Mapping the future: news, data and automation in the Nordic region

Carl-Gustav Lindén
Södertörn University/University of Helsinki, carlgustavl@gmail.com
Aggregated data from four projects

• Between a robot and a zombie: The radical transformation of news work (2015)
• Robot journalism in local contexts (2016)
• New business models in Nordic media (2016)
• Immersive Automation (IA): automated storytelling, audience engagement and user experience in a news ecosystem (2017-2018)
Evolving research

1. Mapping state-of-the-art: interviews in United States (Narrative Science, Automated Insights), France (Syllabs), Spain (Narrativa), UK (Yseop), Germany (AX Semantics)

2. Nordic interviews with Mittmedia/United Robots, Journalism++Stockholm, Nyhetsbyrån Sirén, Nyhetsbyrån TT, Textual Relations, Schibsted, Bonnier, Amedia, Aller, NTB, Ritzau Finans, Alma Media, Sanoma, Keskipohjanmaa


4. Creating R&D and learning project with five media companies, two tech companies, three research partners in Finland, collaboration with Swedish companies
Surprising spin-off

Forskningsprojekt ledde till mångmiljonbusiness

Aamulehtis redaktion. Bild: Suvi Tenner/Aamulehti

Ett forskningsprojekt inom ramen för det redaktionella laboratoriet vid Svenska social- och kommunalhögskolan vid Helsingfors universitet har utmynnat i ny affärsverksamhet värd nästan fem miljoner euro.

Det handlar om projektet "Robot journalism in local contexts", där forskaren Carl-Gustav...
The theoretical frame

- The social construction of technology (SCOT): Social action is a central element in technological development (Pinch, Wiebe, 1984)
What do we want to know?

• RQ1: How has current practices been developed?
• RQ2: Where is development in the field heading and how far has Nordic companies come compared with state-of-the-art?
• RQ3: Why should media and journalism researchers be involved in this development?
RQ1: How has current practices been developed?

• Very slow start compared with other industries
• Computers writing texts since early 70’s
• Narrative Science and Automated Insights setting the trend
• Individual data journalists finding an interesting opportunity to try something new, some later working with computer scientists
Different applications

Algorithms editing, aggregating, publishing, and distributing content:

• Automatically generated texts
• Monitoring data sources, creating news alerts
• Automatically generated first pages
• ...and all ”normal” automation that a newsroom is full of, especially software for content editing and management
User analytics automation

• “When we discuss or mention it is more like, building the algorithm that allows to understand what your interests are. So we could give you more of the content that we know that you're interested in. And less of stuff we know or think you're not that interested. So when we have your attention, we can make sure that we spend that, those seconds, minutes, on stuff that really matters for you”

Manager, Schibsted
Why do publishers want automation?

- Save resources and reduce costs
- Faster and less prone to errors
- Scalable: thousands of stories about same or similar events instead of one unique story about one event
- Easy to implement
- Potential to expand beyond traditional media business
- Almost indistinguishable from human written texts
- Multilingual solutions
Saving journalism?

• "Journalists have a lot to win with automation. Above all, we see this as a possibility to free resources for deepening journalism since reporters don't need to waste their time on manual tasks“

Journalism++Stockholm
• RQ2: Where is development in the field heading and how far has Nordic companies come compared with state-of-the-art?
Mittmedia: Rosalinda

• Writing football stories automatically before, during and after a game
• News angle dependent on user: home or away?
• Slack notifications to journalists about possible stories derived from the data that could be written manually
Sports a natural start
Some findings

• Automated text generation is an emerging field of data journalism (?) in the intersection between journalism, statistics and computer science.
• Current practices are rudimentary and dependent on templates manually produced by journalists.
• In the next phase deep machine learning with text mining and natural language generation (NLG) will be used to automatically produce texts.
• Learning algorithms are fed with massive amounts of journalistic articles to find out topic selection, story structures, news angles and presentation style
• …or comparing data and output to find out how articles are written
Long list of concerns

• Lack of resources and competence in journalism
• Algorithmic authority and accountability, manipulation
• User data restrictions
• Lack of expansive business model for automation, used mainly for cost savings
Data and analytics is key

- Acquiring, validating, normalising data
- Sports, weather, traffic, financials, monthly statistics (employment, crime), real estate prices all easy cases
- Unstructured data still difficult, but machine learning has solved the problem with irony and förkortningar
RQ3: Why should journalism and media researchers be involved in this development?
The future of news automation

- News natural field of trial for media companies
- Updates, insights and alarms
- Sensor data
- Data from monitoring human activity, especially intentions
New dimensions

- Robotisation in society means huge opportunities for media companies, creating systems of communication based on engagement and immersive interaction
- Multilingual applications
- But we also need journalists who can analyse and report on complex decision systems that are crucial to society
Stay tuned!

• International conference on news automation in Helsinki, March 2018
• For both academia and industry
• More information (alive 7 February) on www.immersiveautomation.com

Carl-Gustav Lindén
carlgustavl@gmail.com