Schibsted

EU Whitepaper on AI

Reply to EU consultation

Schibsted is a leading media and digital consumer brands group in the Nordics. Rooted in democratic values, Schibsted has a long heritage of contributing to society. Millions of people interact with Schibsted brands every day to become empowered in their daily lives, e.g. through reading our newspapers, purchasing second hand goods on our marketplaces or utilising our financial services.

Schibsted has developed its values through 180 years of free and independent media, founded on a strong awareness of the social responsibility that characterizes publishers' activities. A key guarantor for upholding these values is the principal owner, Tinius Trust, which is also making sure that the values are brought forward into a complex digital world.

Introduction

Schibsted wants to contribute to the AI debate by providing the perspective of a modern, data-driven European media company. We work with various aspects of machine learning and primarily focus on use cases within computer vision and Natural Language Processing (NLP) with AI services in internal processes as well as user facing products and services across our ecosystem.

Examples of current AI efforts in Schibsted include:

- **Newsroom tooling**
  Schibsted newspaper Svenska Dagbladet (SvD) has created the newsroom tool 'Oraklet' which helps editors decide when to put content behind the paywall. Oraklet analyzes published articles, identifies when certain reader interest levels are reached, and indicates whether the article is suitable to be put behind the paywall (and thus for ads to run which would recruit new subscriptions) through a notification to editors. The final choice of putting content behind the paywall is made by human editors.

- **Editorial insights**
  Schibsted newspaper Bergens Tidende (BT) is using computer vision to better understand who they are telling stories about, and specifically, what people the imagery on their website depict. By estimating the age and gender of faces used in an article's imagery, BT's application of computer vision enables insights into how their news coverage relates to the demographics of their audience.

- **Ad category suggestions**
  At our marketplaces, we are using computer vision to identify what category an uploaded

ad image should belong to. If a user uploads a picture of a chair, we use AI to know that the ad belongs to the category of 'furniture' and should be tagged with 'chair'. This makes the ad insertion process smoother for the user, but it also makes the overall quality of the service better as more ads will be categorised in the correct category.

- **Content moderation**
  Another example of AI in our marketplaces is found in content moderation. To further protect our users' safety, we are supporting our human content moderation team with computer vision solutions that can identify content that contains sexually explicit imagery or images where text is inserted in the uploaded image (which tends to be a proxy for fraud; malicious actors want to bring users away from established platform and into their private phones, e.g. if they are trying to sell illegal goods).

  We are currently exploring the use of AI models to identify harassment, toxic and/or offtopic comments to our news articles. Identified comments can then quickly be evaluated for compliance with our debate participation guidelines by our human moderators. Comment moderation is essential for constructive debate where harassment and toxic comments are minimised, and curbing such will allow wider sections of the public to participate in the public debate.

- **Distribution route optimization**
  Schibsted-co owned company Distribution Innovation employs AI to optimize the delivery route and ensure that the most efficient path is taken when delivering parcels and newspapers. This AI-powered planning tool is particularly important for the delivery routes of cars and trucks, as the optimal route will minimize fuel use and reduce co2 emissions.

- **Print prediction**
  Every day newspapers are sent from our distribution centres to various retailers such as Seven Eleven. If there are unsold papers at the end of the day, these are sent back to the central and retailers get a refund for the returned copies. Through an ML solution that predicts how many papers a specific store will sell during a given day, we can now better predict how many papers to print and distribute to specific stores to avoid selling out and to minimize waste.

AI is proving its value across our portfolio. While still a young field, we are convinced that AI has immense potential to strengthen journalism and news practices. If AI is used in good and responsible ways we believe these new tools can e.g. prevent filter bubbles, increase diversity in media representation and contribute to sustainable financing of media operations.

Issues for the Commission to consider

In the following section we present a number of issues for the European Commission to consider in regards to AI and its impact on the media industry.

The global nature of AI

The Commission writes that there is uncertainty as regards the allocation of responsibilities between different economic operators in the supply chain. The Commission also writes that many

actors are involved in the lifecycle of an AI system. Related to this, we believe it to be important for the Commission to note that the data employed for AI often comes as a result of global community efforts. The open source community (often in some connection to the academic world) is building training data sets that are used by companies across the world (e.g. Open Images).

No industry or geographical region is isolated in the age of AI. We are all interconnected, and isolating European players (whether as companies, industries or nations) is not not the answer. Building alliances and working together is the way to really take lead in the age of AI. While we support the notion of increased regional collaboration, not least in order to sustain competition from the technology giants, we must stress the importance of voluntary participation.

In the Whitepaper on AI, The Commission writes that it will 'strive to export its values across the world'. In reality, however, much of AI is built upon work (algorithms, machine learning models, datasets, etc.) stemming from primarily American or Chinese organisations (e.g. Amazon, Google, Facebook, Yahoo, or Tencent models, processors or datasets). With this background, our belief is that the European ambition should be not to *export* our values but rather to identify strategies on how to *safeguard* them as we implement non-European AI in important functions across the European community.

Nuances of AI

It is important that the Commission uses the appropriate definitions of AI. In policy and industry settings, 'AI' is often used as a catch-all term for various forms of advanced data analytics. The EU needs to clearly indicate when it is discussing the full umbrella of AI, or when the discussion refers solely to machine learning, deep learning, algorithms and other relevant terms. This is to ensure uniform application and legal certainty across Europe.

In the Whitepaper it is noted how AI can be 'used by online intermediaries to prioritise information for their users and to perform content moderation' and how this in turn may create issues regarding free speech, privacy etc. While this is true, the very case of content moderation offers an example of AI systems such as computer vision used to protect and empower digital consumers/users. These tools can be used in ways that truly empower digital consumers, and one such example can be found at Schibsted marketplaces as described above.

The misuse of computer vision technologies such as facial recognition and surveillance are central democratic issues. In future efforts for safeguarding European citizens' privacy and wellbeing, we want to urge the Commission to ensure that any potential future legislation in the field of computer vision pinpoints the particular challenges related to use of such technologies. Furthermore, any future legislation - on top of existing privacy legislation - should be specific enough to not stifle European innovation in general.

Broad AI literacy is needed

The Commission notes how education and reskilling is needed to advance the European workforce. We believe an addition is needed in this debate regarding educating consumers and

citizens. All Europeans will be affected by AI, and in order to enable an 'ecosystem of trust' we must ensure that not just the workforce but also e.g. children, digital laggards and consumers broadly have a basic understanding of AI and its impact on their lives.

Global technology companies are currently providing teaching materials offering the general public insight into the basics of AI. This is double-edged; widespread information is needed, but in an attempt to safeguard European interests we should be careful about letting private (often American) companies shape how the European public understands - and subsequently uses - AI. We suggest that the Commission considers methods for coordinating European educational efforts.

Data fit for AI

The Commission writes that machine learning and deep learning is 'characterised by limited interpretability, the need for a large volume of data to train the models and learn through correlations'. According to our experience this statement does not necessarily show the full picture. What is important to have in mind is that the key factor contributing to the success of a machine learning model is not the amount, but the *quality* of the training data.

This being said, access to data will of course be important in order for European businesses to be able to build and utilize AI. Related to this we would like to make clear our positions on three points related to data collaborations, data sharing and privacy legislation:

- **Providing room for companies to collaborate with other players around data is essential.**
  In this respect we particularly believe that *public sector* data from across the EU should be opened and made available so that smaller players can experiment and build new (AI) services based on such data. Furthermore, it is crucial that this data is provided in a usable format and is of high quality to ensure that its use can be optimized.

- **We are strongly against mandatory data sharing.**
  We also believe that it will be important for AI purposes that the EU should serve to facilitate and make arrangements for the *voluntary sharing of data between private sector entities*. It is important that any such arrangement provides actual value to everyone that contributes. Furthermore, we believe that there must be room within competition laws for individual players to collaborate with data where practical and valuable. We believe, however, that measures that would mandate sharing of private sector data are drastic, and are opposed to any stipulation that would obligate private companies to share data with other private entities against its will. For an organisation like Schibsted, e.g. language data (i.e. our archives of articles) is a highly competitive asset that we must be allowed to safeguard. Collaboration and privacy legislation is of utmost importance, but it must ensure a level playing field.

- **We need better regulation and enforcement to ensure we have predictable access to our own data/data about our users.**

Schibsted

Enabling European digital players to manage data fit for AI might require clarifications regarding how the General Data Protection Regulation framework should be interpreted in regards to AI innovation. At Schibsted, we believe that the notion of 'regulatory sandboxes' discussed in the European Data Strategy (2020) may be a fruitful path to explore in regards to this. Furthermore, focus should be put on *enabling* use of data for AI purposes in a good responsible manner where in particular data subjects' rights regarding transparency and control are safeguarded and that privacy by design is embraced as a leading principle.

For further insight into our AI efforts, please watch this webinar: https://bit.ly/2UrMXKY