Primavera De Filippi & Melanie Dulong de Rosnay, Regulating Distributed Peer-Production Infrastructures, The Internet, Policy & Politics Conferences, IPP2014: Crowdsourcing for Politics and Policy, Oxford Internet Institute, University of Oxford, September 2014.

Regulating Distributed Peer-Production Infrastructures

Introduction

Crowdsourcing designates a production process distributed among a large number of peers, which all contribute with their own resources to a common goal. The process can be either centralized, i.e. when the contributions of a disparate group of peers are coordinated through one central authority, or decentralized, i.e. when peers coordinate themselves in a distributed manner, without relying on any centralized authority.
This paper targets a specific kind of online peer-production platforms - so-called commons-based production platforms (Benkler, 2006) - which implement decentralization both at the level of the technical infrastructure (i.e. with a decentralized, peer-to-peer architecture) and at the governance level (i.e. ownership of both the platform and the output of production is distributed or shared in common among all peers, instead of being controlled by a central entity). It will focus, in particular, on three distributed peer-production platforms: Kune, a federated platform for community management and collaborative production; Twister, a decentralized peer-to-peer micro-blogging platform; and Globaleaks, an anonymous, censorship-resistant, distributed whistle-blowing platform.

After analyzing the benefits they might offers in terms of user's autonomy (Section 1.1), privacy (Section 1.2), anonymity (Section 1.3) and freedom of expression (Section 1.4), the paper will investigate the legal challenges they raise in terms of copyright infringement (Section 2.2), hate speech (Section 2.3) and cyber-criminality, more generally (Section 2.3). The paper will then move on to illustrate the regulatory options available to both policy makers and platforms designers to address these challenges (Section 3).

1. Benefits of decentralized architectures

The architectural design of online peer-production platforms has an important impact on legal rules and policy choices, insofar as they determine how these can effectively be enforced.
Most centralized online platforms raise significant challenges not only in terms of information security, but also – and perhaps most importantly – in terms of user's autonomy, privacy and freedom of expression. With regard to the former, data hosted and controlled by one central authority is more vulnerable to external attacks to the extent that the centralized server constitutes a single point of failure. With regard to the latter, the internal policies of various online service providers might go counter the interests of end-users (see, e.g. Google's intrusive privacy policy, Facebook's experiments with emotion, or the extensive collaboration with

intelligence agencies which have been recently revealed by Edward Snowden) by impinging upon users' civil liberties and fundamental rights.
Decentralized network architectures may help mitigate some of these risks. Indeed, distributing both the control of the technical infrastructure and the management of user data amongst a distributed network of peers makes it possible to achieve a safer, more efficient, and potentially more democratic use of resources, as users not only consume resources, but also provide their own resources to the network (Schollmeier, 2001).
From a technical perspective, this is advantageous because, as the load of users increases so will the overall capacity of the network. Decentralized P2P networks are therefore more efficient than their centralized counterpart, but also more resilient since the system cannot be disrupted through the failure of one or more nodes.
But beyond the technical benefits, decentralized architectures also present important social and legal implications (Elkin-Koren & Salzberger, 2004). Most of these platforms are self-regulated through specific governance models that do not rely on any hierarchical structure or centralized authority for coordination (Moglen, 2010; Aigrain, 2012). Governance rules are directly embedded in the protocols and technical features of the online platforms, which apply equally to everyone. To the extent that they are created by the community and for the community these rules are likely to preserve user autonomy and fundamental rights such as privacy and freedom of expression.

1.1. User Autonomy

Decentralized applications are, by design, more likely to support and promote the autonomy of end-users. By relying exclusively on the computing resources of individual users or peers, it becomes indeed possible to create network environments which are not controlled by any third party operator, but only and exclusively by the actual members of the community.
Kune, for instance, is a community management tool that relies on a federated network of nodes, autonomously deployed and independently governed by a particular community of users.

Yet, it is important to note that power does not disappear with decentralization; it just get transformed and relocated. New sources of power control are likely to appear in decentralized networks (e.g. at the level of governance), which are different than the once present in centralized networks but might nonetheless lead to concentration of power in the hands of a few super-nodes.

1.2. Privacy

Online privacy is directly related to the question of logging and data retention. Indeed, everything we do on the Internet leaves traces. In the case of centralized architectures, these traces are recorded into a central server, whereas, in the case of decentralized, peer-to-peer networks, they are recorded into the local devices of users connected to the network.
In both cases, logging and monitoring is necessary to ensure the functionality of the network (e.g. to coordinate users and regulate online communications), as well as to perform statistical analysis aimed at increasing the overall efficiency and quality of service.

Yet, differences emerge as regards the type of data (or metadata) that is being collected, who has access to that data, and for how long it is being retained. In particular, centralized platforms often rely on cookies, spywares, malwares or other forms of navigation tracking to collect additional information about their user-base, which might be used the purposes of profiling users, to implement better marketing campaigns and extract greater advertising revenues. Most of these practices do not comply with privacy and data protection regulations, to the extent that they impinge upon the privacy of end-users. Yet, enforcement is difficult to achieve in a global and transnational environment as the Internet network.
Re-claiming ownership over the platform's technical infrastructure allows users to more easily control the manner in which and the extent to which content or data stored into the system will be accessed and subsequently exploited by the community (as opposed to third parties). As opposed to traditional client-server applications, where data is controlled by one central authority - which can thus potentially disclose it to third parties - decentralized applications (such as Twister or Kune) enable users to maintain ownership and control over their own (personal) data.
In some countries, such as most EU member states, data retention is actually mandated by law. Directive 2006/24/EC requires all online operators to retain users' data for a minimum of 6 months (and for a maximum of 24 months) in order to facilitate police investigations and judicial reviews. Yet, in the context of decentralized network architectures, to the extent that information is not stored in a central location, but rather on the local devices of individual users, it is extremely difficult for any corporate or governmental party to access such information, even if the event of a warrant.

1.3. Anonymity

Anonymity is important to the extent that it empowers people to manage their own privacy, either by concealing their actual identity or by creating a set of online identities that differ from their real identity - by means of a pseudonymous, for example. The degree of anonymity (or pseudonymity) ultimately depends on whether users' actions and communications can be easily traced back to a particular online identity (and whether such identity is connected to an actual real-world identity).
Most centralized online platforms require users to register before using the platform – and some even implement a real-name policy (enforced through government IDs).
Conversely, decentralized platforms make it easier for users to remain anonymous (or to the least pseudonymous) insofar as they can use the platform without having to register to any centralized authority. Some even take extra steps to preserve the anonymity of their users, by implementing specific mechanisms to anonymise the source and/or destination of online communications. This is the case, for instance, of Globaleaks which relies on the Tor network to make it impossible for third parties to track users' activities consistently over time.
Tor is based on a decentralized network architecture designed to preserve the anonymity of users. Anonymity is achieved by means of a specific routing mechanism (onion routing) relying on several layers of layers of encryption (nested like the layers of an onion) in order to conceal the source and content of communications. Yet, the technical measures employed by Tor to protect the anonymity of users are not sufficient, as such, to preserve the security and

confidentiality of online communications. While the onion routing makes is virtually impossible to trace back the source of communications, if users do not take care to encrypt the content of their communication, people connected to the Tor network as “exit nodes” (i.e. those in charge of sending packets outside of the Tor network towards the actual destination) could try to infer users’ identity by monitoring (or sniffing) the packets transiting through their node. Therefore, privacy-by-design can only be ensured through a combination of tools based on decentralized architectures, including the lack of registration, no data retention, anonymous browsing and encryption of communications, may guarantee.

1.4. Freedom of expression

The right to freedom of expression – including the right to access to information – can be significantly affected by CBPP design choices. Indeed, different types of architectures (more or less centralized) might either support or impede the practices of surveillance, filtering, or censorship from the part of both platform providers and users.

As a general rule, censorship is much easier to achieve in the case of centralized platforms than in the decentralized counterpart, since platform operators have the ability to intervene on the accessibility of information they store in their own servers. As a result, they might engage into various forms of censorship (ex-ante, ex-post, or on-going) by introducing a variety of technical and non-technical measures (i.e. automatic detection of infringing content, manual take-down procedures, removal from search engines, etc.) aimed at limiting the accessibility of specific content according to different criteria (e.g. type of information it contains, user location, etc).

In certain cases, these measures are actually requested by law, which requires online intermediaries to suppress information that might be regarded as unlawful, harmful, or objectionable by certain governments or public authorities.

Oftentimes, online operators – which are not subject to the legal and constitutional safeguards that prohibit (or limit) governmental censorship in several areas of society –implement their own (arbitrary) content filtering policies which often go further than what is actually required by law (see e.g. Facebook, Google, or even Reddit’s policies which explicitly grant platform operators the right to delete content that they be regarded as improper or objectionable).

Conversely, the architecture of decentralized P2P networks (that do not rely on any specific intermediary or gatekeeper) is such as to ensure that communications cannot be blocked, nor filtered by any given third party – be they corporate or governmental entities. In the case of Twister, for instance, its decentralized architecture precludes any possibility for censorship: individual accounts cannot be blocked (as opposed to Twitter) and only the account owner has the ability to edit or remove posts. Similarly, GlobalLeaks is a distributed application that relies on the Tor network’s “hidden service” functionality in order to provide a censorship-resistant whistle-blowing platform, which preserves the anonymity of users contributing to the platform.

2. Legal challenges raised by distributed architectures

In spite of their advantages in terms of user autonomy, privacy, anonymity and freedom of expression, decentralized P2P networks also present a few downsides, mainly with regard to security and law enforcement.

Given their decentralized character, the security of CBPP platforms relying on distributed peer-to-peer networks is often difficult to ascertain. Indeed, many decentralized P2P networks are inherently insecure by virtue of their "open" design. To the extent that anyone is entitled to join the network (either as a client or a relay node transferring packets throughout the network), then anyone connected to that network is also capable of intercepting (or sniffing) the packets transiting through the network. Thus, unless users employ end-to-end encryption, the content of all data or communication can be monitored by third parties. As a result, the open and collaborative nature of decentralized peer-production platforms might actually go counter the security and privacy of their users. Indeed, as opposed to centralized peer-production platforms, operated by a central authority, which is also responsible for managing and securing the network, decentralized networks are operated by a distributed and undefined community of peers. Although many tech savvy individuals are generally involved in the initial set up of these networks, most of the users that subsequently connects to them are unlikely to spend much time securing the network. Thus, if a network is as secure as its weakest node, most of the decentralized peer-production platforms deployed today are likely to be less secure than the vast majority of commercial or centralized platforms.

In terms of law enforcement, in order to properly understand the extent to which the law can effectively regulate online CBPP platforms, it is necessary to look at the infrastructure and the technical features they implement. As general rule, the greater is the degree of decentralization, the harder it becomes to control or to regulate the platform. Centralized architectures are, in fact, easy to regulate because they rely on a centralized entity that essentially dictates the rules to which everyone must abide. Conversely, as a result of their distributed architecture, decentralized P2P networks often enjoy a higher rate of criminal or unlawful activities. Indeed, the regulation of decentralized P2P networks often rely on community governance (or self-governance), making it difficult for any central authority - be it either a firm, an individual or the State - to enforce its own rules on community members (Guadamuz, 2011). We focus here on three category of illegitimate activities that are promoted by decentralized P2P architectures, namely copyright infringement (Section 2.1), hate speech (Section 2.2), and cybercrime (Section 2.3).

2.1. Copyright infringement

Since the early 2000s – following the success of Napster – many decentralized P2P networks have been developed to support and facilitate music and video file sharing. By distributing small technical acts among a large number of peers, liability for copyright infringement cannot, in fact, be easily nor directly attributed to any of these peers. As a result, P2P networks have come to be considered as an important threat to the copyright industry, and the infringement of intellectual property rights has become a prominent argument for condemning the deployment and use of these networks (Elkin-Koren, 2006).

Many governments have thus enacted laws or regulations - such as, most notably, the international Anti-Counterfeiting Trade Agreement (ACTA), the Stop Online Piracy Act (SOPA) and the PROTECT-IP Act (PIPA) in the U.S. - which endow Internet service providers and online intermediaries with the ability (and, sometimes, the obligation) to police the Internet on behalf of the State (McManis, 2009). This is generally achieved through the regime of intermediary

liability limitations, whereby platform operators are responsible for some of the activities undertaken by their users unless they abide to specific monitoring obligations and comply with the notice and take-down procedures stipulated by law (see e.g. the E-Commerce Directive in Europe, and the DMCA in the US). Accordingly, by delegating the task of enforcing the law to private entities, States eventually turned Internet service providers and online operators into private police or information gatekeepers (Hintz, 2012).
In the context of CBPP, these rules apply differently to centralized infrastructures (e.g. Wikipedia) where a centralized authority acts as some kind of editor or publisher to regulate process and output of production, and decentralized infrastructures (e.g. Kune, Twister, or GlobalLeaks) where such entity does not exists. In the case of decentralized architectures, in fact, the removal of central hubs controlling the infrastructure of communication eliminates the possibility for authorities to rely on private actors (online operators or ISPs) to monitor and police online communication. Without a central authority, the only way to assess whether or not a particular piece of content is infringing copyright law is to rely on the community of users - whose social norms are however often incompatible with the provisions of the copyright regime.

2.2. Hate speech and unlawful content

Governments are not only concerned with the repression of piracy, but also, more generally, with the preservation of public order and morality. Hence, from a legal perspective, freedom of expression is generally subject to a series of limitations regarding hate speech, slander, obscenity, incitement to violence, and so forth.
It is common practice that, whenever undesirable content is published on an online platform, both the platform operator (if any) and community members might intervene, requesting that such content be removed.
Yet, even when the community is willing to cooperate, the dissemination of illicit content, such as child pornography or hate speech, remains a critical issue in platforms (such as Kune, Twister, GlobalLeaks) designed to be anonymous and/or non-censorable by anyone, including the original platform provider.

2.3. Cybercrime

Online CBPP platforms based on distributed P2P networks might raise a number of issues with regard to law enforcement and cyber-criminality. Indeed, decentralized technologies present a series of advantage in terms of privacy, autonomy and freedom of expression, which might however turn out to be problematic when used by a particular group of ill-intentioned individuals.
Many people look at decentralized architectures as an opportunity to counteract the regime of surveillance and control that is emerging on the Internet, bypass network restrictions and arbitrary censorship, or expose governmental wrongdoings. Indeed, decentralized P2P networks can support the deployment of socially-valuable applications - such as Tor to protect user's privacy or anonymity, Twister to promote freedom of expression, GlobalLeaks to support and protect whistleblowers in the context of authoritarian regimes, or Kune for grassroots community management and federated online communications.

Yet, any technology that is sufficiently secure to protect an activist is also sufficiently secure to protect a terrorist. In view of the difficulty for governmental authorities to control or monitor user's activities and online communications in online decentralized architectures, people fear that decentralized networks (and particularly those that rely on specific technologies of anonymization such as Tor) might also be used by malicious users in order to reveal secret and confidential information, escape from governmental control, or even to engage into criminal activities, such as copyright infringement, pedophilia, or hate speech, but also money laundering, identity theft, fraud, hacking, etc.
This could constitute an obstacle to the effective deployment and use of these platforms, or might even encourage the enactment of new laws, aimed at furthering the power of the States and discouraging the use of decentralized technologies on the grounds that they are likely to disrupt public order and morality, or that they could potentially jeopardize national security. For example, in the U.S. the war against terrorism legitimized the adoption of texts such as the USA Patriot Act or Foreign Intelligence Surveillance Act Amendment (FISAA), relying on the notions of "cyber-crime", "terrorism", "pedophilia", to justify the enactment of draconian regulations, which might impinge upon the fundamental rights of internet user.

3. Regulation of decentralized architectures

3.1 Law regulating code

Instead of relying on traditional regulatory mechanisms based on legislative tools and ex-post mechanisms of enforcement, States that find themselves unable to extend their territorial sovereignty into cyberspace started experimenting with specific technologies to ensure compliance with the law by means of ex-ante transjurisdictional regulatory mechanisms.
Different policy choices will lead to different regulatory policies that mandate certain technical features and forbid others. During the cryptowar, for instance, the U.S. government tried to forbid the implementation of specific encryption mechanisms or anonymisation techniques, while forcing manufacturers or application developers to incorporate compulsory identification mechanisms, or backdoors into their products. Some countries even went one step further, by entirely precluding the deployment and use of a particular technology (e.g. Bitcoin forbidden in Russia and Thaïland), or by deploying themselves the technology necessary to achieve their goals (see e.g. China, Syria or North Korea where all Internet communications are not only monitored by the State, but are also filtered or censored by national blacklists or firewalls). Conversely, policies aimed at protecting online civil liberties might require applications or device manufacturers to embed specific technologies (e.g. end-to-end encryption) and privacy-by-design principles directly into their products, so as to protect citizens from pervasive monitoring.

3.2. Law regulating peers

In addition to regulating the underlying code or technology of online platforms, legislators can also try to regulate the behavior of individual actors -- be they either the platform owners or users themselves.

The former solution is easier to achieve in the context of CBPP platforms based on a centralized governance model (e.g. Wikipedia). It is generally implemented through intermediary liability limitations regime requiring that the platform owner keep track of all users' activities (through logging and data retention obligations) and/or act expeditiously to bring any infringement to an end (through notice and takedown procedures).
Yet, many CBPP platforms - including as Kune, Twister and Globaleaks - do not rely on a centralized architecture, but rather on a decentralized infrastructure governed by a distributed network of peers, which does not easily lend itself to regulation. In this context, law enforcement is more difficult to achieve, because it requires assigning or sharing responsibilities between all peers connected to the platform. Such a model has been implemented in France, for instance, with the introduction of a three-strike legislation for copyright infringement aimed at discouraging unauthorized file-sharing, along with a specific administrative sanction for characterized negligence in securing one's Internet connection (see, the French HADOPI2 law of 2009). Yet, given that these rules apply at the level of the individual, they are generally difficult to apply without relying on alternative mechanisms of enforcement (by technical or social means).

3.3. Policing the network

As a complement to the former two types of regulation, the legislator might implement a series of preventive measures to discourage deviant behavior online. This can be achieved, for instance, through a regime of mass surveillance (general monitoring) to identify potential threats or divergences from the rules defined by law. Police is already using monitoring software to track social networks according to the use of some specific keywords. The regulator might also establish a specific task-force in charge of infiltrating a number of online platforms or communities (moles, or collaborators), or for policing the Internet against malicious individuals or suspicious behaviors (cyber-patrolling).
Besides, when traditional mechanisms of ex-ante or ex-post law enforcement fail, it is possible to rely on community self-governance to establish a system of guidelines (soft power) or social norms (peers regulating peers) to effectively enforce community rules, or even the rules of law. For instance, while guidelines might be used to promote lawful behavior (e.g. a best practice notice against copyright infringement or other illicit behaviors), social norms often play an extremely important role in regulating community activities (e.g. counter-speech used to limit the effects of hate speech).
Finally, when it comes to policing the network, specific policies might also be implemented to encourage civil society to refer any suspicious activity of to the police through the procedure of "delation". This is done by requesting community members to voluntarily monitor the network, or even develop a system of punishment and incentives to either reward or punish other community members, according to their current or past behaviors.

Conclusion

Analysing the implications of the infrastructural design and technical features of online peer-production platforms has shown that decentralized architectures are likely to be more compliant and more respectful of users fundamental rights, such as the right to privacy and freedom of

expression. With regard to the former, given that most of the data is stored locally on the users' devices, it is harder for any third party to spy on their users and to surveil online communications. The task can be made even harder by means of specific technologies such as encryption, anonymisation, or other privacy-enhancing-technologies (PETs). With regard to freedom of expression, similarly, given that there are no centralized information gatekeepers that can preclude access or censor specific types of information, users can express themselves more freely. This is even more true in the context of platforms that allows for anonymity, where users can express themselves without fear of being subsequently retaliated upon.

Yet, while the design of the technology has a significant impact on what can or cannot be done on the platform (code is law), it is difficult to control the manner in which the technology will be subsequently used by users. Hence, to the extent that they promote anonymity and freedom of expression, decentralized peer-production platform can potentially be used by malicious users in order to engage into criminal activities. Decentralized architectures (such as mesh networks, Bitcoin or Ethereum) eliminate the centralized gatekeeper, making it increasingly difficult to monitor online communications or track individual infringers. Thus, these platforms become very attractive to people engaging in illegal activities, such as copyright infringement, hate speech, pedophilia, etc. A perfect example is the rapid growth of online piracy, which has been considerably facilitated with the deployment of many peer-to-peer file-sharing platforms. But certain Internet users also rely on these platforms in order engage in much more critical and/or criminal activities, such as it has been recently illustrated with the case of Silkroad, where users were relying on the anonymity provided by Tor and Bitcoin in order to reduce the likelihood of being identified and incriminated for selling drugs or weapons.