

# If You Build It, They Will Come

---

Pietro Ortolani

2022-11-07T09:27:11

Content moderation is not only an Internet governance problem; it is also, unavoidably, a form of *de facto* adjudication. Online platforms make determinations that affect individual rights, whenever they decide whether to remove content, suspend/terminate accounts, or impose other restrictions. This is true not only for the user posting the content, but also for third parties (including vulnerable or marginalized groups) seeking remedies against online harm. As a result, platforms are routinely required to balance legal entitlements against each other. Such a balancing test, traditionally considered as [part and parcel \(although not a monopoly\) of the judicial function](#), is now carried out by private actors, with a frequency that no judicial authority could (or should be required to) sustain. To be sure, the platforms' decisions do not limit the users' ability to seek redress in court: content moderation, after all, is not a form of arbitration. Nevertheless, since platforms control the infrastructure enabling the [self-enforcement](#) of their own decisions, content moderation procedures end up being the main avenue through which a wide range of parties seek redress. The outcome of those procedures will often not be reviewed by any State court.

Over the years, some platforms have expressly acknowledged the para-judicial nature of content moderation decision-making: the most prominent example is the one of Meta, which set up the [Oversight Board](#) precisely for the purpose of developing a body of precedent and guidance ([not unlike a sovereign willingly subjecting itself to judicial scrutiny](#)). So far, however, the choice whether to embrace the adjudicative nature of content moderation has been largely left to the platforms themselves. As a result, even though content moderation has progressively mutated into a form of private adjudication, access to these *de facto* private adjudication fora has been scattershot at best, [with platforms prioritizing certain categories of complaints over others \(e.g. disregarding certain unfair commercial practices\), and providing insufficient transparency over their decision-making procedures and substantive standards](#).

## Observing the DSA through the lens of access to justice

This state of affairs is partially about to change with the Digital Services Act (DSA). The DSA has been described as marking a "[procedural turn](#)" in European lawmaking: rather than setting forth any bright-line substantive rule on the limits of online freedom of expression, the new Regulation creates a series of procedural obligations and redress avenues. The DSA's "procedure before substance" approach is reminiscent of international investment law, [where dispute resolution procedures were devised at a time when no consensus existed as to the substantive standards of investor protection](#). Hence, it makes sense to observe this new instrument through

the lens of [access to justice](#), to evaluate whether the DSA effectively enhances the possibility for aggrieved parties to obtain redress within platforms, as well as outside of them. But the thorny issue of access to justice is not only interesting for those affected by harmful content. Already in 1986, [Mirjan Damaška](#) urged us to study systems of justice as a way to understand how a State conceives of its own authority and officialdom. Today, conducting a similar exercise on content moderation and the DSA can show us how the EU lawmakers conceive of the public/private divide in the European digital space: as usual with procedural law, the big question is “who gets to do what?”.

The remainder of this contribution will briefly reflect on whether the “procedure before substance” approach of the DSA can indeed contribute to enhancing access to justice in the field of content moderation. What role do the different dispute resolution avenues of the DSA play? How do they interact with each other, and with the pre-existing framework of European civil procedure? To what extent can the EU lawmakers solve some of the many content moderation problems, by setting forth procedures (rather than substantive rules)? These questions would deserve a much longer discussion than a blog post allows. This contribution, thus, is a mere first attempt to “scratch the surface” of DSA procedures, shortly considering selected provisions of this new Regulation.

## **Access to justice within platforms**

Article 16 of the DSA requires hosting service providers (including platforms) to put in place a notice-and-action mechanism enabling “any individual and entity” to point out the presence of allegedly illegal content. [Practice](#), however, shows that certain categories of harmful content may be not outright illegal, but nevertheless incompatible with a platform’s terms and conditions. For these types of harmful content, the availability of a notice mechanism depends on the platforms, which remain free to determine the purview of user affordances.

From an access to justice perspective, importantly, notices prevent platforms from claiming ignorance about the presence of illegal content (as long as the notice enables a diligent service provider to identify the illegality without a detailed legal examination). This, in turn, excludes the platform’s immunity from liability, thus opening the door for possible liability claims by affected parties, if the illegal content is not removed expeditiously (Article 6).

Furthermore, Article 44 of the DSA promotes the standardization of the electronic submission of Article 16 notices. Such a standardization could have an important impact on the practical usefulness of notice-and-action mechanisms as a tool for access to justice. More specifically, standardization of notice affordances may help avoid dark patterns, and ensure that affected parties have equal access to the mechanism, irrespective of the type of illegality they are reporting. This may help overcome the current status quo, in which platforms facilitate the reporting of certain categories of illegal content, while failing to do the same for others ([e.g. “advertisements” and other unfair commercial practices](#)).

Under Article 17, if platforms take content moderation measures (including not only take-downs or account terminations, but also, for example, [deprioritizations or demonetizations](#)), they are obliged to provide a statement of reasons to the affected users. Interestingly, the DSA does not require such a statement in cases where a platform *refuses* to take moderation measures, following a notice. Despite the somewhat one-sided scope of application of the provision, Article 17 enhances transparency in some meaningful ways, obliging platforms to disclose for example the nature and scope of the measure (thus minimizing the grey area of “shadow bans”), as well as the legal or contractual ground relied upon. From this last point of view, the DSA draws a sharp distinction between moderation of illegal content, and moderation on the basis of the platform’s own contractual terms and conditions. Interestingly, this dichotomy is not entirely consistent with the approach taken by the Oversight Board, which [frequently interprets Meta’s community standards in light of international human rights law](#), rather than simply on the basis of the applicable contract law. In sum, despite some important limitations, the statement of reasons under Article 17 should provide insights into what the decision amounts to, and why it was taken. This information, in turn, can inform the future dispute resolution strategy of the affected parties.

Article 20 of the DSA requires platforms to put in place an internal complaint-handling system, [partially modeled after](#) the [Platform-to-Business Regulation](#). This system is accessible both in cases where the platform has taken a moderation measure, and in situations where it has declined to do so; thus, both users posting content and parties submitting a notice can access the complaint-handling system. Article 20 sets forth some basic (and rather vague) guarantees. The system must be available electronically and free of charge for at least six months after the platform’s decision. While the provision requires the system to be “easy to access” and “user-friendly”, no real procedural standardization is required here: the platforms remain largely free to decide how to organize their complaint-handling system, and the requirements of Article 20 can potentially be met by a wide range of different mechanisms, spanning from “appropriately qualified” human moderators to a highly judicialized body such as the Oversight Board. In any event, the platforms are obliged to reverse their original decision when sufficient grounds exist, and they are prevented from handling complaints solely through automated means. In practice, the lack of detail in Article 20 may prove detrimental to the possibility for internal complaint-handling mechanisms to ensure effective access to justice: the experience of international arbitration, for instance, demonstrates that the success of an alternative dispute resolution mechanism hinges (among other factors) on the [availability of a predictable procedure, which remains comparable across different service providers](#).

## **Access to justice outside of platforms**

As already noted, the unprecedented volume of content-related disputes cannot be effectively dealt with by state courts. In order to guarantee access to justice, thus, it is necessary to provide any affected party with cost-effective and reasonably fast alternatives, [as the experience of high-volume online dispute resolution has been showing for over two decades now](#). To this end, Article 21 of the DSA foresees the

possibility to access out-of-court dispute settlement mechanisms, where the content moderation decisions made by platforms can be reviewed. In a similar vein, the European lawmakers have already attempted to meet the dispute resolution needs of consumers, by encouraging alternative dispute resolution with the [Alternative Dispute Resolution Directive](#) and the [Online Dispute Resolution Regulation](#). Article 21 of the DSA, in particular, enables the Digital Services Coordinators of each Member State to certify dispute settlement bodies established on their territory (according to a procedure which only partially resembles Article 20 of the ADR Directive). Once certified, these bodies can offer dispute settlement services to all parties seeking redress against a platform decision: not only users at the receiving end of a content moderation measure, but also parties that have filed an unsuccessful notice under Article 16, and users that were unable to obtain redress through a platform's internal complaint handling mechanisms. In other words, the DSA aims to enlarge the market for dispute resolution, with the complainant being able to choose among different (private, and sometimes public) certified dispute resolution bodies.

The [experience of the European ODR Portal](#) demonstrates that alternative dispute resolution risks becoming a [paper tiger](#), if the traders (or, in the case of content moderation, the platforms) have no incentive to participate in the dispute resolution procedure and comply with its outcome. From this point of view, the original DSA proposal was bold: [platforms would be bound by the decisions taken by the certified bodies](#). The final text is, from this point of view, much less demanding: platforms must inform the users about the possibility to appeal to a dispute settlement body and must generally engage in good faith in the procedure, but have no obligation to comply with the outcome (Article 21(2)). This, however, does not automatically make out-of-court dispute settlement ineffective. The cost structure of these procedures remains extremely attractive for users when compared with court litigation, and platforms have a transparency obligation (under Article 24) to disclose “the share of disputes where the provider of the online platform implemented the decision of the body”. Furthermore, compliance with the outcome of these out-of-court procedures may become part of the risk mitigation measures of very large online platforms (VLOPs) under Article 35. In sum, even if out-of-court dispute settlement has been significantly watered down (compared to the original proposal of the Commission), the overall framework of the DSA does recognize a meaningful role for these procedures, and VLOPs will not be able to systematically ignore the existence and outcomes of out-of-court dispute settlement. In practice, the impact on the protection of marginalized groups will also depend on what type of bodies will obtain certification, and what the purview of their expertise will be. At the very least, the information obligations of Article 21(4) will provide some transparency in this respect.

Finally, in addition to the possibility to lodge a complaint with the competent Digital Services Coordinator (Article 53), court litigation is never precluded under the DSA: the dispute resolution options described so far never impair the possibility for affected parties to initiate court litigation, seeking e.g. the removal or reinstatement of online content. Furthermore, the right of the service providers to compensation for infringements of the DSA is expressly enshrined in Article 54. Nevertheless, court litigation will often remain inaccessible in practice for many affected parties, and [the](#)

[costs and duration of proceedings will vary dramatically across the Area of Freedom, Security and Justice \(AFSJ\)](#). These factual obstacles often preclude effective access to justice, especially for marginalized groups and impecunious litigants. In addition, the current European framework for content moderation-related litigation is fraught with doubt, concerning *inter alia* jurisdiction. Despite the fact that litigation involving very large platforms will often be cross-border in nature, the DSA does not enshrine any special jurisdictional rule, so that claimants will need to resort to the [Brussels I bis Regulation](#) to establish jurisdiction before an EU Member State court. This, in practice, may turn out to be challenging: some claimants, for instance, may [fail to qualify as consumers](#), and thus be unable to establish jurisdiction in their home court. Furthermore, the application of the traditional tortious grounds of jurisdiction to Internet-based harms leads to a potential [splintering of jurisdiction all over the AFSJ](#), thus hampering legal certainty.

A final layer of doubts concerns the possible role of collective redress: could class actions become a tool for the protection of marginalized or vulnerable groups, affected by harmful online content? From this point of view, the DSA introduces some important innovations. First of all, Article 90 amends Annex I to the [Collective Redress Directive](#), thus enhancing the possibility ([already existing in some Member States](#)) of class actions for content moderation disputes. Furthermore, Article 86 expressly enables recipients of intermediary services to mandate a representative body to exercise their rights on their behalf.

## Conclusion

When observed in detail, the “procedure before substance” approach of the DSA leaves many questions unanswered. The final text of the Regulation contains compromises (e.g. concerning out-of-court dispute settlement), and blind spots (e.g. the absence of jurisdictional grounds for moderation-related litigation). However, the DSA also brings about important procedural improvements, concerning e.g. notice-and-action mechanisms and statements of reasons. Looking at the allocation of powers across these different dispute-management and dispute-resolution avenues, there seems to be a growing expectation that platforms (especially very large ones) will contribute to law enforcement in Europe, and will apply legal standards when engaging in decision-making (concerning e.g., whether content is illegal, or incompatible with the platform’s own general terms and conditions). However, many questions remain open. As far as access to justice is concerned, one of the most urgent ones is how EU Member State courts can deal with the growing challenges of the European digital space, while relying on a jurisdictional framework that dates back, in its overall architecture, to the [1968 Brussels Convention](#). Furthermore, to what extent can the procedural innovations of the DSA address the challenges of content moderation, in the absence of any major harmonization of the substantive law applicable in this very broad and porous area? In the 1989 drama *Field of Dreams*, a mysterious voice whispers to Kevin Costner, “If you build it, they will come”. The DSA has built (or, at least, enhanced) a procedural framework for content moderation disputes. Will legal certainty and access to justice follow? Only time will tell.

