How do YouTube algorithms calculate value? An analysis of the production of value for digital music videos using the social logic of the derivative[1]

Como os algoritmos do YouTube calculam valor? Uma análise da produção de valor para vídeos digitais de música através da lógica social de derivativo

LEONARDO DE MARCHI[a]

University of the State of Rio de Janeiro, Graduation Programme in Communications at UERJ. Rio de Janeiro – RJ, Brazil

ABSTRACT

The article analyses the production of monetary values for music videos embedded on YouTube based on the theory of the social logic of the derivative. Using the value gap debate as a starting point, the hypothesis is that the method used to create monetary value for digital videos is the same technique developed in the derivative market. This phenomenon results in the generation of inconstant monetary values which entails, in turn, serious consequences for the social order of the record market in the digital age.

Keywords: YouTube, value gap, financialization of everyday life, social logic of the derivative

RESUMO

Neste artigo, analisa-se a técnica aplicada à produção de valores monetários para vídeos de música hospedados no YouTube à luz da teoria da lógica social de derivativo. Partindo do imbróglio sobre transferência de valor (*value gap*), a hipótese do trabalho é que o método utilizado para a criação de valor para vídeos digitais toma emprestadas técnicas desenvolvidas no mercado de derivativo. Esse fenômeno resulta na geração de valores monetários variáveis, o que acarreta, por seu turno, graves consequências para a organização social do mercado fonográfico na era digital.

Palavras-chave: YouTube, transferência de valor, financeirização da vida cotidiana, lógica social de derivativo

[1] A preliminary version of this paper was presented at the XXVI Congress of the National Association of Graduate Programs in Communications (Compós), São Paulo, Brazil, 2017.

[a] Visiting Professor at the School of Communications at the University of the State of Rio de Janeiro (UERJ) and permanent member of the Post-Graduation Programme in Communications at the same University. Orcid: http://orcid.org/0000-0001-5654-8938. E-mail: leonardodemarchi@gmail.com

INTRODUCTION

THE THEME OF the financialization of the everyday life has received significant attention in recent years, as it became clear a systematic application of the operational logic of the financial market in the conduct of social life. With the adoption of policies designed according to the precepts of neoliberalism, social security institutions were dismantled, liberalizing both capital and labour. In a highly individualistic, competitive society, wherein each individual is defined as a company, connected by digital communication technologies, the notion of risk has become a fundamental ideology for social control, serving to justify both social mobility and social exclusion (Beck, 2006; Lash, 1997). Therefore, it became necessary to create devices that can transform uncertainties (situations in which the agent has no information to base probabilistic calculations) of the lifeworld (*Lebenswelt*) into risks (situations in which probabilities can be attributed) to be assumed by the society.

The financial market has become a privileged locus for the development of technologies that can be applied to social administration (Sassen, 2016). Particularly, the derivative market gains prominence, not only because it is a trade based on risk management of economic activities, but also due to the digitization of the trading floor and advances in probabilistic theories. In this market, the technologies developed led traders to virtually fragment any underlying asset (commodity or future event) into different attributes, to use them in the creation of some new composed financial product (CDO, ABS, among others) to be traded on secondary markets. The detachment between productive activity and its risk has reached such a degree of abstraction that financial products have become independent of the original underlying asset, being a business of buying and selling virtual risks per se.

As different sectors of everyday life are exposed to market logic and, thus, estimated in terms of risk probability, derivative market techniques are of obvious interest to other sectors of the economy. Progressively, contracts for labour, education, or health services are more and more tailored based on the fragmentation of the underlying asset (the service itself) and its composition by several other products, priced based on a scale of risk probabilities. The objective is to reduce operating costs for the service provider by transferring the risks of the contracting activity to third parties (the contractors). Such dissemination of the logic of the derivatives into ordinary economic relations and, by extension, to the social relations in which individuals are conceived as micro-enterprises (as entrepreneurs, not workers), allows one to speak of a social logic of the

derivative as an implicit rule for the regulation of social life (Arnoldi, 2004, Arvidsson, 2016, Bryan, Rafferty, 2014, Martin, 2013).

One of the sectors of the economy that has been using the derivative technique on a massive scale is that of the digital platforms. In their search for new business models that can profit from lived processes that unfold in their interfaces, digital companies have sought to fragment the activities of their users into attributes to be lately gathered in some abstract informational unit, to which their algorithms will assign a specific monetary value. Social media, sports betting sites, peer-to-peer ridesharing systems, intimate relationships mediators, among others, can offer a cascade of new services, information and/or advertising, relying only on likely scenarios created by their proprietary algorithms. This practice creates conflicts not only of a legal character (violations of the privacy right of users) but also of economic order, especially regarding the pricing of virtual compounds. After all, how do algorithms calculate these values for attributes derived from processes experienced on digital platforms? This information is not always available, as algorithms are considered trade secrets.

In this article, I want to contribute to this discussion by presenting a case study of the production of monetary value for music videos hosted on YouTube (Google Inc.), a video streaming platform. YouTube is a suitable object for analysis because it is at the centre of one of the major disputes in the current creative destruction of the music industry. Despite having developed techniques that allow the company to pay royalties for the use of copyrighted material embedded on its plataform, the company has been repeatedly questioned over the disproportionality between the number of views (which can reach the scale of millions) and the low amount of money paid to copyright owners. This dispute became known as the value gap imbroglio.

This clash between the music industry and Google is based on a fundamental misunderstanding: how do YouTube algorithms value the views of videos hosted on the platform? After all, a view does not have a precise equivalence in terms of cash. So, what kind of calculation do the algorithms perform? My hypothesis is that the method used to create monetary value for such digital content has been borrowed from the derivative market. In this case, the concept of viewing a video is transformed in its very definition: rather than referring to the visual experience of an individual concerning a specific digital content, it is about the gathering of distinct attributes, produced by several users on the platform, that form an abstract information unit (a view), which will be priced by an algorithm. As a result, the values assigned to each digital content can vary vastly, because they depend on the specific combination of several attributes and valuation rules

that have been programmed. Although YouTube resembles cases already studied, such as sports betting sites (Bryan, Rafferty, 2014) or Facebook (Arvidsson, 2016), the conflict between music market agents and the digital company allows us to see how the use of derivative logic can cause disagreements between economic agents, generating disputes whose results will be decisive for the reconstruction of the music industry in the digital era.

This analysis presents considerable challenges regarding research techniques. First of all, it would be necessary to have access to the programming of the platform's algorithms to know how they perform the assignment of value, which is not possible because this is a proprietary algorithm, whose programming is a trade secret. Secondly, Google employees have a confidentiality clause in their contracts, preventing them from giving interviews about how the company works. To circumvent such constraints, I made the analysis based on: bibliographic review, using from articles and academic books on the topic to newspaper articles on the value gap; visualization of video monetization tutorials made by independent video producers (youtubers), as well as of YouTube's manuals on the conditions for the monetization of the videos published on the platform. It is important to stress that I do not want to guess what kind of programming is prescribed for the algorithms but rather to understand what kind of technique the company uses for assigning value to the videos.

Regarding the theoretical framework, I use three trends in contemporary social theory: the theory of risk society (or reflexive modernity), the actor-network theory, and the social logic of the derivative. Despite its simplistic evolutionism, the authors who defended the idea of reflexive modernity have correctly identified that this new stage of contemporaneity presupposes the phenomenon of individualization, i.e., the progressive liberation of human agency from social structures (Lash, 1997), in which there is the perception that the experienced world becomes a management of innumerable "risks" (Beck, 2006). The actor-network theory understands social life as a result of the connection between human and non-human agents, dissipating any criticism of technological determinism when dealing with markets whose functioning can only be realized in virtual space (Beunza; Stark, 2004, Knorr-Cetina, Bruegger, 2000). On a critical perspective, the discussion on the social logic of the derivative has been more successful in explaining the unequal power relations that are performed through devices such as algorithms, since the reason of this technique is the transfer of risks between agents endowed with unequal power.

Though recognizing discrepancies between these theories, I understand that the analysis of the social logic of the derivative can only be achieved by

recognizing the ideology of risk as a fundamental *motif* in social life and that it is not possible to think of contemporary sociability only by taking as a parameter of analysis the symbolic relations between humans or considering nonhuman artefacts from a merely instrumental or symbolic perspective, to the detriment of their materiality.

The article is divided into three parts. In the first, I describe what has been called the social logic of the derivative, emphasizing the role of algorithms in contemporary economy and society. Then, I apply this discussion to the creation of value in digital platforms in general. In the third part, I analyse the case of music videos on YouTube, discussing the value gap imbroglio, followed by a brief comment on the mechanism called Content ID, to argue that this technology transforms copyright holders into the agents most exposed to risks in this business model. As concluding remarks, I discuss some consequences of applying the social logic of the derivative to the digital music market.

THE SOCIAL LOGIC OF THE DERIVATIVE: THE FINANCIALIZATION OF EVERYDAY LIFE IN A RISK SOCIETY

> If we have lost the hope of assigning a site to value creation (labor, land, etc.) and must now live with long cascades of derivation, how do we set the criteria for good and bad cascades? (Lépinay, 2011: XXIV).

One of the themes that have been receiving more attention in economics, as well as in economic sociology, is that of the financialization not only of the economy but also of the very social relations, what has been labelled as the financialization of the everyday life. As Natascha van der Zwan (2014) notes, it is about the systematic application of technologies developed in the financial markets in other fields of life, questioning how an increasingly autonomous domain of global finance alters the underlying logic of industrial economy and even the functioning of democratic societies.

The due understanding of the interest in the financialization of different spheres of the experienced world is only possible if it is framed in a broader context of political, economic, social, and cultural transformations – which have been labelled neoliberalism, both as an ideology[2] and as politics. In his detailed analysis of neoliberal literature, Michel Foucault (2008) observes that the worldview held by these authors is based on the individualization of society and the commodification of all spheres of life. Unlike the classical liberals, for whom

[2] For ideology, I adopt the meaning assumed by Luc Boltanski and Ève Chiapello (2009: 33) of a "set of shared beliefs, inscribed in institutions, implicated in actions and therefore anchored in reality."

the rational individual had a determined space to act (the market), neoliberals profess that the individual must be understood as an enterprise, managing a human capital (a set of capacities, knowledge, competencies, and attributes of an individual that allows her/him to produce economic value from her/his subjectivity), and that all spheres of the experienced world can be commodified.

To realize this worldview, however, it became imperative to dismantle regulatory institutions of economic and social life, whether of the Welfare State or of the Socialist State (terms that, for the neoliberals, are synonymous), and to develop another way of life, based on the entrepreneurship ideology. More than a deontology tailored for the business world, entrepreneurship is an ethic that transforms a whole way of life in an immaterial resource for the market, that is, the subjectivity of each person becomes a source of value (Lazzarato and Negri, 2001). The work of Luc Boltanski and Ève Chiapello (2009) is precise in demonstrating how the literature on business administration has incorporated several critiques to capitalism made by the May-68 generation (especially what these authors call "aesthetic criticism," which preached greater freedom for the individual to make their work less bureaucratic and more creative), and transformed them into the ideal of individualized labourers, who define themselves as an entrepreneurs (no longer as workers), and must therefore adopt a proactive stance in order to increase their income in a market that is conceived as constantly changing.

Gilles Deleuze (2010) presents a brilliant synthesis of this new spirit of capitalism in his article on the society of control. On Foucault's assumption about the decline of disciplinary societies, which were based on body confinement and ordering of time and space, Deleuze points to the emergence of a new regime of power, in which individuals are atomized, becoming responsible for their constant insertion into a flexible labour market. This is what would define the fundamental difference between the factory and the enterprise: in the factory, there was a circumscribed space, with regulated time, to discipline the workforce; in the flexible enterprise, the boundaries of the workplace are removed, and the workforce needs to adjust to an ever-changing market, what means that it needs to keep working even when the working hours are over. What allows the functioning of this new social organization are the networked digital information technologies: devices displaying flexible forms and functions, whose decentralization and automation are most adequate for a society that lacks social structures and demands an incessant actual and virtual productivity.

In all these interpretations, the idea of risk is lurking. If some new stage of modernity presupposes the phenomena of exacerbated individualization

(*Freisetzung*) by promoting political changes that dismantle institutions dedicated to social security, forcing capital and labour to meet their needs according to the law of supply and demand, the main problem of social administration is to avoid some ontological insecurity that may results in anomie (Beck, 2006; Lash, 1997). Therefore, the ideology of risk is mobilized as a way of controlling the competition between social entities conceived as companies of different scales. The use of the idea of risk can be interpreted in a similar way as the use of theories of purity, as discussed by Mary Douglas and Baron Isherwood (2006: 82), tests of fitness for social mobility at the same time as they function as selective social exclusion techniques. To paraphrase these authors, a direct relationship between risk, success, and competency is established so that it becomes a rule in social relations.

This is only possible, however, when uncertainty (situations in which the agent has no information on which to base probabilistic calculations) becomes risk (situations to which probabilities can be attributed to) through techniques dedicated to interpreting the contingencies of the experienced world in terms of probabilities[3]. In other words, technical solutions are sought to make the risks that accompany the production of goods (energy supplies, biogenetics, deterioration of the environment, artificial intelligence, social inequality, among others) to be distributed, avoided, controlled and/or socially legitimized (Beck, 2006). To do so, risks must be transformed into isolated, standardized, and interchangeable entities, which allows their prediction through probabilistic theories. The locus in which uncertainties are transformed into risk is the financial market (Sassen, 2016).

[3] On the development of the uncertainty and risk concepts in both Economics and Sociology, cf. Beckert (1996).

Since the deregulation of financial activities and the digitization of the trading floor, financial markets developed into virtual networks of transactions that became global. The intensive use of algorithms provided a certain alibi of impersonality and precision to the laws of the market, expanding negotiations in both qualitative and quantitative terms (Muniesa, 2003; Lépinay, 2001). If negotiations were previously geographically restricted and dependent on the capacity of individuals to circulate information in isolated markets, with digital information technologies these negotiations could be freed of such constraints and carried out at a concrete universal point: the computer screens (Knorr-Cetina and Bruegger, 2000). This means that markets are no longer represented on the computer screens of digitized trading floors; they can only be, that is, become presented, in the cyberspace. The use of this technology has allowed the reinvention of several financial products, making these markets a source for credit, speculation, and management of business risks, both for companies and individuals (Beunza; Stark, 2004; Farhi, 1999;

Krippner, 2005). Among the renewed financial products, derivatives took the centre stage.

Technically, the derivative is a sort of risk insurance, being a contract in which future payments are established a priori, based on an underlying asset, that can be the price of a share, a commodity, a financial instrument, or an event[4] (Carneiro et al., 2011; Farhi, 1999). Its purpose is to protect economic agents against price fluctuations over time, transforming uncertainty into risk. Thus, economic agents who want to protect themselves from the risks to which their productive activity is exposed (market, credit, interest, currency, and/or operational risk), establish a contract through which they pass them on to a bank or financial institution.

If traditionally derivatives were simply instruments of insurance, since the deregulation, the digitization of financial markets and the creation of new formulas for calculating prices based on advances in probabilistic theories (notably the Black-Scholes-Merton model[5]), they have become a device that allows not only to decouple the risk from a specific productive activity, but also to make the risk itself another product to be reproduced, regardless of the underlying assets that were in the original contracts (Bryan; Rafferty, 2014).

This is possible because the derivative contracts were standardized so that they could be chopped into several attributes and then bundled together into packages containing parts of different contracts, each one displaying different degrees of risk. Finally, they can then be traded on secondary markets under obscure titles such as Collateralized Debt Obligation (CDO), Credit Default Swaps (CDS), or Asset-Backed Security (ABS). As Bryan and Rafferty (2014: 893) summarize,

> Stripped of mathematical formalism, the idea of derivatives is quite simple. They involve deconstructing a 'thing' (and we use a bland term intentionally) into a set of constituent elements or attributes, and configuring those attributes in a way consistent with quantification. These quantified attributes can be interpreted through the lens of risk and risk-trading in a way that it is unlikely the underlying, original 'thing' will be.

As a result, derivatives no longer relate to an underlying asset or to the realization of a future event, but only to the transfer of risks among tertiary agents, so that they can be traded without limits. This created an arrogant feeling in the financial market that derivatives were a product that, although based on "risk," were "risk-free" for traders (Varoufakis, 2016: 35).

[4] The history of derivatives dates back to the epoch of modern capitalism. However, they were a limited and specific form of financial product. The current derivatives market has, as founding marks, the opening of the Chicago derivatives exchange in 1972, and the publication of the Black-Scholes model in 1973. There are four basic types of derivative products: forward, futures, options, and swaps. However, since they are only mathematical formulas, such products can be combined with each other.

[5] Published in 1973, the equation was originally developed by the economists Fisher Black and Myron Scholes. It is a mathematical model of the market in which the price of an asset is calculated in a stochastic process (a group of random variables representing the evolution of a value system over time). Years later, the economist Robert C. Merton published an article in which he put the equation in a mathematical form. This equation made it possible to calculate the option price based on the (future) volatility assessment of the underlying asset, demonstrating that the underlying price follows a record and a predictable random path. Its formula projected an image of stability for any derivative markets. As MacKenzie and Millo (2003) argue, the success of this model lies in addressing several economic situations that had the characteristics of options, such as business decisions and corporate debt assessments, becoming the mathematical paradigm of contemporary economics. In 1997, the three economists were awarded the Nobel Prize for Economics, as their pioneering formula for stock valuation allowed the creation of new synthetic financial products.

Once an underlying asset gives way to a derivative product, the challenge is to assign value to it. The way in which it can be done is the core of the arbitrage market. In an ethnography of an arbitrage office[6], Daniel Beunza and David Stark (2004) describe how technicians determined the intrinsic value of an asset, as opposed to its market value, not from traditional data (revenues minus expenses, amounts of assets and debts, stock value) but from an evaluation created by probabilistic formulas from individual attributes:

> Value investing is the traditional 'buy low, sell high' approach in which investors look for opportunities by identifying companies whose 'intrinsic' value differs from its current market value. They do so by studying a company's annual reports, financial results, products, and executives; they then compare the intrinsic value that emerges from this analysis with the market price of the company. […] Value investors map the many aspects of a company by translating them into abstract variables – e.g., return, growth, risk – and collapsing them into a single number ('value') with the use of formulae such as discounted cash flow. (Ibid: 375).

[6] Arbitrage is to identify market opportunities in the difference between the intrinsic value of an underlying asset (companies, countries, products, currencies, among other possibilities) and their market value. This is possible because, through algorithms, the value of an asset is established from unusual correlations between certain derived qualities, that escape traditional measures, such as investments, equity, or the number of employees. Insofar as there is a gap between the difference of the market value, as measured by the stock exchange, and the so-called intrinsic value, the risk of doing business is assumed (Beunza, Stark, 2004, Arvidsson, 2016).

This means that the intrinsic economic value of an asset is no longer based on an objective (work) or subjective (utility) quality, but rather on mathematical formulas that extract values from a bundle of independent attributes. This results in a wide range of possible values, which makes the choice between good and bad cascades of derived value the fundamental challenge of value production nowadays, as Vincent Lépinay (2011) suggests at the opening of this section.

These virtual transactions are possible due to a device: the algorithm. This is a set of rules that demands a specific result from a well-defined and unambiguous mathematical formula so that it can be translated into computer language and executed by it (Knuth, 1980). It is a kind of narrow artificial intelligence, programmed to perform different scenarios for all predicted probabilities, looking for optimal combinations (O'Neil, 2016). Since they do not depart from any previous knowledge, algorithms divide the information into distinct fragments or attributes, draw possible scenarios, and perform calculations in nanoseconds. From the first processing, the algorithms use the matching result as the basis for their next one, making the accumulation of the previous results the reference to the subsequent ones (Neyland, 2015). Their calculations consider probabilities based on past events, which may happen again in a probable future, therefore, not necessarily referring to 'what is' but 'to what can be' inferred from past proclivities (Arvidsson, 2016). It is

this calculated probable scenario that makes its decisions entail what Fabian Muniesa (2003: 288) classifies as "accuracy effect" (*effet de justesse*), in the technical sense of the word, as when it is said that the result of an arithmetic operation was correct.

Notwithstanding such a sense of security, it is critical to understand that the derivative always deals with the transmission of the risk(s) of one (or more) business(es) to a third party(ies). Although the technology of the derivative market had drastically reduced risks to traders, the 2008 financial crisis revealed that risks continued to exist – though those who would bear the consequences of a failure would not be the agents initially thought. The rescue of banks and risk agencies made by national governments using public funds, and the lack of punishment for those involved, even dismissing a more effective regulation of the financial market, revealed that the agents who would bear the risks of the derivative business would be the individuals who took on debt to pay mortgages, instalments for consumer goods, and the ordinary taxpayer (O'Neil, 2016, Sassen, 2016, Varoufakis, 2016). This means that risk is assumed by agents who would like to be protected from them: a perverse inversion of the derivative *raison d'être*.

Despite the systemic crisis that the derivatives inflict on the economy, its technique has been used by other economic sectors to create value for different goods and services. Increasingly, the design of contracts for a plethora of services, such as employment, housing, education, or health has been following a model of agreement in which the service itself is broken down into different *attributes*: health insurance only for heart or liver disease; educational contract with the possibility of choosing disciplines *a la carte*; temporary employment contracts for specific functions, and so on. This movement allows, at once, (i) the risks of the business to be passed on to contractors (workers, tenants, students, patients) and (ii) to use the contracts themselves as underlying assets for new synthetic products in the derivative markets, insofar as they represent a safe harbour for endless risk speculations (Bryan, Rafferty, 2014 and Sassen, 2016). This spreading of the logic of decomposing underlying assets has affected different levels of social life in societies that define themselves as a microenterprise gathering. Hence, social theorists can label it as part of a social logic of the derivative (Arnoldi, 2004, Arvidsson, 2016, Martin, 2013). Another sector of the economy that has relied on the derivative technique is that of digital platforms.

THE VALUE OF DIGITAL CONTENTS IN THE PERSPECTIVE OF ALGORITHMS: USE OF THE SOCIAL LOGIC OF THE DERIVATIVE BY DIGITAL PLATFORMS

All business models developed by digital platforms share the same fundamental challenge: how to designate a monetary value to the uses of their abstract services? The characteristics of the digital economy demand vast networks of users. This can be achieved by providing a high number of files, through sophisticatedly designed and user-friendly interfaces. However, this results in a risky gamble: high investment in information technology for the provision of sophisticated services, but for a minimal monetary cost (or even for free). When it concerns services of digital content, it is necessary to consider that the enjoyment of a song, news, or audio-visual material is fragmented and distributed, that is, the user can access parts of the files (sometimes just for seconds), share them with their peers (even without access the content), comment or enjoy the file on different devices, among other possibilities that break with the cognitive conventions of analogical cultural goods. The result of this chaotic use is the production of an immeasurable amount of data to be converted into some monetary value. The question is: how to do so?

To address this challenge, digital platforms have been relying on derivative techniques. Firstly, proprietary algorithms are programmed to fragment an underlying asset into different attributes, using parameters that allow them to compare entities that are seemingly incompatible. Then, these attributes are recomposed in a new compound: an abstract informational unit. As in the arbitrage market, some monetary value is attributed to this product. Thereafter, other compounds will be valued for their own intrinsic value. Thus, a wide range of social practices becomes amenable to codification as they are standardized to be interchangeable.

A paradigmatic example of this procedure can be found on sports betting sites. As Dick Bryan and Michel Rafferty (2014: 893) describe, usually the purpose of this type of bet was about the result of a game of a certain sport (football, basketball, cricket, chess, etc.): it was important to know who would win, what would be the final score of the match or which player/team would be the best or the worst among other binary choices. The digital betting platforms, however, adopted the method of fragmenting matches into attributes and creating probable scenarios for the games. With this move, they were able to break up any kind of match, of any sport, into a variety of equivalent attributes so that the bettors could set up a *portfolio* (according to financial jargon).

The same user can bet on what will be the score of a basketball match at the end of the first quarter of time, which player of an "x" football team will score a goal until determined minute of a match, how many strikes a baseball player will miss during one season, among a variety of other attributes that can be put together in a betting package. As happens in derivative markets, this technique increases the odds of winning for the bettors with the lowest possible risk (in this case, if the player of the team "x" does not score the goal until the minute expected, the whole bet is not lost but only a small fraction). There are platforms that even provide room for users to bet against each other, as happens in securitization markets, since odds change due to market forces (red cards, injuries during a match, among other possibilities). The outcome of the game itself (the underlying asset) becomes just a minor attribute to be added to the portfolio.

A similar case is found in the advertising on Facebook, as demonstrated by Adam Arvidsson (2016). Instead of using concrete socio-economic parameters to evaluate their users (social class, age, gender, occupation, place of residence etc.), Facebook algorithms work with attributes that are derived from information-sharing practices and interactions of its users on the platform. These attributes are recomposed so that they are related to the attributes of other users of the network. This method of data-mining is called social graph.

The social graph collects information from all user activities on Facebook (as well as in applications that belong to the company, such as WhatsApp, Instagram, Messenger, among others) to transform such lived processes into abstract informational entities, labelled *objects* and *edges*. The *object* category comprises various information that may vary from published content to the user's own personal data. The *edge* category refers, in turn, to interactions between users, or to interactions between users and content. So, when someone makes a post, it characterizes an object, and when a user likes a picture, it constitutes an edge. Both categories are connected to each other by algorithms, which form several possible scenarios from a topological representation of these data.

This system allows the company to evaluate the potential of attention (and, therefore, of purchase) of each user. Its logic assumes that user A may be interested in the product (a shoe, for example) that the user F (who is on the friends list of A) bought through the internet, or that user A may be interested in a news of a particular newspaper, because user T (who is a friend of the user M, who is, in turn, a friend of user F) enjoyed it and so on. The platform performs such assessment by comparing these implicit and potential trends in the same way that the Black-Scholes-Merton equation allowed for pricing decisions in

derivative markets, taking the inferred volatility of different comparable asset prices. Once more, the inference here is not based on the content of the post itself, nor on any affective bonds (to be a friend of somebody) in real life. These are projections that are based on past proclivities to create future scenarios of interests or affinities and, thus, to establish a value of their users for the advertising market.

Since there is no need to relate to the underlying empirical reality, the measurement of values is dynamic: each time a value is assigned for a set of attributes, the algorithms begin to search for new compositions and, therefore, reach different values. This is a most important feature as it will make the amount of money paid to the owners of the underlying assets to vary enormously. Because the algorithms are proprietary, however, their programming cannot be made public. Therefore, one cannot be sure that there is a precision effect, i.e., that the algorithms are proceeding within parameters that reach a consensus between those providing the underlying assets and the digital platforms. Such opacity opens space for several disputes in the markets where such platforms operate. This is particularly true in the case of the value gap imbroglio between the music industry and YouTube.

THE IMBROGLIO OF VALUE GAP AND YOUTUBE

In the introduction to a report on recorded music consumption practices, the CEO of the International Federation of the Phonographic Industry (IFPI), Francis Moore, makes the following statement:

> Research on video streaming and the dominance of YouTube is instructive for the ongoing debate on the Value Gap. User upload services, such as YouTube, are heavily used by music consumers and yet do not return fair value to those who are investing in and creating the music. The Value Gap remains the single biggest threat facing the music world today and we are campaigning for a legislative solution (INTERNATIONAL FEDERATION OF THE PHONOGRAPHIC INDUSTRY, 2017b: 3).

These words formalize the position of the music industry lobby against YouTube, which has been called by the specialized media the value gap problem. Put simply, this dispute refers to the apparent disproportionality between the large number of views of copyrighted music videos embedded in Google's video platform (that can reach millions) and, as considered by

the interested party, the low and inconstant amounts of money paid for the copyright holders.

The basis of this disagreement lies in a divergence on how the value of digital content is generated. Copyright holders assert, on the one hand, that the money YouTube earns comes from the creative work embodied in the content produced by composers, performers and producers (music publishers and record companies) in the music industry. Logically, they consider that the value returned by the platform is unfair and that the work of others is vicariously appropriated by YouTube. It is an essentialist conception of value since it is understood that value is generated from human labour, which transforms something useless into useful. On the other hand, the company states that if it was not for its ability to attract users and to calculate the lived processes there (such as likes, views, comments, and shares) in the form of monetary amounts that will be paid to copyright holders, the use of such content would occur only through informal channels (peer-to-peer network sharing programs), which would not generate any return to copyright holders (RBB Economics, 2017).

In this perspective, the value does not lie within the good itself, but in the capacity to generate access to the digital content and to be able to develop technologies that transform the lived processes that unfold in the platform into monetary quantities. It is a relational conception of value: the value of the contents can only arise from the control that is exercised over the connections established between nodes of a network; the question would be how to choose the value between the cascade of values that the algorithms produce. It is true that there are different vested interests lurking behind the scenes[7]. However, the key question remains: what is the method used by the platform to calculate a monetary value for its contents? In the case of the music industry, this is an issue of utmost importance.

After a long period of decline in record sales, since the 2010s there was an increase in the trade of digital music due to different business models, from paid download to streaming services (De Marchi, 2016, Herschmann, 2010, Kischinhevsky, Vicente, De Marchi, 2015). In 2016, such activities accounted for about 50% of the total revenue of the international music industry (INTERNATIONAL FEDERATION OF THE PHONOGRAPHIC INDUSTRY, 2017a).

Among such business models, YouTube stands out as one of the most remarkable experiences. According to the IFPI (2016: 10) market survey, about 82% of average YouTube users access music primarily (the number increases to 93% among users aged 16 to 24 years). A recent version of the same research indicates that 55% of access to music via streaming services is made through

[7] According to journalist Glenn People (2016), the claims of copyright holders seek to achieve some complementary objectives. First, major songwriters and music publishers make public their complaints against YouTube as negotiations begin with Google for the renewal of digital content licenses. Second, such complaints serve to garner support among politicians to support changes to Article 512 of the Digital Millennium Copyright Act (DCMA), which is the so-called Safe Harbour institute. Because YouTube fits into the definition of an "internet service provider," since its content is produced by users, it can use copyrighted videos without negotiating, first, with the copyright holders (blanket licenses). Third, there is an effort to restrict the development of business models based on user-generated content to prioritize content provided by communication and culture corporations and made available through closed streaming services.

videos, with 46% of this consumption made through YouTube (Id., 2017b: 5). A survey published by the RBB Economics (2017), commissioned by Google, claimed that YouTube paid more than a billion dollars to the music industry for copyright royalties in 2016 alone.

Such excellent performance is due to the platform's sui generis business model, as observed by Pedro Francisco and Mariana Valente (2016: 274). After all, this is a company that provides infrastructure for other digital companies (multi-channel networks) and depends on its users for the provision of contents, flowing between amateurism and professionalism (Burguess; Green, 2009; Cunningham; Craig; Silver, 2016). The visualization of videos is also singular: a user can view a file or a sequence of them but can stop it or change it in seconds; she can also select what type of advertising to see, among other possibilities of interaction with the content. This type of use blurs the traditional definitions of public display and private consumption, making it difficult to demand payment or exemption of certain tariffs[8]. This signals the importance of the platform not only to relay broadcasted programs (TV and radio) but also as an experience of new types of musical products (Sá; Holzbach, 2010).

Notwithstanding its remarkable capacity to attract users, or rather for this excellent performance, doubts about how the company assigns value to its digital content gain attention. After all, the traditional television model of selling advertising time is discarded in favour of other factors. But how is this done?

It is true that there is an abundance of video tutorials on *how to monetize* (the jargon of this market *to generate money*) the use of videos. Nevertheless, the so-called youtubers (persons who are specialized in producing videos systematically for the platform) cannot explain precisely how a specific value is calculated for a certain number of views. Nor does the careful reading of company manuals clarify the situation[9]. In effect, it is explained that a user must (i) create her own channel, (ii) contract Google AdSense service, creating an account to receive money from monetization, (iii) choose the types of advertising she wants to display before, during and/or after the videos. As the channel reaches 10,000 views, Google will validate the page as eligible for earning money from advertisers[10]. When the channel accrues a number of views equal to $100 dollars in cash, Google will carry out the transfer.

Despite the explanations, one does not know the equivalence between the number of views and the amount of money to be received. How many views are needed to make $100 American dollars? The company makes clear that the relationship between the two terms is not precise. On the contrary, it depends

[8] For this reason, there is a global effort made by collective copyright management societies to enforce the legal acceptance that YouTube fits in public viewing category. Thus, the platform would be obliged to pay royalties to copyright holders. Cf. Francisco and Valente (2016).

[9] Available at: < https://bit.ly/2LtYzXE>. Access in: 2 jan. 2018.

[10] This rule was established in the second half of 2017 as a measure to avoid transferring money to political and religious extremists. Before that, every time the channel reached a thousand "monetizable" views (Cost per Mile, CPM) Google made the money transfer.

on the interaction of a variety of factors: the viewing time, average duration of visualization, audience retention, the type of advertising displayed, numbers of likes and dislikes, the money paid by advertisers at a given moment of a year, the geographic location in which the user is, among others[11]. But how are these variables combined?

The recommendation system used by the platform can help answer that question. In an article dedicated to the theme, YouTube engineers revealed a technique that seems to fit the video pricing as well (Davidson et al., 2010). The uses of the videos are divided into two broad classifications: explicit and implicit activities. Explicit activities include rating a video, liking it, or subscribing to a channel. Implicit activities refer to simple views of content. What the algorithms of the company do is to relate explicit and implicit activities of several users (and not just one of them) to predict what would be the related videos that a user would like to see next[12]. It should be noted that it is about establishing relationships between singular, standardized and independent uses among each other, regardless of the effective taste of the user.

Could such a technique also be applied to price the uses of videos? The analysis of the YouTube monetization explanations suggests a positive answer. For example, how do you measure an attribute as unusual as *audience retention*? YouTube explains that this index is built on: (i) the average view duration for all videos on a channel (absolute retention), (ii) the main videos or channels listed by display time, (iii) the audience retention data for a specific video in different periods and (iv) the relative audience retention of a video compared to YouTube's average for similar videos. From these criteria, the algorithms will gather data from other variables (such as display time or demographic data) to calculate a specific amount of money for the uses of a given video.

The description of audience retention is revealing. It should be noted that the factors from (i) to (iii) decompose the traditional notion of visualization: it is not only considered regarding the length of time a user accesses a video but becomes a composite of attributes of a video relative to the attributes of other videos on a channel. Factor (iv) enlarges this scale, linking the use of a video to several others in the platform. Considering that audience retention is only one of the factors that help to compose the value of a video, it becomes clear that the concept of view is a set of independent attributes taken from an underlying asset (the observation of a video itself), just as it is in the derivative market. Therefore, audience retention itself, one of the most important categories on the YouTube scale, does not explain in itself the value to receive for each view of a video. The main problem with this

[11] The explanation can be found at: <https://goo.gl/PtBPqw>. Accessed on: 19 dez. 2016.

[12] The determination of which are the related video makes use of a technique known as association rule mining and co-visitation counts (Davidson et al., 2010: 294). In this model, Google algorithms define similar videos departing from what content a user access, after seeing the initial video. For a set period of time, the algorithm counts how often each pair of videos was watched. Besides, there are other data to be computed in the system (such as metadata, designated by whoever inserts it into the system).

breakthrough technology is that there is no transparency in the process of generating monetary value for the videos since the YouTube algorithm formulas are trade secrets. As a result, the *effet de justesse* is completely missing, creating a problem of legitimacy for an algorithm-based market. So, an economic paradox emerges: at the moment when the algorithms should deliver perfect transparency of market relations, they provoke opacity and distrust in economic relations.

It should be noted that the problem does not end here. After all, if YouTube is really based on the social logic of derivatives, it is fundamental to understand who assumes the risks of such business. In the case of ordinary videos, even if made by youtubers, the risk is entirely borne by producers themselves, as such agents calculate that it is worth trying to make some money from a video that could be displayed without generating any monetary return initially. For copyrighted content, the scenario is different. Not only is there a legal issue regarding the payment of royalties for public performance, but also the risk that YouTube could be sued for secondary liability in copyright law violations. To reduce it, Google developed a content monitoring system to track the contents published by users on the platform, called Content ID. Despite the appearance of a problem-solving tool, this device opened another clash with copyright holders.

Who watches the watchdogs? Content ID and copyright holders as agents exposed to risk

Content ID is a monitoring system that follows videos uploaded on YouTube in search of copyright infringement[13]. Based on attributes that identify music and images (a melody, the tempo of a song, an image, among other technical details), provided by the copyright holders to YouTube database, the algorithms search all the videos embedded in the platform.

[13] About Content ID: <https://bit.ly/2ILN3sP>. Accessed on: 10 Dec. 2016.

The original idea was to prevent Google from being exposed to an even greater number of legal proceedings, reducing the confrontation with media corporations. Subsequently, the company created a secondary use: videos could be monetized, even without prior authorization of the copyright holders (thus avoiding legal proceedings for censorship against YouTube by third parties). This was possible because the system allows copyright holders to choose what should happen to videos published without their authorization. As explained by Francisco and Valente (2016: 320), the system can (i) deactivate the audio of copyrighted songs, keeping the images of the video; (ii) block the video completely; (iii) make the transfer of advertising money to the copyright

holder of the original song (not to whom uploaded the video); or (iv) track video view statistics.

Content ID was most welcomed by the music industry at first. Shortly after, however, a new exchange of accusations began concerning the accuracy of the device. According to Google, its Artificial Intelligence (AI) would detect about 98% of infringing uses. This number is contested by copyright holders, who continue to force the company to admit their direct interference in the monitoring of infringements.

This complaint is not undue. If the efficiency level proclaimed by the company is actually achieved, it would not only cease to be sued by music publishers and record companies but would become responsible for the application of copyright law. This would be something of great importance: it would remove all liability from the company, only leaving to the copyright owners the burden of registering the metadata correctly to the database. This means that the Content ID can act as a risk transfer clause for content producers in the digital music business, what is consistent with the logic of the derivative.

The problem becomes evident in borderline cases between infringement of the copyright law and legal exceptions and limitations. That is the case of the gamers, who have many of their videos suspended or presented without sound (music and the voices of the commentators) for using copyrighted material as soundtrack. These content producers contend that incidental music is not a relevant feature for their videos, so the use of copyrighted music is legally protected by the fair use doctrine. A similar claim is made by composers of derived works that use copyrighted material as a source. Take the case of the compositions called *mashups*. These are pieces of music produced by DJs who overlay two recordings of known songs, creating a new piece of music. In such cases, YouTube algorithms tend to direct the payment of royalties to the original copyright holders of the two original songs; not for the producers of the mashups. The algorithms do not recognize the third song as a new composition, but as an unauthorized use of two copyrighted materials. Although the original songs are indeed the initial attraction to listen to the hybrid music, is it not the third work that we want to be listening to? Should not their creators get a fair reward for their creation? While jurists discuss answers to such questions, the money generated by creative uses of musical works goes to cataloguers, not to those who generate new contents.

The point in the application of this monitoring system is that it transforms algorithms in automated judges which take decisions about what can and cannot be used in videos, at least in the first instance. Even considering that the user can submit a dispute request, it will be reviewed by Google's

lawyers, what takes time and money from anyone who wants to monetize their videos, as the monetization of the play is suspended until a final decision of the company or of the courts. It is also the algorithms that determine who should or should not receive the money for the use of copyrighted material. Once again, the opacity of the algorithms brings insecurity to the producers of digital contents, who become the agents most exposed to risks in this business model.

CONCLUDING REMARKS

Although each digital company has its own strategies to designate a monetary value for experiences that unfold in their platforms, the comparative analysis of sport betting sites, Facebook or, as done in this article, YouTube, confirms the systematic adoption of techniques developed in the derivative market. The case of YouTube is special because it presents another face of this phenomenon, which can be classified as the paradox of the algorithmic rationality: the inability of its algorithms to provide transparency and legitimacy to market players.

The value gap imbroglio plays an important role in a set of disputes that will determine a profound change in the ways of valuing phonographic products in the digital era, regardless of its outcome. As applied by YouTube, the derivative logic points to a promising yet controversial technical possibility. On the one hand, the multiplicity of uses of digital video requires a more dynamic and fragmented way of allocating monetary amounts to the uses of digital content. On the other, the central role the AI assume requires that traditional agents of the record market adapt to its operating ways, which means that the volatility and opacity of economic operations may become a constant feature. This may have several consequences for the shape of the digital music market.

As the designation of the money to be distributed by the algorithms is fundamentally volatile, it tends to affect the economic inequality among the agents of the record industry in a structural way. Making money requires considerable effort from content producers, but payment rules are not clear. This may even be surpassed by artists who easily get millions of views, but what about those who get only a few hundred? In addition, the volatility of wealth distribution makes it impossible to plan long-term production, which will affect the type of enterprise appropriate to the digital music economy. Finally, as the Content ID suggests, it is possible that copyright holders are more exposed to risks in this new business model.

It is evident, therefore, that the social logic of the derivative implies changes in the power relations of the record market. The traditional social order based on a stable relationship between dominant (major labels) and dominated (independent labels) players, which interact with each other due to the stability given by institutions (copyright laws, record stores, mass media) melts into cyberspace. A market controlled by algorithms makes impossible a stable market organization, pointing to some new social order that is not even outlined on the horizon yet. M

REFERENCES

ARNOLDI, J. Derivatives: virtual values and real risks. *Theory, Culture and Society*, London, v. 21, n. 6, p. 23-42, 2004. DOI: https://doi.org/10.1177/0263276404047414

ARVIDSSON, A. Facebook and finance: on the social logic of the derivative. *Theory, Culture and Society*, London, v. 33, n. 6, p. 3-23, 2016. DOI: https://doi.org/10.1177/0263276416658104

BECK, U. *La sociedad de riesgo:* hacia una nueva modernidad. Barcelona: Paidós, 2006.

BECKERT, J. What is sociological about economic sociology? Uncertainty and the embeddedness of economic action. *Theory and Society*, Dordrecht, v. 25, n. 6, p. 803-840, 1996. DOI: https://doi.org/10.1007/BF00159817

BEUNZA, D.; STARK, D. Tools of the trade: the socio-technology of arbitrage in a Wall Street trading room. *Industrial and corporate change*, Oxford, v. 13, n. 2, p. 369-401, 2004. DOI: https://doi.org/10.1093/icc/dth015

BOLTANSKI, L.; CHIAPELLO, È. *O novo espírito do capitalismo*. São Paulo: Martins Fontes, 2009.

BRYAN, D.; RAFFERTY, M. Financial derivatives as social policy beyond crisis. *Sociology*, London, v. 48, n. 5, p. 887-903, 2014. DOI: https://doi.org/10.1177/0038038514539061

BURGUESS, J.; GREEN, J. *YouTube e a revolução digital*. São Paulo: Aleph, 2009.

CARNEIRO, R. M. et al. A quarta dimensão: os derivativos em um capitalismo com dominância financeira. *Texto para Discussão*, Campinas, n. 199, p. 1-33, dez. 2011.

CUNNINGHAM, S.; CRAIG, D.; SILVER, J. YouTube, multichannel networks and the accelerated evolution of the new screen ecology.

Convergence, London, v. 22, n. 4, p. 376-391, 2016. DOI: https://doi.org/10.1177/1354856516641620

DAVIDSON, J. et al. The YouTube video recommendation system. *RecSys*, Barcelona, v. 26, n. 30, p. 293-296, 2010. DOI: https://doi.org/10.1145/1864708.1864770

DELEUZE, G. *Post-scriptum* sobre a sociedade de controle. In: DELEUZE, G. *Conversações*. 2. ed. São Paulo: Editora 34, 2010. p. 219-226.

DE MARCHI, L. *A destruição criadora da indústria fonográfica brasileira, 1999-2009:* dos discos físicos aos serviços digitais. Rio de Janeiro: Folio Digital, 2016.

DOUGLAS, M.; ISHERWOOD, B. *O mundo dos bens:* para uma antropologia do consumo. Rio de Janeiro: Ed. UFRJ, 2006.

FARHI, M. Derivativos financeiros: *hedge*, especulação e arbitragem. *Economia e Sociedade*, Campinas, v.13, p. 93-114, 1999.

FOUCAULT, M. *Nascimento da biopolítica*. São Paulo: Martins Fontes, 2008.

FRANCISCO, P. A., VALENTE, M. G. *Da rádio ao streaming:* ECAD, direito autoral e música no Brasil. Rio de Janeiro: Azougue Editorial, 2016.

HERSCHMANN, M. *Indústria da música em transição*. São Paulo: Estação das Letras e Cores, 2010.

INTERNATIONAL FEDERATION OF THE PHONOGRAPHIC INDUSTRY. *Global music report*: state of the industry overview 2016. London: IFPI, 2016.

_____. *Global music report:* state of the industry overview 2017. London: IFPI, 2017a.

_____. *Music consumer insight report September 2017*. London: IFPI, 2017b.

KISCHINHEVSKY, M.; VICENTE, E.; DE MARCHI, L. Em busca da música infinita: os serviços de *streaming* e os conflitos de interesse no mercado de conteúdos digitais. *Revista Fronteiras*, São Leopoldo, v. 17, n. 3, p. 302-311, 2015. DOI: https://doi.org/10.4013/fem.2015.173.04

KNORR-CETINA, K.; BRUEGGER, U. Markets as an object of attachment: exploring postsocial relations in financial markets. *Canadian Journal of Sociology*, Edmonton, v. 25, n. 2, p. 141-168, 2000. DOI: https://dx.doi.org/10.2307/3341821

KNUTH, D. E. *Algorithms in modern mathematics and computer science*. Relatório técnico. Stanford Department of Computer Science, Stanford, jan. 1980. p. 1-30.

KRIPPNER, G. The financialization of the American economy. *Socio-Economic Review*, Oxford, v. 3, n. 2, p. 173-208, 2005. DOI: https://doi.org/10.1093/SER/mwi008

LASH, S. A reflexividade e seus duplos: estrutura, estética, comunidade. In: BECK, U.; GIDDENS, A.; LASH, S. *Modernização reflexiva:* política, tradição e estética na ordem social moderna. São Paulo: UNESP, 1997. p. 135-206.

LAZZARATO, M.; NEGRI, A. *Trabalho imaterial:* formas de vida e produções de subjetividade. Rio de Janeiro: DP&A, 2001.

LÉPINAY, V. A. *Codes of finance:* engineering derivatives in a Global Bank. Princeton: Princeton University Press, 2011.

MACKENZIE, D.; MILLO, Y. Constructing a market, performing a theory: the historical sociology of a financial derivatives exchange. *American Journal of Sociology*, Chicago, v. 109, n. 1, p. 107-145, 2003. DOI: https://doi.org/10.1086/374404

MARTIN, R. After economy? Social logics of derivative. *Social Text*, Durham, v. 31, n. 1, p. 83-106, 2013. DOI: https://doi.org/10.1215/01642472-1958908

MUNIESA, F. *Des marchés comme algorithmes:* sociologie de la cotation électronique à la Bourse de Paris. 2003. 462 f. Tese (Doutorado em Socioeconomia da Inovação) – École Nationale Superieure des Mines de Paris, Paris, 2003.

NEYLAND, D. On organizing algorithms. *Theory, Culture and Society*, London, v. 13, n. 1, p. 119-132, 2015. DOI: https://doi.org/10.1177/0263276414530477

O'NEIL, C. *Weapons of math destruction*. Nova York: Crown, 2016.

PEOPLE, G. War of words: labels and trade groups target YouTube's "value gap". *Billboard Magazine*, 4 mar. 2016. Available at: <https://goo.gl/8PWpNt>. Access: 28 set. 2016.

RBB ECONOMICS. *Value of YouTube for the music industry*: Paper I, Cannibalisation. Technical report. London, may 2017.

SÁ, S. M. P.; HOLZBACH, A. D. #u2youtube e a performance mediada por computador. *Galáxia*, São Paulo, n. 20, p. 146-160, 2010.

SASSEN, S. *Expulsões:* brutalidade e complexidade na economia global. Rio de Janeiro: Paz e Terra, 2016.

VAN DER ZWAN, N. Making sense of financialization. *Socio-Economic Review*, Oxford, n. 12, p. 99-129, 2014. DOI: https://doi.org/10.1093/ser/mwt020

VAROUFAKIS, Y. *O minotauro global*: a verdadeira origem da crise financeira e o futuro da economia global. São Paulo: Autonomia Literária, 2016.

Article received on October 28th 2017 and approved on January 11th 2018.