

Prosjekt 3 i Distribuerte Informasjonssystemer

RSS-Nyhetsleser

Prosjektet er utført av:

Holtmark, Rolf Krog, s113435, 3AA
Kaata, Knut Ola, s113430, 3AA

Slik gjør du for å starte applikasjonen:

Legg .war fila inn i tomcat sin webapp katalog. Den vil da pakke seg ut automatisk når tomcat kjører. Så går du til <http://<host>/<pakke-navn>/> Der vil du bli redirected til servleten.

Du trenger følgende programmer og moduler:

rsslib4j-0.2 er inkludert i pakken.

Vi har benyttet følgende teknologi:

Startet med å se på htmlparsere, men det gikk vi fort vekk fra.
RSS Parser / XML
Servlet
Cache midlertidige filer til disk.

Vi valgte denne tekologien fordi:

Den gjør at vi kan hente informasjon/nyheter fra flere sider ut fra et standard format. Dette gjør av vi slipper å ta hensyn til for eksempel layout og andre problematiske temaer rundt html parsing. Det som denne tekologien krever er at man benytter en rss informasjonskanal (eks: www.itavisen.no/rss.php).

Hva som kan gjøres med teknologien:

Applikasjonen vår laster ned RSS feeds fra forskjellige websteder og samler disse i en enkel og pen portal. Med vår applikasjon kan man greit lese de siste nyhetene uten masse reklame og bli oppdatert. Når man ser en artikkel som er intressang, kan man lett klikke seg inn på den og lese hele nyheten. RSS er et standard format som de fleste større seriøse nyhets-aktørene tilbyr.

Vi har lært følgende i dette prosjektet:

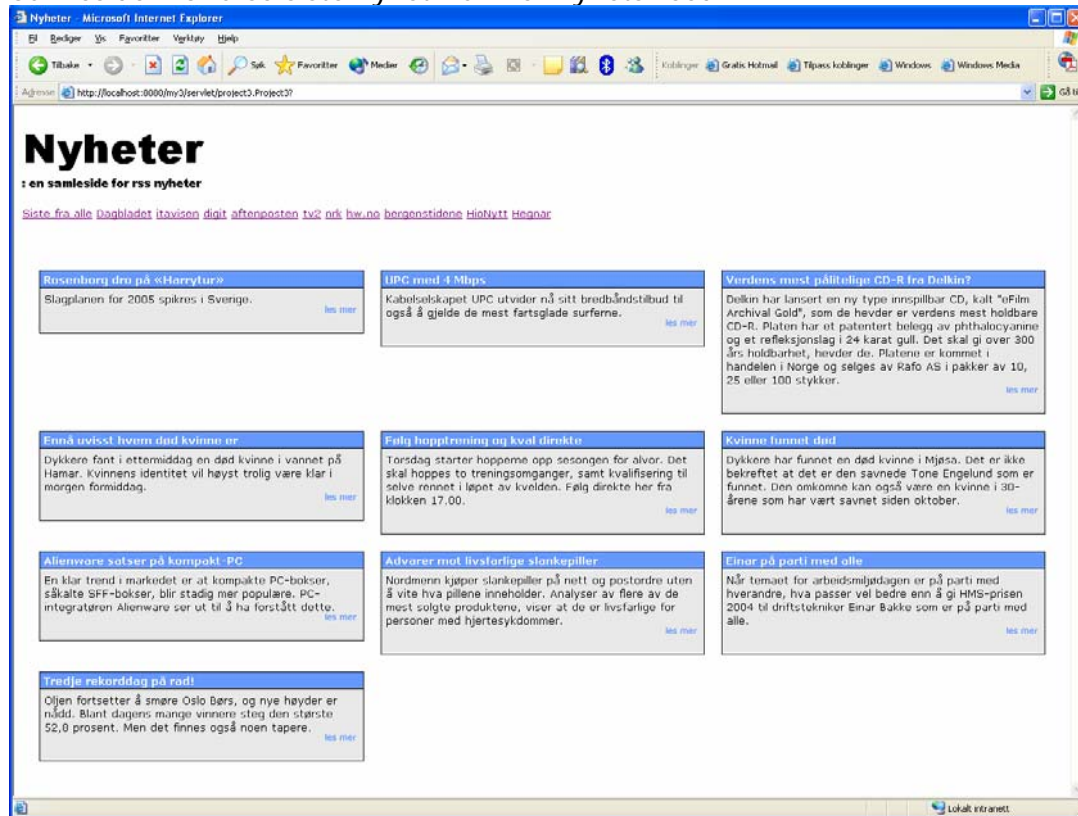
Først startet vi med å se på mulighetene for å parse html websider, men fant ut at det ikke er særlig gunstig siden design og layout på sidene forandres stadig. Da kom vi opp med at det er smartere med å parse wap sider (du ser noe rester av dette i prosjektet: Downloader.java). Etter en liten stund kom vi over rss-teknologien. Den er helt genial i forhold til å hente nyheter fra mange sider og lese de på en kjekk måte. De fleste større nyhets-websidene benytter denne teknologien, og i vår demonstrasjon av RSS nyhets parsing viser vi hvordan man lett kan hente ned og benytte RSS informasjon i en samlet nyhets side.

Det som var spesielt vanskelig med denne teknologien:

Det å finne fram til hva som egentlig var best å bruke for å parse rss nyhetene. Vi kunne ikke se at Jbuilder/Java hadde noen slik funksjonalitet fra før. Den hadde noe XML parsing, men lite med direkte RSS parsing. Vi brukte derfor en del tid på å finne fram til en grei og god library som kunne gjøre disse funksjonene. La inn en en funksjon for å cache nedlastingen slik at nyhetene lagres som midlertidige rss filer på disken. Det fikk applikasjonen til å gå raskere fordi den da slapp å laste ned nyhetene hver gang.

Screenshots:

Samlesiden for å se siste nyhet fra hver nyhets-feed:



Viser alle nyhetene som dagbladet har lagt ut i sin nyhetsfeed:

