

upgrade
democracy

Impulse #2

May 2023

Potentials of small platforms: What can we learn from them?

Cathleen Berger, Charlotte Freihse, Vincent Hofmann,
Matthias C. Kettemann, Katharina Mosene

Translation by Tim Schroder

Legal notice

© Bertelsmann Stiftung, Gütersloh
May 2023

DOI 10.11586/2023078

Bertelsmann Stiftung

Carl-Bertelsmann Straße 256

33311 Gütersloh

www.bertelsmann-stiftung.de

Upgrade Democracy

www.upgradedemocracy.de

Authors

Cathleen Berger, Charlotte Freihse, Vincent Hofmann, Matthias C. Kettemann, Katharina Mosene

English translation

Tim Schroder

Responsible for content

Cathleen Berger

Design

nach morgen

About the Impulse Series and this Publication

In a collaboration between the Bertelsmann Stiftung's Upgrade Democracy Team and the Humboldt Institute for Internet and Society, we are organising a five-part Impulse Series on "**Digital Platforms: Design Proposals and Alternatives**" from April to September 2023. The focus of the series is an in-depth examination of current challenges and problems on existing social platforms and the identification and discussion of alternatives. Regarding the democratic design of dominant social platforms, special attention is placed on the topics of participation and platform governance as well as on questions of how to make decision-making processes more oriented towards the public common good. The individual impulses intertwine thematically and aim to develop ideas, action and policy recommendations for sustainable platform and content governance in digital spaces.

All impulses take up ideas from expert workshops, in which provocative hypotheses and central questions are discussed in small, intimate groups under Chatham House rules. Following each event, an impulse paper summarising the most important aspects of the discussion is published.

The series was conceived by Cathleen Berger, Charlotte Freihse, Matthias C. Kettemann, Katharina Mosene and Vincent Hofmann.

Participants during our discussion on May 16, 2023

Conversation Starters:

- Christina Dinar, Leibniz-Institut für Medienforschung | Hans-Bredow-Institut
- Luca Thüer, Liquid Democracy e.V.

Experts:

- Benjamin Fischer, Alfred Landecker Stiftung
- Josefa Francke, Leibniz-Institut für Medienforschung | Hans-Bredow-Institut
- Daniela Heinemann, nebenan.de
- Jerome Trebing, Amadeu Antonio Stiftung
- Felix Sieker, Bertelsmann Stiftung
- Falk Steiner, freelance journalist
- Sven Winter, gutefrage.net

Moderation:

- Cathleen Berger and Charlotte Freihse

1 What are small platforms and what potential insights do they offer?

The academic and public discourse on platforms focuses mainly on the industry's major digital sites, which are therefore the target of regulatory activities (such as the Digital Services Act, or DSA, at the EU level). What is lacking, however, is a debate about platforms that goes beyond Facebook, TikTok and the like. It is worth taking a look at smaller platforms, since not only is it easier to observe how content is moderated there, such sites are also exciting and promising spaces for experimentation when it comes to democratic innovations in platform governance.

Yet which platforms are considered “small”? To date, there has been no systematic and comprehensive survey comparing small and medium-sized sites. Yet it makes sense to look at smaller platforms, since they are where new, creative, customer-centred and inclusive approaches to responsible content moderation can be found. For example, unconventional and community-based methods of moderation and rule enforcement exist that promise a fair and more democratic balance between the freedom of speech of as many participants as possible, on the one hand, and the interests of platform operators, on the other.

Existing regulations such as Germany's Network Enforcement Act (Netzwerkdurchsetzungsgesetz, or NetzDG) or the special provisions laid out in the DSA for very large online platforms (VLOPs) do not apply to sites with smaller numbers of users. In addition, the landscape of small and medium-sized platforms is very diverse, extending from neighbourhood forums like nebenan.de, question-and-answer sites like GuteFrage.net, and LGBTQIA+ networks like planetromeo, to alternative networks like Jodel. Other examples can be found in the political sphere. Germany's Pirate Party, for example, used its own platform to organise debates, present its positions and meet its own above-average standards of transparency. The platform contained transcripts, minutes of meetings, and draft legislation as a result. Smaller platforms often offer a space for exchange that is geared to certain interests – a small, protected, non-competitive area where users can spend time, express themselves and flourish. They are smaller because they are dedicated to specific issues or geographic locations. How does this diversity influence the platforms' design? Is it possible to generalise from the insights they offer? And can their successful approaches to moderation be scaled?

2 Smaller platforms: Content moderation and design

Previous research has shown that smaller platforms rely on the involvement of their own community for content moderation. Depending on the extent of the involvement, this can encompass all steps in the moderation process, from individual decisions about specific posts to how community guidelines are created. This contrasts with the approaches major platforms take to moderation, which are generally supported by technology and handled by large teams of full-time staff. When it comes to content moderation, smaller platforms usually have a vested interest in gearing their rules and practices to the needs and wishes of their own community. Their success lies in the fact that the (relatively) small number of users feel at home on the platform and thus continue to use it. Against this background, we have formulated two hypotheses, which we have examined together with recognised experts:

- **Human or manual moderation in dialogue with a platform's own community creates a high level of identification with the rules used on the site and greater acceptance of the standards set.** Involving everyone in making the rules not only leads to better moderation practices, but also to fewer violations, since users identify more with the rules governing the platform's moderation.
- **Smaller platforms voluntarily adapt their guidelines and rules for moderation to the regulatory requirements large platforms must meet.** In doing so, they contribute to how those requirements are understood and to their expected enforcement.

These hypotheses lead to four key questions:

1. How can we gain a structured overview of the rules and strategies used for content governance on smaller platforms?
2. Which methods of content moderation that have been successfully implemented by smaller platforms can be transferred to large platforms?
3. What role do smaller platforms play in democratic discourse? Do best practices exist here that increase the visibility of marginalised groups and their needs?
4. Can smaller platforms influence the implementation of regulations through the approaches and strategies they adopt?

2.1 Content governance on smaller platforms: What we know and what we don't (yet)

Smaller platforms are diverse both in their design and in the content they offer. They thus create spaces for communication that can be very instructive for public regulators of online sites as well as for the major platforms. While the VLOPs that fall under the jurisdiction of the European Commission have been identified for the DSA, lists of all other platforms have yet to be assembled. Smaller platforms are currently under the remit of the national Digital Services Coordinators (DSCs). In Germany, a survey of digital services is already being planned by the Federal Network Agency (Bundesnetzagentur, or BNetzA), which will assume the role of DSC. A few examples:

Example 1: Adhocracy+

Liquid Democracy e.V. is an interdisciplinary team that aims to create a democratic culture in which active participation is a given for everyone. To achieve that goal, it uses a variety of approaches: developing tools such as adhocracy+, designing processes for digital participation together with public

administrators, policymakers and civil society, and organising research projects (e.g. KOSMO in cooperation with Heinrich Heine University in Düsseldorf).

- The **Adhocracy+** platform has been online since 2019. More than 300 organisations use the site, and 700 participation projects with ca. 12,000 users have already been carried out. Data are free of charge and publicly accessible on GitHub. The 10 modules for participation available on the platform range from simple brainstorming and surveys to interactive events and structured debates. The implementing organisations can determine which participants – and moderators – are included and they themselves are responsible for what happens in the modules.
- Compared to VLOPs, there is significantly less hate speech on Adhocracy+, which could be the result of each organisation designing its discussion spaces itself, something that makes the spaces inherently more consistent. An “**activating**” **content moderation** seems to have a particularly positive impact: This means moderators interact with users' posts, ask questions, give feedback and sometimes highlight especially constructive contributions for all users. All activities must be transparent and are – in contrast to VLOPs – not aimed at increasing time spent on the platform, but in encouraging users to enter into issue-related dialogue. This is undoubtedly labour-intensive and requires that moderators be integrated into the relevant discussions to a high degree.

Example 2: gutefrage.net

Founded in 2006, this German-language question-and-answer platform gets 70,000 new posts each day. A total of 13 people work as content moderators, with three on duty during each shift, ensuring moderators are present from 8 am to midnight.

- The platform takes a complementary **four-pillars approach** to content moderation: (1) Any **user** can report any type of issue. (2) Permanent moderators receive training every two months. (3) The site involves **user-moderators**, who flag a large number of posts and have a high success rate. (4) **Pre-sorting technology** is used to mark posts that are likely to be flagged. An algorithm is deployed here that accesses data on past deleted posts to predict whether a contribution violates community rules.
- These procedures were developed over time in dialogue with the community, and they have been continually fine-tuned in light of the resources they require. While the user-supported moderation functions quite well, in-depth discussions and current events can occasionally result in especially challenging situations.

Example 3: nebenan.de

Launched in 2015, the platform nebenan.de focuses on exchanging services that promote the building of neighbourhoods and strengthen cohesion within them. The four people who work as content moderators (full- and part-time) manually view the content posted by the site's 2.6 million users that has been flagged. The number of reported posts is relatively small, at less than 1 percent of all content on the platform.

- The platform has the clear goal of making neighbourhoods more cohesive and, in that context, promoting the well-being of its communities. Against this background, the platform has been able to quickly adapt its moderation practices in light of current events, such as the Covid pandemic and the war in Ukraine, and it has set comparatively strict rules that limit the space for non-neighbourhood-related issues.

- Technology-supported moderation could unleash untapped potential for the platform, especially in the form of an early-warning mechanism. In such cases, users would be asked again if they really want to post certain content and if the content is actually relevant to the neighbourhood.

A look at these three examples of smaller platforms reveals two things: First, more research is needed, especially to gain systematic insight into good practices and current challenges. Second, considerable potential exists to learn more about content moderation from existing practices, to further improve these practices and to explore possibilities for transferring them to large platforms.

2.2 Content moderation on smaller platforms: Transferable to large sites?

We would ideally like to give a simple answer here in the form of a list of recommendations for large platforms. Our discussions have made clear, however, that practices cannot be transferred from one group to the other in a one-size-fits-all manner. Nonetheless, we would like to highlight four ideas:

1 User-moderators

[Article 22 of the DSA stipulates](#) that “online platforms shall take the necessary technical measures to ensure that notices ... submitted by trusted flaggers ... are processed and decided upon without undue delay”. These trusted flaggers thus play a special role in the system of content moderation envisaged in the DSA. From the perspective of smaller platforms, however, they represent only one of many possibilities for involving civil society in moderation activities. As a rule, content moderation requires contextual knowledge, and the experience of smaller platforms shows that discussions are perceived as healthier and more constructive when users or the community are involved in content moderation. This is one building block from which large platforms could benefit when it comes to content moderation. In such models, a site’s moderation is organised along different levels or degrees of involvement, commitment and willingness to take responsibility. Thus, users can volunteer to moderate local groups or topics, for example by using activating methods that de-escalate situations or highlight constructive posts. At the same time, we would like to mention here our [first Impulse on the fediverse](#), which also discusses the use of volunteers. This approach has its advantages, but also requires considerable awareness of the problems arising from volunteer activities: little representation of marginalised groups, a lack of resources, a need for training and the necessary processes of supervision and conflict resolution, to name but a few.

2 Technological support and partially automated moderation

As soon as content is automatically captured, filtered or categorised, certain legalities apply. This can create obstacles for smaller platforms. Nonetheless, pilot projects suggest that technological support mechanisms could prove helpful for content moderation, and could even be a good addition to large platforms. For example, the prototype of the KOSMO tool groups comments into four categories, pre-sorting them with the help of artificial intelligence.

As part of this automated pre-sorting, comments that are marked as needing moderation can be blocked on the platform – and only become visible again after a moderator has approved them. This is meant to help moderators work through individual posts and lessen the pressure of deciding what to approve, what to block and what to delete.

Similar pre-sorting tools are deployed by other platforms, which train their algorithms using previous moderation decisions made on their own site. What is crucial for developing such tools is a precise recording of the reasons why a decision was made, so that classifications are transparent and comprehensible – and can be applied again to new developments. Given the dynamics of democratic discourse, however, there is general agreement that such classifications should only be supported by technology and should not be fully automated. Human moderators still have a key role to play in these decision-making processes.

3 Guidelines for current developments

Smaller platforms are creative and diverse in terms of their content moderation methods. They often develop their systems to reflect the needs of their users or community, quickly adapting them to reflect current developments (e.g. elections, the Covid-19 pandemic, the war in Ukraine).

Guidelines for such events can be assembled within a few days, as soon as a controversy arises. What are helpful here are close contacts among platform operators and a high degree of willingness within the community to react proactively and quickly to unfolding developments. This bottom-up approach could be deployed by large platforms for specific groups, local offerings and the like.

4 Differentiating between informational offerings and moderation

When informational or educational offerings are available – such as support for de-radicalisation efforts, suicide prevention counselling or the identification of developmentally damaging user behaviour – it is crucial that these activities not be considered content moderation or assigned to moderators as one of their responsibilities. A healthy, attentive and activating culture of discussion is needed if such offerings are to be made available.

If technology is used to pre-sort content on (large) platforms to an undue extent, it is possible for posts to fall through the cracks that pose valuable pedagogical questions or that contribute to the on-line discussions. Community-based approaches are advantageous here, since they can take the relative contexts into account.

2.3 Smaller platforms and the potential for democratic discourse

As shown above, thanks to their specialist nature, smaller platforms have considerable potential when a discussion is limited to a specific location or issue. Each of these platforms strikes its own balance between freedom of speech and undesirable content.

Nevertheless, discourses also take place on smaller platforms that either reference debates occurring within society as a whole or have a direct impact on those debates. This is especially evident in the run-up to political events such as elections. Certain smaller platforms test effective methods for identifying problematic patterns, trolls, bots or similar phenomena on their site in order to counteract them. These are only individual cases, however; until now, no systematic exchange or sharing of experiences has taken place.

Smaller platforms would undoubtedly benefit if more was done to pool synergies and learn from each other. Coding could conceivably be shared, for example, along with systems for flagging log-ins, recognising patterns, etc. There seems to be great potential on the national level, or between platforms

operating in the same language, to develop democratically embedded good practices that respond rapidly to unfolding events and, once they are developed, to share and evaluate them. After all, specialised platforms not only contribute to democratic discourse, they also play a key role in shaping that discourse within their communities.

2.4 Smaller platforms: Regulation and implementation

Only in individual cases have smaller platforms made it onto policymakers' radar screens. They are greatly underrepresented in the regulatory discourse as a result. Moreover, they lack political influence.

When standards are set without due consideration, smaller platforms can be saddled with accountability or moderation-related obligations that they cannot easily fulfil, given their decentralised or user-based systems of moderation. The same is true when a platform is required to verify or control content or behaviour – obligations that make sense for large platforms, but which can have a crippling effect on smaller sites and which have led or can lead to the closure of business locations.

It is thus imperative that smaller platforms be involved in regulatory issues and in efforts to determine how platforms can effectively meet their legal obligations. The DSA has already been adopted and is now in its first phase of implementation. Smaller platforms must be heard soon by the national DSCs if their perspectives and needs are to be integrated constructively as implementation proceeds.

2.5 Observations and recommendations for content moderation on all platforms, small and large

Our discussions with the experts during the workshop and beyond have given us inspiring insights into all four areas covered by our questions. The landscape is diverse, the perspectives controversial, the discourse engaged and open to development. Despite all the diversity, it is still possible to provide a number of observations and recommendations for the future of platform governance – both for the examples of smaller platforms given above and for their transferability to large platforms, as well as for the potential smaller platforms offer democratic discourse and their impact on regulation and implementation. Here are the elements we feel are most important when designing content moderation mechanisms:

- **Nuance:** When it comes to smaller platforms, the rules must be nuanced. It's good that smaller platforms must meet less extensive requirements. At the same time, it is important to consider the risk that a platform's rules might inherently violate fundamental rights. The obligation to be transparent is appropriate and expedient here, even though, when a platform's content moderation is being evaluated, it must be borne in mind that mechanisms which involve users follow different rules and logics. To be specific: The transparency reports provided by all platforms must include a range of information, including how often posts are flagged, blocked, deleted and restored. The experience of smaller platforms shows that decisions are reversed more often when a large number of users are involved in the site's moderation. That means the balance between participation and transparency must be carefully considered. This is the only way to ensure that a democratically designed process is not seen as being impacted by a lack of effective moderation.

- **Visibility and contact person:** There should be at least one contact person for smaller platforms who is aware of how diverse platforms function and what their needs are. Whether this role is assumed by the DSC, e.g. the Federal Network Agency in Germany, or another entity is secondary. It is important to recognise that the DSA generates considerable work for smaller platforms and can even lead to a standstill in technological development if it ties up too many resources. This main contact could also engage more with issues like guidelines for responding to current events, such as the pandemic, the war in Ukraine and the like. Smaller platforms in particular would benefit if guidelines were developed and made legally binding across sites so that fewer individual resources were required.
- **Responsibility and protocols for conflict resolution:** Such mechanisms are sometimes supported by community management, in that specially designated users directly involve the community in resolving conflicts. The goal is to promote a feeling of responsibility for and identification with the platform among users. Activities often take place locally and globally in a way that is sensitive to both language and context, thereby mediating between different perspectives and interests. A compelling example here is the fediverse, which can be seen more or less as a hybrid platform: With its decentralised moderation, it can contribute to the good practices of the smaller platforms. At the same time, given its interoperability, it has excellent potential for growing into a very large platform in terms of the number of users.
- **Advisory and support services:** More resources will be required – on all platforms – as regulatory efforts proliferate and awareness grows of content moderation’s importance. Smaller platforms in particular would benefit – and could better achieve their goal of serving specialised communities – if advisory and support services were firmly established and if they received government funding. Legal counselling could be offered, guidelines could be developed, and an exchange could be organised between smaller platforms on their experiences and successes with the processes they have adopted. Another possibility would be to make tools and well-functioning, technologically supported methods available to smaller platforms at no charge as free and open software.
- **Strengthening the professional profile of content moderators:** The discussion has shown that manual moderation remains indispensable if various contexts and rapidly changing nuances are to be properly addressed. While technical mechanisms can provide support, moderators remain responsible, especially for content moderation that engages users and for pedagogically sensitive interventions. This means, on the one hand, that we foresee a growing need for formally trained content moderators and, on the other, that the profession must become much more visible and be accorded much more respect. The German trade association Berufsverband für Community Management has called for the same change, whose necessity is made all too clear by the very questionable conditions under which moderators at large platforms are known to work.

One compelling question remains open for us: When is the limit reached after which the interplay of automatic and manual moderation no longer works?

An idea that we have not been able to discuss conclusively in this regard is that of “platform councils”. There are some interesting possibilities and tensions emerging here that must be explored in greater detail. Our next impulse paper will look explicitly at this topic. Can this approach be used to strike an effective balance between community-based moderation and global discourses?

3 What else is there to know about smaller platforms?

The following are recommended readings on small tech and related developments. If you have any suggestions to add to the list, please let us know!

- A good overview of the issues discussed here can be found in the post by Christina Dinar from the year 2021: <https://leibniz-hbi.de/de/blog/von-den-kleinen-plattformen-lernen>
- In connection with the most recent event in our series and the benefits offered by small and medium-sized platforms, we recommend: Tagesspiegel Background, 16 January 2023, Maximilian Henning, “Was Mastodon von kleinen Plattformen lernen könnte”, <https://background.tagesspiegel.de/digitalisierung/was-mastodon-von-kleinen-plattformen-lernen-koennte> (€)
- On how small and medium-sized platforms are involved in regulatory efforts: Tagesspiegel Background, 12 November 2021, Friederike Moraht, NetzDG, <https://background.tagesspiegel.de/digitalisierung/netzdg-co-kleinere-plattformen-wuenschen-sich-mehr-gehoer> (€)
- Many people talk about the important role that large platforms play in elections. But what role did smaller and medium-sized platforms play in Germany’s “super election year” of 2021? This question is discussed by Sven Winter from gutefrage.net, Felix Kartte from Reset – Internet for Democracy, and Daniela Heinemann from nebenan.de: YouTube Tide Event, 23 September 2021, “Kleinere und mittlere Plattformen im Superwahljahr”, <https://www.youtube.com/watch?v=mwRtwqPrX3M>
- Volunteers play an important and positive role as moderators on smaller platforms. Yet it is important to consider the sometimes problematic structures used for volunteer activities, a point discussed by Berger/Freihse/Hofmann/Kettemann/Mosene in their impulse paper “Decentralization as Democratization: Mastodon instead of Platform Power”, <https://upgradedemocracy.de/impulse/dezentralisierung-als-demokratisierung-mastodon-statt-plattformmacht/>
- How the platform adhocracy+ from Liquid Democracy e.V. works and why it’s so important to make participation accessible for as many people as possible is discussed here by Frederik Wegener (2020): [Newsletter NBB \(netzwerk-buergerbeteiligung.de\)](https://newsletter-nbb.netzwerk-buergerbeteiligung.de)
- What does Liquid Democracy mean? In this interview from 2022, Moritz Ritter explains the theoretical concept, implementation strategies and technical tools used by the nonprofit Liquid Democracy e.V.: <https://democracy-technologies.org/industry-news/fluid-representation-improved-deliberation-and-liquid-democracy>
- Small platforms are very important for queer discourses. For an introduction to the role of the queer perspective on technology development (and the significance of AI in creating the digital constellation), we recommend: Michael Klippahn-Karge, Ann-Kathrin Koster and Sara Morais dos Santos Bruss (eds.), Queere KI Zum Coming-out smarterer Maschinen, <https://www.transcript-verlag.de/978-3-8376-6189-7/queere-ki/?c=313025272>
- Cathleen Berger has put together a summary of the role played by the Digital Services Act, including definitions and the act’s applicability to different platforms [Your Cheat Sheet for the Digital Services Act: What and for Whom - Upgrade Democracy](#)