

SKUP Metoderapport

1. Journalist

Anders Brenna

2. Prosjektnavn

Gravemaskinen.no: Kulturdepartementet somler med journalføringen

3. Publiseringer

Kulturdepartementet somler med journalføringen

<http://blogg.abrenna.com/kulturdepartementet-somler-med-journalf%C3%B8ringen/>

12. oktober 2010

4. Redaksjon

blogg.abrenna.com

v/Anders Brenna
Solhaugveien 77
1337 Sandvika

Mobil: 90077860
abrenna@gmail.com
twitter.com/abrenna
blogg.abrenna.com

Gravemaskinen AS

PB 227
1300 Sandvika

abrenna@gravemaskinen.no
gravemaskinen.no

5. Kontaktinfo

Anders Brenna
Solhaugveien 77
1337 Sandvika

Mobil: 90077860
abrenna@gmail.com

6. Arbeidet

A: Slik startet det

Ulovlig fildeling på nettet er et stort problem for innholdsbransjene (musikk, film, programvare etc) og de ønsker derfor både en strengere lovgivning og tilgang til flere virkemidler for å få bukt med problemet.

Mange av disse tiltakene kan imidlertid enkelt misbrukes dersom de blir implementert som en del av norsk lovgivning, og det finnes mange internasjonale eksempler på at dette ikke bare er en teoretisk problemstilling.

Meningene er derfor sterke, og partene står steilt mot hverandre i samfunnsdebatten rundt hva som kan og bør gjøres.

Fredag 8. oktober 2010 sendte borgerrettsorganisasjonen Elektronisk Forpost Norge (EFN) ut en pressemelding der de vedla et notat sendt fra "Dele - ikke stjele"-kampanjen til Kulturdepartementet. De beskrev det som et "Hemmelig retthaversk notat" som vil "amputere person- og rettsvernet".

Bransjenyhetsnettstedene Computerworld og digi.no fulgte opp med pressedekning. De intervjuet notatets avsendere - som nektet for at det var forsøkt hemmeligholdt på noe som helst vis.

Undertegnede bestemte seg for å sjekke om notatet var tilgjengelig på Offentlig Elektronisk Postjournal (OEP.no). Det var det ikke, men etter å ha søkt forgjeves på flere forskjellige relevante nøkkelord ble det etterhvert klart Kulturdepartementets journal ikke var særlig oppdatert. Det var på det tidspunktet ingen dokumenter som var nyere enn en måned gamle.

Noen stikkprøver på journalene til andre departement viste at enkelte departement hadde relativt ferske oppdateringer av sine journalføringer, mens andre var tregere. Det var imidlertid ingen av de andre departementene som var så dårlig oppdatert som Kulturdepartementets journalføringer.

Konklusjonen ble derfor at "Dele ikke stjele"-kampanjen ikke hadde hemmeligholdt notatet, og at

det var den trege journalføringen som forårsaket et unødvendig ekstra konfliktnivå i en allerede betent samfunnsdebatt.

Utover å informere om dette på Twitter og gjennom ytterligere nettdebatt ble det ikke laget noen redaksjonell sak om selve hemmeligholdet.

B: Den sentrale problemstillingen

Den interessante redaksjonelle problemstillingen lå i å sjekke hvorvidt treg journalføring var et systematisk problem eller en tilfeldighet, samt om det var store forskjeller mellom hvordan departementene håndterte journalføringen.

Hver journalføring har lagt ut informasjon om virksomhet, sak, dokumenttittel, saksnummer, dokumentnummer, dokumenttype, avsender, dokumentdato, journaldato, publiseringsdato, unntaksgrunnlag og kontaktpunkt med telefonnummer og e-post. Denne informasjonen er publisert som en webside, og den er ikke tilgjengeliggjort i et regenark eller i noen form for maskinlesbart dataformat.

Det betyr at man må besøke hver eneste side, klippe ut hvert enkelt informasjonsfelt og legge det inn i et regneark. Ettersom det legges ut mange tusen journalføringer på OEP [.no](http://oep.no) ville det vært en uoverkommelig jobb å gjøre arbeidet manuelt.

C: Dokumentasjon av systematisk journalsommel

Resultatet av det redaksjonelle gravearbeidet ble en oversikt som avslørte at Kulturdepartementet i snitt brukte 28,3 dager på å offentliggjøre sin journalførte kommunikasjon. Det var nesten 3 ganger så lang tid som departementenes snitt på 9,2 og markant mye lengre enn Miljøverndepartementets snitt på 3,5 dager.

Flere skal tidligere ha påpekt at det somles med journalføringen i diverse departementer, men det redaksjonelle arbeidet som ble gjort i dette graveprosjektet ga ubestridelig dokumentasjon på departementenes sviktende journalføringspraksis.

Med faktagrunnlaget fra blogginnlegget fikk Presseforbundet endelig ubestridte fakta som Kulturdepartementet måtte forholde seg til. Dette ble da også fulgt opp i et radioinnslag på NRK Kulturnytt hvor undertegnede primært bidro med å forklare hvordan datagrunnlaget ble innhentet og analysert.

Etter blogginnlegget og dette graveprosjektet har Kulturdepartementet skjerpet sin praksis. Stikkprøver viser at de ligger på rundt en ukes forsinkelse nå, og ytterligere dokumentasjon på dette vil foreligge i løpet av våren når datasystemet Gravemaskinen blir klar med kontinuerlig analyse av journalføringen i tilnærmet sanntid.

D: Gravemaskinen

Som nevnt ville det vært en uoverkommelig oppgave å samle inn og strukturere informasjon om alle journalføringer manuelt. Det ville også vært for tidkrevende å lage et nytt dedikert dataprogram eller et automatisert script fra bunnen av, ettersom det er snakk om både en komplisert innsamlingsprosess og en ressurskrevende analyse.

Graveprosjektet var imidlertid et perfekt pilotprosjekt for Gravemaskinen.no, et redaksjonelt IT-system som utvikles for journalister av undertegnede i samarbeid med tidligere ansatte i Google som har erfaring fra utvikling av avansert søketeknologi.

Gravemaskinen.no er en redaksjonell søkemotor som finner, identifiserer og strukturerer faktaopplysninger i store tekstmengder. Mens tradisjonelle søkemotorer som Google, Bing m.fl. viser de 10 mest relevante dokumentene som inneholder faktaopplysningene man leter etter, presenterer Gravemaskinen de mest relevante faktaopplysningene som finnes i disse dokumentene.

Gravemaskinen er delvis utviklet som en generell søkemotor for å kunne tilby basisinformasjon når journalister søker etter faktaopplysninger på tvers av datakilder, men er primært utviklet for å raskt kunne tilpasses journalistiske graveprosjekter i utvalgt kildemateriale. Dette innebærer blant annet å lage automatiserte journalistiske algoritmer som leter på samme måte som en journalist.

En tradisjonell søkemotor bygger en indeks med alle ordene som finnes i tekster for så å la brukerne søke raskt gjennom alt etterpå. Slike søkemotorer vet ikke hva de leter etter, men legger i steden tilrette for brukerne som vet hva de søker etter.

Journalister derimot, vet hva de ser etter når de manuelt leter etter informasjon i en tekst. I dette tilfellet ville en journalist lett etter alle journalføringer på OEP.no der journaldato og publiseringsdato er oppført for så å klippe ut informasjonen og legge den inn i et regneark.

Dette arbeidet ble i steden overlatt til Gravemaskinen gjennom å raskt lage noen spesialiserte digitale "journalister" som hentet ut informasjonen og la den inn i et regneark. Totalt tok det noen få timer å konfigurere systemet, samt å hente ut informasjonen fra OEP.no.

Deretter ble dataene manuelt analysert i et regneark. Det innebar stort sett enkle grep som å sortere datoene pr. departement, samt regne ut snitt og andre relevante statistiske beregninger.

Ettersom langt fra alle er komfortable med tabeller og regneark, og med tanke på at de er visuelt kjedelige, ble det eksperimentert med en rekke forskjellige grafer for å illustrere faktaopplysningene om journalføringen som ble dokumentert. Det inkluderte blant annet søylediagrammer, speedometerdiagrammer og såkalte radardiagrammer.

Hvilke grafer som egnet seg best er fortsatt usikkert, men resultatet ble solid dokumentasjon presentert på en enkel visuell måte som er lett å forstå for alle som leser blogginnlegget.

Ettersom dette var et pilotprosjekt med ny datateknologi ble det brukt mye tid på å gå gjennom deler av datasettet manuelt for å sjekke at ikke programvarefeil forårsaket feil i datagrunnlaget.

Det ble ikke utført intervjuer i forbindelse med dette graveprosjektet. Målet var å finne frem redaksjonelt interessante faktaopplysninger om departementenes journalføringspraksis ved hjelp av datastøttet journalistikk, og det lyktes.

7. Spesielle erfaringer

Hele prosjektet må i seg selv anses som en spesiell erfaring. Foruten det er det to momenter som undertegnede ønsker å trekke frem. Det første går på redaksjonell interesse for datastøttet journalistikk, og det andre går på Offentlig Elektronisk Postjournals tekniske praktisering av offentlige journaler.

Før blogginnlegget ble publisert ble flere redaksjoner tilbudt saken på frilansbasis, og etter publisering på bloggen ble det sendt ut tips til en rekke redaksjoner. Foruten Kristine Foss i Norsk Presseforbund - som fant blogginnlegget på eget initiativ og fulgte opp med eget innlegg på offentlighet.no - var det ingen interesse.

OEP.no er laget for å bidra til meroffentlighet, men gjør det på en lite hensiktsmessig måte. Det burde ikke være nødvendig å måtte bygge et produkt som Gravemaskinen for i det hele tatt å kunne gjøre en systematisk analyse av OEP.no. Det ligger jo til og med i navnet at dette er en "offentlig" postjournal.

8. Sted & dato

Anders Brenna
Sandvika, 15. januar 2010