Module 2 - News Algorithm guest speaker Carl-Gustav Lindén, Media Journalism Researcher

Hello everyone! Welcome back to week two of the MOOC.

I'm very happy to have Carl-Gustav Lindén with us today. Carl-Gustav is a media journalism researcher at the University of Helsinki and also a lecturer at Södertörn University in Stockholm. At the University of Helsinki he led a project called Immersive Automation which was all about using automated production of news content.

Carl-Gustav, hello, it's great to have you with us via Skype today. Thanks so much for sharing your time and knowledge with the MOOC. So, first off I wanted to kind of ask if you could fill us in a little bit more on this Immersive Automation project on news automation that you've been working on. What is that all about? and what did you sort of accomplish in that project?

Yeah great. Yes, This is my favorite topics. I started researching news automation, some call it robot journalism. Meanwhile, I will tell you why I don't like the word or metaphor robo-journalism, but we started starting to do work on news automation a few years ago and then realizing that this is something really exciting and we put together a group of media companies, tech companies, and researchers from the University of Helsinki to create a project named immersive automation and got funded for that to work for 17 months, which in research terms is quite speedy sort of timescale, but we actually get a lot done in a very few few months.

What we did was that we started with creating a prototype or writing news from the local elections in Finland in 2017. Election books and that's was really excellent in three languages. One of the aims of the project was to create multi-language applications so we had this in Finnish, Swedish and English.

And Finnish is in no way related to either Swedish or Finnish so that was the big issue, how to sort of develop a really system for languages which is rarely used and has no sort of support in computer science. So, we did that and the books worked very well.

We did a news, some updates and we also tested the sort of texts written by the bots on our real audiences and I had them read text by journalists and text written by their bots to see sort of the differences. And, then we actually get funded again for a new products that is now continuing as a EU project called EmbedIA.

Last week we had a kickoff meeting in Ljubljana in Slovenia, 50 people in the room so that is a much much bigger project. So yes, and during the product I sort of travelled the world to meet everybody sort of in the business of news automation, meeting service providers, journalists who work with news automation, researchers and so forth.

So I think they have a pretty good picture of what's being done, and I think they both shot this immersive automation and our bot, it is sort of the modularity, I mean we have sort of put together when you usually

have these templates or texts are written by journalists, but a nicely written template, a long text they used to fill in in the numbers.

Our system puts together sentences and is quite autonomously deciding what text should be like. You have some some basic rule, but otherwise, I mean, it is really a system which is more free to sort of decide by itself, and I think that this is something which is [skip in video] because everything I see about news automation is template based, text based.

I was very happy last year when I heard that a German company had produced 8000 templates, and I thought well now it's happening finally. So, I called them and nobody had repeated every one of these eight thousand templates by hand. It was a huge job. So we are far away from real automation, this is sort of very much handicraft, I mean, this is done by journalists. It's ,very little intelligence involved in the system.

So, you mentioned that you did some evaluation in the context of your project. I mean what were you testing? and what did you find in those evaluations?

One sort of key issue was is if there's a text written about something happening close to you, your party, your [video skips] than people are much more accepting the sort of data, the text might not be as up as good as they've ever written by a journalist, and the differences is much smaller than what you would think.

And sort of the people trust text written by these systems almost as much as journalists, so this is not the problem either, but the one thing with the credibility of automated texts is that it's culturally contested.

I mean, if you do this and say [video skips] you get totally different results than for instance in Germany or in France or in the UK. So you have something called the Hollywood robot syndrome which means that how do you feel about robots or is it a picture you get from about robots from say Hollywood films, quite much decide what you would think about something like this, so these are things you need to think about.

Yeah so, given your experience kind of working on, and building, and studying these these systems, what do you sort of see as the main benefits of using automated content in newsrooms?

Of course, these texts have been done by journalists as a routine, news, texts, small pieces of news, which takes a incredible lot of time to write and with little influence on the audience. I think today's journalism needs more elaborate journalism, more quality, more investigative journalism, this is what people are prepared to pay for. I don't think people are prepared to pay for these pieces of sort of you know "who won a game last night" type of stuff.

So I think journalists need to move up the scale, up the sort of ladder up the quality scale anyway and this is just one way of sort of making sure that there's still all this small pieces of news which are important for somebody, but which takes a lot of time to do that somewhere they are produced.

And I think that is sort of freeing up and sort of the time and energy for journalists to do things that really matter for them. That's sort of their greatest benefit here.

And so, on the flip side, I mean, what are maybe some of the drawbacks to using this type of content for newsrooms?

Of course, the publishers are always looking for a way to save money, and certainly you can save money by doing this, but I don't think it's a really good business model. I think, as I said, that people are prepared to pay for quality news, for unique news, news that are relevant to them, and not sort of a billion stories about dumping not that meaningful.

So I think that, you can save money of course. You can produce more stuff and still having a few journalists, and it might be a trade off. At least have some journalist staying if there's a problem with the business model of news media. This might be a way of sort of dealing with that, but of course, I mean, as I said, publishers want to save money because it's paying journalists is very expensive.

Right. So, to pick up on that a little bit more, I mean, do you see automated content impacting jobs in some kind of substance way? What have you seen automated content already putting journalists out of work?

Well, we have had automation in newsrooms for decades. I mean, you have photoshop, you have the word spelling check programs, you have all these kind of softwares, what used to be called artificial intelligence until you had to be your hands, actually working with it. So of course, I mean, darkroom assistant stuff disappeared from newsrooms.

I don't think journalists think this is such a big deal. Journalists are doing more and more stuff themselves with this new tools. So, in a way I mean everybody else sort of disappeared from the newsroom, but journalists are still there working. So, I think they should be happy about that. I don't see any impact from active automation on newssrooms per se.

But then again of course, the positive sort of spin is of course around the efficiency of journalism having more to smart do is telling you what's happening there. What is relevant to your audience. I mean metrics in any way. And I think those those sort of tools are making the jobs off of journalists more efficient. And it might be that generally you're not able to sort of. Work in that environment you might get laid off. I don't know if this might be an actual problem for some.

Yeah I mean I think in the in the U.S. I've seen in fact a few job like new jobs created in order to kind of manage automation or oversee it or develop it. I mean are you seeing any of that going on in the in the automation projects that you've studied or are people actually hiring in order to maintain or supervise automation.

I think we've lost the voice. Your voice I could repeat the question.

Yeah. So in the in the US context I've actually seen a few companies are hiring people so automation editors, supervisors, people to maintain, people to create data and so on for these systems. Are you seeing any of that in your studies of the systems that you've come across.

Yes of course there's new and new sort of categories in the newsroom but these people are very few. I mean maybe don't see a massive employment of let's say computational journalists. These are our special skills needed for a few but I don't think this will affect the sort of skills of the newsroom in a bigger picture.

What I was referring to was maybe a more sort of skills to use used new tools but maybe not to develop new tools. And I think journalists I mean I've journalists there are big big users of new technology and it needs sort of the [video skips] adaption sort of processes for journalists to be able to sort of gain gain from the development in this software business.

As interesting as one of the things that we're talking about this week in the MOOC is how to use a tool called Arria Studio, which is very much you know a template writing program as you know. So hopefully that's a chance for some of the students taking the course to start to get some of those skills using these tools.

So I wanted to kind of you kind of alluded to something earlier and I wanted to kind of come back to that. So you know going back several years now automated content production was originally kind of framed or popularized under the term of robot journalism.

And I know you've written some things and you've been sort of critical of that framing. Can you kind of explain that a little bit more. You know what's your view on this this whole robot journalism idea.

Yes a former journalist and I would say as an academic is I really want people to use the right wording for what they're talking about. We have a lot of problems with misunderstanding each other depending on the metaphors we use, the stories we rely on. I mean there's so much especially in innovation is that such a would I say VR or hype prone to hype and I think it makes our lives much easier if we use the proper words to describe define what we're talking about. And with robot journalism is that we have no mechanical parts. I mean these are software systems that are no robots coming to take the jobs but still illustrators of this piece of news or these analysts send these articles always have a robot there with the keyboard but it just doesn't make sense.

And and and there's several problems. One is as I said the Hollywood robot syndrome. Journalists tend to tend to be afraid of these sort of robot journalism thing. I think that these are going to take the jobs and if you look at the illustrations that actually seems to be the way even though these are not these intrusive systems.

And of course I mean that with regard to robots overall. I mean you have had some problems with that. I tried to sort [video skips] success I have so far I must say. We introduced a washing machine. Where I see

what journalists are doing they are sort of sorting the laundry, they are sorting the data, cleaning the data according to the sort of color fabric and you know what what temperature you should have in the washing machine and then use the right program and push the button and start button. And that's sort of I think in a way captures all the work that journalists did before and now can just have the machine to do. But as I said you know. I discussed this on many conferences, people laugh. I think this is funny but I see the robot journalists coming back everywhere unfortunately.

So we should be talking about the washing machine journalists is that the better metaphor I guess?

In your class you could have sort of the topic of trying to define, define a better metaphor. I don't think that washing machine is that good either. I would prefer something neutral like news optimization but let's see. Please let your students think about this.

So what I mean what do you think is the role of things like drones and I mean. I mean will we ever see sort of true robots coming into and coming into journalism like real physical robots or is that you know 50 or 100 years away.

Yeah I think it's probably a 1920s something. I wouldn't see why would we have robots in journalism I mean we have sort of software devices probably all kinds of I mean bots that do interviews for us, collect information, drones they get pictures just following something. Is a drone a robot? Not quite sure. But I believe and actually now habit having worked with sort of news automation for for years as a computer scientist. I'm really really interested in what makes journalists human. So my new research project is actually about passion in journalism. What is passion and how is that sort of an passion as part of the sort of everybody is talking about engagement. And engagement in the business [video skips] for media, which is important because you need to cut get the right sort of connection to your audience and make them have to pay for what you produce and not having Facebook and Google in between taking the money. And I think true authentic passion is one important part of the business model but it's never been discussed in journalism studies. It's just recently come up as something you know well we all know we've been working as journalism. Newsrooms are really passionate. I mean people are screaming almost crying every day. And but this is not covered in research. I think it's because I think is it would be an interesting way of [video skips] insights.

So that that's a really interesting sort of area for future research. I mean in terms of the automated production of content I mean what are you excited about there in terms of future research or are there interesting directions that you're thinking about? I mean where is this all going?

Yeah I mean this it's really interesting with 5G technology, mobile qualitative pathology and sensors everywhere producing information. And sort of how do we get this information? What is what is relevant for us? This really interesting project in research product in Germany called Ambient News in Hamburg around there where it sort of gets you know get on your smart mirrors in your bathroom for instance get news or from social singles. And so you sort of you're always updating with what you need to know about what's happening right now weather or traffic or whatever it is. I think those types of it's not really news but it's part of the new sort of repertoire. And these kinds of things I think would be really important with

alerts, alarms, updates from from these sensors from you know maybe an app moderated by a human or something.

Interesting. And so I mean what do you think news organizations should be investing in then? I mean where should they be. Yeah. So where and where do you think news organizations should be investing? I mean what should they be. Where should they be putting there are innovation funds and there's sort of innovation data science teams and computer scientists. What should they be building next?

It's a really good question. Fortunately very few media companies have these resources to build new stuff from scratch which just maybe needed. A sort of [video skips] providers and service providers for this. And it's really hard to say what the focus will be, but I would say that a focus on actually real journalism I think that's really an area which we tend to forget when you talk about metrics and analytics and data whatever. You need real journalists to do real journalism. That's really important. That's sort of the basis for the business model and as we add this new software services and algorithms and automation to [video skips] assistance. We need to make journalists even more human.

And this is something I think that that the CEO of Apple has said in books that you know is not concerned about robot machines becoming humans. He's concerned about humans becoming machines and I think we should think about that much more.

Yeah that's I think that that's a really interesting point. It's come up and a variety of contacts I'm studying as well as sort of you know how do we push back against people becoming too mechanical and sort of their role in these systems.

Well so stepping back is there any kind of I guess final parting advice you would offer the students of the MOOC as far as you know as they start to learn more about these tools and technologies and and even try some of them on in terms of like writing an Arria Studio. Yeah. Any any final advice for them.

Yeah I think it's important to think about the systems as interesting exciting but you might not need to have the ambition to become an expert on developing this stuff. But you need to be an expert in understanding what they do. What they're for. What's the logic behind? How to decide and debate, discuss these because as you're an expert on algorithmic decision systems overall and I think that sort also something we need to teach journalists, journalism students to understand how these systems work. Not just in journalism but everybody in society. And I think that's sort of a broadening the view from not just journalism and news media.

It sounds like definitely the stuff that we're trying to cover in the MOOC so that's so that's great. Well this has really been fascinating Carl-Gustav. Thanks again for spending some time with us and for sharing your knowledge with the students in the MOOC. It's was great to have you here.

Thank you for having me. It was nice. Thank you.