



HAL
open science

Automation

Marc-Antoine Pencolé

► **To cite this version:**

Marc-Antoine Pencolé. Automation. Susanna Lindberg; Hanna-Riika Roine. The Ethos of Digital Environments: Technology, Literary Theory and Philosophy, Routledge, pp.151-160, 2021, 9780367643270. 10.4324/9781003123996 . hal-03670665

HAL Id: hal-03670665

<https://hal.parisnanterre.fr/hal-03670665v1>

Submitted on 18 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Automation: Between factuality and normativity

Pencolé Marc-Antoine

Abstract:

Moral or ethical decisions, involving values and norms of a given community, necessarily have to be made, as it seems, by subjects, characterized in particular by their intentionality and autonomy. Thus, delegating them to an automated apparatus would consist either in eliminating their moral dimension, or at least limiting it to the morality of its original designer.

Such classical conception hinges on strictly non-dialectical notions of factuality and normativity. Our claim is that the delegation of a moral or ethical decision to an automaton – complex and powerful enough to be assimilated to a « decision-making » machine – may indeed be considered a dispossession yet also, under certain circumstances, be the very condition of its effectivity. Indeed, as exposed through the hegelian tradition and particularly thoroughly by Habermas, certain normative demands could require, to be fulfilled, the collective power that is being cast into an institutional or a technical system – an algorithm for instance. To be normatively satisfactory, concrete interactions imply a motivational, cognitive and organisational load that the subjects involved have to bear: without being mediated by factuality, such as that of informational machines, these normative demands would remain ineffective.

We will use the case of Wikipedia, and of diverse and ambivalent peer-to-peer files exchange communities to back our claim and help shift the debate about the morality of « decision-making » machines toward the question of the different forms of delegation to automata: what are the intrinsic or contextual elements that make such delegation a successful effectuation of collective norms or a sheer dispossession of our autonomy?

Keywords:

Autonomy, morality, effectivity, Hegel, Habermas, Peer-to-peer, Wikipedia

To cite this article :

Pencolé Marc-Antoine, « Automation: Between Factuality and Normativity », in *The Ethos of Digital Environments: Technology, Literary Theory and Philosophy*, Susanna Lindberg et Hanna-Riika Roine (ed.), New York, Routledge, 2021, pp. 151-160.

Automation: Between Factuality and Normativity¹

[151]The use of automated machines started being worrying when they became complex and powerful enough to impose their own norms to the human activity. The inflexibility of a factory system certainly strips its workers of most of the control they could have on their labour, its pace, its form and its meaning and destination. In the last decades, the problem arose under quite a new guise with the development of digital automata and their increasing capacity to calculate, evaluate, organize and anticipate human activity. Should we see it as an inevitable deepening of the dispossession of our prerogatives as subjects in favour of an extending network of cold and impersonal algorithmic administrators?

Our collective existence is being regulated through different ways. First – and it was indeed for long the main topic of classical political philosophy –, laws, decrees and rules formally organize social life and shape congruent subjectivities. Then, another powerful regulator consists in the normativity immanent to any social activity: these bundles of common norms and expectations that determine what is a successful and respectful interaction – and it explains, for instance, why people spontaneously queue up in a certain order in front of the bakery without any law to prescribe it. Finally, technology may also act as a regulator, materializing certain norms within fixed processes, architectures and artifacts. Technological normativity hasn't always been a problem – beyond the workplace at least – : concern arose when the development of computer science extended the scope and depth of technological regulation over domains where rules were traditionally set by law.

The power of artifacts rests on their very factuality: the factual properties of cement and stone determine the capacity of a certain architecture to organize visibility and circulations; the diversity and complexity of softwares, functions, libraries and databases on one side, and the miniaturization of chips and processors on the other, is what grants digital devices the power to take very adequate (« intelligent ») decisions. Although, as soon as it comes to effects that can't be reduced to physical phenomena but tends to be assimilated to pseudo-decisions, similar to the kind of behaviours manifested by intelligent subjects, then the possibility arises that actual subjects may hand the exercise of their defining faculties over to factual objects, and thus end [152]up being dispossessed of some aspects of their subjectivity. Indeed, a subject is traditionally seen as a being that is – at least partially – able to reflect itself, and act freely, thus whose most proper domain is that of moral and political deliberations and decisions: delegating decisions about the good life and justice to impersonal machines would mean to alienate oneself from one's very autonomy and responsibility as a subject. Now, we would like here to question the inevitability of such a trade-off between the power offered by digital automata and the preservation of human autonomy and responsibility.

The normativity of technical artifacts.

First, we must examine different conceptions of the normativity of technical artifacts to specify the precise conditions of the trade-off. The most common one may be the *use theory of norms*, according to which the norm is extrinsic to the artifact and stands entirely in the diverse ways subjects choose to use them. Artifacts would then be no more than neutral instruments. This applies quite well to abstract tools: for instance, a screwdriver may be used as a weapon as well as a repair tool or any other purpose we can think of that requires a short and rigid metal rod. It covers also the normativity of indeterminate technological fields, understood as a sum of knowledge beyond any determined application, like informatics in general, which can lead to an indefinite variety of actual devices. Yet, it doesn't account for the specific normativity of concrete « technical individuals » upon their environment².

1 This presentation reframes and extends the reflections elaborated in a previous work : Pencilé, M.-A. (2017), « Nos algorithmes peuvent-ils être plus justes que nous ? », *Revue Française d'Ethique Appliquée*, n°5.

2 Simondon, G. (2017). *On the Mode of Existence of Technical Objects*, Minneapolis, Univocal Publishing, p. 63.

An opposite theory that we could call the *substantial theory of norms* pretends to account for concrete technical individuals by stating that the norm is intrinsic to the artifact. Indeed, certain technical beings tend to channel, facilitate or forbid certain behaviours, hence imposing their own norms to the actors. For instance, the now famous Long Island bridges were designed at the beginning of the XXth century in such a guise that only low vehicles could drive under their archs; since cars were still rare and merely owned by wealthy residents, the Black and Puerto Rican poor, mainly travelling in tall public buses were practically forbidden the most convenient access to the island³. Although, such normative imposition appeared later to be highly context dependant, revealing the extrinsic conditions of « intrinsic » technical norms: indeed, driving habits evolved along the century and today the rich drivers of bulky campers happen to be the ones impeded from enjoying the sea shore⁴.

These two theoretical frames are either too partial or partially contradictory, and any combination of both will also partially carry their flaws. Their shared presupposition, from which stem these limitations, consists in the idea that the opposition between the subject and the object is a very static one. The modern concept of subject posited it as absolutely autonomous and transparent to itself, casting the object as its radical other, some pure exteriority [153] or unreflected stillness, an extended and unthinking thing. Philosophy of technology recently remembered that old critic of a too naive subject-object divide and proposed a *mediation theory* of technical factuality and norms⁵. In order to overcome binary reasonings about technology, the mediation theory suggests a dynamic frame, where a technical process or artifact doesn't just stands in front of the subject and her norms, but participates in their constitution; it doesn't simply connect the poles that are intermediated, but transforms them. To give a trivial exemple, phones, video-conference devices, etc. aren't mere communication instruments, neutral and passive, since they transform the message itself, and ultimately what it means to communicate. Certain emotions or nuances can't be carried through a flattened voice message or a visual frontal display with no eye contact. More interesting : Verbeek showed that the very existence of ultrasound-scanners alters what is deemed as reasonable, cautious, or irresponsible during pregnancy, because having missed some serious pathology affecting the child will now be considered the result of a conscious decision, where before it wasn't but blind fate anyway. The parenthood norms mediated by the factual power of such a modern detection tool undergo a real transformation: scanner aren't neutral instruments, obviously, but they aren't imposing some fixed hardcoded norm onto parents either.

Surprisingly, this ambitious move toward a more concrete and dynamic theoretical ground in philosophy of technology boils down to the refusal of any critical discussion of technical phenomena. Indeed, according to Verbeek, the dependancy of the content of the evaluation standards on the mediating activity of the object that is being evaluated makes it absurd to pretend to assess a technology that hasn't been widely adopted yet: it would be tantamount to evaluate a phenomenon from a normative standpoint we already know we will probably disapprove afterwards. Let's then try to substantiate our concept of mediation by digging into the hegelian tradition.

Hegel used the concept of mediation to account for, and extend, the idealist intuition according to which the subject participates in the constitution and the position of her objects. It became in his philosophy the general form of reflection, the negative moment of thought, where the immediate distinctions set by the understanding – between the thing and its others, what the thing is not – are blurred and overcome to eventually lead to a more comprehensive apprehension of the thing. Let us narrow down the perspective from logical generalities to the question of technology: here, the mediation of human activity by technology covers all the manners of effectuation, actualization, or degradation and loss of the self and his or her activity through the objectivity of tools, know-hows and procedures. Human beings exist and reflect themselves through the products of their activity: their subjectivity is given and confirmed to them through the objective matter-of-

3 Winner, L. (1980). « Do Artifacts Have Politics? », in *Daedalus*, Vol. 109, n°1, pp. 121-136.

4 Verbeek, P. P. (2005). *What things do. Philosophical Reflections on Technology, Agency and Design*, Pennsylvania State University Press, p. 117.

5 It has mainly been promoted by Don Ihde, in *Technology and the lifeworld: from garden to earth*. Bloomington: Indiana University Press, 1990 ; and Peter-Paul Verbeek, *op. cit.*

factness of their products. It doesn't simply mean that something, once [154]mediated, isn't the same as before – this seems to be the logical observation underlying Verbeek's argument and it could be deemed rather obvious and slightly vague –; the becoming of the mediated subject may in fact be one of objective realization of inner pretensions, actualization of virtualities, or on the contrary negation of oneself, or inability to confirm one's value in objectivity. The mediated being isn't a whole new mysterious thing, but remains the same in its difference with itself, as for example an impulse of kindness or envy remains itself once it has been manifested outside despite the new mode of existence it has acquired. Hence, Hegel offers us a more complex frame to describe normative phenomena in the technological realm as well as to criticize irrationalities or emancipatory perspectives.

Mediation and norms effectivity.

The most encompassing exploration of the relations between factuality and normativity isn't proposed by Hegel in his few lines on technology but in his study of the mediation of social relations by law, in which we might nevertheless find precious insights for our topic. According to him, our shared norms of morality and justice require the mediation of the legal system, its codes and institutions, to be effectuated and actualized, to pass into the stuff and structure of society. Indeed, norms would remain mere wishes, inner standards with no objective reality, if they weren't to shape somehow social relations and weigh on them. So norms get more actual, morality and justice become more adequate to what they are supposed to be, when passing into the inert factuality of texts, institutions and established procedures.

On another side – and here is Hegel's most decisive argument –, these norms are in their turn affected by their passing into factuality, and not merely in the sense they would become more actual. Indeed, as facts, laws necessarily bring their share of contingency and rigidity, what may appear as irrational and contrary to the ideal of justice they were supposed to embody. Although, in court, a decision has to be taken, even if partly contingent⁶. The worst that could happen is that no judgement is made and that justice is denied, be it because of the inability to qualify a specific act, or to practically understand the intricacies of the texts and procedures. So the law has to be simple enough to handle well yet also be encompassing enough to cover every possible case: it is then impossible to legally code the perfectly moral and just answer to every transgression since nobody would be able to actually learn and apply it to concrete cases – and precisely, actualization of the ideal is the essence of the law. We must then acknowledge that a factual legal system, partially inspired by the norms of justice but also partially contingent and arbitrary, is *normatively* superior to the pure ideals themselves, *because of its very factuality*. The factuality of actual law is normative. So the norms need to become facts to be proper norms as well as the facts tend to become themselves norms in the process.

[155]Habermas later developed these insights⁷. The reason that norms as mere subjective demands are weak and that factual systems may bear their own kind of normativity is that, for the self, to stand alone at the origin of ethical decisions about justice or the good life, and to align her activity on these norms, and to recognize herself in them, represents an immense demand. As an actual and empirical thing, every subject is finite. Thus, she faces subjective factual limitations in her will to accomplish her ideals, so that the individual subject, the true bearer of these norms, responsible before them, also appears dramatically separated from them because of three determinations of her finitude.

First, realizing the norm, tearing it away from its abstract ideality, means applying its general prescriptions to concrete cases. Now, the universality of the norm opens a rift, between itself and the always too particular cases, that demands some intellectual effort to be crossed: realizing the norm in complex situations demands to face a huge *cognitive indeterminacy*. Everybody experienced the toil of exploring, analysing and evaluating the moral issues of everyday

6 Hegel, G. W. F., *Outlines of the Philosophy of Right*, §213 add.

7 Habermas, J. (1996). *Between Facts and Norms*, Cambridge, MIT Press, pp. 114-118.

life: to be sure that what we do is moral, outside trivial situations, we need to collect informations and put a lot a thought into it, weighing the right and wrong, sometimes even deliberating with others, before being convinced that a certain behaviour really is the right course of action. The more persons are involved, and the more extended the situation is, the more difficult it will be for the arbiter to cognitively relate the many singular and interconnected facts to the unity of a simple general ideal. This is why a proper law posits more or less clear borders between, for instance, what qualifies as a harmful negligence and what is but an unexpectable accident. All the work of customs and legislators sedimented in tradition and positive law relieves the present arbiters from the immense burden of crossing the cognitive gap between the particular and the universal. The counterpart of enabling the judge not to be struck by doubt is of course a certain part of contingency in the determination of the application rules.

A second burden consists in the *motivational uncertainty* moral actors may experience. Knowing what is just is one thing, but willing it is another. Situation will often rise where the right course of action happens to be detrimental to the particular interest of the actor. There, we can expect she sometimes lacks the motivation demanded to eventually choose the general interest over her own. Since the legal system is factual, the subject in position to judge as well as the individual subjected to the law are strongly encouraged to act according to the norm – because they can expect a reward or at least avoid some harsh punishment. Factuality of an effectuated norm may be seen as a promise once made to oneself to later comply to principles whose general validity has been recognized.

Finally, moral action faces the question of the *organizational (or accountability) indeterminacy*. The subject is traditionally defined by her freedom thus her responsibility, and yet, beside trivial situations, nothing is more indeterminate than the extension of the domain of what can be reasonably [156]imputed to her. Assisting the person in immediate danger in front of her is obviously her responsibility, but is it still her duty to go and help the same way when the person lives in a completely different region of the world, though she knows she could have made a difference, would she had committed herself to it? Beside a procedure of application and a series of incentives, a legal system also consists in a complex system of agreements about where everybody's responsibility ends, that makes possible to practically live under the rule of just norms.

Such unburdening of the demands of moral action onto the factual power of a legal system can be deemed virtuous as long as the community ruled by its law recognizes itself into it. As soon as it gets deeply heteronomous, we may speak of a dispossession of autonomy, which ends up in the position of an inert opposit, an oppressive system of unilaterally imposed norms. Couldn't then a powerful technical system endorse through its factuality the same role as positive law in effectuating norms of morality and justice?

Digital automata and the effectuation norms.

Complex technical systems, and among them, algorithmic regulations, usually appear as impersonal mechanisms tearing off actors' autonomy and forcing extraneous norms upon them, but this isn't an essential property of technological mediation as such – in hegelian terms we could brand them as insufficiently rational, as an oppressive state would be, if ruled by a particular class in its own interest. Let's try to identify two exemples of relatively valid digital mediations : the BitTorrent protocol and sharing communities on one side, and the Wikipedia encyclopedia on the other.

The BitTorrent system is a technical assemblage allowing the decentralized sharing of files on the Internet between peers. A key component is the BitTorrent protocol – which has been instanciated in many different softwares –, a simple set of technical rules designed by Bram Cohen to ensure the effectuation of an certain ideal of sharing: altruistic behaviour (hosting files for others to download) had to be encouraged, while free-riding (downloading while never uploading back) was to be avoided as much as possible, but narrow bandwidth users also had to be able to access a significant share of the downloads even though they couldn't contribute as much as broadband

peers⁸. Another crucial part is that of the platforms, composed of a web search engine – similar to some sort of directory or catalog – and of a tracker – the server that informs a given user’s software about the other uploaders who may send her the requested file. Admittedly, many of these platforms aren’t much more than profit organization pushing in direction of the basest consumerism, but some of the most prominent institutions of torrent-sharing are built as well ordered communities: chats and forums are dedicated to mutual help, the imperative to keep hosting the file for others once downloaded is repeated [157] *ad nauseam* by the users themselves, the users have to register and the tracker keeps count of the total amount of data they transfer, blocking downloaders who don’t take their share in uploading, and finally every users participates in the evaluation of the safety and quality of each new file they download⁹.

Cognitive and motivational loads weighing on the subjects and relieved by the protocol and the platforms are easily recognizable. Knowing how fast to download, which users to upload to, and which files to avoid sharing – because they’re corrupted or poisoned by malwares – to respect the general principles of the community, all this would obviously require hundreds of complex calculations every couple minutes. What the human mind can’t handle, a network of computers can. Besides, there is a strong incentive to share and fill the needs of others peers with the automatic effect of the protocol, the download over upload ratio, and the whole esteem economy enforced throughout the system, so that everyone has an interest in being a truly altruistic contributor.

The third burden needs to be slightly redefined to fit the specificity of technological mediation since the scope of the regulation doesn’t reach far enough for imputability problems to arise. Habermas related the problem of accountability to the material cost of being moral: if accountability stretches too far, almost nobody will own the means necessary to accomplish any significant moral action. Here, we can see that normative advantage of a factual algorithmic system is even more obvious than in the legal domain: it would be materially impossible to share so many cultural works by burning disks and dispatching them through the postal network or any other means. So building huge digital libraries, curating collections and gathering millions of daily sharing peers around them wouldn’t even be conceivable without digital technologies, while here it doesn’t require more than a few servers and a handful of animators and technicians.

The case of peer-to-peer file sharing isn’t pure, not every platform will qualify as a virtuous mediation, but it illustrates the fact that such communitarian values of sharing would remain but wishful thinking, weren’t they taken in charge by the factual power of a certain device. Even though far from perfect, they’re effective at least to a certain extent. Yet, every assemblage isn’t that successful, and the conditions of such a rational mediation musn’t be neglected. There are two risks of dispossession of users’ autonomy: the verticality of the the organization, whose administrators are formally accountable to nobody – even though one might observe diverse forms of protest emerging from the community –; and the pressure of culture industry’s capital, that has been trying to shut down the platforms through legal means. Here again, the ambivalent factuality of the system helps understand why it could have hold so long against heteronomous constraints: the cheap infrastructure and little labour that needs to be put in to make it operational allowed it to be repaired or built again elsewhere after every serious blow.

[158] Wikipedia is a collaborative encyclopedia launched in 2001, two years before the invention of the BitTorrent protocol, and based on a few principles: neutrality in the presentation of the different points of view, free access to reading as well as participation, respectful collaboration, and finally debatability of every rule. Before these quite general norms, instead of resting solely on spontaneous impulses toward the preservation and improvement of common knowledge amongst its dozens of millions users per day, the project developed a complex of technical mediations, ensuring everyone isn’t required to be selfless and absolutely dedicated to keep it working. We could show it through the way the platform architecture elicits the distribution of tasks and

8 Cohen, B. (2003), « Incentives build robustness in BitTorrent », *Workshop on economics of Peer-to-Peer Systems*, Berkeley, California

9 Dagiral, E. & Dauphin, F. (2005). « P2P: From File Sharing to Meta-information Pooling », *Communications & Stratégies*, vol. 59, n°3.

cooperation¹⁰, but let's focus on the more strictly technical part of its inner regulation – the bots¹¹. Wikipedia is humming with task-specific software activity: thousands operate constantly (the majority in the English speaking sphere), the total production of some amounting to millions of edits. Their tasks, among many others, may be to automatically detect vandals and repair their deeds, or to quickly identify copyright issues, they may participate in data structuration to standardize certain elements in articles or assist in the labelling and distribution of the remaining work to be done by humans, they also often clean pages from broken code and dead links.

Without the help of these tiny automata, substantially contributing to Wikipedia could end up being wearing given the numerous and complex conformity rules between which to arbitrate. Bots offer some assistance by relieving the users from having in mind all the exact details of every procedure, for example, to alter a strongly debated article or to call for a second opinion about a contested deletion. The motivational assistance consists here in punishing deliberate rules transgression, like patterns of obvious vandalism which are automatically detected and end up in banishment and the lockout of the affected articles, but again, it also implies a strong esteem economy, where generous contributors are respected and praised. Finally, the organizational burden lies on one side in the tedious effort and numerous hours spent in standardization of the text and code of each article, and on the other in the tremendous material cost such a worldwide cooperation would imply if conducted through analogical ways; both of which are spared by the efficiency of automated informational procedures.

Beside a certain social homogeneity of most active contributors, Wikipedia's socio-technical system achieved a remarkably horizontal, open, and yet robust common institution of knowledge. Admittedly, its global infrastructure is terribly expensive but the software and all the produced content are freely reproducible, and the project managed to ensure its independancy by regularly calling for donations and thus creating a non-commodified bubble to isolate it from the influence of private property and the market. Yet another significant condition of its success is certainly that the [159]principle of debatability of every rule has been instanciated in the algorithmic automata themselves: of course, every bot may be subjected to a public discussion and redefined by the community, but even individuals may temporarily yet very quickly contest and interrupt the course of a certain bots by inserting specific labels in the article before it is automatically scanned. Users' activity may be assisted, controlled, complemented by automata, but one merely needs to type a few letters to get back her hold onto the processes previously handled by the machine.

Conclusion.

It appears that once equipped with a mediation theory of technical norms, the normative value of factuality in technical systems may be properly assessed: the effectivity of our norms about the good life or justice depends on them passing into the factuality of a technical mediation, given the colossal weight of the cognitive, motivational and organizational demands adressed to the subjects of any community that ambitions to live along their shared norms of justice.

Yet, the study of actual communities mediated by digital automata – the BitTorrent networks and Wikipedia – reveals that their factual organizational advantage isn't just a normative requirement to the effectuation of the ideals of sharing and open cooperation, but also contributes on another level to their consolidation against the risks of dispossession. Such systems are technically reproducible with no insurmountable obstacles and thus particularly difficult to shut down. But above all, their design tends to encourage or at least allow collective appropriation. Contrary to cumbersome machines and bureaucratic administrations, algorithmic mediations are characterized by their *openness to communitarian appropriation* and the difficulty to establish a

10 Nicolas A. (2009). « De Linux à Wikipedia : Régulation des collectifs de travail massivement distribués », in *L'évolution des usages et des pratiques numériques*, FYP Editions.

11 Halfaker, A. & Riedl, J. (2012). « Bots and Cyborgs: Wikipedia's Immune System », *Computer*, vol. 45, n°3; Geiger, R. S. (2017). « Beyond opening up the black box: Investigating the role of algorithmic systems in Wikipedian organizational culture », *Big Data & Society*, vol. 4, n°2.

private monopoly over them. So even the delimitation between what is positively and negatively normative in technological mediations may also be established by their factual determinations.

Once established these logical groundings, claiming that impersonal automata in general are bound to dissociate the subjects and their essential prerogatives and duties is as abstract as pretending that since norms and technology are linked, no consistent critic can be made about the latter. A critical assessment of technological delegation is actually possible, but only on the level of concrete and situated socio-technical assemblages. Now, collective autonomy seems to be the general condition to determine whether a given mediation grants effectivity to the norms; certain factors are of course extraneous to the mediation itself (as is capital or market's pressure in the aforementioned cases), yet it remains possible to identify in the very factuality of some systems a form of flexibility and openness to communitarian appropriation that seems to limit phenomena of dispossession: this is finally the only characteristic of machines, considered in abstraction from any concrete social embedding, that would make sense of the idea of machine's morality.

Pencolé Marc-Antoine