

Title Mikä on robottijournalismi - miten sitä voi hyödyntää ?
Author(s) Södergård, Caj
Citation Uudet mediapalvelut, median uudistuvat muodot ja ansaintamallit,
7.12.2016, Helsinki, Finland
Viestintäalan tutkimussäätiö (2016), 14 s.
Rights This presentation may be downloaded for personal use only.

VTT
<http://www.vtt.fi>
P.O. box 1000
FI-02044 VTT
Finland

By using VTT Digital Open Access Repository you are bound by the following Terms & Conditions.

I have read and I understand the following statement:

This document is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of this document is not permitted, except duplication for research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered for sale.



Mikä on robottijournalismi - miten sitä voi hyödyntää ?

Prof. Dr. Caj Södergård

Median uudistuvat muodot ja ansaintamallit
seminaari

Helsinki 7.12.2016



Aikajana – automaattinen tarinakerronta

- 2005 Chicago Crime – Google map mashup
- 2010 Los Angeles Times. Uutiset henkirikoksista: 10 % -> 100 % kattavuus
- 2011 Los Angeles Times. Quakebot kirjoittaa maanjäristyksestä
- 2013 Wikipedia – Lsjbot (Sverker Johansson, Ruotsi)
- 2014 AP. Tilinpäätös analyysit. Automatic Insights (USA)
- 2016 Botit tuottivat 15 % USA-vaalien some-postauksia
- 2018 20% of business uutisista koneiden kirjoittamia (Gartner)
- 2030 90 % artikkeleista koneiden kirjoittamia (Narrative Science)



The First News Report on the L.A. Earthquake Was Written by a Robot



The L.A. Quakebot helped to write a headline for the Los Angeles Times.



AP News
McCormick beats Street 3Q forecasts
 October 02, 2014

SPARKS, Md. (AP) — McCormick & Co. (MRC) on Thursday reported profit of \$122.9 million in its fiscal third quarter.

The Sparks, Maryland-based company said it had profit of 94 cents per share. Earnings, adjusted for non-recurring costs, came to 95 cents per share.

The results topped Wall Street expectations. The average estimate of analysts surveyed by Zacks Investment Research was for earnings of 81 cents per share.

VIDEO: Adobe 3Q Sales Miss Estimates Amid Shift to Cloud

The spillover and seasonings company posted revenue of \$1.04 billion in the period, which also topped Street forecasts. Analysts expected \$1.03 billion, according to Zacks.

McCormick expects full-year earnings in the range of \$3.30 to \$3.37 per share.

McCormick shares have declined nearly 5 percent since the beginning of the year, while the Standard & Poor's 500 index has increased slightly more than 5 percent. The stock has increased 1 percent in the last 12 months.

VIDEO: H&M 3Q Gross Margin 58.9%; Est. 58.4%

AP 12 kertaa enemmän artikkeleita tilinpäätöksistä

- Tuottaa 3700 artikkelia vuosineljänneksessä
- Käyttää noin 30 uutismallia
- ”Before journalists felt themselves as robots, now are robots doing the work” (Tom Kent)
- Vapautti raportointiaikaa – ei irtisanomisia, vähemmän virheitä
- Nyt myös urheilu-uutisia ja versiointeja

15/12/2016

AP News

McCormick beats Street 3Q forecasts

October 02, 2014



SPARKS, Md. (AP) — McCormick & Co. (MKC) on Thursday reported profit of \$122.9 million in its fiscal third quarter.

The Sparks, Maryland-based company said it had profit of 94 cents per share. Earnings, adjusted for non-recurring costs, came to 95 cents per share.

The results topped Wall Street expectations. The average estimate of analysts surveyed by Zacks Investment Research was for earnings of 81 cents per share.

VIDEO: Adobe 3Q Sales Miss Estimates Amid Shift to Cloud

The spices and seasonings company posted revenue of \$1.04 billion in the period, which also topped Street forecasts. Analysts expected \$1.03 billion, according to Zacks.

McCormick expects full-year earnings in the range of \$3.30 to \$3.37 per share.

McCormick shares have declined nearly 5 percent since the beginning of the year, while the Standard & Poor's 500 index has increased slightly more than 5 percent. The stock has increased 1 percent in the last 12 months.

VIDEO: H&M 3Q Gross Margin 58.3%; Est. 58.4%

This story was generated automatically by Automated Insights (<http://automatedinsights.com/ap>) using data from Zacks Investment Research. Full MKC report: <http://www.zacks.com/ap/MKC>

AP Sports is using “robot” reporters to cover Minor League Baseball

Posted Jul 3, 2016 by Lora Kolodny (@lorakolodny)



The AP wants to use machine learning to automate turning print stories into broadcast ones

The experiment is part of a larger effort by the news agency to incorporate automation into its journalism.

By JOSEPH LICHTERMAN @ylichterman Oct. 31, 2016, 10:51 a.m.

Lsjbot: 10 % Wikipediasta

- Sverker Johansson on kirjoittanut *Lsjbot* ohjelmallaan
 - Yli puolet Ruotsin Wikipedian artikkelista
 - 10 % koko Wikipediasta
 - 10.000 artikkelia päivässä on tyypillistä



WIKIPEDIA
Den fria encyklopedin

- Huvudsida
- Skriv en ny artikel
- Deltagarportalen
- Byrunnen
- Senaste ändringarna
- Slumpartikel (-bot)
- Ladda upp filer
- Stöd Wikipedia
- Kontakta Wikipedia
- Hjälp
- Skriv ut/exportera
- Skapa en bok
- Ladda ner som PDF
- Utskriftsvänlig version
- Verktyg
- Sidor som länkar hit
- Relaterade ändringar
- Specialsidor
- Permanent länk

Inte inloggad [Diskussion](#) [Bidrag](#) [Skapa konto](#) [Logga i](#)

Artikel [Diskussion](#) [Visa](#) [Redigera](#) [Redigera wikitext](#) [Visa historik](#)

Latuka [redigera | redigera wikitext]

Koordinater: 4.64041°N 32.55208°Ö﻿ / ﻿4.64041°N 32.55208°Ö﻿ / 4.64041; 32.55208

Den här artikeln har skapats av Lsjbot, ett program (en *robot*) för automatisk redigering. (2015-09)
Artikeln kan eventuellt innehålla språkliga fel eller ett märkligt bildurval. Mallen kan avlägsnas efter en kontroll av innehållet (vidare information)

Latuka är ett **stamområde** i **Sydsudan**.^[1] Det ligger i delstaten **Eastern Equatoria**, i den sydöstra delen av landet, 110 km öster om huvudstaden **Juba**.

Omgivningarna runt Latuka är huvudsakligen **savann**.^[2] Runt Latuka är det glesbefolkat, med 8 invånare per kvadratkilometer.^[3] **Savannklimat** råder i trakten.^[4] **Årsmedeltemperaturen** i trakten är 28 °C. Den varmaste månaden är februari, då medeltemperaturen är 34 °C, och den kallaste är juli, med 23 °C.^[5] Genomsnittlig årsnederbörd är 985 millimeter. Den regnigaste månaden är juli, med i genomsnitt 183 mm nederbörd, och den torraste är februari, med 2 mm nederbörd.^[6]

Klimatöversikt (förklaring)											
J	F	M	A	M	J	J	A	S	O	N	D
42	45	43	40	35	31	27	27	28	29	30	36
20	23	25	23	21	21	19	20	20	20	20	21

Latuka	
Stamområde	
Land	 Sydsudan
Delstat	Eastern Equatoria
Höjdläge	554 m ö.h.
Koordinater	4.64041°N 32.55208°Ö﻿ / ﻿4.64041°N 32.55208°Ö﻿ / 4.64041; 32.55208
Tidszon	EAT (UTC+3)
GeoNames	371514


Botit ja totuudellisuus

- 15 % Twitter postauksista oli botin tuottamia USA:n presidenttivaaleissa 2016
- ”Totuusanalysointoreita” tekeillä (mm. [Le Monde](#))
- [Faktabaari](#) Suomessa (crowdsourcing)

Social Media ‘Bots’ Working To Influence U.S. Election

November 7, 2016 1:27 PM

Filed Under: [Bots](#), [Election](#), [Social Media](#)



(ALAIN JOCARD/AFP/GettyImages)

KCBS
ALL NEWS
740 AM | 106.9 FM

SAN FRANCISCO (KCBS) — The way we vote on Tuesday might have been shaped by forces that humans can't detect, including the influence of bots in at least one of our social media feeds.

Computer science researchers at the University of Southern California created sophisticated algorithms to figure out, during a one-month period including presidential debates, how many tweets about Donald Trump and Hillary Clinton were generated not by humans, but by software.

“Looking at the content itself, but also the timing of the activity, the network structures, how information spreads from one account to another and so forth. Eventually we realized that about 15 percent of the accounts on the Twitter feed discussing elections in the U.S. are bots,” Research assistant professor Emilio Ferrara told KCBS.

More importantly, Ferrara says, these bots generated about 20-percent of the entire conversation.

Kuinka uutisrobotti toimii ?

- Esimerkki: USA:n koripallo- ja finanssi-uutiset
- Automated Insights:n *Wordsmith* ohjelma
- <http://www.bbc.com/news/technology-34204052>

Sports reporting

Short headline: **UNC** beats Louisville 72-71 on late Paige basket

Long headline: Led by a Paige game-winner, **North Carolina** defeats Louisville 72-71

CHAPEL HILL, N.C. — **Marcus Paige scored with nine seconds remaining** in the game to give North Carolina a **72-71 lead** over Louisville. **The Heels** held on to win by that same score following a **missed 3-pointer by Wayne Blackshear** and an **unsuccessful second-chance attempt by Terry Rozier.**

The Paige basket capped off a **13-point comeback for the Tar Heels,** who trailed 63-50 after a Blackshear 3-pointer with 8:43 left in the game. UNC finished the game on a 22-8 run to

Lukija ei
välttämättä pysty
erottamaan
ihmisen
ja koneen
kirjoittamat
uutiset toisistaan

Users' perceptions of automated content

Christer Clerwall

The advent of new technologies has always spurred questions about changes in journalism—its content, its means of production, and its consumption. A quite recent development in the realm of digital journalism is software-generated content, i.e. automatically produced content. This paper seeks to investigate how readers perceive software-generated content in relation to similar content written by a journalist. The study utilizes an experimental methodology where respondents were subjected to different news articles that were written either by a journalist or were software-generated. The respondents were then asked to answer questions about how they perceived the article—its overall quality, credibility, objectivity, etc. The paper presents the results from an initial small-scale study with findings suggesting that while the software-generated content is perceived as descriptive and boring, it is also considered to be objective although not necessarily discernible from content written by journalists. The paper discusses the results of the study and its implication for journalism practice.

Suomalainen projekti ”Immersive Automation”



Uutistuotanto automatisoituu – metatoimittajat tulevat

6.10.2016



Uutistuotanto muuttuu, kun rutiinomaisempaa toimitustyötä automatisoidaan. VTT ja Helsingin yliopisto kokeilevat, kuinka korkeatasoisia ja kiinnostavia uutisia tietokone pystyy tuottamaan ja millaisia käyttäjäkokemuksia niistä saadaan.

Palvellakseen yhä vaativampia yleisöjä useissa digitaalisissa kanavissa mediatilat pyrkivät automatisoimaan rutiinomaisimman toimitustyön. Näin toimittajat voisivat jatkossa keskittyä haastavampien erikoisjuttujen tekemiseen ja antamaan yleisölleen eläytymisen mahdollisuuksia yhä henkilökohtaisempien uutiskokemusten myötä.

Helsingin yliopisto ja Teknologian tutkimuskeskus VTT tutkivat alkavassa Tekes-hankkeessa automaattista uutistuotantoa, jossa datalähtöisyys ja koneoppiminen mahdollistavat yksilöllisemmän uutiskokemuksen. Keskeistä tälle ovat hyperpaikalliset uutiset ja yleisön osallistaminen.

”Käytännöksi muodostuu semiautomaattikka, jossa toimittaja viimeistelee koneen tuottamaa tekstiä ja määrittelee automaattiohjelmille uutismallit. Tulevaisuudessa kaikki toimittajat ovat jossain määrin tällaisia metatoimittajia”, uskoo tutkimusprofessori **Caj Södergård** VTT:ltä.

Automaatioaste nousee vähitellen

Tähän mennessä automaattikkaa ovat kokeilleet uutistuotannossaan isot toimijat, kuten yhdysvaltalainen uutistoimisto AP (*Associated Press*), esimerkiksi tilinpäätösanalyysien kirjoittamiseen. Talousuutisen lisäksi myös urheilu-uutisia tehdään jo automaattisesti maailmalla.

”On odotettavissa, että myös muita uutistyyppejä voidaan automatisoida tiettyyn pisteeseen saakka, riippuen tietyn datan saatavuudesta. Vaativampi journalistiikka – kuten pääkirjoitukset ja syvälle luotavat artikkelit – jää kuitenkin ihmisten tehtäväksi”, toteaa journalistiikan tutkija **Carl-Gustav Lindén** Helsingin yliopistoon kuuluvasta Svenska social- och kommunalhögskolanista.

”Helsingin yliopisto tutkii, miten datatiedettä voidaan soveltaa uutistoimintaan ja sen automatisointiin. Kehitämme tekstin louhintaan ja koneoppimiseen perustuvia työkaluja, joilla toimittaja voi helpottaa omaa työtään”, kertoo professori **Hannu Toivonen** Helsingin yliopiston tietojenkäsittelytieteen laitokselta.

OTA YHTEYTTÄ

Södergård Caj
Research Professor
+358505539356
caj.sodergard@vtt.fi

ASIAKASPALVELU



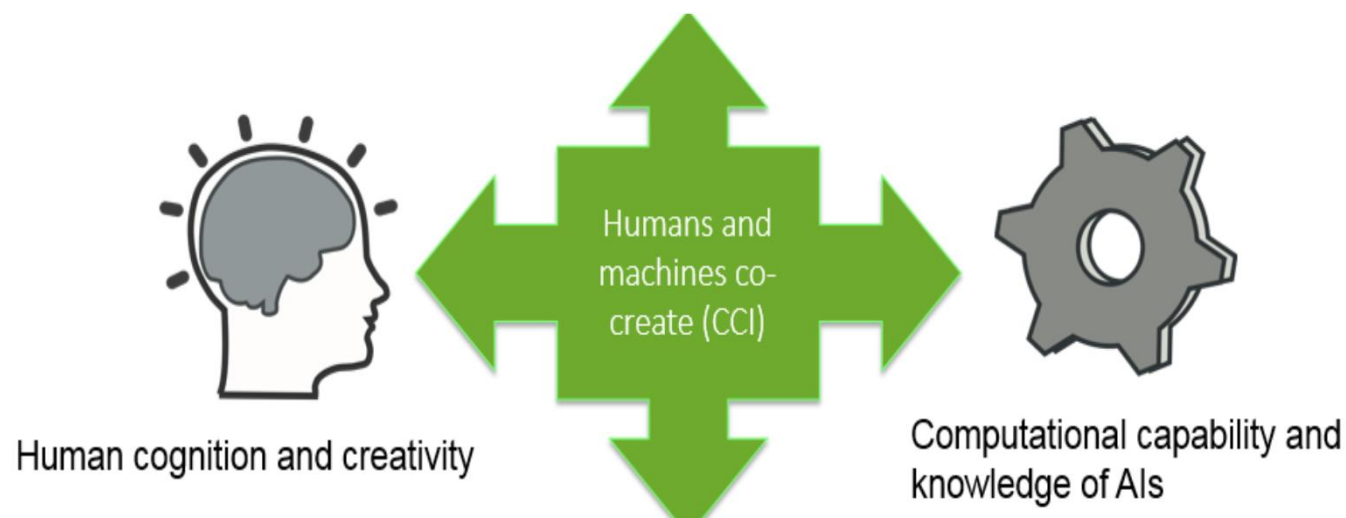
Sähköposti info@vtt.fi
Puhelin [020 722 7070](tel:0207227070)
Avoinna arkisin klo [9.00 - 11.00](tel:09009112405) ja
12.00 - 15.00

Lisätietoja:

Helsingin yliopisto
Hannu Toivonen,
tietojenkäsittelytieteen professori
Puh. [050 911 2405](tel:0509112405)
hannu.toivonen@helsinki.fi

Immersive Automation : tavoitteet

- Luo tiekartan ja demonstraation uutisten ekojärjestelmästä, joka perustuu
 - automaattiseen tarinakerrontaan
 - lukijoiden innostamiseen
- Tutkii, johtavatko dataan perustuvat tarinat vahvempaan uutiskokemukseen sisällön **kohdistamisen** kautta.



Projektitietoja

- Koko: 1,1 milj. € (93 henkilökuukautta)
- Kesto: 1.11.2016 – 31.5.2018
- 3 tutkimustahoa ja 9 yritystä ja organisaatiota
- Päärahoittaja: Tekes
 - Ohjelman "Media Remake" puitteissa

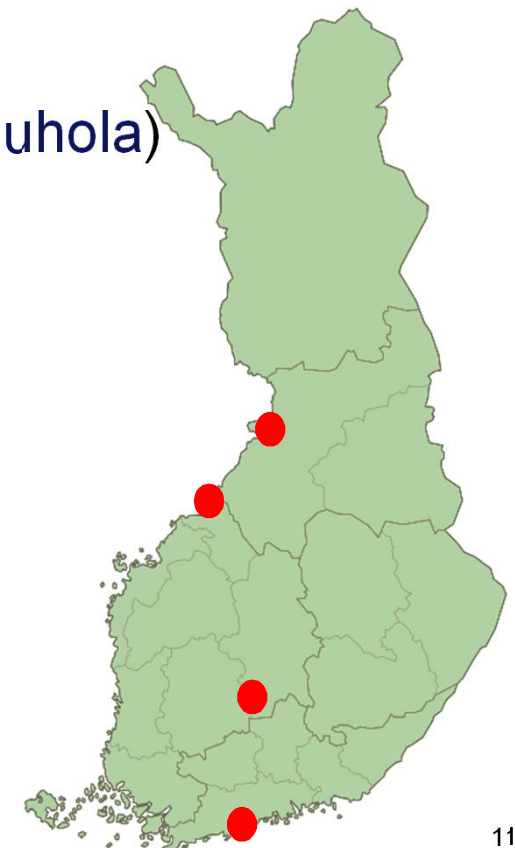
Osallistujat

■ Tutkimus

- Helsingin yliopisto, Social och kommunalhögskolan (Gusse Linden, projektipäällikkö, Klas Backström)
- Helsingin yliopisto, Tietojenkäsittelytiede (Hannu Toivonen)
- VTT Data-driven Solutions (Caj Södergård, Asta Bäck)

■ Teollisuus

- Viestinnän tutkimussäätiö (ohjausryhmän pj Helene Juhola)
- Sanoma (Vesa Lindqvist)
- Alma Media (Santtu Elsinen)
- Keskipohjanmaa Kirjapaino (Juha Erkkilä)
- Kaleva (Niiles Nousuniemi)
- KSF Media (Sami Kallinen)
- Svenska kulturfonden (Leif Jakobsson)
- Conmio (Tero Hämäläinen)
- Steamr (Nicke Nylund)



Yhteenveto: Potentiaalisia hyötyjä

- Nopeus ja skaala
- Tarkkuus
- Tekee toimittajan työn kiinnostavammaksi
- Voi tuottaa hyvin kohdistettuja uutisia ("long tail")
- Uudentyyppinen journalismi = Datajournalism 2.0

Yhteenveto: Potentiaalisia haittoja

- Toimii toistaiseksi vain rajatuilla aihealueilla (yritysanalyysit, urheilu,...)
- Vaatii kallista opettelua
- Läpinäkyvyys ja totuudenmukaisuus. Kirjoittiko ihminen vai kone ?
- Kulttuurinen vastustus
- Koneet erehtyvät myös (vrt. Netflix 2015)



TECHNOLOGY «» FOR BUSINESS

