Arcom

Le régulateur de la communication
audiovisuelle et numérique

Public consultation on access to data from online platforms for research purposes

May 2022

Public consultation on access to data from online platforms for research purposes

1. Access to data from platforms for research purposes : a key issue in a changing world

1.1. Recent developments on social media and online habits are redefining how we access information

Search engines, video sharing platforms and social media are redefining how content, especially information, is consumed and shared.

These sources of innovation have led to **new ways of expression and have accelerated certain forms of citizen participation**. However, they can also be subject to misuse and abuse, including the phenomena of information manipulation and online hate.

Today's information environment is no longer defined by the addition of sectors with hermetic borders: audiovisual broadcast and digital; traditional media (television, radio, press) and new content consumption services (social media, apps); traditional modes of reception and future devices; national, European and international media. Conversely, the overlaps are now increasingly substantial. This is resulting in a redistribution of the amount of time spent on media and the sources chosen, which reinforce the **pivotal and growing role of the internet in the access to information**. **Internet usage now competes with the one of traditional media**[1].

In addition to this role of providing information access role, there is also a general impact of the internet , and of social media in particular, on opinion formation. Increased exposure to content that is close or similar to users' known opinions is, for example, one of the main features of news feeds on social media.

1.2. The research community has a key role to play in understanding online usage

In this context, it is **crucial for research to able to study these new dynamics and develop independent tools and approaches in order to shed light on them**. The idea is to gain collective knowledge of phenomena whose potential effects can be harmful to our societies.

The development of a framework enabling the study of online behaviours and their effects should help to **protect and strengthen the independence, autonomy and analytical capability** of research, and enable it to play its role in accompanying and understanding contemporary societal changes.

In order for the research community to fully grasp these issues, **it is therefore necessary to reflect on public actors' role in facilitating the research**. This facilitator role must be

---

[1] According to the latest Kantar/La Croix media barometer, French people rank the internet as the second most important information medium (32%) after television (48%) but ahead of radio (13%) and the written press (6%). Nevertheless, the trust placed in these different media outlets is not positively correlated with their use: as such, radio and the written press are considered the most reliable information media at 49%, just ahead of television (48%). From this point of view, the traditional media are still largely retaining their users' trust. By contrast, just 24% of French people believe that credible information can be found on the internet.

expressed more specifically in the **use and analysis of data from social media or online platform services, which govern the development of knowledge specific to digital environments**. The challenge of using this data properly is twofold: it involves both perpetuating a dynamic, effective and sustainable research ecosystem capable of generating knowledge for the benefit of all (**scientific production**) and contributing to the regulator's expertise in its assessment of the measures implemented by platform operators to meet their obligations, such as the moderation of hate content (**transparency regulation**).

2. Why Arcom intends to play a role in the access to platform data for research purposes

2.1. In compliance with the GDPR, the regulator needs to be a facilitator of data access for the research community

Created by the merger of the French audiovisual council (CSA) and the high authority for the dissemination of works and the protection of rights on the internet (Hadopi) on 1 January 2022, **the Authority for the regulation of audiovisual and digital communication (Arcom) was created to support the major transformations of the audiovisual and digital landscapes.** Regulation is one of the responses to these challenges, which have been well identified by public institutions. Arcom is notably in charge of protecting the creation and its actors, monitoring economic balances in the audiovisual sector, supervising the means used by online platforms to protect audiences while guaranteeing freedom of expression, and ensuring political pluralism on the air. More broadly, it aims to protect all audiences in the audiovisual sector and online.

In addition, the systemic regulatory powers of online platform operators (as defined by article L. 111-7 of the French Consumer Code) **entrusted to Arcom by the legislator have been continuously strengthened since the end of 2018**. They apply mainly to social media (Facebook, Snapchat, etc.), search engines (Google, Bing, etc.) and video sharing platforms (Dailymotion, YouTube, etc.), and exclude, for example, video on demand subscription services (Netflix, Amazon Prime Video, etc.). Nevertheless, it is within a broad definition of "platforms" that Arcom intends to conduct this consultation, in order to open up the debate to all participants in the digital information ecosystem and to encompass new categories of actors that could emerge in the short or medium term and fall into the "platforms" category.

This new paradigm, which complements its regulatory model, gives Arcom a new position in an extensive and polymorphic ecosystem. The Authority supervises the media used by operators, who have a duty of cooperation and transparency[2]. The research community is conducting work to shed light on the understanding of these phenomena. Civil society as a whole participates in these actions through its analyses, feedback and alerts. These different fields of action complement each other and form a **feedback loop** in which the regulator is one of the various actors to identify, analyse, assess, question and, if necessary, propose mechanisms to respond to the risks identified.

It is also important to stress that this approach is aligned with the European legal framework of the General Data Protection Regulation (GDPR) for users of online platform services. The GDPR is intended to apply to the vast majority of personal data

---

[2] Within limits that must be duly justified, for example, in terms of the security of their services.

processing carried out by each of the actors. As the anonymization of data from platforms is technically complex to implement in practice and can have an impact on the definition of research questions, it is important to take into account the personal nature of the data. The French Data Protection Authority (CNIL) has also conducted a public consultation with researchers on how they access data and with regard to the GDPR. This initiative has led to the **publication of various resources for these actors**: presentation of the challenges and the rules to follow, reminder of the tools available for compliance, etc[3]. **The issues of access to data on online platforms therefore form part of this framework of protecting users' rights and freedom to conduct research on these services, given the societal issues they raise**. **The problems of access to data on online platforms are therefore part of this framework of protection of users' rights, particularly the right to control the data by the persons concerned**[4]**.**

2.2. Online platform operators' current practices in terms of giving access to their data are very diverse

In order the above-mentioned to fulfil their different missions, the identification of the problems arising on online platform services should not rely solely on operators' initiatives. In addition to the data these players make available, incidentally in very diverse ways, the research community must also be able to access high-quality data in ways that are not defined by platforms alone. As such, a **transparency regulation needs to be rolled out**, in the framework of which Arcom must be able to draw on the various actors' contributions while at the same time playing a role in enabling these stakeholders to act.

Indeed, **access to data from online platforms is currently complex, notably due to the lack of a unified framework or a common availability policy between platforms, at a national or supranational level.** This situation is highlighted by initiatives such as the European Digital Media Observatory (EDMO)[5]. Established in 2020 and led mainly under the guidance of the European University Institute in Florence (EUI), this group of experts from academia, the media and government aims to shed new light on the issues of online disinformation. To this end, one of the EDMO's objectives is to contribute to the reflection on the use of data from online platforms, **in particular by supporting the competent authorities in their regulatory efforts**[6].

Currently, access to data is mainly provided by platforms on a voluntary basis, focusing the research work on the most active services in this area. While these initiatives are to be welcomed, it is clear that research is mainly concentrated on Twitter,

---

[3] https://www.cnil.fr/fr/recherche-scientifique-hors-sante

[4] The identification of the roles and responsibilities of each actor with regard to the RGPD, particularly with regard to the transparency due to the persons concerned, must enable people to exercise their rights. Cf. « air2021 : entre partage et protection, quelle éthique pour l'ouverture des données ? » ("air2021: between sharing and protection, what ethics for open data?"), CNIL

[5] https://edmo.eu/

[6] The second objective that appears in the 2021 EDMO activity report is as follows: "Creating a governance body which ensures public trust regarding the work of the platform and establishing a framework to provide secure access to data of online platforms for research purposes." (Source: https://edmo.eu/wp-content/uploads/2021/09/EDMO-Public-Report-June-2020-%E2%80%93-March-2021-2021.pdf)

which offers various APIs, including one dedicated to research[7]. This openness has enabled many projects to be launched, particularly through automated content collection. One example is the initiative by *the Institut des Systèmes Complexes de Paris Ile-de-France* (ISC-PIF, CNRS laboratory), which since 2016 has brought together a team of researchers and engineers to use the data from this social network. The scientific work of data processing and analysis has, for example, made it possible to implement the *Politoscope*[8], a visualisation tool for the general public that aims to give insight on current political issues and how they evolve[9]. **Other social networks and search engines have chosen a more restrictive policy on access to their data, including for researchers.**

2.3. Arcom aims to play a key role in the framework set up by the Digital Services Act (DSA), which addresses the most topical issues while raising operational questions

To meet the challenges posed by online platforms, the need for action at a European level has gradually become clear. This is expressed in particular by the consideration of issues concerning the emergence and consolidation of new digital markets, with the Digital Markets Act (DMA), and those around the circulation of data between companies, with the Data Governance Act.

In addition to these initiatives, the EU Digital Services Act (DSA) aims to ensure users' safety and the protection of fundamental rights online. Arcom, together with ERGA through its different position papers on the DSA, welcomes this regulatory development. In particular, the DSA proposes a model for the **systemic regulation** of online platforms that addresses some of the most important informational disorders of our time, while preserving one of the internet intrinsic characteristics – the provision of a space for exposure and expression. For the very large online platforms[10], additional obligations are anticipated to further increase the transparency of their actions, particularly with regard to the functioning of their content moderation approaches, their advertising services and the algorithms they use on their services.

More specifically, **article 31 of the DSA** aims to regulate researchers' access to the data of these very large platforms in order to contribute to the assessment of the systemic risks that their services may pose. The DSA positions itself with a view to overhauling the relationship between platforms, authorities and users and could lead to the emergence of a **new regulatory model**[11]. As such, the research community would be closely involved in gaining a better understanding of the **socioeconomic, political and cultural dynamics** emerging

---

[7] It should be noted, however, that more generally in terms of research, platforms may conduct work internally or directly commission external researchers. These initiatives remain at actors' discretion and do not require the creation of permanent data access mechanisms.

[8] *Politoscope Project, CNRS Institut des Systèmes Complexes Paris Ile-de-France* (ISC-PIF), http://politoscope.org

[9] The *Politoscope* example is not intended to serve as a model for a research mechanism that would be favoured by Arcom: it is used here to illustrate how the automated collection of data from a social network has resulted in a scientific use that has generated a contribution to the public debate in the form of a tool made available to the public.

[10] The "very large online platforms" (VLOP) category includes services that reach at least 45 million users in the EU per month. See in particular: "Digital Services Act Briefing", European Parliament, 2021. URL: https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/689357/EPRS_BRI(2021)689357_EN.pdf#:~:text=The%20DSA%20proposal%20is%20a%20horizontal%20instrument%20putting,and%20Digital%20services%20act%20(DSA)%20draft%20asymmetric%20obligations

[11] Regarding the DSA's ambitions and their possible repercussions on the international debate on platform regulation and the organisation of transparency, see for example Schiffrin (2021), who highlights the consequences that the DSA could have in the US: https://www.cjr.org/business_of_news/europe-regulates-big-tech.php

in this new informational ecosystem. Arcom hopes to contribute at its level to the reflection on these issues of access and the construction of an innovative model at a European level.

However, article 31 of the DSA raises the question of its full operationality given the objectives pursued:

- The position of the intermediary between researchers and platforms: the *"Digital Services Coordinator"* is one of the two intermediaries, together with the Commission, between the stakeholders. The definition of its role will therefore be particularly pivotal.

- The data concerned by this access: the scope of the data concerned encompasses the *"identification and understanding of systemic risks"* as defined by the DSA. These risks should, as the text currently stands, cover three categories in particular: the potential manipulation of the platform services, in particular to disseminate illegal content or for economic purposes; the impact of these services on fundamental rights such as freedom of expression, particularly with regard to the algorithmic systems used; and the intentional manipulation in order to disseminate information on a mass scale that could have an adverse impact on public health, electoral processes or security. It should be welcomed that these fields cover the most urgent issues among the informational disorders already identified by research. Nevertheless, questions may arise regarding the relevance of an eventual more encompassing approach, particularly from an interdisciplinary research perspective. In addition, it is still crucial to be able to identify new risks in the future that have not yet been observed but that the research could detect.

- The status of researchers authorised to access these data: article 31 gives access to data subject to certain conditions. This provision would therefore provide a clear framework for researchers wanting to study the phenomena covered by the DSA, without prejudice to the GDPR. Future delegated acts will specify the conditions under which such access would be provided to researchers who request it. At this stage, it seems useful to question the risks that excessively strict eligibility criteria for data access (administrative or financial capabilities of the applicant structure, relative work previously carried out by one or more members of the research team, effective possibilities for interdisciplinarity, etc.) could pose in terms of limiting side effects. For example, valuing technical expertise could substantially favour academic researchers who have already produced numerous articles on the topics targeted by the DSA. As such, a whole swathe of research could be excluded from the access mechanisms: young researchers, journalists, NGOs, etc. This question also raises the issue of the possible trade-off between openness to a large number of actors and the risks in terms of the use of personal data for the persons concerned. The qualification of scientific research in the sense of the GDPR may indeed be more restrictive than a strictly scientific evaluation of projects.

2.4. Arcom intends to use an open and contributory framework to establish the model for accessing data from online platforms.

As such, Arcom is launching this public consultation on access to data from online platforms for research purposes and in relation to the issues over which the Authority has jurisdiction: combating information manipulation and online hate.

Through five themes – sharing experiences of using data from these services (A), governance (B), construction of scientific projects (C), data protection and technical considerations (D) and feasibility of access and incentives (E) – this public consultation aims to question all relevant stakeholders. The aim is to learn initial lessons regarding the implementation of an operational framework for accessing data from online platforms and thereby contribute to the various stakeholders' general reflection concerning these issues, particularly researchers and the public sphere. Academia, online platforms, public authorities and associations are therefore invited to share their ideas and contribute to the public interest through research.

The elements collected by Arcom will then be summarised to provide input for existing debates on research-related access to data from online platforms; this work may lead to new reflections at French, European and international levels. All responses and the summary will be made public[12].

---

[12] The publication of responses for transparency purposes does not, however, prevent respondents from asking for some of their responses to be treated as confidential.

3. Arcom intends to base its reflection on the answers to five main themes of questions

A. Sharing experiences of using data from services related to the theme

➢ *Questions for all actors interested in the study and research work related to online platforms:*

The interest in issues related to platforms and the study of online activities have become part of the research agenda of a growing number of disciplines. These fields of study are very diverse, ranging from **natural sciences to computer science and social sciences**. This therefore requires data processing based on various **protocols and methodologies** and implies the consideration of possible disciplinary specifics that would make certain access and study methods more appropriate than others depending on the research issues. In addition, some services have **a policy of opening up their data to researchers**, notably through the provision of APIs while, conversely, access may be limited or even subject to strict control in other cases.

The following questions aim to gain a better understanding of **respondents' experiences with platform data in their research projects**, the **difficulties** they may have faced and any technical or legal **constraints** that may have influenced the construction of their research projects.

> A.1. Have you ever conducted **research using data** from one or more online platforms? If so, how did you collect it (e.g. using APIs, crowdsourcing, etc.)?
>
> A.2. Did you encounter any **difficulties** in collecting this data? If so, of what kind? Please provide examples.
>
> A.3. If so, have you ever **abandoned all or part of a research project due to the inability to access data** from online platforms? If so, was this the consequence of access being refused? Please provide examples.
>
> A.4. If not, which **factors** do you think enabled you to successfully collect this data? Did you have the possibility to **cooperate with the platform** studied to access this data? If so, how did this materialise? Please provide examples.

➢ *Specific questions for online platforms:*

The policies for making data available for research differ significantly from platform to platform. The following questions aim to gain a better understanding of **their respective policies** and the determinants of these policies: nature of the service, technical or legal specificities, or assessment of specific risks that data sharing could pose.

Arcom
Le régulateur de la communication
audiovisuelle et numérique

Access to data from online platforms for research purposes

A.5. Do you have a **policy for data sharing** with third parties for research purposes?

  i) If so:

  - **how long** has it existed?

  - does it concern one or more specific **categories of recipients** (researchers, NGOs, businesses, etc.)?

  - are there any **criteria for selecting** these recipients? If so, which?

  - what **type(s) of data** does this policy cover?

  - does it include **a control or monitoring component** regarding the use of the data provided?

  ii) If not, for what **reasons** have you not initiated such a policy? These may include legal, regulatory, technical, financial and other risks. Specify your assessment of these risks resulting in the decision to not open up your data.

B. Governance

➢ *Definition of actors:*

Access to data useful to society raises the question of **opening it up to all research-related stakeholders**. While the academic world appears to be the main beneficiary of a more open access, the contribution by **think tanks, journalists and civil society** to the knowledge of issues related to online platforms deserves reflection and consideration[13]. The question of **actors' neutrality**, given the funding they may receive from certain platforms, also arises.

B.1. Should we **define and possibly consider limiting the types of actors** that can receive access to data: researchers, journalists, NGOs, think tanks, civil society, etc.?

  i) If so, based on what **criteria** (possibly combined with the nature of the research or the objectives pursued)?

  ii) Should they have the **same access possibilities** or should these differ according to the type of actor?

B.2. Should there also be a **minimum level of access for the general public** (or a broader category of recipients than academic researchers), such as the mandatory provision of a certain amount of anonymised data in an open data format?

---

[13] One of the modes of these contributions is participatory science and research. These are *"forms of scientific knowledge production in which civil society actors participate, alongside researchers, in an active and deliberate way"* at all stages of the research continuum, such as data collection, analysis and interpretation of results (Source: Participatory research · Inserm, Science for Health).

➤ *Modes of granting access to data:*

The modes of granting access and the possible **criteria** based on which research projects would be selected should also be taken into account. Indeed, although the legitimacy of the use of data for research purposes is not under discussion here, the implementation of this principle raises many issues. For example, the **respective roles of European or national institutions** that might be involved in selecting research projects needs to be considered.

B.3. In your opinion, is a data access model based on providing access to a **trusted third party** relevant?

 i) If so:

 - should this trusted third party be a **European or national** public actor? In this case, what would be its **interactions with other authorities**, for example those responsible for personal data protection?

 - What could be the terms of organization for a **targeted and supervised** data access protocol?

 - Should the terms of involvement of the trusted third party be defined according to the **level of risk** associated with the data?

 ii) If not:

 - for **what reasons?** These can be diverse: legal, academic, logistical, etc.

 - Do you think a model of **direct interaction** between the platform and researchers is preferable? If so, why?

B.4. In the eventuality of a mode of regulation that would involve the intervention of a **trusted third party** in opening up the data for research projects:

 i) who would be responsible for **monitoring the implementation** of the application protocol?

 ii) what **safeguards** could be put in place to ensure access to data that satisfies the need expressed?

 iii) how should the decisions **transparency** by access protocol organisers be guaranteed?

 iv) what position and roles should each of the **stakeholders** have, especially the platforms?

 v) do you identify any **risks inherent** to this model? Which ones?

C. Construction of scientific projects

Recent and future transformations of online platforms raise the question of **researchers' ability to identify their data needs** in order to shed light on a social, economic, political or cultural phenomenon. The risk **of information asymmetries** between researchers and

platforms is high, and **support for a scientific project by an external committee or regulator** could be a way to facilitate the development of research protocols.

> C.1. When preparing their access request(s), how **researchers' knowledge of the data** from platforms they might reach out to for their studies be fostered?
>
> C.2. Who would define the **scope of the research projects** and their connection to one or more missions of public interest and preside over the identification of the data to which access would be necessary? Should the data concerned be restricted to particular fields of research? If so, which ones? For example, combating information manipulation, hate and online piracy.
>
> C.3. How would requests for access be **formulated** by interested researchers? For example, through calls for project tenders on predefined and/or ad hoc themes, after identifying relevant study topics?

➤ *Assessment of access requests and granting criteria:*

The questions in this section are based on the assumption that research projects requiring access to data from online platforms have been defined through formalised requests (e.g. to a trusted third party). The question of assessing their **scientific quality** arises. The level of innovation and **their level of contribution to the scientific literature** are aspects that could influence the modalities of data opening. Examining requests in light of these issues would require **the involvement of independent expert committees** to assess requests, based on a clear protocol and transparent criteria. These could take different forms depending on the discipline, **while remaining within a previously defined theoretical authorisation framework**.

> C.4. Do you think it is appropriate for **a committee to assess and monitor** access requests?
>
> i) If so, how should this **assessment committee** be composed (e.g. an international scientific committee)? Should one or more **regulators** have a position and role on it and, if so, which?
>
> ii) If not, why not? What mechanisms would you consider more able to meet researchers' access requests?
>
> C.5. To what extent would the more or less **binding** nature of the **obligations for platforms to open up their data** require their presence on the assessment committees? Should platforms also have a **right of return** in relation to researchers' requests or even a **right of refusal**?
>
> C.6. What would be the **criteria for granting access**? For example, is it necessary to have a research project involving interdisciplinary teams, possibly from structures located in at least two EU countries, in order to be selected?

> C.7. Should a **time dimension** be included in the assessment of calls for project tenders, so that only those with a short or long duration are selected?

➤ *Scientific production and showcasing:*

In order to inform the public debate, research projects that will use platforms data to answer scientific questions are intended to be **published in scientific journals**. If award committees and platforms should not interfere in the results and conclusions drawn by researchers in order to **guarantee their independence**, the valorisation of the work could be recognised, for example through **certification protocols**. These aim to confirm that the data was used in accordance with the regulatory framework in force, for example the model of the *cascad* certification from the Centre for Secure Access to Data (CASD)[14].

Moreover, the criteria for publication in social sciences are evolving, particularly with regard to quantitative studies, and now they better integrate the **principle of *reproducibility* of the results** by other researchers. As such, the analysis protocols that led to particular results should be available for **study, critique or as a basis for further work**. This principle requires the provision of data and resources (codes, scripts, etc.) and may raise particular difficulties in the case of sensitive data collected from online platforms.

> C.8. Should the work resulting from the analysis of these data be **externally certified?** If so, what form might this take?
>
> C.9. What precautions should be taken concerning the **publication of the studies carried out**, for example concerning the sensitivity of the data that would have been used? How can the implementation of these precautionary measures be reconciled with the fundamental principle of **researchers' independence**?

D. Data protection and technical considerations

➤ *Identification of relevant data and construction of materials:*

The term "data" can cover a very wide field (content, users, archives, etc.). Defining its scope is therefore a real challenge to ensure **consistency between study subjects and the changing characteristics of platforms**. In addition, each original research question may require specific formatting of the study databases to match an analysis methodology. For example, the degree of **granularity of the variables**, the **composition of certain aggregates** and the **ability to match the data with additional databases** from other sources should be taken into consideration to avoid the pitfalls of a **one-size-fits-all model** that would not allow certain issues to be addressed from certain angles.

---

[14] The CASD is a system for accessing secure data, particularly from French administrations (INSEE, ministries, etc.), via the provision of an "SD-box" to previously approved parties involved in a study project (universities, authorities, etc.). cascad-CASD certification enables researchers to report to their peers the reproducibility of their research on confidential data hosted at the CASD.

D.1. Given that research projects relying on platform data may favour an angle of analysis that would require a specific database format (variables, granularity, etc.):

    i)   how can we enable the **creation of specific or unique databases** that would be built to meet specific needs?

    ii)   to what extent would certain research projects enable the **construction of innovative indicators or measures** that could contribute to collective knowledge on the issues studied?

D.2. Can and should data access be **jointly constructed** on an equal footing between governance actors, researchers and platforms, based on the model of INSEE's CASD[15]?

D.3. How can the **data access framework** – governance, types of data identified in relation to missions, etc. – **be made long-term** to ensure it remains suitable to the regular innovations of and changes to platforms?

> *Terms of access and storage:*

In addition to the formulation of data access requests, there are **technical considerations** regarding the terms of access and their implementation. Indeed, the mechanisms for making these resources available and sharing them must **be secure and reliable**. Models for accessing data via secure boxes have already been tested by data producers such as the INSEE. Other **ways of accessing and storing data** could be considered.

D.4. **What terms of access** should be preferred for online platforms data? What are their different advantages and disadvantages? Should these differ according to the data collected? If so, why?

D.5. How can a **secure access mode** be guaranteed, particularly when the data is **not anonymised** and/or concerns **business secrecy** issues?

D.6. How should these data be **stored** to ensure the **protection of personal data** and, where appropriate, **business confidentiality**?

D.7. What would be the role and scope of intervention of **data protection authorities** (national and the European Data Protection Centre) in assessing the risks associated with access to data?

D.8. Should research projects receive **support** from the structure granting access, e.g. of a technical, financial or other nature?

E. Access feasibility and incentives

> *Support for researchers:*

The construction of research projects based on the use of platform data raises a number of **risks relating to inequalities between disciplines or research teams**. Indeed, some may not be able to offer analysis protocols due to limited resources (technical capabilities, staff, etc.). In addition, **the lack of knowledge of access protocols** could act as a **disincentive** to smaller actors, for example those less well funded or less able to respond to national or European calls for tenders.

> E.1. How can **researchers be supported** in building their research projects and complying with the GDPR and the standards set by the mechanism?
>
> E.2. What mechanisms could be used to mitigate the **funding and technical capability differences** between academic institutions, which could lead to a small number of research teams capturing projects?

> *Platform incentives:*

Researchers' access to data from online platforms aims to improve the understanding of socioeconomic, political and cultural dynamics, and therefore **could justify platforms' participation in, for example, a scientific knowledge contribution system.** They could also benefit from the results of the research carried out, which would help facilitate their dialogue with the research community.

> E.3. How can effective and balanced **incentives** be put in place to ensure that platforms are part of the open data dynamic? How can these actors be integrated into the system in a coherent way and how can best practices be promoted?
>
> E.4. Would the involvement of an **external audit committee** be relevant:
>
> > *i)* upstream, in the assessment of approval decisions based on the CESP model in the field of statistical surveys in France, for example?
> >
> > *ii)* downstream, in the review of the platforms' responses to access requests?
>
> E.5. What procedural safeguards could be put in place in relation to **business secrecy** issues?