Bundesverband Digitalpublisher und Zeitungsverleger

Medienverband der freien Presse

Bundesverband Digitalpublisher und Zeitungsverleger e. V. (BDZV)
Medienverband der freien Presse e. V. (MVFP)

Contribution on
Competition in Virtual Worlds and Generative AI

11.03.2024

I. Competition between press publisher and AI

While AI has the potential to optimise the quality of work processes and reduce the workload of journalists, it also has the potential to threaten press and media freedom on the internet.

Both in the operation of AI-based language models and in the creation of press-like content, AI systems rely on and access content provided by professional media. Currently, the legal and technical means available to prevent this are inadequate and place the burden of proof on rightsholders to prove abuse, which is almost impossible.

Moreover, these new generative AI systems, including potential "AI press", can already enter and compete in the same primary markets as the press and media by using costly and professionally made editorial content for an artificially generated competing products at a marginal cost close to zero and within seconds**. There is thus a risk of direct unfair competition and the distribution of revenues from media content could shift further away from press publishers and other media services towards competing AI:** while the press bears the cost of editorial content, the latter will be primarily monetised by AI providers using publisher content to create automatically generated competing content and publications.

Ground rules to tackle the issues arising from this competition between humans and machines are urgently needed. The mistakes of the past with regards to inadequate regulation of platforms must not be repeated regarding AI providers. Instead, we need a fair framework.

In relation to competition concerns and the AI field, Thomas Höppner and Luke Streatfeild, highlighted in a paper the fertile ground that the AI market provides for competition concerns: *"Regarding competition, the characteristics of the AI stack have all ingredients for winner-takes-all battles with significant collateral damage affecting third party businesses: (i) Vertical integration in closely interrelated markets, with (ii) upstream dominance and downstream value generation, (iii) unequal access to proprietary resources, (iv) issues of interoperability, (v) data portability, (vii) non-transparency, (vii) IP licensing, and (viii) platform usage fees and conditions vis-á-vis dependent business users, etc. The incentives and abilities of incumbents to exploit, leverage and further entrench their market positions are evident. And this is not even considering the particularities of specific AI applications such as the exploitation of creative industries for generative AI (that uses their content to create competing services) or the tacit collusion within harmonized AI models deployed by competing companies."*[1]

In addition, and irrespective of general concerns about human-machine competition, providers that are also digital gatekeepers, such as internet browsers and search engines, already have particularly advanced AI models. The concentration of market power and (by then at the latest)

[1]Höppner, Thomas and Streatfeild, Luke, ChatGPT, Bard & Co.: An Introduction to AI for Competition and Regulatory Lawyers (February 23, 2023). 9 Hausfeld Competition Bulletin (1/2023), Article 1, Available at SSRN: https://ssrn.com/abstract=4371681 or http://dx.doi.org/10.2139/ssrn.4371681

the power of opinion creates a variety of risks for media diversity and democratic opinion-forming.

AI-press or media in the hands of gatekeepers would be a perfect breeding ground for non-transparent steering of opinion and information, as well as the misuse of professional editorial content: What content is fed to these AI offerings and how it is weighted will be entirely in the hands of the gatekeepers. Fine-tuning of the algorithms will always be possible and hardly verifiable. This is already the case with search engines. **The algorithmically controlled creation of content, from editorial to educational to entertainment, creates a new dimension of danger for the free and diverse formation of opinion**. The risk increases because the impact on the public is neither verifiable nor discernible. Professional offerings from publishers and media professionals are therefore more important than ever.

Against this background, the following points are essential:

1. No self-preferencing of AI media or "prompt journalism"
AI media and generative AI content raise specific competition concerns for the free press and media. It is therefore important to recognise this competitive threat at an early stage and to address the resulting concerns. Specifically, the concerns can be divided into two scenarios:

a. The **prohibition of self-preferencing must also apply to AI media or AI media of selected cooperation partners**. For example, the ban on self- preferencing would not prove effective should it not apply to cases of preferential treatment of AI media over professional press and media content, on the wrongful assumption that AI and "human media" are different products, and that AI media can therefore be favoured accordingly.

b. The **ban on self- preferencing must also apply to so-called "prompt journalism"**. AI-generated answers to readers' interest in news or other editorial content through for example gatekeeper platforms directly replace and therefore compete with professional press and media content. The argument that prompt results are not a medium in their own right and therefore cannot replace professional press and media content does not correspond to reality and must be rejected. In other words, answers generated by GenAI systems must duly be examined from a DMA or competition law angle in the way in which they compete with traditional media, without the need for the creation of a separate AI newspaper or magazine.

2. DMA provisions and AI-Systems of gatekeepers

a. First, it is crucial that the European Commission examines **whether and to what extent Generative AI services owned and operated by gatekeeper platforms**, such as Gemini (Alphabet), Llama 2 (Meta) or the forthcoming "AppleGPT" (Apple), constitute core platform services (CPS) of the gatekeepers in their own right and, if necessary, updates the list of CPS without delay.

b. Second, there is an urgent need **to carefully examine and assess how and to what extent other "traditional" CPS of the gatekeepers are only strengthened in their dominance and restrict competition by the additional and powerful integration and use of AI**. We are referring to services such as Google Search (Alphabet, search engine), Bing (Microsoft, search engine), AI Sandbow (Meta, advertising services). The DMA provisions, for example those on self-preferencing, must apply accordingly.

c. **Fair and non-discriminatory access and treatment:** Both when training AI-based language models and when creating editorial or other content, AI systems also access and rely

on the offerings of professional media. The means of preventing or proving this use still seem inadequate. Moreover, there is also the risk, particularly in relation to gatekeeper platforms, that the reservation of rights, the request for a licence and the payment of a fair price for the use of content, etc. will have repercussions, for example in the form of exclusion from or downranking in search results or similar. There is also a risk that access to and treatment by the gatekeeper platform will be linked to granting access to content for AI purposes. It is therefore important to ensure that access to and treatment conditions by gatekeeper platforms are fair and non-discriminatory.

3. The rights and intent of the professional press and media must be respected.

In addition, and even if this question does not directly fall within the scope of competition law**, AI systems must not use content against the will of the rightsholders**. This is a necessary precondition and requirement. Should AI systems be allowed to use (editorial) content against the will of the rightsholders, competition is no longer possible.

Against this background, it is important to ensure that publishers and the media have an exclusive right over their content, that a right to remuneration is provided and that verifiability is guaranteed. After all, any rights against AI providers are meaningless if the rightsholders do not or cannot check whether the content has been used despite their exercise of the reservation of rights. The obligation to provide a "detailed summary" of the training data used, as foreseen in the AI Act, is a first, well-intentioned step, but lacks in meaning and potential for necessary enforceability. If rightsholders are to be put in a position to know whether their content has been used or not, AI providers must be required to provide a complete, comprehensive and up-to-date list of all content used for training, input or other purposes.

II. Commission Questionnaire

1) What are the main components (i.e., inputs) necessary to build, train, deploy and distribute generative AI systems? Please explain the importance of these components.

Generative AI (GAI) models depend on the processing of data and other relevant information. Training such models requires high quality, verified and/or creative content (e.g. editorial, scientific and artistic content). In other words, GAI models are only as good as the content and data on which they are trained. Currently, rightsholders are generally not remunerated for the use of their content for training purposes. Following the training process, the GAI model can produce unlimited output based on prompts, queries, etc.

2) What are the main barriers to entry and expansion for the provision, distribution or integration of generative AI systems and/or components, including AI models? Please indicate to which components they relate.

At the heart of any GAI system is the model.

And it is here that the challenges of data collection begin. Obtaining high quality, comprehensive and representative data is crucial, with careful consideration given to both purchasing/sourcing and the legal aspects of data use. Data protection rules, such as the GDPR, have an important role to play here. With the database in place, there are a number of technical challenges that follow. Training the model requires powerful hardware and significant computing time. Such processes are costly, particularly if frequent iterations are required to optimise the model.

For deployment, scalable and reliable infrastructures need to be put in place, both for the technical delivery of the model and for user support. Legal compliance, in particular with regard to data protection and copyright, must be ensured throughout the process, as violations can have serious consequences and hinder or even prevent market access.

Additional challenges entail, *inter alia*, the distribution and integration of the AI system into existing markets and/or systems. Interfaces and compatibility with different technologies and platforms need to be considered.

Market dynamics are a significant risk factor for survival or expansion. The rapid evolution of AI technology is forcing companies to remain flexible and cost-effective to keep pace with new trends and competitors. To remain relevant and competitive in the long term, continuous development and adaptation of the AI system is essential.

Beyond the model, however, an AI system requires other components to support it, each with its own risk factors:

- Data infrastructure (storage, management, processing)
- Computing infrastructure (server, network, CPU/GPU)
- User interfaces (UI/UX)
- Integration tools (libraries, APIs)
- Security components
- Compliance and governance tools (laws, regulations, industry-specific rules)
- Support and maintenance systems (customer service, system maintenance)
- Training and assessment tools (monitoring, assessment, optimisation)

3) What are the main drivers of competition (i.e., the elements that make a company a successful player) for the provision, distribution or integration of generative AI systems and/or components, including AI models?

From the perspective of content providers, a copyright law fit for the AI age is essential for transparent and fair competition. Such a copyright law with comprehensive transparency rules (detailed summaries are not enough!) would create a fair legal framework and enable the fair remuneration of rightsholders whose content is used by AI models.

Additionally, AI providers, which are often part of large platforms, can further increase their market power through the output generated, so that they can create a competing product based on the content of rightsholders. (see specific concern in I.1 and I.2)

4) Which competition issues will likely emerge for the provision, distribution or integration of generative AI systems and/or components, including AI models? Please indicate to which components they relate.

From the perspective of content providers, existing copyright rules seem inadequate. As a result, rightsholders currently find it difficult to participate in the value creation process of AI providers. Furthermore, it is unclear whether an opt-out (reservation of rights with regard to text and data mining) is associated with a disadvantage in search engine indexing. It is therefore important to clarify that a reservation of rights with regard to text and data mining cannot not lead to a disadvantage in search engine indexing (prohibition of tying).

In addition, large digital gatekeepers are already using acquisitions and exclusive agreements with AI providers to prevent new players in the AI field from jeopardising the current market dominance of the gatekeepers.

5) How will generative AI systems and/or components, including AI models likely be monetised, and which components will likely capture most of this monetization?

GAI systems that are trained with high quality professional content can be integrated into other systems via interfaces or use rightsholders' content for training purposes for their own AI systems. By training and/or regularly crawling websites for content, including content behind paywalls, AI systems can be put on the market to directly compete with the offerings of rightsholders, e.g. online newspapers or magazines.

These AI products, which are enabled and based on the editorial, artistic etc. offerings of rightsholders, threaten to drive them out of their primary market. AI products can use the same monetisation options as rightsholders (subscription fees, advertising revenues and licensing fees) while being able to generate the content for a marginal cost of zero.

6) Do open-source generative AI systems and/or components, including AI models compete effectively with proprietary AI generative systems and/or components? Please elaborate on your answer.

In general, open-source GAI systems and components can compete effectively with proprietary systems. However, the extent to which they are competitive depends on the specific requirements of each project. While currently the largest and most powerful models are often proprietary, open-source systems offer advantages in areas such as pricing, privacy, customisation and scalability. Community-driven development of open-source models can lead to faster innovation and bug fixing. Examples of successful open-source AI systems underline this potential. Nevertheless, there are challenges to these systems, such as limited resources and financial support. In the long term, trends may shift towards greater use of open-source AI models, especially if privacy and customisation become more of a priority.

7) What is the role of data and what are its relevant characteristics for the provision of generative AI systems and/or components, including AI models?

There is no GAI model without data. GAI models depend on high quality, verified data (e.g. editorial and creative content). GAI models will be only as good as the data and content used to train them.

8) What is the role of interoperability in the provision of generative AI systems and/or components, including AI models? Is the lack of interoperability between components a risk to effective competition?

Interoperability is generally not a major issue as LLMs are text based and have a very accessible interface. Other modalities such as audio/video/images are more complex. This is because they require more specific interfaces and data formats, making the integration of different AI systems more challenging. While a lack of interoperability in these areas could theoretically raise barriers to entry for new players, in reality there is no significant risk to effective competition.

9) Do the vertically integrated companies, which provide several components along the value chain of generative AI systems (including user facing applications and plug-ins), enjoy an advantage compared to other companies? Please elaborate on your answer.

Companies that are vertically integrated, offering different components of the value chain of GAI systems, may indeed have advantages. Often more complex and optimised using proprietary data, their processes may be harder to replicate and enable more efficient and higher quality production. This can lead to faster response to market changes and greater control over the final product. However, such an advantage does not automatically translate into dominance in all areas. In some cases, it may be advantageous to spread the

infrastructure across several companies in order to exploit specific strengths. This can increase flexibility and reduce risk. Also, the challenges of vertical integration, such as high initial investment and potential difficulties in adapting to technological change, should be taken into account.

10) What is the rationale of the investments and/or acquisitions of large companies in small providers of generative AI systems and/or components, including AI models? How will they affect competition?

Providers of AI have the potential to partially disrupt the gatekeepers' business models (see Google Search vs. Chatbot Search). The attempts of gatekeepers to acquire AI providers or to engage them in exclusive contracts run the risk of further consolidating their gatekeeper role. Conversely, AI providers depend on very large computing capacities to train their models, which often are only available to the gatekeepers.

11) Do you expect the emergence of generative AI systems and/or components, including AI models to trigger the need to adapt EU legal antitrust concepts?

The gatekeepers' market power is still intact; there is currently no level playing field. Gatekeepers are search engines and AI providers all-in-one and benefit from their network effects. As long as there are no clear AI-proof competition and copyright rules, the gatekeepers can be expected to further increase their market power and fair competition will remain at risk. The past has shown that regulation needs to adapt to new digital business models. In the same way that search engines, ad tech and online marketplaces required specific new responses from authorities, AI business models may pose particular challenges. There is therefore a strong case for close monitoring and rapid response.

12) Do you expect the emergence of generative AI systems to trigger the need to adapt EU antitrust investigation tools and practices?

Yes, competition law and the DMA, as well as national legislation, need to adapt to AI developments. It must be acknowledged that the most powerful AI models will be both developed and operated by the current gatekeepers. In particular, problems similar to those in several digital markets which are currently the focus of the DMA could arise, where the competition procedures were too lengthy and not able to remedy the situation. In the meantime, markets had already tipped and competitors had to leave the market before the investigations could be completed.

For further info, please contact:

Philippe Meistermann
Head of Brussels Office

Bundesverband Digitalpublisher und Zeitungsverleger e. V. (BDZV)
meistermann@bdzv.de

Transparency Register ID: 201132116888-24

Alexander Nitz
Head of Brussels Office

Medienverband der freien Presse e. V. (MVFP)
alexander.nitz@mvfp.de

Transparency Register ID: 876509717508-94