Introduction

It is not without reason that our current age is sometimes referred to as the media age. According to Mediebarometern 2020, the average Swede consumes 7 hours of media daily. The most frequently accessed media types are radio, tv, music, and social media (video games were not included in the study). At the same time, media lends itself well to AI applications: There is a huge abundance of digital data, a large part of which consists of text and images, two modalities where generative models such as GPT3 and DALLE have proved particularly successful. There are also strong commercial interests to motivate investments in new products and formats. It stands to reason then, that if AI comes to be widely adopted by the Media industry, and we spend almost half of our waking time on media, then advances in media AI are likely to have a pronounced impact on our daily lives.

In this Community Reference Meeting (CRM), we discussed how AI is being put to use in different media sectors, what helps and hinders the technological uptake and what are potential consequences for people and for society. Similar to previous CRMs, a number of themes were discussed in parallel at different roundtables. It was a hybrid event, with discussion tables spread over three locations: In Gothenburg, where the themes were AI for Text Generation, and How AI Transforms the Production and Consumption of Arts, Music, and Media. In Stockholm, where we held discussions on AI for Gaming, AI for Movie Production, and AI for Music, and in Umeå where the focus was on AI in Politics. Mark Harrisson, CEO of the Digital Production Partnership, kickstarted the event with a talk on AI for Video Production which was broadcasted live to all locations.

Important takeaways from the discussions included:
- leveraging the experience of Nordic countries as early adopters of automation tools, towards a positive use of AI in the sustainable transformation of the current global media landscape.
- how to deal with built-in, conscious and unconscious societal and ethical effects of AI, including biases, and lack of inclusion and diversity.
- the impact and importance of regulation of AI, as a superstructure of critical analysis for the use of AI in (social) media.
- addressing the complexities of semantic and contextual correctness of text generation, and how hybrid methods (rule-based and data-driven) can potentially be used for this effect.
- the potential of AI for game content generation, driving gameplay, and facilitating social media experiences in future computer games, including immersive experiences (VR, AR or metaverse).
- and in general, across all tables, how the sheer scale of AI-driven developments will transform media industries and our global connectedness.

Acknowledgement

This CRM was organised in collaboration with WASP Research Arenas – Media and Language. We would like to thank the individuals and organizations who generously shared their time, experience, and materials for the purposes of this event. The summaries of the round table discussions presented in the report are collaborative work and would not be possible without the engagement of all attendees representing different sectors and their reflections on the topic.
AI for Video Production

Mark Harrison, CEO, The DPP; Johanna Björklund, Codemill AB and Adlede AB, Umeå University

Main Challenges

- Maintaining the human voice and the sense of authenticity despite increasing automation
- Providing AI-enabled creative tools to a fragmented and heterogeneous community
- Finding sustainable combinations of technological solutions and business models, in particular with regards to data sharing throughout the media supply chain

AI in its many incarnations - speech recognition, object detection, video summarisation, etc - has come and gone in turns from the broadcasting industry’s hype radar. There seems to be a widespread disillusionment in the industry as to what AI has promised, compared to the actual impact that it has delivered. At the same time, AI research is progressing at a break-neck speed, with generative models such as GPT3 and DALLE pushing the limits of what is possible in terms of content generation. The following is an attempt to answer where these solutions could contribute, but also why they are not making their way into the market more rapidly. The text is largely based on Mark Harrison’s keynote on the practical impact of AI on the broadcasting industry, complemented by ideas and examples put forth in the following round-table discussion.

In his talk, Harrison observed that the industry has traditionally been supply-side driven, with broadcasters releasing their content at their own pace, but that OTT solutions are causing a shift towards an ecosystem driven by the demand side, where a transactional relationship needs to be established between broadcasters and viewers. The media supply chain can broadly be divided into three parts: Production, packaging, and monetization. AI has the potential to contribute to all of these. Looking first at production, AI can be used to improve the user experience of premium content, e.g., by providing better special effects, visualisations, and ad placements. At the other end of the spectrum, we have high-volume, short-turnaround content. Here AI can add speed, allowing local sports or influencer videos to reach their audiences more quickly. However, the middle segment, which makes up the greatest volume and accounts for the largest revenue, is fragmented into isolated productions, which has here-to-made it difficult for technical innovations to gain traction. Many producers are still struggling with the most basic forms of automation, so an important challenge is to design and serve broadly applicable AI solutions in an accessible way.

Moving on to packaging, we are looking at tasks such as compliance, preparation, and regionalisation. Compliance is vital to adhere to local regulations, but is time-consuming and may incur heavy fines if violated. This points towards AI-based support systems for human operators, to allow them to work more efficiently with maintained accuracy, rather than fully automated systems. Preparation then ensures that the content can be consumed through the target channels. This includes transcoding and translating the content to meet various formats and standards. A promising area for automatisation, but more likely using rule-based approaches as opposed to machine learning. Lastly, we have versioning, where the content is adapted for different audiences and locales. Here AI can contribute with automated dubbing, lip-syncing, and semiotics.

The final step in the production process in monetization, where the aim is to translate content into revenue. AI could in theory provide valuable tools for understanding audiences and informing future production projects. However, the platforms over which the content is produced is seldom owned by the content producers, which may therefore have to resort to buying data back about their own productions. Organisations are slowly becoming more willing to share their data, as the insights and business models that this unlock promise to out-weigh the perceived IP risks.

The Nordic countries have traditionally been early adopters of automations tools, as this is often a necessity to cater to small audiences without overstressing the production budget. In general, innovation tends to happen where there is little money, for example, the coverage of college sports or regional news reporting. Over time, this can come to have a transformative impact on the media landscape, and if we’re lucky, provide more sustainable business models that balance demand- and supply-side interests. The main challenges remain: How do we create AI solutions that scale, are trusted, and don’t take away from the viewer’s sense of authenticity?
The roundtable brought together a diverse group of participants from different academic backgrounds, ranging from the social sciences to mathematics and engineering, to address issues related to the use of AI-driven automation in political contexts online. This was both in terms of strategic and large-scale campaigns by parties and governments and from a more grassroots-oriented perspective, focusing, for example, on hacktivist uses of social bots. Representatives from the business sector were also present. The below paragraphs summarize the starting point for our discussion and our key conclusions.

The roundtable raised several issues relating to critical perspectives on AI and politics. We discussed how critical perspectives should be positioned in relation to ethical AI and how the vocabulary used when discussing critical aspects of AI and how you approach it is heavily influenced by your field of research and academic background. Furthermore, how terms such as ethical, fair, transparent, and responsible sometimes act as synonyms to ‘critical’ but also make it possible to avoid truly critical perspectives. Critical analysis of AI and ethical AI do not exist in a vacuum independently of each other. Critical studies on AI may instead begin where ‘ethical AI’ ends. A more profound difference between the ethical and the political dimensions is that ethical AI has traditionally been focused on AI in relation to the individual, not society and collective interest, in which critical analysis would take its starting point.

Further, we raised the concept of regulating AI, as a superstructure of critical analysis of AI, and how that could be achieved in a democratic way. This would have to proceed from a holistic point of view as it is closely related to politics driving AI. Although monetary forces from the industry are a central part driving the politics of AI, there is also a strong belief in technology from politicians and society as a collective. Tech is seen as non-political, as “success,” and even as a form of “collective success” connected to the national identity. From this emerges a driving force to be at the forefront of the technological arms race, where AI is the term you want to be associated with. AI has become the “it-word” word of our time and is closely associated to funding and flows of money. To get funding for research or in the private sector, it is often beneficial to have AI ‘in the title’ or claim to use AI in some way. In the academic setting, this has tended sometimes to research conducted in the borderlands around AI having to reformulate or compromise on initial (‘critical’) ideas in order to get funding. In the private sector, things have sometimes gone so far that people have falsely pretended to use AI while outsourcing these very tasks.

Politics in AI can be at least two separate things. Politics driving AI and politics within AI. Politics driving AI refers to the underlying ideology behind AI and the forces contributing to its continued development. Examples of such ideologies are the Californian ideology, Silicon Valley evangelism/innovationism or other forms of technological rationality, meaning; solving tech-oriented problems with more tech. Politics within AI, on the other hand, refers to how AI may function as a vehicle for ideology. AI may have built-in, conscious as well as unconscious biases: racism, sexism, erasure of non-binary identities to name a few, which have real-life impacts on healthcare, policing, and other parts of the public and private sectors. We discussed how research should approach these issues critically without condemning AI. Can the use of participatory design be a way around the problem? Should it be the job of AI to solve racism? Is it ‘simply’ a medium? Or is it a paradoxical loop: Racism in society -> racism in data -> racism in (“correctly” trained) models -> racism in society.
AI for Text Generation

Aarne Ranta, Chalmers University of Technology and University of Gothenburg; Joel Jonasson, Blast Bit Enterprises AB; Måns Magnusson, Uppsala University

Main Challenge

- Combining semantic correctness, fluency, and productivity. Different applications may prioritize correctness and fluency differently. Productivity is particularly relevant for languages that lack resources both for symbolic and data-driven methods; these are often also the languages that would profit the most from the availability of generated text.

Automatic text generation provides a cheap and fast way to create content in natural language. Recent systems based on machine learning, such as GPT-3 and BERT, can generate text that is fluent but whose content is difficult to control. In contrast to this, classical rule-based methods, such as text robots, can be reliable in rendering content, but often in a clumsy and repetitive style. The round table will gather researchers developing methods of both kinds, as well as users of text generation from industry and society. We will discuss the choice of optimal methods for different use cases, as well as ideas for combining them in fruitful ways.

What are your most important projects in text generation, if any? Projects ranged from task-oriented applications (in the sense of applications directed to specific tasks such as games, e-commerce, journalism, and Wikipedia articles) to the creation of large neural language models. A cross-cutting topic was text generation in low-resource languages.

What are the hardest problems in text generation, in your experience? A common problem in task-oriented applications was guaranteeing the semantic correctness of the generated text. Another problem was the amount of work and expertise required. For large neural models, the problem of energy consumption was mentioned. However, this is a one-time cost when the model is created and does not affect subsequent uses of the model.

What are the most promising application areas, in your opinion? The task-oriented applications were all seen as promising since there is so much that can be done to make the dissemination of information more efficient. Cost saving is one obvious advantage (as manual translation is not needed), but in many cases, speed is even more important. Therefore, even if it is costly to build systems that are reliable enough not to require post-editing, the speed aspect can justify the cost. In robot journalism, stylistic post-editing is tolerated as long as the facts in the text can be assumed to be correct.

What are the most promising technologies for your purposes? Task-oriented applications are dominantly addressed by rule-based methods, which can guarantee semantic correctness. On the other hand, machine learning methods can give better fluency, and they can also help bootstrap rule-based methods by automatically suggesting rules.

The concluding question of the roundtable was: what did you learn from the others? One repeated answer was that rule-based methods were so widely supported, at the same time as machine learning had the most publicity.

Interestingly, similar conclusions are made in a wider context in a recent survey "What Do NLP Researchers Believe" by Michael et al (2022), under the term "false sociological beliefs". An example was the statement Q5-1, "Currently, the field focuses too much on scaling up machine learning models". 72% of the participants supported this statement, but the average belief of the amount of support in the community was 58%. The statement Q2-2 "Linguistic structure is necessary" was supported by 50% but believed to be supported by 38%.

Combining semantic correctness, fluency, and productivity. Different applications may prioritize correctness and fluency differently. Productivity is particularly relevant for languages that lack resources both for symbolic and data-driven methods; these are often also the languages that would profit the most from the availability of generated text.
A Shift in Culture Through AI

Kıvanç Tatar, Chalmers University of Technology; Petter Ericson, Umeå University; Kelsey Cotton, Chalmers University of Technology

Main Challenges

• What are current AI technologies for music, visual arts, and media?
• What does it mean to have infinite content? What does it mean to have large scale content created by humans vs created by machines? How does this shift in data affects the status-quo of cultural capital, and labour put in data generation?
• How does the new paradigm shift brought by advancements in AI challenge the status-quo in artistic practices? Which technologies are easily accessible, to whom? Who are the stakeholders? Who is involved in decision-making in technology creation? How do the roles of audience and artists shifting? What are the consequences of these technologies in the working conditions of artists? How is the understanding of intellectual property shifting?
• What kind of changes would we like to see, that is brought by AI technology in music, arts, and new media? What kind of considerations should the (AI/+artistic) practitioners have in this new era? How do we initiate cross-discipline and cross-community communication? Where does this new art paradigm situates in our daily life?

Increasing ubiquity of computing resources has enabled many to try, run, and observe AI technologies for music, visual arts, and media. The lines and borders between the audience and the artist are becoming less clear as AI content generators have become more accessible — as simple web pages for novice users, open-source code, and open-access models for experts. The explosion of content generated by such AI models has initiated a cultural shift in arts, music, and media, where roles are changing, values are shifting, and conventions are challenged. The readily available, vast dataset of the internet has created an environment for AI models to be trained on any content on the web, whether that content is collected by scraping the web, or via ready-made datasets and APIs provided freely by researchers, organisations and practitioners. With AI content generator models shared openly, and used by many, globally, how does this new paradigm shift challenge the status-quo in artistic practices? What kind of changes will AI technology bring into music, arts, and new media?

Data Labour Shift through Tools. The increasing accessibility of AI tools for generating cultural content, encompassing music, image, and video, has provoked substantial debate within popular media discourse regarding the decreasing societal value of skilled and artisanal labour in cultural industries. In this discourse, the perception is that AI generation is sufficiently human-like in its production for a variety of tasks, thereby posing a substantial-enough threat to the livelihoods of many artistic practitioners, who express concerns about the potential for the wider community’s perception around the value of human-created art to change. In that discourse, we should be mindful that at every point of AI technology creation a human is involved, and the AI technology is in close relationship with the conceptualisations and values of their makers.

Data, AI, Sustainability, and Ethics. Fundamentally, the accessibility of AI tools has driven a paradigm shift in regards to data. Data— as both an act of labour and form of cultural capital—implies processes of collection, surveillance, and categorisation, which raised philosophical discussion around the need for artist’s to “know our own profit”, and how our “profit” is utilised in non-reciprocally profitable contexts. A discussion on societal aspects of AI is entangled with the concepts of sustainability, ethics and values. The broader framing of these issues-focused on the evolution of culture, and impact upon culture industries—has implications on the environmental and climate impact of leveraging AI technologies within artistic practices. Within discussions around sustainability are also broader concerns with social
sustainability of artists, and the sustainability of the role of art within society. A core concern in this regard is how advancements in AI technologies deployed within artistic contexts may adversely impact the cultural economy of artists in society, and further how their role in society is implicated by the wider usage of AI tools. It is known that AI models reproduce problematic stereotypes in the data, which also may appear in artworks that are generated using these models. In the case of tools such as Dall-E/Midjourney/Stable Diffusion, it is worth noting the difference in the quality of the images generated when the prompted text used for generation refers to images, symbols or art that belongs to the hegemonic culture (undoubtedly of easier access) when compared to the outputs of prompts that relate to more “obscure”, underrepresented or non-hegemonic cultural manifestations. Cultural hierarchies could indeed be extrapolated from how much detail, quality or “realism” (or lack thereof) the generated images finally contain.

**What kind of changes would we like to see, that are brought by AI technology in music, arts, and new media?** Although still in a period of intense formulation and development AI technologies will fundamentally change society and culture. To this end, there are several key propositions and areas that we would like to see positive and transformational change within. Primarily, we wish to see fundamental power shifts and inclusive distribution of power in regard to the decision-making of all AI system designs. An important aspect of art and life in a digital world that is further highlighted and further highlighted and complicated by the introduction of widely available and powerful AI technologies is that of copyright. The exact reforms necessary are not entirely clear, but discussions within e.g. various pirate parties, digital rights groups, hacktivist organisations, open-access global initiatives of technology replications (such as GPT-NEO, GPT-J, StableDiffusion and DALL-E 2 PyTorch replication) are significantly more instructive than the various rear-guard actions fought by e.g. the IFPI and various other industry organisations. What is entirely obvious is that there is an immense need for reform of the current regime of “might makes right”, wherein large corporations can infringe copyright with absolute impunity. The discrepancy between corporate copyright transgressions in the creation of enormous datasets as input to AI systems with no knowledge or consent from the copyright holders, as permissible, and the automated “copyright” takedowns of private individuals’ meticulously researched fair-use remix and commentary works on platforms such as YouTube, Amazon, etc. highlights the differences in power and punishment across private versus industrial utilisation of AI technologies. Legislations, communal guidelines, and ethical dimensions of AI technologies for artistic practices are ongoing societal discussions. It is long overdue, and the perfect moment in time, to include artists and practitioners proactively in these discussions, accommodating and fusing their voices and knowledge in the reformation of equity, equality, and justice in AI technology creation. Art is for everyone, and the tools we use to make art, especially AI tools, should enable and empower just and equitable creation.
This roundtable discussed Game AI and the opportunities and challenges of using AI in computer games. The group at the roundtable gathered experience from many different perspectives, from academic research, the game industry to independent SMEs in media.

The call to the roundtable questioned if Game AI in today’s game could be somewhat deceptive. Most characters and interactive worlds in computer games today, like the enemies in a first-person-shooter, or the Sims are not examples of learning agents, i.e. advanced artificial intelligence. They’re scripted boots that give the perception of smart behavior on the screen; however, they don’t learn or develop new behavior over time as we imagine in today’s AI research.

A consensus from the roundtable was that classical Game AI is mainly used to operate NPC (Non-playable Characters) and behavior in games, not modern ML. Today’s game platforms also use AI/ML for game and player data analytics. However, also here, most commercial systems still mainly use classic heuristic models and underutilize modern AI’s power to learn and adapt behavior over time. Analytics should be better in utilizing modern AI to drive game creation and play.

One of the critical points that Mark Harrisson made in the keynote: The Impact of AI on the Broadcasting Industry, was on how traditional methods, such as creating content and special effects, don’t scale and point to the emergence of AI solutions. The roundtable noted similar trends in Games. However, putting AI in operation and scaling content creation with AI is a major challenge. Beyond technical challenges, legal issues, versioning, recommendations, culture, and tool-making are significant obstacles beyond the technical challenges.

Another topic for discussion was that much attention is now put on generative models for visual content, and there is a gap in other media, such as music and animations. For example, how can AI be used to control sounds and make soundscapes that don’t repeat and are perceived as more natural. Another big topic is the use of language models. Aside from the apparent use in creating dialogues and conversational agents, language models will also significantly improve other models by reasoning and deduction.

The roundtable then discussed how the gaming industry still misses the human-in-the-loop perspective. This will ultimately show how AI in games affects engagement and emotions in gameplay, but also more imminently helps game designers with better tools.

Another point raised was that some studies suggest that routine tasks leverage creativity, and by removing these craft skills, AI tools might have the opposite effect and hamper rather than drive creativity. In the end, games need to be fun, which comes from a highly creative mindset driven by game developers. The fundamental question here is whether AI can mimic imagination or is better suited to support the creative process.

The roundtable also discussed the ramification of legal issues that few know how to navigate, and there is an urgent need for more clarity. For example, music has standard practices on how many beats can be re-used. These guidelines are different for visual content.

Despite all the above challenges, the roundtable agreed that we are in a good place. Given the progression in AI and the needs in the game industry, there are clear opportunities. We need to come together better as an industry since several of these matters can only be solved by some players. We also need to identify what can be done here and now to accelerate the uptake of AI in computer games.

In conclusion, we see clear and exciting trends in using AI for content generation, driving gameplay, and facilitating social media experiences in future computer games. The call for the roundtable also mentioned trends, such as Metaverse, VR, and AR, a bit further down the line, and we need more time and opportunities to discuss these aspects in our future dialogs.
AI for Music

Bob L. T. Sturm, KTH Royal Institute of Technology; Andre Holzapfel, KTH Royal Institute of Technology

Main Challenges

- Issues of ownership and control over the resulting creations
- Enabling co-creation instead of push-button creation
- Economic problems for professional artists

This roundtable focused on the involvement of artificial intelligence (AI) in a variety of music practices, from listening, to composition and performance, to analysis and criticism, and even legal aspects. The motivation for this theme comes from the increasing application of AI in generating personalized music in the commercial sphere. This is sure to have a large impact on how music and musicianship are viewed and valued.

The following challenges resulted from the roundtable discussion:

1. AI in music can enable access to and create opportunities for creation. It can lower the “skills bar” and invite and empower people in creative activities. It can also enhance expression. However, there do exist problems with ownership and control over the resulting creations. For businesses offering personalized music for use in commercial works, this means they must somehow guarantee their customers ownership over the musical material generated by the company’s music system. This means building into the system “plagiarism detection” – which is an outstanding challenge that involves engineering, music perception, and law.

2. A major focus of work applying AI to music should be on co-creation, and not push-button creation – where a finished work results from minimal user effort. We should try to limit automation, which is a perspective that may not find much appreciation in the commercial realm where reducing costs is an objective.

3. Even outside the application of AI to music and art, there exist serious economic problems for professional artists: their ability to live to do their art becomes far more difficult as their wage-earning activities are automated or made redundant. The consolidation of power in the hands of few via AI is also a concern.

The following question was asked at the conclusion of the session: “Do we live in a unique time? Aren’t we asking the same questions that were asked of book printing, photography, computer aided design, and 3-D printing?” Yes and no: the uniqueness of our moment may arise from the sheer scale of these industries, and our global connectiveness.

References

The vision of the Wallenberg Artificial Intelligence, Autonomous Systems and Software Program – Humanities and Society (WASP-HS) is to realize excellent research and develop competence on the opportunities and challenges of artificial intelligence and autonomous systems with a strong investment in research in humanities and social science.

The WASP-HS program is planned to run 2019 – 2028 and will form an independent and parallel program to WASP, The Wallenberg Artificial Intelligence, Autonomous Systems and Software Program.

Request or more details: contact@wasp-hs.org

**Authors**
Johanna Björklund, Department of Computing Science, Umeå University; Co-Founder, Codemill AB and Adlede AB
Kelsey Cotton, Department of Computer Science and Engineering, Chalmers University of Technology
Virginia Dignum, Department of Computing Science, Umeå University
Petter Ericson, Department of Computing Science, Umeå University
Mona Forsman, Research & Development, Adlede
Mark Harrison, CEO, Digital Production Partnership (The DPP)
Andre Holzapfel, Division of Media Technology and Interaction Design, KTH Royal Institute of Technology
Joel Jonasson, Blast Bit Enterprises AB
Simon Lindgren, Department of Sociology, Umeå University
Mårn Magnusson, Department of Statistics, Uppsala University
Aarne Ranta, Department of Computer Science and Engineering, Chalmers University of Technology and University of Gothenburg
Henrik Sigurdh, Department of Sociology, Umeå University
Bob Sturm, Division of Speech, Music and Hearing, KTH Royal Institute of Technology
Kivanç Tatar, Department of Computer Science and Engineering, Chalmers University of Technology
Konrad Tollmar, Research Director, Electronic Arts

**How to cite this report**

**Design and typesetting**
Tatyana Sarayeva