Den er med andre ord rettet mot folk som markedsfører søkemotorer, Internettbaserte katalogtjenester og vevsjefer (Den delen som retter seg mot vevsjefer kalles "A Webmaster's Guide to Search Engines") som vil lære hvordan man bygger en side for at den skal være mest mulig synlig for søkemotorer. SEW publiserer omtaler, statistikk, produktlanseringer både innenfor kommersielle og tekniske aspekter og ulike former for kritikk og omtale av "search engine controversy". Mange av artiklene er holdt i en kritisk tone, mens mange artikler handler direkte om hvordan de ulike søkemotorenes teknologier for å rangere og indeksere treff og hvordan dette påvirker jobben til vevsjefer. Vi må imidlertid være klar over at SEW opererer i et marked der det er trangt om plassen, og er underlagt samme regler som alle andre i Internett-bransjen; utgifter kan ikke overstige inntjeningen. Dermed kan det oppstå påstander som hevder at man kjøper positive omtaler hos SEW. Vi tviler på at dette har skjedd, all den tid SEW har publisert flere kritiske artikler om de fleste søkemotorene.

¹ Det er visstnok grunnleggeren av WebCrawler, Brian Pinkerton, som først ble sitert på denne allegorien.

² Google.no gir 2,3 millioner treff på dette ordet.

Ulike temaer rundt søkemotorer er fra tid til annen gjenstand for offentlig debatt eller avisoppslag. Dette gjelder for eksempel oppslagene som kom da det ble offentlig kjent at når man søker etter ordene "miserable failure" på Google og trykker "jeg prøver lykken" så kommer man direkte til George W. Bush' offisielle biografi³, lagt ut av Det Hvide Hus. I dette tilfellet hadde SEW offentliggjort at dette også gjaldt Inktomi. SEW hadde også utført dette søket i forkant av disse oppslagene, og kunne dokumentere utviklingen til dette søket og redegjøre for årsakene til at fenomenet oppsto i tillegg til å presentere mer nøyaktig og utfyllende informasjon. Dette bidrar til å øke troverdigheten til SEW, fordi vi antar et nettsted som har eksistert lenge (siden 1995)(SEW [uten dato]) over tid bygger opp kjennskap til hvordan ulike søkemotorer virker og også har kontakter i bransjen og kan lettere verifisere og utdype informasjon. Dette er ikke en garanti for troverdig informasjon, men et indisium. Årsaken til at nettstedet oppsto er at personen bak nettstedet, Danny Sullivan, publiserte en side på Internett og gjorde ulike undersøkelser av hvordan søkemotorene fant den og hvordan de behandlet. Deretter eksperimenterte han med ulike nettsider og erfarte hvordan søkemotorene oppførte seg.

Vårt kanskje viktigste argument for at SEW er til å stole på er at foreleserne i DRI1002 har vist dit i forelesningene. Vi er også kjent med at tidsskriftet Bibliotekaren, tidsskriftet til Bibliotekarforbundet, tidligere har omtalt SEW.

Videre ser vi at siden oppdateres relativt ofte. Oppdateringshyppighet er viktig når man vurderer troverdighet. SEW har både nettdagbok og daglige nyhetsbrev. Vår erfaring er at hovedoppslagene fornyes relativt ofte. Vi ser også at ulike statistikker er av relativt ny dato. Artikkelen vi baserer oss på er imidlertid ikke ny. Den ble publisert i 2003. Teknologier og løsninger forandrer seg raskt i internettbransjen, og dette dokumenteres av stadig nye artikler publisert av SEW. Vi har imidlertid ingen grunn tro at artikkelen er utdatert, men mange mange av detaljene kan i dag være annerledes.

Kort om kildebruk

Det meste av omtalen av SEW er basert på erfaringer samlet gjennom lang tids bruk og er således et resultat av mange av artiklene som har vært publisert der. Videre har vi hentet stoff fra Googles sider om hvordan man kan tilpasse siden sin så den er synlig for Google. Rådene Google gir er i tråd med W3C-standarder og Google linker dit slik at man kan få mer informasjon og verifisere opplysningene. Til slutt supplerer vi med informasjon fra Digma, et lite konsulentbyrå som driver med rådgivning innenfor søkemotoroptimalisering og lignende. Vi ville normalt vært forsiktige med å bruke en slik kilde, men eier og eneste ansatte, og ansvarlig redaktør, er Håkon Styri, som benyttes som foreleser på DRI1002. Vi tillegger derfor informasjonen han publiserer relativt stor troverdighet.

Vårt eget nettsted⁴

Vi har valgt å begynne med generelle søkeord for å se hvor mange treff vi får. Deretter har vi brukt flere søkeord for å avgrense søket mer. I tilfellene søkte vi på norsk. Fordi søkemotorer ikke nødvendigvis forstår at "dri 1002" er norsk har vi brukt de norske speilsidene eller valgt søkespråk på norsk der det er mulig. Når flere ord står mellom " " innebærer det at vi søker etter eksakt frase – alle ordene skal forekomme i denne rekkefølgen. Der kun et ord står

³ Se blant andre Aftenposten(2003): *Bush er "ynkelig taper" hos Google*. Morgenutgaven, 08.12.2003.

⁴ Fordi vi ønsker å beholde obligatorisk oppgave nr 2 konfidensiell er den sperret for indeksering hos søkemotorene. Se vedlegg.

mellom “ ” har vi søkt etter alle sider hvor det forekommer. Vi har brukt ti treff som stoppunkt. Ved så få treff er lett å undersøke relevansen til alle manuelt.

Søk	Søketerm	Kommentar
1	”dri 1002”	
2	”dri 1002” ”oblig 2”	Antar at alle gruppene har disse ordene på siden sin
3	”dri 1002” ”obligatorisk oppgave”	Forkortelsen i forrige søk kan ekskludere de som ikke forkorter
4	”dri 1002” site: folk.uio.no	Sider på serveren for personlige hjemmesider for ansatte og studenter på UiO

Tabell 1 Våre søkeord

Google

Søk 1 gir 30 treff. Dette er for mange. Vi ønsker å redusere antall treff til under ti, slik at vi bare får en side med treff. Vi prøver å avgrense til sider med hvor fagets navn og ”oblig 2” forekommer, fordi vi antar at alle gruppene på kurset har nevnt dette i oppgaven vår. Søk 2 gir fire treff. Dette ble for spesifikt, og det er mulig at forkortelsen ”oblig 2” ikke brukes. Vi bytter ut ”oblig 2” med ”obligatorisk oppgave”. Søket gir fem treff, men de fleste er på sider vedlikeholdt av universitetet sentralt. Etter at disse søkene ikke ga mye utbytte skifter vi strategi og prøver et mer avansert søk, hvor vi angir serveren treffene skal ligge på. Fordi vi antar at alle gruppene legger oppgaven på webserveren for personlige brukere ved universitetet, avgrenser vi med søk 4 som gir åtte treff. Problemet med dette søket er nettopp antagelsen om at studentene utelukkende bruker universitetets server. Som vist over påvirker likevel ikke dette søket antall relevante treff i noen grad.

Yahoo

Det første søket vi gjør avslører en vanlig feil ved søkemotorene. Resultatsiden angir 22 treff, men når mange treff ligger på samme server vises ikke alle. Når man velger å vise også disse, viser det seg at det kun var 18 treff. Med søk 2 indekserer Yahoo 38 sider, men ved å vise de skjulte resultatene (som ligger på samme server) viser det seg det i realiteten er fire treff. Det er få nok treff i henhold vårt krav, men vår side er ikke blant dem. Når vi tar hensyn til antagelsen om at noen ikke bruker forkortelser, får vi latterlige to treff. Søk 2 gir dessuten bare treff på UiOs studiesider. Søk 3 gir tre treff.

Kvasir

Selv om Kvasir bruker Google som leverandør av fritekstsøk er Kvasir bedre på norske sider (Bertnes 2003). Vi prøver for sammenligningens skyld, men ikke med alle søkene vi planla å utføre.

Søk 1 gir syv treff. Uten familiefilter vises et treff i tillegg, mens søk 3 gir ti treff. Vår side er ikke blant dem. Disse interessante treffene gjør oss nysgjerrige. Hvordan kan et søk som gir 30 påståtte treff i Google bare gi åtte hos Kvasir? Vi mistenker at dette dreier seg om feil i forbindelse med telling av treff som ligger på samme server, og som utelates ved normalvisning. Man også tenke seg at litteraturen vi har er utdatert.

MSN

Microsoft har nylig laget sin egen robot. Man kan derfor anta at basen ikke er særlig omfattende såpass tidlig. Av kommersielle årsaker er det imidlertid risikabelt å lansere tjenester som ikke er ferdig. Folk kan bli skuffet og vende seg til andre søkemotorer

Våre treff er kanskje en illustrasjon. Søk 1 gir åtte treff, søk 2 og 3 ingen. Dette overraskende resultatet gjorde oss nysgjerrige. MSN har funnet veien til folk-serveren, og indekserer 54000 sider, men ingen dokumenter som inneholder ”dri 1002”. Google finner 185000 dokumenter på folk, mens UiO sin interne søkemotor (teknologien leveres av Inktomi) gir 116670 treff.

Alle søk på MSN filtreres gjennom et innholdsfilter. Det går ikke klart frem hva slags innhold som filtreres ut, men det dreier seg om innhold som umyndige ikke skal se. Vi måtte love pent og hellig med kors på halsen at vi er myndige.

Hvordan bli mer synlig?

Vi vil i denne delen beskrive hvordan man kan gjøre det lettere for søkemotorene å finne og rangere en side.

Det er mange grunner til at webredaktører vil gjøre siden sin lettere tilgjengelig for søkemotorene. De fleste kommersielle nettsteder selger reklameplass og får som regel betalt på to måter: Enten per visning(hvor mange har sett annonsen) og per klikk(antallet som fulgte linken som annonsen inneholdt). I dag er det felles standarder for hvordan antall visninger telles. Antall klikk måles enten ved å lage sider som bare vises for de som fulgte linken(klikket) eller ved å brukes webstatistikken og se hvor mange referers som kommer fra siden hvor annonsen sto.

Ikke-kommersielle vil ha noen av de samme betraktningene. Det er riktignok ikke slik at man får betalt per treff, men for å forsvare ressursbruken er det nødvendig med et visst antall treff. Ikke-kommersielle nettsteder vil derfor ta i bruk mange av de samme virkemidlene som kommersielle for å få flest mulig treff, skape lojalitet/avhengighet og få flest mulig av brukerne til å bruke mest mulig tid/ besøke flest mulig sider på nettstedet.

Søkemotoroptimalisering

Hvordan man blir mer synlig for søkemotorene er med andre et spørsmål om vinn eller forsvinn for mange vevsjefer. En avansert og omfattende nettside er svært dyr både i oppstarts- og driftsfasen, se bare på de mange dot-com selskapene som tok sin død etter å ha opprettet en nettside, men ikke klart å få verken trafikk eller annonseinntekter. I dag har de fleste innsett at antall treff er veldig viktig, og annonsesalg er viktigst, men de to er avhengig av hverandre. I denne delen skal vi omtale ulike måter å bedrive søkemotoroptimalisering. Selv en amatør kan søkemotoroptimalisere svært mye, uten å risikere å bli anklaget for juks fra søkemotorenes side og uten å bruke dyre konsulentfirmaer.

Synlighet for en nettside styres på flere nivåer. Det er flere tiltak vevsjefer og innholdsleverandører kan ta for å gjøre siden sin mer søkbar. Som vi har vist i tidligere obligatoriske innleveringer har kampen om treff og sidevisninger bidratt til en hel industri av firmaer som selger råd om hvordan man får treff. Noen av disse går også så langt som til å garantere treff og hevde et ”special relationship”(Google 2004) til de store søkemotorene, slik at man får en god plassering hos de store søkemotorene.

Enhver som vedlikeholder en nettside har flere veier til selv å gjøre siden synlig og lettleselig for søkemotorene, uten å bruke omstridte metoder søkemotorene ser på som juks.

Meta-informasjon

Metainformasjon om en nettside er informasjon som ikke er direkte synlig gjennom en nettleser (Bing 2001), men inneholder beskjeder om hvordan innholdet skal vises og behandles og brukes både av nettleser og søkemotorer. På denne måten kan man gi direkte beskjed til søkemotor om hva siden inneholder, både i form av stikkord og beskrivelse. Dette gjør det lettere å oppnå relevante treff, fordi man selv kan definere hvilke søkeord som skal gi treff på siden. I tillegg kan man eksplisitt be besøkende søkemotorer om å indeksere denne siden og følge linkene på siden, eller om å la det være. Videre kan man bruke HTML-koder og innhold aktivt. Flere av de store søkemotorene prøver dessuten å gjette seg frem til hvor i dokumentet det står ord som kan beskrive innholdet på en god måte, og antar at vevsjefer bruker overskriftstiler i HTML-standarden. Dermed kan fornuftig bruk av H1, H2 osv gjøre siden lettere søkbar (Digme 2002), i den forstand at man kan oppnå høyere plassering der søkeordet stemte med ord i dokumentet du la ut.

Her er jeg!

Alle disse måtene er passive og server-side. Uansett hvor hardt man jobber med å gjøre seg selv synlig må man pent vente på at en søkemotor banker på døren til din <HEAD>-tag. De store søkemotorene tilbyr i tillegg en add-url funksjon, hvor vevsjefer tilbys å legge adressen til sin side inn i en interessedatabase, en oversikt søkemotorene vedlikeholder frem til de har mulighet til å besøke. Det er ikke gitt hvor fort, eller i det hele tatt *om* søkemotorene kommer, men det er den eneste aktive måten vevsjefer har til disp for å fortelle at siden fins (Google 2004).

Den siste måten, relativt utbredt og med mange ulike løsninger avhengig av søkemotor, innebærer å betale seg til treff (Pay for Placement (P4P)) (Digme 2002). For oss er dette av ulike grunner uaktuelt, men det er altså en mulighet. Det er ganske vanlig å kjøpe gode plasseringer hos søkemotorer, men er nok mer kontroversielt i dag enn bare få år siden. Det har vært mye oppmerksomhet rundt dette temaet, fordi det ikke alltid har vært enkelt å skille betalte fra ikke-betalte treff.

Internett og makt

En definisjon på anarki hentet fra Øyvind Østerud (1997) er: "Fravær av makt, mangel på styring". Videre sier beskrivelsen at begrepet gjerne blir brukt nedsettende om forhold man synes er preget av lovløshet eller kaos. Elisabeth Staksrud skriver at i begynnelsen på 90-tallet hadde de aller fleste av de *vestlige* landene en slags laissez faire-holdning til innholdet på Internett (Staksrud 2002). Laissez faire kan bety noe a la "la det skure". Med laissez faire, slik Staksrud benytter begrepet, kan vi anta det ikke var fokus fra myndigheter eller andre på hva som ble publisert på Internett og hva man kunne finne av informasjon. Det var "fritt frem", og det var foreløpig ingen veldig bevissthet rundt innholdet. Siden har det derimot vært en voldsom økning i bruken av og tilgang til Internett, og dette har ført til økt fokus på hva som er tilgjengelig. Staksrud viser til at "mange lands myndigheter opplevde press fra ulike samfunnsgrupper som var bekymret for hva de så som skadelig flyt av pornografi, vold, og annen type uønsket informasjon." (Staksrud 2002:68).

Internett er betegnelsen på mange ulike individuelle nettverk som er koblet sammen. Dette verdensomspennende kommunikasjonssystemet er bygd opp av flere lag (Hannemyr 1999). OSI-modellen er en teoretisk tilnærming på hvordan nettverksapplikasjoner og nettverksprotokoller fungerer sammen (Linuxguiden 2005). Internett fungerer omtrent på samme måte. De to nederste lagene består av fysiske kabler, sendere og annet teknisk utstyr.

Vi kommer ikke til å beskrive disse lagene videre i oppgaven. Lag tre og fire tar seg av nettverk og transport, mens i lag fem og seks finner vi tjenestene (Jansen, 2005). Baktanken med denne lagdelingen er at bestemte funksjoner kan isoleres i det bestemte laget. På den måten blir det lettere å holde oversikt, forstå og vedlikeholde de ulike delene ved at de er avgrenset. En viktig konsekvens er at nettet fungerer mer stabilt.

	Funksjon	Eksempel på protokoll/ Standard	Kontrolltiltak	Aktører
Lag 5-7	Tjenester	http, ftp, ulike www-standarder	Filtrere innhold	Myndigheter Teleselskaper Innholdslev.
Lag 3-4	Nettverk/transport	TCP/IP	Brannmurer, filtre	Myndigheter Teleselskaper ISP-er
Lag 1-2	Fysisk infrastruktur	Ethernet, ADSL	Fysisk kontroll av kabler, backbone servere	Myndigheter Teleoperatører

Tabell 2 OSI-modellen. Lagdeling av Internett. Kilder: Se tekst.

Hva skjer i de ulike lagene?

I lag tre til fire finner vi henholdsvis Internett protokoll (IP) og transmission communication protocol (TCP). Ved hjelp av TCP og IP kan ulike maskiner kommunisere med hverandre. For at dette skal være mulig trenger de ulike nettene ha en port eller gateway. Dette kan for eksempel være en større maskin som sørger for overføringen mellom informasjon inn og ut av nettet. De ulike portene benytter TCP for å sende informasjon (Liestøl og Rasmussen 2003). TCP sender informasjonen ved å dele den inn i ut ulike selvstendige enheter i form av pakker. Det finnes ingen egen datalinje mellom mottaker og avsender. Hver pakke sendes for seg selv fra punkt til punkt i nettet. Dersom det oppstår en feil ved at en pakke ødelegges eller forsvinner, holder det med å sende det ene elementet en gang til. Denne formen for kommunikasjons utveksling omtales som pakkesvitsjing (Hannemyr 1999). For at de ulike pakkene skal vite hvem som er mottakeren, trenger de en adresse. Alle maskiner som er knyttet opp mot Internett er registrert med en unik IP-adresse (Hannemyr 1998).

I de tre øverste lagene finner vi ulike tjenester. Her er det snakk om menneske til menneske kommunikasjon i form av godkjenning av tekst, bilder og lyd med mer. I lag fem finner vi sesjonen som kommuniserer mellom prosessene i de ulike maskinene. Lag seks består av presentasjon, og fungerer som en oversetter for data til nettverket. Disse dataene blir sendt av applikasjonslaget, som er det øverste laget og fungerer som et vindu for brukerne og applikasjon prosessen for å få adgang til nettverkstjenester (Microsoft Press 1996).

Anarki eller styring på Internett?

På begynnelsen av 1990-tallet var det som nevnt litt "la det skure"-holdning til innholdet på Internett. Vi antar at man stort sett kunne publisere hva man ville; de ulike vestlige lands myndigheter hadde ikke helt styring over hva Internett inneholdt. I hvert fall ikke i den vestlige verden var det ennå noe form for innholdsfiltrering på ISP-nivå eller tilbud om "KidSurf" eller lignende kommersielle programmer for sensurering. Vi kan anta at de nasjonale lovene gjaldt like fullt da også – men hadde noen styring på hva og hvem som publiserte?

De tekniske aspektene ved Internett ble i utgangspunktet utarbeidet for å sikre full åpenhet og tilgang til informasjon. Hensikten med world wide web var å lage et standardisert brukergrensesnitt mot store mengder lagret informasjon. I følge Berners-Lee var en av baktankene ved Internett at det skulle være like lett å legge ut informasjon, som å hente den. Videre var det en forutsetning at de ulike lagene og innholdet i det øverste lagene i OSI-modellen måtte være uavhengige av hverandre. Dette skulle føre til at ingen kan sette tekniske hindre som er i strid med lover og normer for å forhindre tilgjengelighet av innholdet (Liestøl og Rasmussen 2003).

Berners-Lee og hans W3C samt andre organisasjoner jobber for å fremme de mål nevnt over. Dette for å standardisere retningslinjer og protokoller for Internett. Dermed er ikke Internetts arkitektur overlatt til tilfeldighetene eller at alle jobber mot "sin egen" standard. Som Liestøl og Rasmussen peker på, kan konsekvensen være at de store organisasjonene får for mye kontroll. De fleste av oss er ikke så interessert i jobben de gjør, og har ikke nok kunnskap til å engasjere oss. Dermed blir beslutningene overlatt til noen få – "kunnskap er makt".

Det som fortsatt er det samme i dag som på begynnelsen av 90-tallet, er at man til alle døgnets tider kan publisere hva man vil på Internett, og man kan finne informasjon om hva som helst. Internett er internasjonalt, men det finnes ingen internasjonale lover som kontrollerer nettet. I disse dager pågår det en debatt om dette. De ulike lands nasjonale lover gjelder, og de blir i noen land håndhevet i form av innholdsfilter og brannmurer. Dette kan begrense anarki på internett, men i noen land som for eksempel Norge kan man publisere og laste ned det man ønsker. Dette avhenger av kunnskap, nettleverandører og om man overholder lov og rett. Videre er det viktig å understreke at kontrollmengden er så enorm at det er umulig å vite hva "alle" gjør.

Hvem bestemmer hva på Internett

Over har vi sett på OSI-modellen som viser hvor og hvem som kan sensurere på Internett, og hvilke kontrolltiltak som finnes på hvert av lagene. Lag en og to kan kontrolleres med fysiske linjer i form av tilknytning. Aktørene er teleoperatører og myndighetene i form av lovgivning og kontroll tiltak. De to neste lagene kontrolleres med tilgangskontroll og tjenestenekt, i form av egenkontroll eller som pålegg fra myndighetene. I lag fem-sju finner vi aktører som myndigheter, innholdsleverandører etc. Myndigheter kan, i følge modellen, være aktør i alle lag. Vi skal under se på myndigheter og andre grupper som kan sies å ha kontroll over Internett – og hvilke kontrolltiltak de kan benytte.

De ulike lands *myndigheter* har som nevnt over et ønske (i forskjellig grad) om å kontrollere informasjon på Internett. Dette kan være kontroll med nettsider som omhandler ulovlig/uønsket innhold. Liestøl og Rasmussen skriver at "Generelt gjelder prinsippet om at det som er ulovlig i samfunnet, er ulovlig på nettet" (Liestøl og Rasmussen 2003:127). Ut i fra dette kan vi altså si at myndigheter forsøker å utøve kontroll av informasjon på Internett ved hjelp av lover. Myndighetene har også mulighet, ved for eksempel kampanjer, og oppfordre folk til "fornuftig" bruk av Internett, og også gi andre typer veiledning I denne omgang tar vi for oss eksempler på lover som kan gi myndigheter kontroll over bruken av Internett. Lovene kan ses på som et maktmiddel. Som nevnt over er det de nasjonale lovene vi må se på, da det ikke finnes noen internasjonale lover for bruk av Internett (men EU har gitt direktiver på noen områder, for eksempel angående elektronisk handel (Liestøl og Rasmussen 2003)).

Som vi også beskrev i obligatorisk oppgave nr. 1 regulerer flere paragrafer i den norske straffeloven også innholdet på Internett. Et eksempel er salg av pornografi (Liestøl og Rasmussen 2003).

Åndsverkloven er til beskyttelse for de som skaper et åndsverk; "Den som skaper et åndsverk, har opphavsrett til verket" jfr. Åndsverkloven §1. første ledd. Videre sier §1 annet ledd at "med åndsverk forstås i denne lov litterære, vitenskapelige eller kunstneriske verk av enhver art og uansett uttrykksmåte og uttrykksform". Eksempler listes opp, som musikkverk og datamaskinprogrammer. Den som har rett til å benytte et dataprogram, har lisens, kan lage sikkerhetskopier i den grad det er påkrevet for å få benyttet programmet (Liestøl og Rasmussen 2003 og jfr. Åndsverkloven §39h). Nå er et forslag til endring av åndsverkloven på trappene, og vi kan i Ot.prp. nr. 46 "Om lov om endringer i åndsverkloven m.m." lese blant annet at adgangen til å kopiere for eksempel musikk til privat bruk "foreslås videreført". Men dette gjelder ikke om man må omgå tekniske beskyttelsessystemer for å få kopiert. Tekniske beskyttelsessystemer beskrives av Liestøl og Rasmussen (2003:147) som DRM (Digital Rights Management Systems) og er "ulike tekniske løsninger for å hindre kopiering eller avspilling av digitalt innhold som programmer...".

Andre lover kan sies å trygge vår mulighet til å ytre oss. Retten til å ytre seg finner vi i Grunnlovens § 100 og gjennom De Internasjonale menneskerettighetene. Vi finner dette beskrevet i blant annet Den europeiske menneskerettskonvensjon av 1950 (EMK), art 10.

Det er også flere organisasjoner som jobber med ulike sider ved Internett, og som forsøker å påvirke og opplyse både myndigheter og brukere av Internett generelt. Enkelte av de utvikler standarder for Internett – som vi mener kan være med på å gi de makt.

W3C er som nevnt en organisasjon som har Internett som område. De har Tim Berners-Lee som direktør, han er best kjent som WWWs "far". World Wide Web Consortium sin misjon "er å lede world wide web til sitt fulle potensial ved å utvikle protokoller og retningslinjer som sikrer langvarig vekst for nettet" (W3C 2005a). De mener at for å få nettet til å nå sitt fulle potensial, må de mest grunnleggende nett-teknologiene være kompatible med hverandre og la all software og hardware som brukes til å aksessere nettet med jobbe sammen. Ved å publisere åpne standarder for nettspråk og nettprotokoller, ønsker organisasjonen å unngå markedsoppdeling og videre at nettet eventuelt deles opp fysisk (ibid.). W3C støttes av medlemmer, forskningsstipender og andre offentlige og privat kilder. De utvikler altså standarder for Internett, og W3C samarbeider med eksperter og andre organisasjoner for å videreutvikle teknologien. Nettsidene deres sier at medlemskap er tilgjengelig for alle organisasjoner – men er særlig aktuelt hvis din organisasjon jobber legger mye ressurser i å utvikle teknologier for nettet (W3C 2005b). Som medlem kan man være med å påvirke, og medlemslisten deres er offentlig og ligger publisert på nettsiden. Eksempler på medlemsorganisasjoner er America Online Inc og American Express (W3C 2005c).

ISOC (The Internet SOCIety) er en tredje organisasjon som kan nevnes. ISOC arbeider med temaer som gjelder Internettets fremtid. Den er moder-organisasjonen for gruppene som er ansvarlig for Internettets infrastruktur standarder, som for eksempel Internet Architecture Board (IAB). Organisasjonen har både andre organisasjoner og enkeltindivider som medlemmer, og det er slik den finansieres (ISOC 2005). "ISOC har ingen formell makt, men anbefaler åpne standarder etter lange diskusjoner" (Liestøl og Rasmussen 2003:34).

IETF (The Internet Engineering Task Force) er en organisasjon som er åpen for alle som er interessert, i følge deres nettsider. De sier om seg selv at de er ”et stort åpent internasjonalt samfunn av nettverksdesignere, forskere etc. som er opptatt av utviklingen av Internettets arkitektur” (IETF 2005).

Den siste organisasjonen vi nevner er ICANN (Internet Corporation for Assigned Names and Numbers). Kort fortalt er ICANN et teknisk koordineringsorgan for Internett, opprettet i 1998. ICANN koordinerer fordelingen av IP-adresser, Internett domene navn og protokoll parametre og port numre. ICANN er en non-profit organisasjon blant annet dedikert til å opprettholde den operasjonelle stabiliteten til Internett (ICANN 2005).

Ser vi på organisasjonene over, gir de inntrykk av at ”alle” kan bli medlemmer av disse organisasjonene, og kan vi da si at de har noen egen makt over Internett? Hva slags type maktmidler har de i så fall? IETF er en åpen organisasjon, ISOC har også enkeltmedlemmer mens hos W3C kan organisasjoner være medlemmer. Så det er for eksempel mulig å melde seg inn i IETF og på denne måten være med å påvirke gjennom e-post og diskusjonsgrupper. Noen kan sikkert det – men på langt nær alle. Har organisasjonene et maktmiddel så er det ”kunnskap er makt”. De fleste av oss har overhodet ikke nok kompetanse til å være med på å bestemme over Internettets standarder eller protokoller. Så kanskje vi bare skal være glade for at noen orker å engasjere seg i dette for oss?

Det er mulig via de ulike organisasjonenes nettsider og lese om hva de jobber med, noen legger ut møtereferater etc. På denne måten har man mulighet til å kontrollere hva de gjør til en viss grad. Og målet til flere av disse er å jobbe mot *åpne* standarder og retningslinjer for alle. Men disse og andre organisasjoner er ikke etablert av myndigheter eller demokratisk valgt, de er organisasjoner som har valgt seg selv på grunnlag av en interesse for Internett. Men hvis det ble besluttet at noen av de viktigste standardene etc. skulle det stemmes over – alle kunne stemme - ville folk hatt nok kunnskap?

Vi kan også tenke oss at organisasjonene er ”vaktbikkjer” overfor hverandre og de store kommersielle selskapene som kan tenkes å ønske mer makt over nettet. Et problemet for oss ”vanlige” nettbrukere er at det er mange organisasjoner som jobber med Internett, og selv om man har lyst til å finne mer ut om hva de jobber med, kan det være vanskelig – hvem jobber med hva og hvordan kan en vite at man får all nødvendig info via nettsidene deres?

For å igjen komme tilbake til lagdelingsmodellen – på lag 3-4 finner vi aktører som ISP-er. ISP står for Internet Service Provider og alle firmaer som tilbyr internettilgang til oss (hjemme eller på jobben) er såkalte ISP-er. Dette er en gruppe som kan sies å i hvert fall ha *muligheten* til å utøve kontroll på Internett (i betydning den informasjon vi har tilgang til). I Norge samarbeider for eksempel ISP-en Telenor med Kripos ved utvikling av barnepornofilter (beskrevet i vår obligatoriske oppgave nr. 1), og dette filteret kan gjøre at også andre sider blir blokkert ut (men vi vet ikke hvilke). I noen land er ”ISP-ene forpliktet til å drive aktiv kontroll av innholdet på sine tjenermaskiner” (Liestøl og Rasmussen 2003:127). I Norge er det ikke slik, men muligens så er Telenors barnepornofilter en sped start på mer sensurering av Internett?

Brukere og misbrukere

Vi som brukere av Internett er en gruppe som ikke nødvendigvis hver for oss har kontroll over Internett, men vi alle er med på å legge ut informasjon (og med det er med på å bestemme hva andre får tilgang til) og vi henter ned informasjon vi ønsker å lese. De nye digitale mediene

har ført til at vi som brukere har flere valgmuligheter til å finne kunnskap. Som brukere har vi fått flere informasjonskanaler og det forandrer hvordan vi forholder oss til den informasjonen vi tilegner oss. På den annen side har de digitale mediene skap et stort rom for at brukeren kan manipulere informasjon, og derfor styrer vi selv hvordan vi søker informasjon og hvilken mening vi tillegger den (Liestøl og Rasmussen 2003). Tilbud og etterspørsel gjelder her som på så mange andre områder.

Internettet er svært sårbar for sabotasje og ødeleggelse. Alvorligheten av dette fikk verden føle i 1988, da studenten Robert Morris jr laget et program som overbelaster de maskinene som det ble kjørt på. Programmet var også programmert slik at det kopierte seg selv til andre maskiner som var koblet opp til nettet. Resultatet av dette programmet, og ikke minst frykten for det førte til at hele Internettet kollapset og var nede i flere dager.

En konsekvens av denne hendelsen ble Computer Emergency Response Team (CERT) etablert. Deres formål er å forhindre ødeleggelse av nettet i form preventive tiltak, i tillegg har de ansvar for logistikk under krisesituasjoner (Hannemyr 1999).

Problemene rundt spam og virus kan virke som et problem som er kommet for å bli. Det er viktig å ikke overse "fordelene" av slik sabotasje, fordi det fører til utvikling av nye sikkerhetsløsninger.

Internett @UiO

Vi har valgt å se på hvordan fellesadministrasjonen og fakulteter ved Universitetet i Oslo utnytter de mulighetene IKT og spesielt Internett har å by på sett fra en students ståsted.

En organisasjon kan ifølge Liestøl og Rasmussen (2003:155) beskrives som den samlede kommunikasjon mellom medlemmene. Det vil si at hvordan og hvor kommunikasjonen finner sted vil være selve smøringen for en organisasjons maskineri, jo bedre olje jo mindre motorproblemer. Med fokus på nettopp denne viktigheten av lett tilgjengelige kommunikasjonsfora, ser man hvordan Internett har revolusjonert samfunnets oppfatning av god kommunikasjon. E-post har erstattet brev, world wide web har overtatt funksjonen til oppslagstavler og rundskriv, og hvorfor ringe når man kan chatte online?

Majoriteten av alle organisasjoner har omfavnet den tekniske revolusjonen innenfor tekstbehandling og kommunikasjon. Dette har økt effektiviteten både ved intern og ekstern kommunikasjon, og har dermed frigjort ressurser som kan brukes til andre formål. Samtidig har IKT og Internetts inntog banet vei for nye rutiner og regler for kommunikasjon og informasjonshåndtering (Liestøl og Rasmussen 2003:155). I en organisasjons travle hverdag praktiseres utnyttelsen av Internetts mange muligheter på vidt forskjellige måter, som alt fra kundebehandling på e-post til spredning av intern informasjon på intranettet. Et intranett kan beskrives som et lokalnett, og er ofte passordbeskyttet. Dermed faller intranettet inn under Bertnes' (2003:134) karakteristik av "den skjulte weben"

Studenter og ansatte ved Universitetet i Oslo har også adgang til intranett, et eksempel er Classfronter som fungerer som en nettbasert kommunikasjonskanal mellom elever og lærer. Classfronter er ikke åpen for allmennheten, startsidene fungerer som en innloggingsportal hvor man blir bedt om å oppgi brukernavn og passord. Ved hjelp av Classfronter kan man blant annet sende e-post, levere oppgaver og få løpende informasjon angående fag og studier. Dette kan i realiteten erstatte gamle løsninger, som personlig oppmøte, innlevering av oppgaver i faglærers hylle og behovet for fysisk betjente informasjonsskranke. Ikke alle fakultet og fag

er brukere av Classfronter i praksis, så mulighetene programmet tilbyr er kun teoretisk tilgjengelig for mange studenter.

Et annet lukket kommunikasjonsforum er Studentweb, som også er en del av den skjulte weben. Studentweb åpner for elektronisk eksamensopp- og avmelding, betaling av semesteravgift og oversikt over løpende studieplan. Resultatet er reduksjon av papirbaserte skjemaer, og muligheten til å kunne endre sin studieplan fortløpende uten noe fysisk mellomledd. I likhet med Classfronter frigjør Studentweb ressurser, og forenkler hverdagen for studentene ved at de fra sin hjemme-pc kan ha tilnærmet full kontroll over sitt studieforløp. En annen hendig tjeneste for studenter er ”Mine studier”, en side skreddersydd etter den enkeltes studieplan med oppdaterte beskjeder fra de enkelte fag. Denne siden er passordbeskyttet og personlig.

De oppdaterte informasjonssidene på nettet letter hverdagen for studentene ved at de kan kommunisere via sin hjemme-pc, og sådan være elektronisk til stede på skolen hele døgnet. Skulle den allerede produserte informasjonen vise seg å ikke være tilstrekkelig, har man muligheten til å sende en e-post til fakultetet eller faglærer. Strømbrevets gjennombrudd som kommunikasjonsmiddel har myket opp gamle hierarkistiske kommunikasjonskanaler både horisontalt og vertikalt, på grunn av sin uformelle og beskjedne form (Liestøl og Rasmussen 2003:156). En e-post banker ikke på døra, den lager ikke høy ringelyd, og den krever ikke å bli besvart i det sekundet den når mottaker. Hvis man kvier seg for å oppsøke en professors kontor, kan den perfekte løsningen være å sende et skjermbrev.

Det er ikke bare e-post og de lukkede intranettene som har forbedret kommunikasjonen i organisasjonen Universitetet i Oslo. Ved å logge seg inn på universitetets hjemmesider kan man hente ut relevant og sikker informasjon om studier, de forskjellige emnene og aktuelle saker som angår studenter og ansatte. Internett har erstattet mye av den papirbaserte informasjonen vi var vant til å forholde oss til, og dermed gjort Internett til et viktig og uunnngåelig verktøy i hverdagen. Universitetet har på alle plan i organisasjonen valgt å benytte seg av de muligheter Internett og IKT bringer med seg. Et godt eksempel er selve undervisningen, hvor foreleser ofte legger ut notater på nettet til timen på forhånd slik at studenten kan nyttiggjøre seg av notatene både før, under og etter forelesningen. Forelesers presentasjon er som regel basert på PowerPoint, som vedkommende enkelt kan hente fram fra hjemmeområdet sitt eller medbrakt minnekapasitet ved å bruke en stasjonær eller bærbar pc.

Med utgangspunkt i Internetts dominerende rolle i studenter og ansattes hverdag, ser man viktigheten av pc-tilgang og ikke minst pc-kunnskap for å kunne ta del i kommunikasjonen innad i organisasjonen. Uten tilgang til pc eller kunnskap om IKT Internett faller man lett utenfor, og det vil være tilnærmet umulig å følge et studium. Likevel vil vel de fleste konkludere med at datarevolusjonen er opphavet til langt flere fordeler enn ulemper, i en hverdag som både er sulten på informasjon og avhengig av kommunikasjon.

Referanser

Bertnes, Pål (2003): *Faglig Informasjon på Internett. Kvalitet og Kildekritikk*. Abstrakt Forlag, Oslo

Bing, J(2001) Bing, Jon: *Meta-informasjon og Internett: Frem med stigen. I: Bygrave, L.A: Yulex 2001*. Unipub, Oslo

- Digme(2002): *Nettmarkedsføring og søketjenester*. URL: <http://www.digme.no/tips/marketing.html>. Datert 20.12.2002. Uttaksdato 17.03.2005
- Digme(2003): *Annonser på nettet*. <http://www.digme.no/ehandel/nettreklame/annonser.html>. Datert 01.05.2003. Uttaksdato 17.03.2005
- Google(2004a): *Google Information for Webmasters. Search Engine Optimizers*. <http://www.google.no/intl/nn/webmasters/seo.html>. Uttaksdato: 18.03.2005
- Google(2004b): *How Do I Get My Site Listed on Google?* <http://www.google.no/intl/nn/webmasters/1.html> Uttaksdato 17.03.2005
- Hannemyr, Gisle (1998): *Nettverksarkitektur*. Artikkel til kurset IN-INT, Universitetet i Oslo.
- Hannemyr, Gisle (1999): *Begynnelsen på en historie om Internett I*: Braa, K et al(1999): *Netts@mfunn*. Tano Aschehoug, Oslo
- Internet Corporation for Assigned Names and Numbers(2005): *Welcome to ICANN* <http://www.icann.org/new.html>. Uttaksdato: 18.03.2005
- Jansen, A (2005): *IKT og organisering*. Forelesning i DRI 1002. 07.03 2005. Avdeling for forvaltningsinformatikk. Universitet i Oslo
- Liestøl, G. og T. Rasmussen(2003): *Digitale medier*. Universitetsforlaget, Oslo.
- Linuxguiden.no(2005): *OSI-modellen*. <http://www.linuxguiden.no/index.php/OSI-modellen> . Uttakdato: 18.03.2005
- Microsoft Press (1996): *Networking Guide. Technical Information and tools for the support Professional*. Washington
- Ot.prp. nr. 46 (2004–2005): Om lov om endringer i åndsverkloven m.m.*
- Search Engine Watch[uten dato]: *About Search Engine Watch*. <http://searchenginewatch.com/about> . Uttaksdato 17.03.2005
- Search Engine Watch[uten dato]: *Hovedoppslag*. <http://searchenginewatch.com/>. Uttaksdato 17.03.2005
- Staksrud, E.(2002): *Ytringsfrihet og sensur på Internett, I*: T. Slaatta (red) *Digital makt. Informasjons- og kommunikasjonsteknologiens betydning og muligheter*, Gyldendal Akademisk, Oslo
- The Internet Engineering Task Force(2005): *"Overview of the IETF"* <http://www.ietf.org/overview.html> Uttaksdato: 16.03.2005

The Internet SOCIety(2005): *All about the Internet Society*.
<http://www.isoc.org/isoc/>. Uttaksdato: 13.03.2005

World Wide Web Consortium(2005a): *About the world wide web consortium*
<http://www.w3.org/Consortium/>. Uttaksdato: 13.03.2005

World Wide Web Consortium(2005b): "About W3C Membership"
<http://www.w3.org/Consortium/membership.html> Uttaksdato: 16.03.2005

World Wide Web Consortium(2005c): *World Wide Web Consortium (W3C) Members*
<http://www.w3.org/Consortium/Member/List> Uttaksdato: 16.03.2005

Østerud, Øyvind: *Statsvitenskapelig leksikon*, 1997. Universitetsforlaget, Oslo.

Åndsverkloven. *Lov om opphavsrett til åndsverk m.v.* av 12 mai 1961 nr. 2.

Vedlegg

Vi har sperret alle søkemotorer ute gjennom en robots.txt fil og med noindex og nofollow i metainformasjonen, Når ingen søkemotorer finner den har vi altså lyktes. Vi gjengir først robots.txt og deretter meta-kildekoden som ligger på sidene vi har benyttet.

Robots.txt

robots.txt for <http://folk.uio.no/stfjelds/>

```
User-agent: *  
Disallow: /dri/ # This is an infinite virtual URL space  
Disallow: /dri/index.html/  
Disallow: /dri/oblig2.doc  
Disallow: /dri/oblig2.pdf  
Disallow: /dri/*
```

Forklaring:

User-agent kan spesifisere søkemotorer som tillates å se siden. Linjene som starter med disallow spesifiserer hvilke enkeltsider søkemotorene ikke kan se. Denne filen er således smør på flesk, fordi Disallow: /dri/ sperrer hele mappen og alle filene i den.

Meta-kildekode

```
<META NAME="ROBOTS" CONTENT="NOINDEX, NOFOLLOW">
```

Forklaring:

Noindex sier til søkemotorene at siden ikke skal vises i et søkeresultat. *Nofollow* forbyr søkemotorenes crawler å følge linkene på siden.