

Dokumentfangst fra Internett

- Nasjonalbibliotekets innhøsting av nettsider

Kjersti Rustad, seksjonsleder

kjersti.rustad@nb.no

Hvorfor tar Nasjonalbiblioteket vare på norske nettsider?

Pliktavleveringsloven

- Trådte i kraft 1. juli 1990
- Omfatter alle typer dokumenter uavhengig av medium
- Hovedkriteriet for avleveringsplikten er at et dokument er gjort allment tilgjengelig, dvs. at utgiver har en intensjon om å spre dokumentet ut over en privat krets



LOV
OM
AVLEVERINGSPLIKT
FOR ALLMENT
TILGJENGELEGE
DOKUMENT
Av 9. juni 1989 nr. 32

Forskriftene til pliktavleveringsloven

§ 30, 2. ledd

Edb-dokument som er tilgjengelege ved direktekopling gjennom tele-, fjernsyns-, datanett e.l. skal avleverast i to eksemplar etter konkret oppmoding frå mottaksinstitusjonen i det einstilte tilfellet.

 Lovdata

Norske nettdokumenter – hva er avleveringspliktig etter loven?

- Alle dokumenter, dvs. hjemmesider og filer, som ligger på den norske delen av Internett (.no)
- Alle dokumenter på norsk eller dokumenter som er publisert av norsk utgiver på domener utenfor .no (.com, .org, .as etc.)
- Intranett faller utenfor loven!



Hva høster Nasjonalbiblioteket inn?

- Domeneinnhøstinger av det norske nettområdet én til to ganger i året
- Utvalgte nettsider høstes inn hyppigere
- Vitenskapelig publisering på nett, dvs. serier og rapporter (pdf)
- Tidsskrifter på nett og nettaviser



Seleksjonskriterier for innhøsting av utvalgte nettsider

Nettsider som dokumenterer

- Begivenheter av nasjonal interesse. Eksempel:
 - Stortingsvalg, kommune- og fylkestingsvalg
 - Kronprinsbryllupet
 - Finanskrisen
 - Olympiske leker / Paralympics
 - Melodi Grand Prix Oslo 2010!
- Norsk kultur og samfunnsliv. Eksempel:
 - Norske nettaviser
 - Forfatterjubileer



Aftenposten.no



NETTAVISEN

Litt historikk

- Selektiv innhøsting av enkeltdokumenter fra nettet t.o.m. 2004
- Innhøsting av hele den norske delen av Internett f.o.m. 2005, supplert med selektiv innhøsting av utvalgte nettsider og enkeltdokumenter



<http://www>

Antall innhøstede filer (URLer) siden 2005

2005	2006	2007	2008	2009
310 mill.	315 mill.	280 mill.	485 mill.	20 mill.

Juridiske begrensninger for innhøsting og tilgjengeliggjøring av nettdokumenter

Nasjonalbiblioteket må ha konsesjon fra Datatilsynet for innhøsting og tilgjengeliggjøring

Bruk av dokumentene skal være til forskning og dokumentasjon

Brukerne må oppsøke Nasjonalbibliotekets lokaler for å få tilgang til nettarkivet



Internasjonalt samarbeid



International Internet Preservation Consortium (IIPC)

Blant deltakerne er nasjonalbibliotekene i Norden, Frankrike, Italia, Canada, Australia, Storbritannia, Library of Congress i USA og Internet Archive, USA.

www.netpreserve.org/about/index.php

Verktøy for innhøsting

HERITRIX for domeneinnhøsting

<http://crawler.archive.org/>



for tematiske
innhøstinger og innhøsting av enkeltsider

<http://webcurator.sourceforge.net/>

Utfordringer med Web Curator Tool og Heritrix

Begrensninger

- Får ikke med passordbeskyttede sider
- Får ikke med informasjon i databaser (deep web)
- Vanskelig å få komplekse sider helt perfekt

Positive sider

- Web Curator Tool er brukervennlig
- Enkelt å sjekke kvaliteten på innhøstingen



En titt inn i nettarkivet ...

Velkommen til Maihaugen! - Maihaugen - Windows Internet Explorer

http://wct.nb.no:8080/wct/curator/tools/browse/1403224070/http://www.maihaugen.no/

Velkommen til Maihaugen! - Maihaugen

MAIHAUGEN

SØK

Posten skal fram

VELKOMMEN TIL MAIHAUGEN!

Maihaugen på Lillehammer er et av de største museene i Norge og det mest opplevelsesrike! De fem museene og fire fagavdelingene favner temaer som diktning, idrett, håndverk, samtid, kulturelt mangfold og historie.

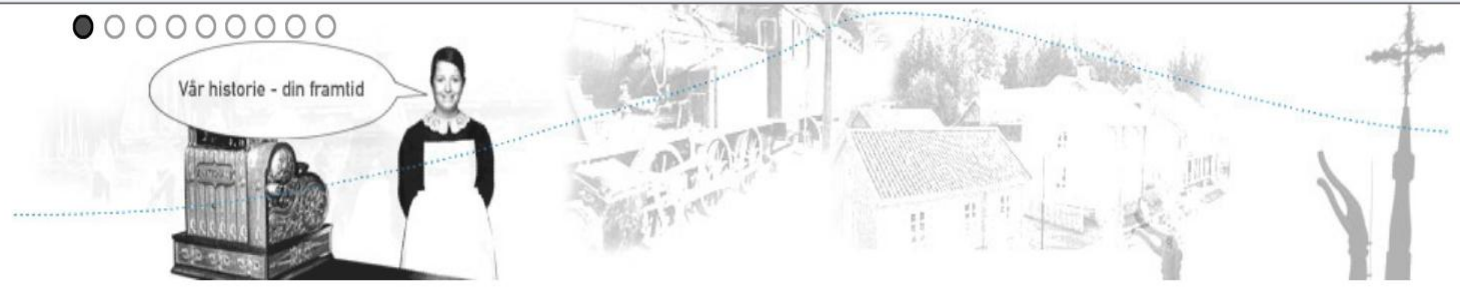
Besøksinformasjon

- Åpningstider og priser
- Fellesbillett
- Vegbeskrivelse
- Kontakt oss

AULESTAD

Bjørnsonåret 2010

- Maihaugen
En reise i tid
- Postmuseet
Postens historie fra 1647
- Norges Olympiske Museum
Norsk idretts utstillingsvindu
- Tid for mat AS
Velsmakende måltider
- Bjørnstjerne Bjørnsons hjem Aulestad
Karoline og Bjørnstjerne Bjørnsons hjem
- Sigrid Undsets hjem
Sigrid Undsets hjem og dikterverksted



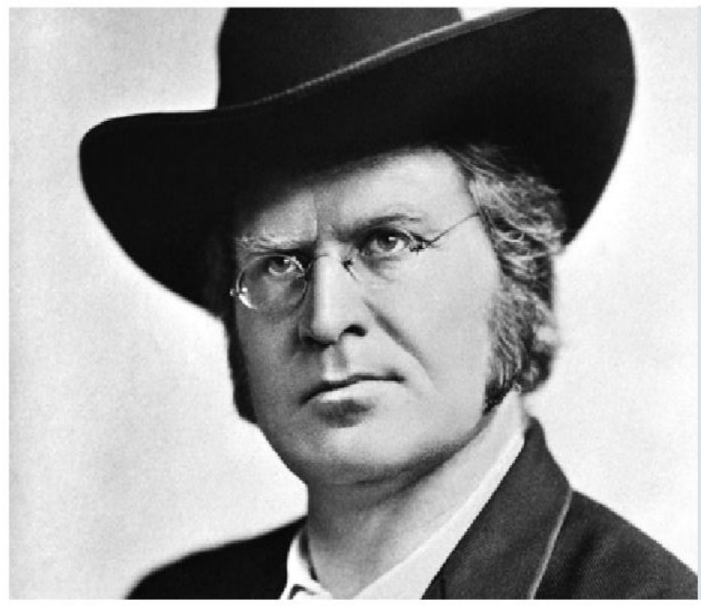
- Maihaugen
En reise i tid
- Postmuseet
Postens historie fra 1647
- Norges Olympiske Museum
Norsk idretts utstillingsvindu
- Tid for mat AS
Velsmakende måltider
- Bjørnstjerne Bjørnsons hjem Aulestad
Karoline og Bjørnstjerne Bjørnsons hjem
- Sigrid Undsets hjem
Sigrid Undsets hjem og dikterverksted
- Norsk handverksutvikling - NHU
Viderefører handverkskunnskap
- Handverksnett

VELKOMMEN TIL MAIHAUGEN!

Maihaugen på Lillehammer er et av de største museene i Norge og det mest opplevelsesrike! De fem museene og fire fagavdelingene favner temaer som diktning, idrett, håndverk, samtid, kulturelt mangfold og historie.

Besøksinformasjon

- [Åpningstider og priser](#)
- [Fellesbillett](#)
- [Vegbeskrivelse](#)
- [Kontakt oss](#)



AULESTAD

Bjørnsonåret 2010

Foto: Warrens Portraits

« Forrige || Neste » 2 / 20

GD.no**SØK I SKATTELISTENE**
Se hvem som tjente hva.**GD PÅ FACEBOOK**
Bli vår venn på Facebook!**SE BIRKEN I BILDER!**
Vått, rått og vilt i årets ritt.**KONTAKT OSS**
Tips telefon: 977 21000
E-post tips@gd.no
SMS/MMS: gdtips til 2097
Kundeservice tlf: 61 22 10 00

Forside

Nyheter

Sport

Kultur

Magasinet

Motor

Meninger

Bilder

GDpuls

Hyttepuls

Søkeord

Søk

TIPS OSS

Kundeservice

Adresser

Bestill abonnement

Bestill annonse

Uteblitt avis

GD i pdf

GD-fondet

Lokale bedrifter

RSS

TIRSDAG 26.10.2010

annons



VÆR

annons



Lesja



Otta



Gausdal



Mediahuset Novas www.telefonkatalogen.biz Firmsøk, bransje, tlf, sted .biz

GD.no **SKATT 2010** **SØK I SKATTELISTENE** Se hvem som tjente hva. **GD PÅ FACEBOOK** Bli vår venn på Facebook! **SE BIRKEN I BILDER!** Vått, rått og vilt i årets ritt. **KONTAKT OSS** Tips telefon: 977 21000 E-post tips@gd.no SMS/MMS: gdtips til 2097 Kundeservice tlf: 61 22 10 00

Forside Nyheter Sport Kultur Magasinet Motor Meninger Bilder Gdpuls Hyttepuls
TIPS OSS Kundeservice Adresser Bestill abonnement Bestill annonse Uteblitt avis GD i pdf GD-fondet Lokale bedrifter RSS TIRSDAG 26.10.2010



TrygVesta Forsikring har skiftet navn til Tryg
Les mer på tryg.no >>

Det handler om å være trygg



annonse

VÆR

- Lesja
- Otta
- Gausdal
- Lillehammer

Skeikampen
Lillehammer
SkiResorts

Over 1,5 millioner lesere hver dag

VG

Oslo 6.8° 1.3 m/s
Endre sted

Google Custom Search SØK
VG Nett Skattelister Google

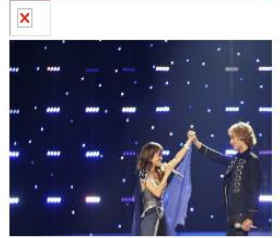
NETT NYHETER SPORT FOTBALL RANPELVS REISE VÆR NYTTE TV-GUIDE LESERNE GØY A til A

vektklubb.no VEST minMote.no LIVE nettbu.no VBLISTA MØTEPLASSEN DINE PENGER KONTAKT OSS



Seierskysset

Rybak klinte til med tyske Lena (19) etter Didriks GP-flopp. **Les saken**
Protokoll: **Gratuler Lena og Tyskland med seieren**



- Bør være krise i Didrik-leiren
VG TV [Se GP-Josteins oppsummering av nedturen](#)
Les også: [Sammenligner Didrik med julenissen](#)



Brøt seg inn på MGP-scenen
Drama på direkten - [Les hele saken](#)
VG TV [Her hopper «Jimmy Jump» rett inn på scenen](#)

LEVER! HJEM!

Klikk her

annonse

http://flash.vg.no/annonser/d/emea.doubleclick.net/adj/N1 cannot be found.

done

Over 1,5 millioner lesere hver dag

VG

RIKINGEN

FØLG MILLIONÆRLIVET PÅ VG TV

Sett VG som startside

TORS DAG 20. MAI 2010

Oslo 21.2° 1 m/s
Endre sted

Google Custom Search SØK

VG Nett Skattelister Google

N E T T NYHETER SPORT FOTBALL RAMPELYS REISE VÆR NYTTE TV-GUIDE LESERNE GØY **A til A**

vektklubb.no VG10 minMote.no LIVE nettbu.no LISTA MØTEPLASSEN DINE PENGER KONTAKT OSS



Eggen tar over Rosenborg

Rosenborg har innkalt til hastepressekonferanse kl. 14.00
[Les hele saken](#) Hva synes du om comebacket? [Si din mening her](#)

TIPS VG mobil: 2200 | 2200@vg.no | 22 00 00 00

Siste nytt Mest lest Mest sett

Alle [Nyheter](#) [Sport](#) [Rampelys](#)

- 13:05 : - Bør koste mer å ta ut penger i skranken
- 12:53 : Nils Arne Eggen tar over RBK
- 12:46 : Pressekonferanse i Brann nå
- 12:26 : Bhatti frifunnet for drapforsøk
- 12:16 : Bataljonen-leder - En katastrofal sesong
- 12:11 : Pakistan stenger Facebook, YouTube, Wikipedia og Flickr
- 12:11 : Malerier for milliarder stjålet i Paris
- 11:58 : Gutt utenfor livsfare etter Alta-ulykke
- 11:43 : Frank Løke må betale 130.000 kroner

SHARIF'S NETTBUTIKK

koste hva det koste vil!

minMote.no

HJEM!

Klikk her

annons

http://flash.vg.no/annonser/doi
emea.doubleclick.net/adj/N18t
cannot be found.

done



Bjørnson
BJØRNSON
 FOR VÅR TID 2010



Forside Om Bjørnson2010 For presse English Kontakt Nyttige lenker

Bjørnson - en innføring

Litteraturen

Samfunnet

Bjørnsons beste

Romaner og skuespill

Dikt og sanger

Månedens sang

Faderen

Lyd og bilde

Galleri

Film

Liv og verk

Bjørnson-året

Oslo

Aulestad

Molde/Neset

Forfatter Bjørnstjerne Bjørnson 1832-1910

- Bjørnstjerne Martinius Bjørnson, født 8. desember 1832 på Kvikne nord for Tynset.
- Dikter og samfunnsdebattant.
- Skrev teksten til Norges nasjonalsang, "Ja, vi elsker".
- Ble tildelt Nobelprisen i litteratur i 1903.
- Døde i Paris 26. april 1910.

Kjempesuksess på Aulestad.

Sist helg passerte Aulestad 20.000 besøkende. Det er mer en tre ganger så mange som de hadde totalt året før restaureringsarbeidet begynte.

[Les mer ...](#)

nyhetsbilde

Og, så kommer du ikke!

I går avsluttet "Det hemmelige teater" sin sommerfestival på Frydenlund teater i Bødal, Vestre Gausdal. På tross av at teateret ligger utenfor allfarvei, har skuespillerne spilt for fulle hus.

[Les mer ...](#)

nyhetsbilde

Synnøve Solbakken

Synnøve Solbakken ble fullført på en hybel i nærheten av Slottet, i det området som den gang var utkanten av Christiania. Bjørnson var nylig vendt tilbake fra sitt første lange opphold i Danmark, der han hadde stiftet kontakt med et rikt litteratur- og kunstmiljø.

[Les mer ...](#)

nyhetsbilde

Bok til ære.

Utstillingen "Måneskinskåpen opunder pol" tar utgangspunkt i Bjørnsons bok "Arne". Nå er utstillingen blitt bok.

[Les mer ...](#)

nyhetsbilde

søk i bjornstjerne.no

[Søk på Bjørnson i](#)

[Nasjonalbiblioteket](#)

Skolesider

Nettutstilling

Arrangementer

Bjørnson-konferanse 2010

Bjørnson i media

[Aursjøvegen - eventyrlig biltur](#)

[Festivalen vil låne stua di](#)

[Vil låne stua di](#)

[KOMISHOW, - Komishow eller ølfest](#)

[Ane Bjørnsgaard](#)

Oppsummering

Nasjonalbiblioteket tar vare på norske nettsider med hjemmel i pliktavleveringsloven

Domeneinnhøsting kombineres med selektiv innhøsting

Innhøsting av nettsider og framvisning av nettarkivet krever konsesjon fra Datatilsynet



Foto: Andreas B. Wilse