

## BIROn - Birkbeck Institutional Research Online

Archibugi, Daniele and Tosoni, N. (2023) Is users rating becoming overpowering? The risks of inappropriate use of digital feedback. Working Paper. CIMR, Birkbeck, University of London, London, UK.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/52614/>

*Usage Guidelines:*

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html> or alternatively contact [lib-eprints@bbk.ac.uk](mailto:lib-eprints@bbk.ac.uk).



## **CIMR Research Working Paper Series**

*Working Paper No. 67*

### **IS USERS RATING BECOMING OVERPOWERING? THE RISKS OF INAPPROPRIATE USE OF DIGITAL FEEDBACK**

by

Daniele Archibugi, Natalia Tosoni

12 October 2023

ISSN 2052-062X

**IS USERS RATING BECOMING OVERPOWERING?  
THE RISKS OF INAPPROPRIATE USE OF DIGITAL FEEDBACK<sup>1</sup>**

**Daniele Archibugi\* and Natalia Tosoni\*\***

\* National Research Council, Irpps, and University of London, Birkbeck College, School of Business, Economics and Informatics, E-mail: [daniele.archibugi@cnr.it](mailto:daniele.archibugi@cnr.it)

\*\* University of Bologna Alma Mater, Department of Sociology and Business Law, E-mail: [natalia.tosoni2@unibo.it](mailto:natalia.tosoni2@unibo.it)

**ABSTRACT**

Digital technologies have made it easy to generate and disseminate feedback on the performance of products and services. This technological advance has contributed significantly to reducing the information asymmetries between producers and consumers, between suppliers of public goods and users, and between politicians and citizens. It would seem that Web 2.0 has enhanced the effectiveness of the "voice" evoked by Albert Hirschman in the 1970s. However, is there a risk that the feedback provided and the connected numerical evaluation, the so-called rating, may become too invasive, so much as to constitute a concrete threat to the confidentiality of individual data? We propose to distinguish between three types of feedback: 1) bottom-up feedback, which occurs when many individuals evaluate and comment on the performance of economic and political organizations; 2) the transversal one, which happens when a series of individuals, in the same hierarchical position, exchange comments, appreciations and reciprocal evaluations; 3) the top-down one, which arises when organizations assign a rating to individuals. Based on our analysis, we also offer some suggestions to moderate the already existing risks.

---

<sup>1</sup> We wish to thank Patrizia Grifoni, Tiziana Guzzo, Carlo Inverardi-Ferri and Mathias Koenig-Archibugi for comments on an earlier version.

## How close is *Nosedive*?

In the episode *Nosedive* of the well-known British TV series *Black Mirror*, the protagonist, Lacie Pound, wishes to improve her social position. To do so, she doesn't have to overcome obstacles of coming from a modest dynasty or a humble social class, or fight for admission to exclusive colleges and universities. Her social ascent depends on a score, or a rate, assigned to her by a ubiquitous social network. Indeed, Lacie, to get access to an apartment in a more elegant condo or to a better paid job, must improve her rating and building her own virtual image based on an aura of success, beauty, and wealth. In this dystopia, the unique and powerful rating is raised or lowered according to the outcome of any social interaction, even the most daily and ordinary<sup>2</sup>.

A well-rated individual finds himself not only with a more prestigious circle of friends but can get also cheaper mortgages, better automobile insurance, hotel rooms with views, and better jobs. If, on the other hand, an individual slips downwards, a chain effect is easily unleashed, making her life increasingly complicated until it pushes him to the margins of society.

Dystopias are successful when they represent a caricature of existing reality and highlight its perverse and latent sides, not yet clearly perceived. That there are an infinite number of assessments to which we are exposed to is something we have learned to coexist with for some time. For example, we are used to our car insurance policy having a class associated to the reimbursement claims made in the past. We are less sure of what happens if we ask for a loan from a bank, but we are not so surprised if the employee in front of us consults an archive that credit institutions share with each other to verify our reliability (which they define "credit rating"). Similarly, many employers carry out an annual assessment of their employee performance, as happens in Amazon warehouses, where the quality of the work performed is measured through quantitative indicators. Those having a commercial activity are familiar with the fact that some portals, such as Tripadvisor, Google, Airbnb and Amazon, allow customers to provide a numerical evaluation, often accompanied by a comment, of the product or service they offer. Not even university professors are exempt from evaluating students, so much so that they are classified on the basis of the numerical score achieved. For what concerns research performance, which academic hasn't consulted the number of citations received by their peers on Google Scholar?

The dystopia of *Nosedive* is insightful because it adds one more piece: in this society, evaluations are unified. Rather than a variety of different ratings achieved by functional areas (automobile insurance, occupational performance, commercial reliability), there is a single rating that is attributed directly to individuals. What would the society we live in be like if it were possible for some reasons to aggregate existing evaluations to the point of providing a kind of behavioural and virtual identity card of the individual? Information technologies run much faster than the political, legal, and social capacity to regulate them, and it would be all too easy for some nerds to find an algorithm to combine the different evaluations to give a score to the individual as such. In *Nosedive*, individuals score each other when they meet in the park, chat at a reception, or even share an elevator ride. Such a social network, still non-existent, perhaps imminent, makes the old saying *vox populi*

---

<sup>2</sup> With the word "rating" we mean a particular type of numerical evaluation through which it is possible to express a synthetic judgment which is, generally and apparently, immediately understandable thanks to its simplicity. Rating, in this elaboration, is configured as a numerical form of the feedback. The latter must be understood as any form of feedback that an individual or organization, both political and economic, can receive when it shows its interest in the matter and activates methods to receive it (as in the case in which a commercial activity registers on TripAdvisor allowing its customers to provide feedback).

*vox dei* (Voice of the people, voice of God) quantifiable and publicly accessible. Therefore, it is not foolish to think that what it describes is much closer to reality than imagined.

Nowadays, all of this is possible because we can easily express our critical opinion on the economic, social, political, and cultural aspects we live. Whether we read a book, watch a movie, rent a car, eat in a restaurant, or stay in a hotel, we can speak our mind and make it public. Before the advent of the Internet, only a few citizens had the ability to do so. It was the professional critics with access to newspapers, magazines and television who led the dance, while the people could only resort to word of mouth, or decide, after a bad experience, to change suppliers, where possible. What are the consequences of this extensive use of rating and feedback? Do we risk moving towards a society where they represent a tangible form of social control similar to a dictatorship?

In this paper, we try to point out the advantages and disadvantages that the new conditions – economic, political, social even before the digital opportunities – have entailed regarding the renewed possibility of people to provide their own opinion, underlying how this great potential, capable of reducing the information asymmetries existing between large organizations<sup>3</sup> and the public, also fearing considerable risks. To this end, although it is not possible to consider all existing forms of feedback, we suggest distinguishing at least three of them:

- 1) feedback that acts with a bottom-up direction, which includes all those situations in which a myriad of individuals, who interface with the same organization, manage to share opinions and information about it;
- 2) transversal feedback, attributable to the possibility that individuals have of being able to interact and express judgments reciprocally within a basically peer-to-peer relationship;
- 3) feedback with top-down dynamics, which instead refers to the ability of organizations to evaluate and classify individuals.

In the following two sections, we start from the famous scheme by Albert Hirschman, dedicated to the three attitudes that individuals can assume towards organizations (Exit, Voice, and Loyalty), then reflect on how it has changed behaviour. In the next section, we address the issue of Web 2.0 and how this has strengthened the possibility for consumers, citizens, and users in general to make their voices heard, also in view of the risks that this form of expression can take on at the moment in which it ceases to refer exclusively to organization and leads instead to assess and comment on individuals. We then dedicate a section to the case of the Social Credit System tested in China, since it overturns the nature of the "voice", as argued by Hirschman, making organizations, and even the government, capable of assigning rating to individuals and administer corresponding rewards and penalties. We then move on to the definition of the different forms in which feedback can manifest itself and, finally, based on this distinction, we make some suggestions for strategic action.

### **Exit, Voice, and Loyalty: the Hirschman's scheme**

Many channels for expressing opinions of users and consumers were already active already in the 1970s. Indeed, in that period, Albert Hirschman (1970) wrote about possible strategies that users could implement, consciously or unconsciously, in response to the specific modes of action of

---

<sup>3</sup> The term organisation(s) is used here in a broad sense to mean both political organizations (such as, for example, parties or trade unions) and purely economic organisations, such as businesses.

companies, political parties and institutions in general. In this regard, Hirschman identified three different kinds of behaviour: exit, voice, and loyalty.

Loyalty can be described as the willingness of the subject to continue to participate in the organizations' activities because she is overall satisfied with what she gets or because she is unable, or not interested, to consider further alternatives. The individual, choosing to be loyal, continues to be a member of the same political party or trade union organization, to purchase the same brand of biscuits, or to renew his policy year after year with the same insurance company.

The second strategy that the individual can implement is exit, and therefore the possibility of expressing one's dissent by interrupting the relationship of fidelity, or loyalty, towards the usual organization. This choice can be absolute or substitute. It is absolute when the individual definitively abandons the field in which the political and/or economic institution operates. As far as political participation is concerned, this means ceasing to be a member of a party; while in the case of economic participation, the individual can completely leave the reference market. An example of absolute exit from the side of political organizations is when the citizen stops going to the polls. On the economic side, it happens when the consumer stops having an auto insurance policy because she gets rid of her car. The substitute exit occurs when the individual continues to participate in political and economic life relying, where possible, on rival organizations and changes party or insurance company.

The third option is that of voice. Citizens, consumers or users, notes Hirschman, not only have the possibility of bovinely continuing to follow organizations (loyalty), nor that of going elsewhere to try to satisfy their own needs (exit), but they can also undertake actions of protest to request an improvement in the service or product that is offered to them by the organizations. Hirschman observes that voice can manifest itself as a simple and harmless grumbling, but also as a real action of violent revolt. Although the use of voice is often practiced to obtain a change in the strategies of the political and/or economic systems of reference, Hirschman's theory also contemplates the protest action undertaken by that subject who, feeling she belongs to the community represented by the organization, wishes to contribute to its improvement. Indeed, for some leaving an organization could be particularly complicated and represent a cost in economic, functional or ideological terms. In this sense, giving voice to one's opinions, suggesting adjustments or blatantly protesting the quality of products and services, their cost or even just employee rudeness, can be a very effective means of helping organizations prevent and avoid what for them is the worst possibility, and so that their constituents, users or customers, opt for the exit without giving reasons (for a review of the effectiveness of online reviews by customers, see Pan, 2023).

From the point of view of the organization, loyalty and exit are diametrically opposed attitudes, which, however, present a paradoxically similar aspect. If individuals could choose only one of the two behaviours, the organization would not intercept the reasons why their adepts decided to be faithful or not. In other words, if it is assumed that the organization is aiming to preserve and grow its base (whether membership, vote, or market share), loyalty and exit, by themselves, do not provide insight on the possible stability of this quota, nor on its potential increase or decrease, with obvious implications on the definition of an organization's strategy.

There is obviously the implicit assumption that organizations wish to act to maintain or increase their electoral base (for political parties) or their market share (for businesses). However, if these operate in a monopoly regime, they will have no penalty if they move in a completely different direction than the ones desired by the individuals, whether they are citizens (in the political process),

users (of public services) or customers (of products and services). In a one-party political regime, for example, leaders have no incentive to question whether their citizens are satisfied with their choices, so much so that they even could force people to go and vote on ballot papers in which there is the symbol of a single party, as happened in Italy during fascism or in the old Soviet Union.

Similar problems arise when public services are provided under a monopoly regime. Until the 1980s, gas, electricity, telephone, water and urban transport were often supplied by a single manager. The effect was that disgruntled citizens could find alternative means of defection (for example, taking the private car and not the bus to go to work), but this did not lead to competing firms entering the market, significantly reducing incentives to improve the effectiveness and efficiency of the organization.

Voice, therefore, is configured as an option of particular interest for all those organizations that operate in a competitive regime and wish to maintain or increase their political and/or market base. Through this option, in fact, the organization understands from its followers what would be better to change to optimize its efficiency and thus attract a greater number of users before they go elsewhere.

Finally, it should be underlined that, in the Hirschman's scheme, the actors involved have very different dimensions and contractual power: on the one hand, the organization, generally large and aggressive, on the other individual, who, if isolated, does not manage to emerge and assert its interests.

### **The impact of Hirschman's scheme**

The Hirschman's scheme, thanks to its simplicity, has had a visible impact and has changed the strategies of many organizations. We do not intend to say that it was only the Hirschman's book that changed reality, but starting from the 1970s there have been radical transformations that can be found both in the economic and in the political spheres, which coincided precisely with the publication of that essay.

As far as organizations of an economic nature are concerned, a paradigm shift has begun to be recorded, especially in the United States, but also in other market economies, relating to the conception of the so-called complaint offices. The latter, in fact, have no longer been considered as a place to send employees to be punished<sup>4</sup> but they were starting to be understood as fundamental company functions for intercepting the information necessary for the improvement of market strategies.

Marketers within the company had to conduct costly consumer and prospective consumer surveys before introducing new products, processes, and services. So why not pay more attention to the information that customers spontaneously sent to complain? In just a few years, the complaint offices have changed their name, introducing ever more captivating names: "customer relations office", "permanent listening line", "problem resolution", etc... The grumpy customer has become, often unbeknownst to her, a precious source of useful suggestions for understanding the company's obstacles and development potential. Not surprisingly, companies operating in a competitive regime were the first to use the "antennas" of the complaint offices. Indeed, the more the customer can make

---

<sup>4</sup> Such as those who, like Daniel Pennac's Benjamin Malaussène, were the scapegoats to be fed to an audience of angry and insolent customers.



a substitute exit without having to bear too many costs, the more fundamental it becomes for organizations to intercept discontent and prevent them from turning to competitors.

An equally important effect can be found within political organizations. In democratic systems, for example, political parties are institutionally called upon to periodically compete for the vote of their electors and are therefore forced to ask themselves what their aspirations are and how to represent them. Marketing techniques, developed originally by businesses, are known to have been increasingly used by political parties, as taught by political marketing textbooks (Lees-Marshment et al., 2019), so much so that it is now common that programs, slogans, and candidates themselves in the electoral campaigns are selected before the electoral competition on the basis of the level of potential approval by voters. Political parties have thus imitated companies and, rather than passively waiting for actual or potential followers to make their voices heard, they too have tried to intercept the aspirations of voters through surveys.

### **The effects of Web 2.0 and information asymmetries**

For almost all goods and services, not all transactions are repeated and there is often an imbalance between the seller and the buyer, what economists called information asymmetries (Stigler, 1961). For example, if you are looking for a restaurant while traveling or if you want to replace a washing machine, the customer is unlikely to have complete information. Collecting accurate information on the quality of a restaurant or on the efficiency of a washing machine involves a waste of resources, in terms of time and money, which the individual may not be willing to invest. For decades, consumers have tried to reduce this knowledge gap and, at the same time, curb the predatory attitudes of companies, using various tools, including word of mouth, one of the mild remedies against information asymmetry (Duan et al., 2008). Travelers of the past relied almost exclusively on the experiences and advice of those who, within their network of real contacts, had already made a given trip.

Web 2.0<sup>5</sup>, by modifying the way platforms and users interact and re-evaluating the position of the latter, makes it possible to partially overcome the impasse and provide a general picture, albeit superficial, on the quality of a specific product or a specific service (Constantinides and Fountain, 2008). The emergence of different platforms, such as TripAdvisor or Booking.com, not only makes word of mouth easier and more immediate by sharing the related experiences in the online world, but also allows people to move within a decidedly wider and more up-to-date network of virtual contacts.

Web 2.0 therefore acts by reducing the information asymmetries and at the same time grants new and more effective ways of expressing and sharing feedback thanks mainly to two innovative characteristics. The first is the active role of users, who configure themselves as truly collaborators in interacting with platforms and in the creation of their contents (Bleicher, 2006). Think, for example, of Wikipedia, the most famous online encyclopaedia, which allows registered users to edit and create pages relating to a specific topic; or to TripAdvisor, a platform based precisely on the sharing experiences, especially of travel, of individual users<sup>6</sup>. The second feature of Web 2.0 to consider in

---

<sup>5</sup> Web 2.0 differs mainly from the original World Wide Web in the user's ability to constitute himself as creator of content and information (for example, Wikipedia) (Bleicher, 2006).

<sup>6</sup> In this regard, it is important to point out that the active role of users in interacting with platforms involves a co-construction of data which does not always require users' full awareness. Sometimes the platform collects information



relation to our study concerns the centrality of the so-called *virtual communities*<sup>7</sup>. In fact, the latter, by having an impact on the reduction of information asymmetries through the sharing of knowledge, make it possible to acquire more information with respect to a given theme or purchase item and thus avoid the risk of running into one of the many "lemons"<sup>8</sup> (as George Akerlof, 1970 would say. In American slang, a "lemon" is a second-hand car with significant defects).

Therefore, the importance of the affirmation of Web 2.0 materializes in the possibility of reducing and, in the best of cases, dropping the information asymmetries that tend to manifest themselves mainly in the forms of interaction between consumers and entrepreneurs through the sharing of related knowledge in a platform easily accessible to users-consumers. Furthermore, it needs to be considered that such an open system in the creation, sharing, and use of the knowledge of others, and therefore sometimes of feedback, influences the reference market itself, rewarding or penalizing entrepreneurs based on the quality of the product they offer.

The possibility of expressing feedback quickly and easily allows the creation of more transparent and competitive markets, thanks above all to the ability to collect information on the price of products and services, their quality and relative reliability.

### **When the rating goes from the organization to the individual**

Therefore, if feedback and rating have been able to reduce the information asymmetries existing between the organization and the individual, where is the problem? Why worry about the risk that it could become too invasive and a form of control in social life? It would seem that the voice of people evoked by Hirschman has finally found, thanks to Web 2.0, the possibility of expressing itself at its best. In the cases mentioned, ratings and feedback are tools used by individuals capable of subjecting political and economic organizations to severe judgment. As long as there is an individual or even a multitude of individuals on one side and a robust organization on the other, it seems that the benefits always outweigh the risks.

Yet, comments are not always about the organizations. When we express an opinion about a novel on Amazon, we are not targeting an organization or the novel itself, but the author. Anyone who decides to publish their works exposes themselves to public scrutiny and, if in the past only the opinions of experienced literary critics in newspapers were known, today ordinary readers have the same possibility. Things are no different when dealing with businesses that identify themselves with the owner, as in the case of a family-run restaurant or a barbershop, or large organizations, often associated with the names of their founders. Today, in fact, names like Bill Gates, Jeff Bezos, Mark Zuckerberg, and Elon Musk represent publicly traded companies such as Microsoft, Amazon, Facebook and Tesla, and it was the same in the past with Henry Ford, Armand Peugeot, Thomas Edison, George Westinghouse and the companies to which they have given their surname.

---

that the individual is not aware of providing (e.g., the time spent on a webpage of a given website rather than another or the frequency with which something is searched in search engines).

<sup>7</sup> Other constitutive elements of Web 2.0 can be traced back to the possibility of using non-hierarchical methods for the arrangement of contents or to the interconnection of applications within the same platform (Bleicher, 2006).

<sup>8</sup> We are referring to the famous article by Akerlof (1970) dedicated to the used car market where, in fact, the seller has full knowledge of what he is selling, while the buyer does not. Hence the possibility of the former to palm off a "lemon" to the latter.

However, the use of feedback is not directed only towards organizations, but, especially in recent times, the expression of similar forms of judgment borders and is also directed towards individuals who are in a subordinate position. The cases in which the worker, and not just the owner, becomes the object of the feedback fall, in fact, in a grey area that needs to be explored. When does it happen?

If we talk about a restaurant on TripAdvisor, is it legitimate to also refer to the services of an employee such as a cook or a waiter? If you leave a negative note on the delivery of a DHL package, can you cite the courier's inefficiency? There are cases of services in which workers are publicly and periodically evaluated by users, such as happens, for example, with Uber or Bolt drivers, regardless of whether they act as freelancers or as employees of a car rental company.

In all these cases, the consumer who expresses an opinion does so without even knowing whether the workers have been put in a position by the employers to do their best. If in a review on TripAdvisor users complain about the quality of cooking and attribute the responsibility to the cook, if they complain to the driver about the delay in delivering a DHL package, customers do not know what is behind the production process and the poor quality of the service: it could depend on the fact that the employer supplied poor quality oil or a broken-down car. Responsibility may not be attributable to the worker, but to the organization of the production process that does not allow them to do better and carry out their professional performance in a safe and efficient manner. Anyone in a subordinate employment relationship cannot be held responsible for the product or service provided, while customer feedback could point to it.

It is more and more frequent that for some services – airport toilets, company canteens, post offices – users are asked to give an instant evaluation by pressing on a smiley with shades ranging from green to red depending on the user's liking. The service provider is asking for the user's collaboration to supposedly improve the service, although the latter has no idea of how his evaluation, which she is voluntarily called to provide, will then be used. It could, for example, be used by contractors (such as airport authorities) to get knowledge of how the service is provided by the contractors, or it could be used by the latter to evaluate their employees. In this way, the unsuspecting user, who presses the green or red button, could find himself defending the users' interests (for example, asking airport authorities to spend more on cleaning the toilets) as well as helping to monitor the rhythms work of employees.

The sense of omnipotence exercised by those who rate it often makes us forget that those who issue the rating are not omniscient and therefore do not have all the necessary information. The company that evaluates the worker cannot know the reasons that reduce her performance; the individual who evaluates another individual is not aware of the reasons that prompted her to behave that way. The consumer evaluating a service may not know the difficulties of the company or what the organization does to maintain those standards.

The most disturbing case of user evaluation is perhaps represented by prostitution, which is also increasingly advertised (and sometimes consumed) on the Internet. There are escort sites where customers even have the right to express opinions on the service they have used. In this way, the sex worker exposes herself, presumably voluntarily, to the judgment of the clients, who can comment on her physical characteristics, the activities she is willing to perform, how much she has been involved in the performance. All, it is supposed, for the benefit of other customers. A practice that demonstrates that the “solidarity” across customers is far more important than that between the customer and the sex worker, just as the customers of a pizzeria are more supportive of each other than they are with

the baker. However, as with feedback provided outside a restroom, the client does not know whether the comments are being used by pimps, who may use them to force slave sex workers to make their services more attractive for actual and potential customers.

As long as the relationships between spouses are inspired by monogamy, we can hope that the same rating system will not be applied within the marriage bond, but it would not be surprising that in other real and virtual communities, such as for example in the social network Tinder, sooner or later it will be introduced systems that allow to evaluate the performance of those who dedicate themselves to the so-called one-night-stands.

### **From reputation to political control: the Chinese nightmare**

Reputation is a vital element in the life of a community, and where individuals and organizations come into contact, the availability of information on potential contractors is of great importance. For example, in Italy, criminal record certificates declaring the absence of criminal charges are still essential for obtaining a job in the public sector. Strangers who intend to enter into a contractual relationship are more likely to resort to a system where their reputation is attested thanks to the endorsement of third parties, as when reference letters are required.

In countries with rooted liberal traditions, such as the United States and Great Britain, reputation can also be a crucial factor in the face of the judiciary. If an individual has a dangerous traffic accident, the courts could consider not only the dynamics of the accident, but also the automotive records and even civic rating of the perpetrator. Thus, even in the face of an offense that has caused a serious accident, the Court could consider as mitigating the points that the person responsible has on the driver's license and the class of motor insurance or whether he has worked as a volunteer in the parish of the neighbourhood or if is a blood donor. This information is not in the public domain until an unfavourable event may induce those in the dock to exhibit it.

As important as the social reputation of individuals may be, in liberal systems it is forbidden to make confidential information available outside the original context. Among the functions of the government there is also that of protecting the confidentiality of citizens' personal data and, already at the dawn of the digital revolution, there were those who, like Stefano Rodotà (2009), warned against the danger that the new technological potential could affect the right to privacy<sup>9</sup>.

In liberal countries, the public authority is committed to preventing the construction of dossiers used for illicit purposes, which would threat individual freedoms, as demonstrated by the case of Facebook/Cambridge Analytica (Manokha, 2018).

However, not all governments share the same spirit. In some countries, the view that citizens' duties are more important than their rights dominate, inducing the various governments to look with interest at the new technological potential. If, on the one hand they can be used to encourage the civic attitude of the population and penalize antisocial behaviour, on the other the question arises of the possible drift towards invasive forms of control, capable of having repercussions on the social, political and even personal life of the individuals.

---

<sup>9</sup> New technologies and new ways of using the online world expose the user to concrete risks with regard to his right to data privacy, so much so that it has become necessary to introduce severe regulations, such as the General Data Protection Regulation (GDPR) of the European Commission.

In a collectivist vision, which some might even define as simply civic, it could be thought that it is advantageous for citizens to have their own rating, elaborated through a sort of observatory that collects and makes available the assessments gathered for various purposes. This would allow, when entering into a commercial relationship (such as a rental agreement) or even simply a social one (such as when welcoming an individual into a club), to have much more extensive information than that provided by a few letters of reference.

In this regard, attention was drawn to the Social Credit System tested in China. The system is still being developed and, at present, combines different forms that are being tested in various parts of the country<sup>10</sup>. In some cases, this is a collection of data like that of the credit rating used by Western banks, in others city-level experiments it has been extended to social behaviour, in other cases it has become so widespread that it has assembled information collection by the bank and of the central government on the economic, ethical, and social behaviour of individuals.

It is not yet clear in which direction the Social Credit System will evolve, but it should be emphasized that, currently, there is no transparency both on the methods of collecting information and on its purpose (Mac Síthighm and Siems, 2019). Large domestic technology companies, such as Alibaba, Baidu and Tencent, the country's central bank and other local authorities, can access and combine data from different sources, including credit history, criminal records, and social behaviour of individuals, without the necessary transparency on the data collected and the purpose for using them (Zhang, 2020).

Thanks to new technologies, including those of facial recognition, the government has the technical possibility and the political will to monitor and aggregate data on individual behaviours by elaborating a synthetic score. Individuals do not always voluntarily submit to this monitoring since the rating can be assigned without their knowledge. Aspects considered may include contractual reliability, social and personal behaviours, as well as interpersonal relationships. In some cities where the system has been introduced, the technology is able to record actions such as crossing the street at a red light, propensity to gamble or spending too much time playing video games.

Virtuous individuals manage to obtain advantages, for example enjoying lower fares, privileged access to transport, as well as access to credit, while individuals deemed vicious are penalized and excluded from using some services, such as access to high-speed rail and air travel. Furthermore, this classification would also seem to impact on an intergenerational level: children of non-virtuous parents could be excluded from private schools and universities. The surprising fact, at least for those who grew up in the liberal West, is how high the consensus is towards the system: most of the Chinese population seems to support such evaluations (Kostka, 2019).

What is important to underline here is that there is no technological obstacle to the creation of a universal social evaluation system, and, in fact, the debate currently underway in China focuses on the advantages and disadvantages of having a single centralized system in the hands of the government, or a variety of parallel systems operated by large private companies.

As repeatedly noted by those who have commented on the forms that the Social Credit System is taking in China, there is a radical difference between Western and Eastern values (Mac Síthighm and Siems, 2019; Zhang, 2020), which hardly leads to mutual understanding. From the perspective

---

<sup>10</sup> Despite having several components, the Chinese Social Credit System, has been scrutinized in recent years, especially by Western authors. For a critical analysis, see Creemers, 2018. It has been noted that we cannot speak of a single system, and that, at present, different evaluation systems have been tested, applied by the private sector, by local authorities and also by the central government (Liu, 2019; Mac Síthighm and Siems, 2019).

of those developing these systems, the intention is to create methods to combine the behaviour of *homo economicus* with *homo moralis*, providing individuals with a sort of identity card which provide a rating for both aspects (Ding and Zhong, 2021). Such an attempt is presented as a way to increase information and, consequently, encourage reliability. The system could give advantages because it would discourage individuals from antisocial behaviours. Those citizens which throw litter on the street may be penalized by paying a higher urban tariff for waste collection, while those who comply diligently with re-cycling instructions may be rewarded with a discount. From liberal perspective, however, the lack of transparency on the purposes of the system makes it a tool for social control rather than a tool for civic participation.

The system applied on a large scale would not resemble the dystopia of *Nosedive* since the inputs for the rating are not provided voluntarily by the citizens themselves, but imposed from above by a government which, moreover, is not even elected. Attending an anti-government protest could lower your rating, as could cultivate friendships with outlandish characters. We will find ourselves in a much more sinister system and close to the old, but no less disturbing, dystopia narrated by George Orwell in *1984*.

We have strayed far from the envisaged possibility of using a rating to allow individuals to use their voice against organizations (public or private) that are called upon to satisfy their needs. We have reached an inverted situation, in which it is the organization, and among them the most powerful and therefore the most formidable, i.e., the central government, that has the possibility of controlling and evaluating individual citizens with a rating.

### **Three different types of feedback**

The analysis conducted so far allows us to advance some considerations and conceptualizations on feedback. This can be understood as any form of feedback that an individual or an organization can receive when they express their interest in the matter and activate specific methods to receive it. However, it is necessary to carefully distinguish between the existing forms of feedback. By applying this definition, it is possible to identify three different types of feedback: a) the bottom-up one; b) the transversal typology; c) the top-down one. Table 1 summarizes the three types.

**Table 1 – The three types of feedback**

	Definition	Power dynamics	Applicability	Policy Options
a) Bottom-up feedback	Evaluation that matches with the category of voice elaborated by Hirschman. The user or citizen can express themselves on products and services for the benefit of other users, which can be useful to the organization to improve its strategies.	The power ratio is contentious and depends on the ability and willingness of the organization to react and consider feedback, as well as the nature of the voice.  There may be abuses or unjustified boycotts.	Web 2.0 has enhanced bottom-up feedback: in economic systems through purchased product reviews and in politics through surveys and approval ratings.	Regulation of the use by educating content-providers about the objectives and the boundaries of the activity, also providing guidelines.  Extended supervision of the comments carried out by platforms.
b) Transversal feedback	Evaluation through peer-rating, between subjects with similar power, as in Nosedive. It allows for a multilateral system of social reliability.	No subject can boast specific prerogatives over the other, such that the methods and consequences of the feedback expressed are arbitrary and without regulation.  Risks bordering on online hate speech.	The applicability of transversal feedback is in its infancy, but it is starting to take shape through hybrid and veiled forms (as in the case of Airbnb, in which it is possible to evaluate hosts and guests). It is not used in purely political organizations or purely economic systems.	Supervision of the comments provided by platforms, but also involvement of consumers associations and focus groups.
c) Top-down feedback	Evaluation based on standards and indicators that	Asymmetrical power dynamics: the organization	It is common in business organizations	Regulation to guarantee the privacy of

	the organization performs towards the user, the worker or even the citizen.	can exercise power over the subject, and it influences the methods and consequences of the evaluation expressed. The individual's ability to react is low.	when employees or customers are evaluated by the company. Rare in political systems.  Danger that evaluations collected for different purposes will be combined and end up being a sort of digital identity card.  If enforced by the government, it endangers individual freedoms.	citizens on the ground of European General Data Protection Regulation (GDPR) and other acts.  Independent authorities overseeing the use and abuse of individual profiles of ratings.  Parliamentary control over the government use of data.
--	---	--	---	---

Source: Authors' elaboration.

**(a) Feedback with bottom-up dynamics.** It is the type of feedback that Hirschman contemplates in his essay and includes all those forms of expression or protest that originate from users and / or citizens towards the products and / or services offered by the reference organization. Feedback with bottom-up dynamics is that we are now most accustomed to. Think of reviews on Amazon, Tripadvisor or the satisfaction surveys of a political candidate.

Basically, the expression of any individual's opinion, whether in the form of a rating or not, brings benefits for others and for the organizations themselves. For the former, because the word-of-mouth mechanism is activated and can confirm the quality of a product and/or service or warn other consumers if the quality is inadequate. For organizations, on the other hand, receiving feedback has a decisive value for directing, shaping, or modifying market strategies or political and electoral ones.

**(b) Transversal feedback.** It is the closest typology to the Black Mirror *Nosedive* episode and is based on a peer-rating system in which all actors have equal power over each other and can evaluate and be evaluated by each other, hopefully on a voluntary basis.

Since all those involved have the same prerogatives, the power asymmetry is reduced, but this does not mean that benefits are distributed for all. In such a system, there is no quality standard, neither implicit nor explicit, in contrast of what happens within the other two types. Therefore, there is no regulation, such that the consequences are arbitrary and strongly dependent on the users' ability to use the tool responsibly and oriented towards the common good. The risk is that the feedback is not expressed with the intention of obtaining an improvement from which the whole community will



be able to benefit, as much as for individual purposes. And it can simply degenerate into the online hate speech that has been brought to the attention of public authorities for several years (see Gagliardone et al., 2015).

This type of feedback does not seem to concern organizations directly but indirectly. Indeed, some online platforms such as Airbnb use this evaluation tool not only to judge the quality of the stay, and so the host, but also the reliability of those who use the service, and so the guest. It still comes in hybrid and veiled forms, but future development could easily evolve in something different than the original format.

**(c) Feedback with top-down dynamics.** It is an assessment that the organization performs based on specific standards and indicators that it has defined itself to judge and categorize the user, the worker or even the citizen. In this case, the power dynamic is highly asymmetrical and decidedly in favour of the organization that uses the feedback to pursue its own interests to the detriment of the user. The field of action of feedback with top-down dynamics is found, to date, mostly in economic systems and, more specifically, within companies to evaluate the workers' performance, such as in Amazon or in banks to assess customers and disburse credit. Both employees and customers can therefore be subject to evaluation. In the economic landscape, top-down feedback can have an impact on the services and products that citizens can access (see credit rating), but it presents limited risks, especially if the spheres are kept separate and it is not allowed, for example, that the Amazon worker rating is shared with banks for their credit rating and vice versa. The alarm bells should go off when it is allowed to unify top-down feedback sources collected for different purposes. Credit rating, auto insurance class, and driver's license points may have a use, as long as they aren't aggregated and used to dispense rewards or fines in different fields.

Far greater concerns arise when individual assessments are carried out in the policy setting.

### **Necessary actions to avoid unwanted effects of feedback**

If we distinguish three possible forms of rating and feedback, it is perhaps clearer what the great opportunities are and where great dangers lurk. Indeed, each of the three options requires different actions to avoid an authoritarian or invasive drift.

Regarding the bottom-up evaluation, the explanation and clarification of its purposes can be helpful. It requires, for example, explaining to consumers who provide inputs that it is their job to evaluate products, processes, and services, not the employees who provide them. Rather than through specific regulations, it is necessary to start a kind of digital civic education that helps consumers, users and citizens to comment and evaluate the services they use in a proper way. It is unrealistic that professional registers of those who provide reviews on websites are created, like the one that Karl Popper (1996) suggested to introduce, moreover with limited success, for mass media operators. We know well how difficult it is to counter hate speech on the web (see, for example, the attempt of the Council of Europe to create parameters for its regulation, Keen et al., 2020). Yet the more instrumental nature of feedback could perhaps make the task of regulating it easier.

It is necessary that the same platforms that collect these comments moderate the comments provided. In many cases, they have even been encouraged to do so by national legislation. If certain terms and judgments are not accepted, it is advisable to explain to those who express them the reasons

why they have been rejected and, at the same time, to give a role to users' and consumers' organisations.

Also, regarding the transversal evaluation, it is necessary to start an adequate digital civic education. It's not just a question of avoiding forms of bullying, racism, incitement to hatred, problems well present in the public debate and against which digital platforms are already acting. It would be equally essential to avoid a more subtle, but equally insidious, form of social classism based on digital identity. As *Nosedive* shows, social networks expose much of our lives, and a clear separation between the private and the public component is unfortunately impossible. A greater awareness of what social networks can determine in life is certainly desirable.

There is a further aspect to consider, namely, to prevent the information available on social networks from being used for purposes other than those for which they were generated. We have already mentioned the scandal of the improper use made by a private company, Cambridge Analytica, of the data made available by Facebook, a scandal which is only the tip of the iceberg on the abuses actually committed and that could potentially be committed. The attempt by the Chinese company TikTok to enter the United States has become famous; the government intervened decisively, even announcing the possibility of a national security problem. These striking cases signal a much deeper need to regulate the use of data (Bria, 2020) and, more specifically with respect to the reasoning developed up to now, it should be avoided that data generated with transversal dynamics are used to create a form of digital individual identity which is publicly accessible.

Finally, as regards the possibility of using digital technologies to create top-down ratings of individuals, this is a risk that liberal democracies must fight fiercely. The problem is not so much when creating specific credit registers or when creating limited penalty systems (such as the one that prohibits some annoying fans from going to the stadium when their favourite team is playing) but when these scores are combined up to providing an overall assessment, with relative rewards and penalties, of the good or bad citizen.

In these cases, it is a question of clearly identifying the limits of the power that the government can exercise over the individual. In recent years the concept of republican citizenship has become more and more popular, demanding a greater individual participation to the public affairs, integrating the rights of the citizens with their corresponding duties. But this should not lead to forms of digital police control over the individual. It is therefore necessary to oppose and prevent the spread of schemes such as that of the Chinese Social Credit System. To do so, it is necessary to avert that a government has the possibility of imposing it, simultaneously invoking forms of control over the work of the government itself such as those exercised by the legislature and independent authorities.

## **Conclusions**

The possibility of providing timely feedback, on the one hand, has given greater power to customers, users, and citizens in the exercise of their relationships with organizations, while on the other hand is fuelling the risk of generating a society in which everyone is constantly evaluated and categorized into virtual social classes, determined by a series of quantitative indicators aggregated on the basis of the real behaviour of single individuals. Digital technologies have greatly enhanced all of this, and it is possible (and extremely easy) to combine evaluations collected by banks, insurance companies, employers, or even extracted from social networks to use them for purposes other than

those for which they were originally generated. Most of our daily actions can today be evaluated individually on the basis of specific and tacit indicators, the sum of which would constitute our personal rating. This, in turn, would outline the quantitative contours of our personality, generating a social stratification that could arrive at dictating the terms of access to goods, services and the labour market.

It therefore becomes essential to distinguish and separate the helpful functions that feedback and its rating can guarantee for the protection of consumers, users, and citizens from those potentially harmful and risky. To avoid them it is necessary not only to strengthen public control over the use of data, but also to create an adequate digital civic education that favours an aware use of the new opportunities.

## References

- Akerlof, G.A., 1970. The market for “lemons”: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3), pp. 488-500.
- Andreasen, A.R., 1985. Consumer responses to dissatisfaction in loose monopolies. *Journal of Consumer Research*, 12(2), pp. 135-141.
- Bashshur, M. R., Oc, B. 2015. When voice matters: A multilevel review of the impact of voice in organizations. *Journal of Management*, 41(5), pp.1530-1554.
- Bleicher, P., 2006. Web 2.0 revolution: Power to the people, *Applied Clinical Trials*, 15 (8), pp. 34-36.
- Bria, F. 2020. *Putting tech and innovation at the service of people and the green transition*. London, UCL Institute for Innovation and Public Purpose.
- Constantinides, E., Fountain, S.J., 2008. Web 2.0: Conceptual foundations and marketing issues. *Journal of Direct, Data and Digital Marketing Practice*, 9, pp. 231-244.
- Creemers, R., 2018. *China's Social Credit System: an evolving practice of control*. Disponibile a SSRN 3175792.
- Ding, X., Zhong, D.Y., 2021. Rethinking China’s Social Credit System: A long road to establishing trust in Chinese society. *Journal of Contemporary China*, 30(130), pp. 630-644.
- Duan, W., Gu, B., Whinston, A. B., 2008, Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems*, 45(4), pp.1007-1016.
- Gagliardone, I., Gal, D., Alves, T. and Martinez, G., 2015. *Countering Online Hate Speech*. Paris, Unesco Publishing.
- Hirshman, A. O. 1970. 1970. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA, Harvard University Press.
- Jennifer Lees-Marshment, Brian Conley, Edward Elder, Robin Pettitt, Vincent Raynauld, André Turcotte, 2019. *Political Marketing. Principles and Applications*. London, Routledge.
- Keen, E., Georgescu, M. and Gomes, R., 2020. *A Manual for Combating Hate Speech Online through Human Rights Education*. Strasbourg, Council of Europe.
- Kostka, G., 2019. China’s social credit systems and public opinion: Explaining high levels of approval. *New Media & Society*, 21(7), pp.1565-1593.
- Lim, E. P., Nguyen, V. A., Jindal, N., Liu, B., & Lauw, H. W. (2010, October). Detecting product review spammers using rating behaviors. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management* (pp. 939-948).
- Liu, C., 2019. Multiple social credit systems in China. *Economic Sociology: The European Electronic Newsletter*, 21(1), pp. 22-32.
- Mac Síthigh, D., Siems, M., 2019. The Chinese social credit system: A model for other countries? *The Modern Law Review*, 82(6), pp. 1034-1071.
- Manokha, I., 2018. Surveillance: the DNA of platform capital—the case of Cambridge Analytica put into perspective. *Theory & Event*, 21(4), pp. 891-913.
- Pacella, G. (2017). Il lavoro nella gig economy e le recensioni on line: come si ripercuote sui e sulle dipendenti il gradimento dell’utenza? *Labour & Law Issues*, 3(1), 1-34.
- Pan, M., 2023. Discussion of online reviews’ impacts on consumers’ behaviors. *Journal of Education, Humanities and Social Sciences*, 13, 219-225.
- Popper, K. 1996. *Cattiva maestra televisione*, edited by Giancarlo Bosetti, Roma, Donizelli Editore.

- Rodotà, S., 2009. Data protection as a fundamental right. In *Reinventing data protection?* (pp. 77-82). Springer Netherlands.
- Singh, J., 1990. Voice, exit, and negative word-of-mouth behaviors: An investigation across three service categories. *Journal of the Academy of Marketing Science*, 18, pp.1-15.
- Stigler, G.J., 1961. The economics of information. *Journal of Political Economy*, 69(3), pp. 213-225.
- Zhang, C., 2020. Governing (through) trustworthiness: technologies of power and subjectification in China's social credit system. *Critical Asian Studies*, 52(4), pp.565-588.