Briefing Paper
Sustaining Trusted Journalism in the AI Age

The News Media Association (the “NMA”) is the voice of UK national, regional and local news media in all their print and digital forms - a £4 billion sector read by more than 47.2 million adults every month. Our members publish around 900 news media titles - from The Times, The Guardian, The Daily Telegraph and the Daily Mirror to the Manchester Evening News, Kent Messenger, and the Monmouthshire Beacon.

Key Points:

- Unique, reliable and high-quality journalistic content is of immense value for Generative Artificial Intelligence (GAI) systems. News is fuel for Large Language Models (LLMs), and it costs money to produce. Big Tech-backed systems cannot simply take such content without compensation any more than they can help themselves to the electricity which powers their services.
- GAI systems are particularly susceptible to disinformation and error, making professionally produced news all the more important. Generative outputs should include clear and prominent attributions, allowing publishers and citizens to understand when trusted content has been used by LLMs.
- Foundation models are often owned or used by global tech giants who will, if allowed to continue unchecked, only entrench their market power and dominance. The Digital Markets, Competition and Consumers Bill will be critical in preventing a few Big Tech-backed players from setting ‘the rules of the game’.

1. From a copyright perspective, generative AI tools impact publishers in two ways; on the input or training side and on the output side. GAI companies are infringing copyright law during the model training stage, building GAI tools on the back of the often unauthorised and unremunerated use of copyright works. Google Deepmind trained its AI using a tool called the CNN/Daily Mail dataset – a million news stories chosen because they share the same logical story construction of headlines and bullet points.

2. The text and data mining (TDM) that trains AI systems is currently only permitted for the purpose of research for a non-commercial purpose. The UK should not be tempted to adopt a more permissive TDM regime. Tech companies are highly likely to use news media content without attribution, reward or opportunities to monetise through advertising. It will be particularly difficult to police because GAI content will be created in response to individual users’ requests.

3. The opacity of LLMs is a major stumbling block when it comes to the enforcement of rights, and ensuring consumer safety. GAI firms must be compelled to make information about systems more readily available and accessible. Generative outputs should include clear and prominent attributions which flag the original sources of the output. This will
allow publishers to understand whether their content has been used, and allow citizens to understand whether the outputs are based on reliable information.

4. If publishers are not fairly compensated for the use of the content by GAI systems and lose audience to them, it will harm publisher sustainability and see less money invested in quality journalism. In turn, less trusted content will be available to train and update GAI systems, harming innovation and increasing the chance that GAI systems produce unreliable results.

5. This represents an immediate threat to our democracy. News media publishers such as The Guardian have found that ChatGPT has fabricated articles purporting to be published by trusted journalists. Experts have warned that next year’s UK General Election could be susceptible to manipulation by personalised disinformation produced on an industrial scale.

6. The rapid growth of GAI systems also raises competition concerns. Foundation models are often owned or used by global tech giants who will, if allowed to continue unchecked, only entrench their market power and dominance. These companies already control most of the news ecosystem but have been reliant, up until now, on publishers for their original content. Once they can create their own synthesised versions of articles and news stories it would be a simple matter to rig the ad markets to reward their own GAI content over the real thing.

7. In transitional moments such as the one currently seen in GAI development, incumbents are prone to panic and block new entrants with anti-competitive and illegal conduct. In the past competition regulators have been too slow and lenient in allowing Big Tech concentration: the Digital Markets, Competition and Consumers Bill - which is currently going through parliament - is urgently needed to ensure these mistakes are not repeated for AI.

8. The Bill will empower the Digital Markets Unit - a new expert unit within the Competition and Markets Authority – to designate the very largest digital firms with substantial and entrenched market power as having Strategic Market Status (SMS). The DMU will be able to enforce Conduct Requirements (CRs) that are tailored to the business models of SMS firms. These CRs will ensure Big Tech firms act in a way that ensures fair dealing, open choices, and trust and transparency in their interactions with the businesses and individual consumers that rely on their services.

9. These principles will be critical if publishers are able to continue to fund the production of trusted information as AI systems grow in significance. Combined with the measures set out above – transparency, to allow publishers to identify when their work has been used, and remuneration for that use – we can ensure the continued sustainability and plurality of UK news media. Failure to do so will leave UK citizens at the mercy of AI-generated misinformation.

Sebastian Cuttill, Parliamentary and Campaigns Manager
October 2023