43ʳᵈ EPRA Meeting
Barcelona, 25-27 May 2016

Working Group 3:
Data Protection & Big Data – What impact on media regulation?
Summary

Moderator: Celene Craig, EPRA Chairperson
Expert: Lucia D'Acunto, TNO (the Netherlands' organisation for applied scientific research)
Content Producer: Maja Cappello, European Audiovisal Observatory

Background

Building on a workshop[1] co-organised in December 2015 by the European Audiovisual Observatory and EPRA on the grey areas between data protection and media regulation, this working group could already build on some preliminary findings, despite being this topic a premiere for an EPRA meeting.

The point of departure for the discussion was the acknowledgment that when touching upon the issue of the impact of big data on media regulation, there are many aspects to consider. Apart from general matters of principle, such as freedom of expression versus privacy, or copyright protection versus data protection, or even data retention obligations for telecom providers versus data protection, there are several domains where the borders between the two sectors are not necessarily clear, both with regard to applicable provisions as to monitoring and enforcements aspects. In many cases nobody intervenes and the underlying rights risk remaining unprotected.

Data protection authorities have been put in place to care about the protection of the data of the users. This explains why data protection issues tend to fall outside the scope of action of media regulators. But there are various aspects where data could become part of the activity of the media regulators as well.

Questions addressed

The purpose of the discussion in this working group was to raise certain awareness as to the issue why media regulators should care about the use of massive data and consider their possible impact on freedom of expression, pluralism of information and editorial independence and responsibility. The following questions were addressed:

1. Algorithms and prominence of European works,
2. Big data and pluralism of information,
3. Profiling and targeted advertising.

---

[1] The summary of the workshop and the presentations of the participants are available at http://www.obs.coe.int/web/obs-portal/workshops/-/asset_publisher/kNG5qM2wH8Kq/content/dli-workshop-obs-epra-the-grey-areas-between-media-regulation-and-data-protection and the "storify" of photo gallery and tweets can be accessed from https://storify.com/EuAVObservatory/workshop-2015.

Key findings of the discussion

The discussion was introduced by **Lucia D'Acunto**, Research Scientist at the TNO[2], who provided the necessary technical toolkit in order to get a proper understanding of the most relevant technical issues underlying exploitation of data in the field of media.

In order to ensure a common understanding of the term, she noted that "Big Data" is an umbrella term for a variety of strategies and tactics that involve massive data sets, and technologies that make sense out of these mindboggling reams of data. The Big Data trend has impacted all industries, including the media industry, as new technologies are being developed to automate and simplify the process of data analysis, and as throngs of data analysts are being trained and hired to meet the demand for the analysis of these data.

As massive use of Big Data can have disruptive effects on media consumption, the four main phases of data exploitation were investigated, namely:

- User tracking, through cookies and web browsing fingerprinting;
- User profiling, working on similarities among items that users watched or liked and among users based on the items they watched or liked;
- Information filtering, based on the relationship between content, so as to select what to present to users;
- Information presentation.

This overview was followed by an immediate Q&A session to clarify possible unclear terms, and issues such as the scarce awareness around the exploitation of cookies, the concern of trusted brands not to lose their credibility and the readiness of young people to give up their data if compared to older people, where touched upon.

The session went then on with short introductions to the three main discussion topics as identified above and the following aspects were highlighted.

Algorithms and prominence of European works

Algorithms are actively exploited by VOD providers in order to ensure that their offer corresponds to the preferences of their viewers. This is especially true in the case SVOD providers, where the primary purpose is the retention of the clients (or "fidelisation"), also by ensuring that the suggestions they get meet their wishes.

By combining the data concerning the viewing habits of their clients, i.e. through analysing the behaviour of their clients and comparing their profiles, providers are able to produce algorithms they can use for the promotion and recommendation of content on the basis of profile similarities between users.

A lively discussion was held on the issue of possibly using algorithms as an implementation tool of the obligation to ensure prominence to European works under Article 13 of the Audiovisual Media Services Directive. The more general issues as data science and governance of recommendation systems were also discussed.

---

[2] TNO is an independent research institute founded by law in 1932 and based in The Netherlands, https://www.tno.nl/en/.

Big data and pluralism of information

In a context of abundance of information, where algorithms help reach desired content, it was recognised that the benefits of algorithms are also accompanied by risks, as they may: challenge the pluralism of information, by interfering in the user's choice (suggestion of content, way of displaying search results, …), threaten fair competition by creating a dominant position of providers able to dictate the access to the content and its findability according to their own economic and political agendas, and also diminish variety and create echo chambers of content that has already been shown so as to create an amplification effect.

Massive exploitation of big data may have a strong impact on the structure of the audiovisual market – both in terms of external pluralism, by strengthening the position of certain major operators, and of internal pluralism, by limiting the choices of the viewers – and on the creation of new content, which is likely to become closer to the already existing preferences of viewers and thus less diverse.

In light of the above, the discussion touched upon issues like data neutrality and information duties towards the viewers.

Profiling and targeted advertising

It is well known practice that data gathering technologies use cookies to collect users' data in order to maximise advertising campaigns, by addressing the right users, with the right information and right tool at the right time.

The result of this matching activity is the provision of tailor-made advertising which is likely to be more efficient in the choice-process of the users. An indirect effect can be the levelling down of the variety of choices and also the possible abuse of the data (mostly inadvertently) provided by the users.

In addition to setting up a cooperation procedure with data protection authorities for the activities related to the assessment of the level of compliance with the regulatory framework on data protection, there was a general consensus around the increasing relevance of media literacy. The understanding of which data are being put together and under which conditions they are exploited appeared as a particularly relevant task to accomplish with by all involved actors in the nearest future.