

## Case/Monica Kleja

Tre vägar att utveckla  
digitalt berättande

Årets Noda-konferens visade upp flera bra exempel på hur medieföretag har vässat det digitala berättandet. IT-skribenten Monica Kleja lyfter fram tre av dessa:

# 1, Helsingin Sanomat: En story kan bli miljontals artiklar

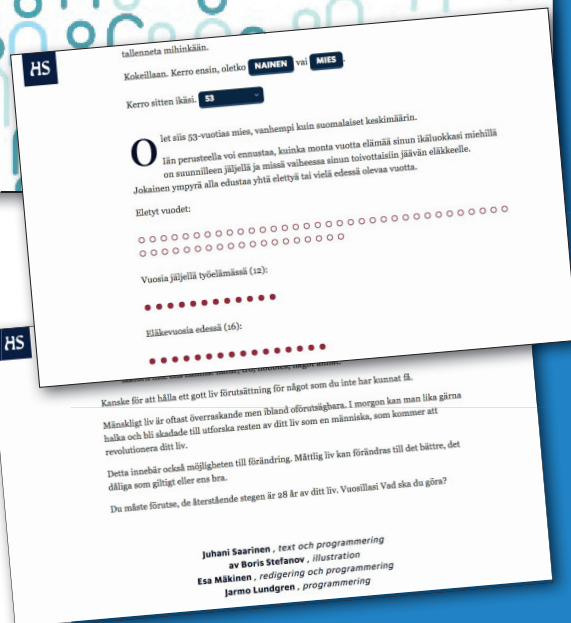
HS



» Hur kan en digital text växa och bli en miljon olika artiklar? Det vet Helsingin Sanomat, som nyligen vann ett nordiskt pris för sin innovativa datajournalistik.

Nordens största prenumererade tidning Helsingin Sanomat knep priset för bästa digitala feature-story när den nordiska datajournalistik-konferensen Noda2017 nyligen ägde rum i Odense. Projektet "Where are you at life? Read a story on family, health, work and happiness tailored to you", tilltalade juryn, vars medlemmar kom från olika mediehus i Norden.

"Detta bidrag har verkligen kommit nära läsarna genom att skraddarsy historier som passar bara dig. Genom att sammanställa ett stort antal datakällor i en läcker design, kan detta bidra till att potentiellt berätta en miljon olika historier", löd deras motivering.



<https://dynamic.hs.fi/2016/valitilinpaaotos/>



MONICA KLEJA

Journalist med ett förflutet på bland andra Ny Teknik och Dagens IT.

Medie  
världen  
Premium



Helsingin Sanomat har 324 000 prenumeranter och når 1,6 miljoner finskspråkiga besökare på webben varje vecka. **Juhani Saarinen**, producent vid mediehusets datadesk, leder arbetet med så kallade adaptiva artiklar: "Vi utgår från en mobilanpassad version eftersom allt fler läsare går in på webben via mobilen eller surfplattan. Och sedan anpassar vi storn till desktop."

Det finska receptet att komma läsarna in på livet bygger på en stor mängd insamlad statistik. I det belönade konceptet har Helsingin Sanomat samlat in befolkningsdata om bland annat finländarnas förväntade livslängd, andel gifta/ogifta och deras hälsotillstånd.

Mitt i webbartikeln ges läsaren möjlighet att interagera och knappa in uppgifter om ålder, kön, civilstånd, bostad, barn/inga barn, arbetsstatus, hushållets inkomst, upplevd hälsa, motionsvanor och livskvalitet och svarar på frågor som hur ofta han/hon träffar sina nära och kära.

När hen sedan pressat ned ENTER formligen växer artikeln med i förväg skriven standardtext som passar ihop med de instoppade värdena. Om jag exempelvis klickat i att jag är gift, svarar artikeln ungefär så här: "Gratulerar, du är en av si och så många kvinnor/män i Finland i din ålder som är gifta..."

"Artikeln blir olika för varje person. Därför visste vi inte själva vad nyheten var och hade svårt att sätta rubrik", berättar Juhani Saarinen, som tillägger att tidningen inte lagrar personuppgifterna. De sparas i besökarnas webbläsare, och går att rensa.

Webbartikeln om lycka och hälsa placerades bakom en mjuk betalvägg (man kan läsa några artiklar gratis), vilket Juhani Saarinen är glad över. Han menar att det underlättade spridning och var en nyckel till framgång.

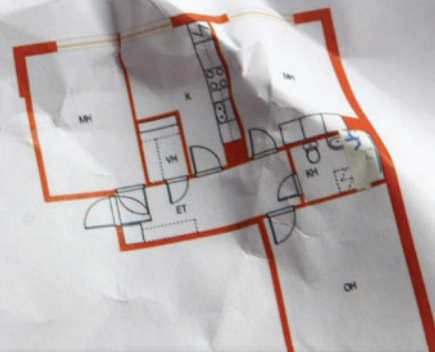
### Hur mäter ni fördelarna med adaptiva artiklar?

"De har inte varit bakom en hård betalvägg så vi mäter bara hur mycket de blir lästa. De har alla varit framgångsrika i det avseendet."

Men ibland behövs prenumeration för att läsa "lättversioner".

"Vi har ett verktyg som vi kan göra en slags light-versioner av adaptiva historier med. Många av dem har varit placerade bakom en hård betalvägg efter beslut i Newsroom, och vi har fått in mängder av provprenumerationer. Så datadriven journalistik påverkar också intäkterna", säger han.

Vad satsningen kostat vill Juhani Saarinen inte avslöja men han uppger att elva personer vid datadesken jobbar med interaktivt innehåll, rörlig och stilla grafik av tidningens totalt cirka 300 anställda. Desken utbildar dessutom journalister på andra redaktioner.



I en liknande expanderande artikel i fjol kunde läsaren kolla hur mycket hans villa var värd. "Artikeln om lokala huspriser var bland de mest lästa historierna i HS förra hösten. Vi fick mycket feedback, folk verkade gilla den och den lästes mer än första sidan, vilket inte är vanligt", säger Juhani Saarinen, som uppger att det handlade om hundratusentals läsare.

En stor del av grundjobbet innebär att bearbeta data i Excel, och att koda i bland annat Javaskript. Varför valde man då att gå denna väg – och låta en story växa utifrån individen?

Juhani Saarinen: "Det var ett nytt koncept som upplevdes som spännande, jag tror inte det gjorts någon annanstans i världen". Pulitzer-priserna kanske dock dröjer ett tag:

"Adaptivt innehåll är enkelt och snabbt att åstadkomma, men erbjuder trots det stort mervärde för våra läsare. I början blev de inskjutna textstyckena stiltiga och stack ut. Nu har vi förbättrat flytet i artiklarna. De är mycket mer elegant skrivna i dag."

Andra artiklar som lockat finländarna har utforskat hur marknadsvärdet på bilar minskat och hur den nya statsbudgeten påverkar din privatekonomi.

### Vilken slags data/statistik letar ni efter?

"Vi letar efter intressanta data som alla har en relation till", säger Juhani Saarinen. "Vi har exempelvis skojat lite om vinklar som: hur ofta byter du dina byxor, och hur många jular har du kvar?"

Kerro postnumerosi:

010|

Mikä on oman asuinalueesi tilanne? Kerro yllä valintasi perusteella. Tietoa ei tallenneta.

00100 un sisältö muuttuu

20100

30100

40100

50100

60100 at nousevat

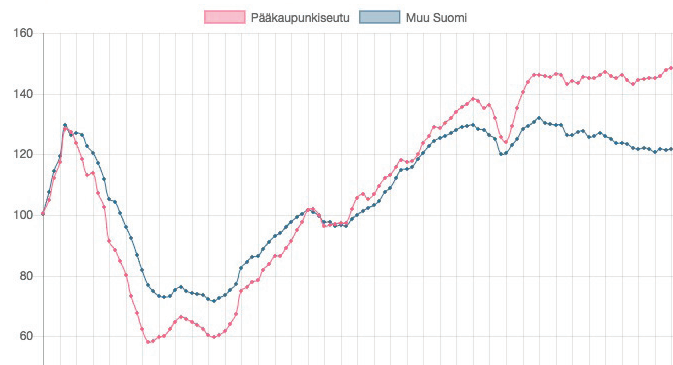
70100 maassa ne ovat

80100

90100

**A** suntomarkkinoilla on käynnissä histor...  
pääkaupunkiseudulla ja muutamassa r...  
laskeneet vuodesta 2010 saakka.

Asuntojen reaalihintaindeksi 1988–2016, 2000=100



# 2, Berlingske: Dataanalys fann miljontals skumma bud

**Berlingske Business** 0,24% OMX C20CAP 1062,29

LIVE

SENESTE NYT

Q

SØG

INDHOLD

LIVE: Investorer ryster i bukserne over Le Pen



Investor

Virksomheder

Privatøkonomi

Karriere

TV

Opinion

DETAILHANDEL

## Ti kunder står bag millioner af bud på Lauritz.com

Torsdag d. 08. september 2016, kl. 21:00

TOPNYHEDER

BREAKING

Medie: Shell klar til at sælge sine aktier i den danske Nordsøolie

ENERGI 6. feb.

Nyt netværk for professionelle kvinder: Skal du være med?

KARRIEREKVINDER 6. feb.

TDC fyrer koncerndirektør

DIGITAL 6. feb.

Dansk trolld og prinsessestøj sikrer Lego-konkurrent rekordår

DETAILHANDEL 6. feb.

Så meget make-up til mænd bliver der solgt

DETAILHANDEL 6. feb.

Gør et hus-kup: Tvangsauktionerede huse handles til under halv pris

BOLIG 6. feb.

LIVE

BUSINESS LIVE

Thulesen Dahl angriber Nykredit-bonus: Hvorfor får direktøren 12 millioner for at blive i jobbet?



Overskrift &lt;



Almindelige forbrugere er oppe mod ekstremt aktive budgivere på Nordens største auktionshus, Lauritz.com. Omfattende dataanalyse viser, at én kunde har budt på 80.000 varer på 13 år.



Signe Ferslev Pedersen  
FØLG +



Philip Dam  
FØLG +

» Efter ett stort datadrivet gräv avslöjade danska Berlingske i höstas att tio kunder stod bakom miljontals bud på auktionssajten Lauritz.com, och att misstänkt falska bud florerade. Nu går man vidare i sin stora granskning.

I två veckor i fjol lät Berlingskes journalister, Philip Dam och Signe Ferslev Pedersen, två datorer skrapa och hämta hem text från cirka 4,4 miljoner undersajter på Lauritz.com, ett av Nordens största auktionshus online och Danmarks största.

Man lade sedan uppgifterna om kundnummer, bud, pris och auktioner i en sökbar databas, som man därefter gjorde djupdyk i. I journalisternas granskning ingår 110 miljoner bud som lagts fram vid 3,5 miljoner auktioner från 1999 och fram till 2016.

**Berlingske** är en av världens äldsta tidningar och har i dag en dagstidning med 76 000 ex i pappersupplaga och 1, 5 miljoner unika besökare på webben.

Medie  
världen  
Premium

**Philip Dam**, som är datajournalist, berättar att upprinnelsen var att en kvinna kontaktade redaktionen i föl, med misstankar om att oegentligheter pågick under internetauktionerna.

“Hon hade undersökt buden under lång tid och hade med sig en stor hög papper med anteckningar från olika auktioner. Då tänkte vi att vi kan göra bättre än så.” Det kunde man.

Granskningen visar att det finns en extrem grupp budgivare, som skiljer sig från den genomsnittlige Lauritz-kunden. En vanlig kund deltar normalt i tre auktioner totalt, enligt underlaget. Tio kunder hade dock lämnat bud på 370 000 online-auktioner (10 procent av totalen). Den mest aktiva – kundnummer 72807 – hade deltagit i över 80 000 auktioner. Samtidigt hade en annan enskild kund budat vid fler än 69 000 auktioner.

Den mest aktiva kunden köpte varorna hen bjöd på i 8, 9 procent av auktionstillfällena. Den näst mest aktiva handlaren i 0,8 procent av fallen, och den tredje mest aktiva bara i 0,3 procent av de auktioner hen var med i.

En het fråga var och är ännu om extremkunderna medvetet medverkar till att driva upp slutbudet, och om ägarna till Lauritz.com är införstådda med det. “I början slog tanken oss att det kunde vara folk hos Lauritz.com som kanske bjöd på varorna, eller att det fanns en robot, men det finns det alltså inget bevis för”, säger Philip Dam.

Mette Rode Sundstrøm, administrativ direktör på auktionsbolaget, uppgav att de tio extremkunderna var kända av auktionshuset och hade löpande kontakt med finansavdelningen och att bolagets larmsystem för kunder som handlar i “ond tro” på Lauritz.com därför inte upptäckt dessa.

Grävresultatet presenterades i Berlingske i ett uppslag i september 2016. Journalisterna fann också ett annat skumt beteende. Några kunder bjuder på varor utan att vilja köpa dem för att pressa upp priset. Om de ser att auktionen slutar “fel”, säljer de om varan – och vissa varor såldes gång på gång. Exempelvis gick en gammal, bleknad kontorstol för 335 danska kronor, såldes om för 625, såldes igen och fick ett drömligt slutpris på 2 300 kronor.

Det ledde till en annan artikel om att misstänkt falska bud florerade i syfte på Lauritz.com. Berlingske hittade 468 liknande exempel. Men så kallad Shill bidding, som företeelsen kallas, är inte olaglig i Danmark.

Flera frågor återstår att besvara, enligt Philip Dam – och nu går reportrarna vidare i sin granskning, rika på nya erfarenheter av datadriven journalistik i stor skala. “För mig var det en överraskning att all data, även historiska data, öppet fanns att hämta. Där fanns källkod, och url-er till varje auktion”, konstaterar Philip Dam.



## Skrap- och analysverktyg

Programvarorna RVEST-package och MySQL användes för skrapning och analys. Det finns olika verktyg för att skrapa hemsidor. Se: [en.wikipedia.org/wiki/Web\\_scraping](https://en.wikipedia.org/wiki/Web_scraping)

Skrapning är inte olaglig, men om man också kan ladda ned exempelvis bilder kan det utgöra ett brott mot upphovsrätten.

## Länkar till Berlingskes avslöjanden:

**“Tio kunder står bakom miljoner bud på Lauritz.com”**

[business.dk/detailhandel/ti-kunder-staar-bag-millioner-af-bud-paa-lauritz.com](https://business.dk/detailhandel/ti-kunder-staar-bag-millioner-af-bud-paa-lauritz.com)

**“Flera hundra misstänkta auktionsbud florerar på Lauritz.com”**

[business.dk/detailhandel/flere-hundrede-mistaenkellige-auktionsbud-florerer-paa-lauritz.com](https://business.dk/detailhandel/flere-hundrede-mistaenkellige-auktionsbud-florerer-paa-lauritz.com)

# 3, NTB: Robot kan få sköta valbevakningen

## Meet Norway's first robot journalist



Foto: Shutterstock / NTB scanpix

**NTB**  
Magnus Aabech, news developer  
NODA 2017

» Norska NTB tror starkt på robotjournalistik. I dag levererar man enbart automatgenererade fotbollsresultat. Men redan till hösten kan en ny robot få hjälpa med att bevaka 2017 års val till Stortinget.

“Vi kommer att få se mer och mer av robotjournalistik”, säger Magnus Aabech, NTB:s robotredaktör, när Medievärlden träffar honom på den nordiska datajournalistikkonferensen Noda2017 i danska Odense.

För nio månader sedan skickade nyhetsbyrån för första gången ut matchresultat från landets högsta fotbollsdivision – med hjälp av algoritmer.

“Vi ser stor potential i detta för oss och våra kunder”, säger Magnus Aabech och beskriver att roboten kan leverera matchresultaten snabbare och vara mer korrekt än en reporter. “Den gör inte samma fel två gånger”.

Samtidigt bygger de maskingenererade referaten på att NTB har frilansjournalister på spelarenorna, som skriver in mindre fakta som mål, straffar, poängställning och målgörare i ett särskilt system. Roboten skriver sedan sin artikel baserad på matchuppgifterna på tio sekunder. Språket styrs av mallar (templates) och ska vara så naturligt som möjligt.

**NTB**, Norsk telegrambyrå, är Norges största nyhetsbyrå och levererar text, bilder, video och grafik till i stort sett alla mediehus i Norge. Ägarna är Edda Media, Schibsted, A-pressen och NRK. Nyhetstjänsten har i nuläget 93 prenumeranter.

16 MAI 2016 MAN 17:18

**Skagestad sikret poeng for Aalesund** ID: 12413995

Aalesund (NTB): Aalesund og Start måtte nøye seg med ett poeng hver på Color Line stadion i eliteserien mandag. Det stoppet tapsrekken til Aalesund, som ikke hadde tatt poeng de siste tre kampene.

Lagene gikk i garderoben til pause uten scoringer. Chidiebere Nwakali ga Start ledelsen etter 66 minutter, men etter 78 minutters spill utlignet Edvard Skagestad for Aalesund.

Kampen endte 1-1.

Aalesunds Bjørn Helge Riise pådro seg gult kort. Aalesund har åtte poeng og Start fem poeng etter mandagens serierunde.

Tommy Skjerven dømte oppgjøret. 6751 tilskuere hadde tatt turen.

22. mai skal Aalesund møte Bodø/Gilmt, og Start møter Sogndal. ©NTB

Levert av NTBs automatiserte artikkeljeneste.

**Medie  
världen**  
Premium

Efter att roboten gjort sitt referat följer frilansjournalisten alltid upp med en mer initierad, längre och bättre artikel om spelet en stund senare, uppger han.

Många på nyhetsbyråns redaktion var skeptiska till en början. "Några kollegor var kritiska till att vi skulle börja med robotjournalistik, men i dag är de flesta ändå positiva."

I nuläget tjänar NTB inga pengar på Bob, enligt Magnus Aabech, som inte vill avslöja vad investeringen kostat. Men fotbollsresultaten är nu en del av dagens tjänsteutbud och det lockar byråns kunder. NTB överväger nu också att erbjuda referat från lägre fotbollsserier, kanske till och med ungdomsserier, till lokala medier.

"Bob" är emellertid inte färdigutvecklad. Det finns utmaningar. Som att språket i texterna kan upplevas som torrt eller torftigt. Och det har hänt att maskinfel uppstått, och att Bob svarat "Error – could not process the article". Hög trovärdighet är nödvändig för en nationell nyhetstjänst. Under alla robottexter står därför varningen: "Levererat av NTBs automatiska artikeltjänst".

Nu ska NTB skapa nya robotar, inhouse denna gång. "Vi ska försöka se om robotar kan göra andra nyheter exempelvis baserade på statistik om trafikolyckor från Statistisk centralbyrå (SSB)".

Den 11 september i år ska Norge välja nya ledamöter till Stortinget och NTB förbereder sig nu för att kunna använda roboten för att bredda bevakningen av antalet räknade röster till fler orter. "Vi gör 50–100 artiklar under valnatten och hoppas att den kan bli ett användbart verktyg, inte minst när det gäller att leverera uppdateringar till lokala medier, säger Magnus Aabech. Det finns också möjlighet att ta hjälp av artificiell intelligens, maskininlärning. Men där är inte NTB i dag. "Vi kommer att komma dit, men är inte där för stunden", konstaterar nyhetsbyråns robotredaktör.

**Example county level:** Mock-up

**HØYRE MEST FRAM I FINNMARK**  
Kirkenes (NTB): Høyre går forbi Arbeiderpartiet og ligger an til å bli største parti i Finnmark, ifølge prognosen etter at 17,8 % av stemmene er talt opp.

**Høyre gjør sitt beste valg siden 1972** og får 33,2 prosent oppslutning. Det gir to nye stortingsmandater. Mens Arbeiderpartiet mister ett og Senterpartiet faller ut.

Et endelig resultat i tråd med denne prognosen vil gi følgende sammensetning av Finnmarksbenken på Stortinget:

- Høyre: Nils Nilsen, Hanne Hanssen, Jenny Jensen, Kari Karlisen og Per Pettersen.
- Arbeiderpartiet: Nils Nilsen, Ola Olsen
- Fremskrittspartiet: Even Evensen

**Det mest utsatte mandatet tilhører på dette tidspunkt Høyres Per Pettersen. Senterpartiets Line Larsen mangler 815 stemmer på å erobre denne plassen.**

NTB

## SLUTORD:

>> Paul Myers, utbildare på BBC, påpekade på Noda-konferensen att det numera finns "briljanta verktyg som kan hjälpa dig att sortera" stora mängder data på nätet och i sociala medier. Helena Bengtsson, redaktör på The Guardians datadesk, noterade också att allt fler svenska medier använder dessa.

>> Men tekniken är inte allt – journalistiska mål måste vara vägledande. Nikki Usher, som undervisar i interaktiv journalistik på George Washington-universitetet i Washington DC, anser att journalister måste inse sina begränsningar. "Visualiseringen av data måste utformas så att den blir begriplig", betonar hon.



**Magnus Aabech**, är före detta sportjournalist som jobbat med utvecklingen av "Bob" – som han kallar roboten – tillsammans med den digitala studion Bakken&Bæck. (NTB har inget officiellt namn på roboten.)

### Editorial robots:

Det finns tre nivåer ifråga om nyhetsrobotar (editorial robots):

- ▶ **Nivå ett** utgörs av en maskin som enbart puttar in siffror i en färdig text.
- ▶ **Nivå två** är den som NTB:s robot befinner sig på. Här jobbar algoritmer med att skapa tidningsprosa med stöd av styrregler och begrepp. Ett litet inslag av artificiell intelligens ingår i språkprogrammet Natural Language Generation som NTB använder.

▶ **Nivå tre** innebär att man använder självlärande smarta robotar. De bygger på en högre grad av artificiell intelligens eller maskininlärning. (Källa: Magnus Aabech)

■ Den som vill botanisera bland bidragen i årets Noda-konferens hittar dem här:

[noda2016.fi/noda-awards.html](https://noda2016.fi/noda-awards.html)

■ Nästa års Noda-konferens blir i Stockholm med Södertörns högskola som värd.

**Medie  
världen  
Premium**