PERSONAL DATA: WHOSE DATA IS IT? WHAT CAN WE DO ABOUT IT?(*)

FERNANDO HERRERA-GONZÁLEZ, PhD

NICHOLAS BLADES

January 2016

In a recent article[1], it was argued that most supposed free markets in Internet, are not actually free. They are considered "free" by most economists and authorities because they consist of economic transactions in which no money is involved. However, the concept of economic transaction goes beyond the interchange of money, even if it has to be acknowledged that monetary transactions has amounted for a very relevant part of economic transactions in the recent history.

It was also shown that those apparently free transactions involve the exchange of Internet services (understood in the widest sense) for time and personal data. In the case of time, there was little more to add: there is no discussion about the scarcity and value of time for people. The only remaining question, from a regulatory perspective, is how long will it take to authorities to accept that in this new world non-monetary transactions may become as important as monetary transactions, and therefore the same may happen to non-monetary markets, including markets in which time is exchanged for services.

However, data (specifically, personal data) presents some features that are not so easy to deal with. Firstly, who is the owner of the personal data? Or, in other words, what does it mean to "own" personal data? Secondly, has personal data any value? If so, what is its value? Some answers have to be provided for these questions before we accept that personal data is what is actually interchanged in those "free" markets whose gratuity we are questioning. And finally, are there any consequences for the individual? Should we take care when exchanging our personal data for services?

Before tackling these questions, it is useful to recall some insight about the concept and origin of property rights and the social function of this institution.

Origin and function of property rights

Individuals have needs, whose satisfaction requires resources to be obtained from the surrounding world. Some of these resources exist in abundance; such amount that every individual can satisfy his/her needs without limiting the satisfaction of the rest of the individuals: air, ice in the poles, or tree leaves in a forest are some examples. These resources pose no problems for interpersonal relations, they are not scarce. It is worthwhile to note that scarcity or not of a

[1]Herrera-González, F. (2015). Are "Free" Relevant Markets Actually Free? *CPI Antitrust Chronicle,* Nov 11.

(*)Authors want to thank Beatriz Sanz and Javier Domínguez all the valuable comments made to this paper.

concrete resource is dynamic and dependent of the needs of individuals in each context.

Arguably, most of the resources we require to satisfy our needs are limited. This implies a high likeliness of conflicts about their use, among those individuals willing to use them, and aware that NOT all needs will be satisfied with the available amount of the commodity.

"The fundamental social function of property rights is to prevent interpersonal conflict over scarce resources. Note that it is the scarcity of the good which makes property rights necessary to prevent the conflict. Conversely, with a dynamic perspective, the very possibility of conflict over a resource renders it scarce."

The institution of property rights provides an easy and possibly fair solution to this issue. It is obviously not the only possible solution, but it is the one that has historically proven to be more effective, as attested by its pervasiveness in the most developed societies.

In summary, the fundamental social function of property rights is to prevent interpersonal conflict over scarce resources. Note that it is the scarcity of the good which makes property rights necessary to prevent the conflict. Conversely, with a dynamic perspective, the very possibility of conflict over a resource renders it scarce. If suddenly two individuals want to carry away all the leaves from the forest, property rights will need to be defined for this up-to-this-moment abundant resource.

The way in which property rights operate is well known: they allocate exclusive ownership of resource to specified individuals (the owners). So, it is for the owner to decide what to do with the owned resource, and not for other individuals. In this way, conflicts are prevented *ex-ante*. Of course, we all know that property rights do not solve all possible interpersonal problems: there still may be conflicts about who is the owner or about the bounds of the property rights, among many other issues.

To be effective, property rights should be discernible and fair[2]. Obviously, property borders have to be objective (interpersonally ascertainable) and unambiguous so that other individuals may avoid using the goods owned by others. On the other hand, property rights have to be seen as fair by those affected; otherwise, they would resort to the use of force or other means and no property rights could be said to exist.

Property rights evolve with the needs of individuals and society, they are not static. One clear example of such an evolution is the definition of property rights for spectrum frequencies. Up to the early 20th century, frequencies were not scarce and thus no property rights were necessary for this "asset". This changed with the evolution of technology, which made possible the use of radio waves for the transmission of information. A whole new kind of property rights has to be

[2] Hoppe (1989)

defined, based on frequency, bandwidth and transmission power[3].

Of course, one last question remains to be addressed: how property rights are acquired. On reflection, this question boils down to how property rights are *originally* acquired, but this is not the place to discuss the issue. For our purpose it seems enough to reckon that it is generally accepted that the property of a good produced from other goods belongs to the owner of the later goods.

With these ideas in mind, let us turn back to the issue at hand, namely the property rights of personal data.

The nature of personal data

This Brief focuses on personal data. As has been said, the issue is of relevance because it seems that a growing number of Internet services are exchanged for personal data of the user. These personal data refers to personal pieces of information, such as name, address or age, and to information about our behaviour, in the Internet (websites visited, contacts, purchases, location...), but also in other areas of our life (by means, for example, of the Internet of Things).

Initially, these data are in an abstract form that makes impossible its exploitation for anyone. In order for the data to be useful, they must be materialized somehow in a concrete physical medium, be it a piece of paper or a database in a disk. Materialization of personal data requires the investment of resources, such as those just quoted, together with labour and time of the individuals that enter and maintain the data.

Once the data is in a physical form and can thus be exploited, it may be valued. The mere fact that someone is willing to invest resources in materializing personal data is a proof that personal data may have some value.

Now that the data is in a physical form and may have value, the question of the ownership of data becomes relevant. While the data was abstract, no property rights seemed necessary: It is clear that the use of my name, my age or the colour of my eyes by me does not exclude using those same name, age or colour by other people. The same could be said of the list of websites visited in my last connection, or of the places I went with my car last week. Nobody says he owns his name, the colour of his eyes or the itinerary of his last trip. Personal data in its abstract form is not owned in the traditional sense of the term, because as it cannot be used, has no value and there is no need of property rights for it.

"Personal data are in an abstract form that makes impossible its exploitation for anyone. In order for the data to be useful, they must be materialized somehow in a concrete physical medium, be it a piece of paper or a database in a disk."

However, things change once the personal data is materialized in a physical medium. Imagine that I write down on a piece of paper some personal attributes as those mentioned above. This piece of

[3] See Coase (1959)

paper with the data may have value for some individuals (possibly due to the data on it). As this piece of paper is a concrete good, scarce, and with clear bounds, conflicts may arise if several individuals want to use it. In consequence, to prevent this conflict, property rights should be defined for it, as is the case with every scarce resource.

In the example above, there should be no doubt that it is the owner of the piece of paper in which the data has been written who owns the asset (doubts could arise if it is another individual the one who writes the data on the paper). But it is clear that the individual whose data is written has in principle no claim on the property of the materialized data: he did not provide any of the scarce resources for the venture. He may just have told the data to the individual who was writing it down, but this did not restrict him of "using" it in the future, for example, to provide it to other parties.

Summing up:

1) No property rights are required for personal data in abstract form

2) In order to obtain value, personal data has to materialize in a concrete physical medium.

3) The owner of the materialized personal data is logically the owner of the resources invested in its materialization (i.e., the medium in which the data has been stored).

4) In principle, the owner of the data could do whatever he wants with his asset, included its possible interchange for other resources.

It is important to note that it has not much sense to say that personal data belongs to the individual that generates it. As has been said, while personal data remains in abstract form, no property rights are necessary as the data cannot be used by anyone.

As a result of the foregoing, it is clear that we do not "own" our personal data, as counter-intuitive as it may sound. It is just not possible to own it while it is an abstract form. However, we have power over it: we can choose to whom to reveal it and whom not, and what kind of data or the level of detail to provide to the interested parties. This is a service that we can uniquely provide to the interested parties. In any case, we should be conscious that once the service is provided, the data is given away and its ownership rests not with us, but of the owner of the medium in which it is kept, the same way we own the files in our hard disk.

The value of personal data

Now that it has been established that personal data may actually be owned, it is time to wonder about its possible value.

There is little doubt about the value that information may have. An indirect proof of this is the huge amount of resources that both enterprises and government are (and have always been) ready to invest in its gathering. As they say, "information is power".

So far, we have just referred to personal data, not to information. For personal data to become useful (and potentially valuable) information, it is

necessary to somehow structure it. Raw data per se has limited value; in fact, the value of personal data will derive from the information that can be deduced from it, whose value in turn will depend on its utility for someone[4].

This need of storage and process capacity in order for raw data to become information explains the sudden importance of the issue. With the generalization of the Internet, the capacity to obtain personal data has multiplied several fold. As this has been accompanied by vast increases in storage and process capacity, it is now possible to obtain information from this raw data at a relatively low cost. Because of the complementarity of processing power and storage capacity, pieces of data that are useless by themselves may become valuable information.

"We do not "own" our personal data, as counter-intuitive as it may sound. However, we have power over it: we can choose to whom to reveal it and whom not, and what kind of data or the level of detail to provide to the interested parties."

One question that cannot be answered by economic theory is what will be the value of this information (which, in turn, defines the value of personal data, and the amount of services Internet provides will be able to give in exchange for it). This is something that can only be ascertained by the effort of entrepreneurs, and will require from them the deployment of plenty of resources. It could also be the case, as with any other entrepreneurial activity, that the obtained information had no value or a value below the resources invested for acquiring it. Only time will tell, but the truth is that there is currently a lot of excitement about the business prospects for so-called Big Data.

Notwithstanding the above, it seems safe enough to identify the following three features as drivers of the value of data:

1) *Amount and scope of data*: the more data available, the more potential information to be extracted from it. So, if the information has value, the more data in the database, the more valuable it will be. This works both in the "vertical" dimension (i.e., number of registers in the database) and in a "horizontal" one (i.e., number of items per register).

2) *Technology for processing the data*: the more effective and sophisticated, the richer and quicker will be the extraction of information. Besides, more types of information will be available from the same set of data, making the database more valuable. The development of algorithms and methods to search and mine the data is one of the area where the entrepreneur can adds more value to the raw inputs

[4] For a more detailed account of how the value of resources is derived from the value of end goods, see section *"Value allocation in the unregulated chain value"* in Telefónica's Regulatory Economics Brief 14-01, available in http://www.telefonica.com/en/about_telefonica/pdf/Regulatory_Economics_brief_1.pdf For a more rigorous description, see Böhm-Bawerk (1891).

3) *Possibilities of use of the data*: The accrual of massive data sets without the ability to search them results in no value creation to the data owner. The more possible uses for the acquired data, more types of information will be possible, and the more value the database will accrue.

This last feature requires a more detailed explanation. If it is accepted that data belongs to the owner of the media in which the data is stored, he should, in principle, be able to do whatever he wants with the data, which is his property. The only limitations could come from the conditions in which the user has agreed to provide the data, in the same way that any other interchange or contract. As will be seen in the next section, these conditions may play an essential role in the acquisition of personal data.

The terms for exchange of personal data

Personal data can be classified into two broad categories according to the way in which they are acquired by a firm.

Implicit data – data generated when using the products and services of a firm

These are the data generated by the user in the process of using web services and required for these to work. Each activity in the value chain generates and possibly stores its specific information, so this implicit data can in turn be classified in app-generated, network -generated, browser-generated, search engine-generated, service-generated, Internet traffic-generated, operative system-generated and so on data can be classified into two broad categories according to the way in which they are acquired by a firm.

Note that as the Internet of Things evolves, the sources of implicit data will multiply and become more near the physical world, making people more aware and concerned about the nature of the data required by these applications.

It seems hard to argue that this category of data should be restricted in its uses by the owner. It is data generated because it is needed to serve the client, which the client need not be aware of, and that is completely generated with the resources of the service provider. If I counted the number of cars travelling the street under my window per hour and annotated it in a paper, nobody would discuss the property of such numbers, and no limits to the use of it by me would be justified. The same appears to happen with what we are calling implicit data.

"Service providers must strive for equilibrium between the unlimited use of the data they acquire and the conditions required by users to supply the data."

The storage and exploitation of implicit data is hardly a new phenomenon. Firms of all industries and economic sectors routinely store data related to its operations in order to increase their efficiency. Up to the present, such activities have not raised any concern for either individuals or public authorities. In fact, the use of such data is clearly welfare enhancing, as the gains in efficiency thus attained are passed on to the market in the form of cheaper products or better services.

Explicit data – information exchanged when joining a service, as a means of payment

"Explicit" data are consciously and expressly provided by the user. For example, when he/she fills out a form, answers a survey, posts an opinion on a hotel, publishes a picture in a website or sends a message to a contact.

When we reveal personal data and allow a service provider to record it as the price of joining a service, no restriction is generally placed on any future provision of these data by us to other parties. We remain free to use such personal data ourselves, as a means of future exchange. Future exchanges with this counterparty may require different data, as the first data set is now known to them and likely worthless as a means of exchange between us. However, providing our personal data is still valuable for other counterparties.

The firm could, in theory, prohibit the consumer from ever exchanging that personal data again, but to enforce such a term it would have to be particularly prominent at the time of exchange. Consumers would understand the impossibility of future transactions on the Internet or the real world, if they could not reveal their name, address, mobile number etc. so it is unlikely that such a service would ever be popular, and/or able to enforce its terms.

At the other end of the spectrum, policymakers have determined a minimum set of restrictions on the use of personal data, something that, as will be seen below, likely reduces the value of that data to the firm, but arguably provides some minimum protection to the customer regarding its future use.

As consumers we may wish to place incremental restrictions on the use of our data when we provide it. For example, if we wish to use a messaging platform we might want to ensure that the content of the messages we send and receive are not read by the platform provider. This is something that is common to all communications systems operated by telcos, for example. Other messaging platforms might provide a cheaper (in money terms) service, but retain the right to read the messages sent by you and received by you[5].

Service providers must strive for equilibrium between the unlimited use of the data they acquire (which would maximize its value) and the conditions required by users to supply the data (which might limit the amount of data on which value can be added, or the level of value add achieved). If Services Providers don't want to compromise in the use of the data, fewer users will provide them personal data; but if they limit themselves too much, it may be the case that there is both less data on which to add value and the value added is limited by the scarcity of the data.

Of course, the point of equilibrium is quite difficult to achieve, if at all, and possibly will change with time and preferences of the users. What is clear is that market forces should have a large part to play in determining the equilibrium.

[5] For example, Google analyzes the content of the emails in Gmail for several purposes. See https://www.google.com/intl/en/policies/terms/

Summary and conclusions

Personal data can be owned but perhaps counter-intuitively generally not by the person it concerns. For a property right to be created these data must materialize into some physical medium. None of us can own his/her personal data in the conventional sense, but what we have is control over the terms under which it is exchanged.

As we have control over personal data exchange, this service may be exchanged for other commodities, including web services (as seems to happen in such apparent "free" markets) – used as a form of remuneration[6]. Once the data has been provided and recorded they may be measured and accounted, so it could be used as unit for measuring for regulatory purposes these relevant "free" markets, in the same way it was proposed for time in Herrera-González (2015)[7], and instead of money, absent in those economic transactions.

"The increased use and valuation of personal data place individuals in a very good position, as each of us is the only possible originator for our personal data."

[6] This possibility has already been recognized by the Courts. See ECJ, 11 September 2014, C-291/13, accepting that remuneration must not necessarily be provided by the recipient of the service himself but can also be provided by "income generated by advertisement posted on a website".

[7] Herrera-González, F. (2015). Are "Free" Relevant Markets Actually Free? *CPI Antitrust Chronicle,* November 11.

The value of personal data, and thus the amount of services we will be able to receive on exchange of it, depends on the possibilities to exploit it. For what we have called implicit data, this will largely depend on regulation, because these data is generated by the providers while serving the customer and for this purpose. For the customer, the option is between using or not the service; but, if he chooses to use it, he has no claim to the implicit data.

However, it is clear that if service providers find additional value for this data, it will sooner or later revert on better service for the end users through the competitive process, if this can run freely. So, it is very likely that regulation on the use of implicit data, as it would limit the possibilities of exploitation of the data and thus its eventual re-valuation, would run against customer welfare.

Regarding explicit data, the situation is simpler. In this case, it is the customer who has to expressly provide the data. This will only be done if the conditions in which the data is provided are attractive enough from his/her perspective. These conditions have to be attractive enough in terms of number of fields required or in the easiness to enter (or delete) the data, but also in the behaviour of the service provider with the acquired data: ways in which is exploited, transparency of the use, in sum the degree of privacy.

This situation opens up several parameters of competition vis-à-vis the customer. And it is through the entrepreneurial activity of trial-and-error that the optimum equilibrium between data value and its privacy can be reached in each moment. *Per se*, this competitive dynamic arguably make unnecessary any regulatory intervention to assure the privacy of data. Those

providers that are not able to offer the privacy required by the end users will be expelled from the market.

The regulation of privacy imposes limitations on the trial-and-error process necessary to attain the correct equilibrium in each moment, likely reducing the value of personal data (for example, by banning concrete uses not actually relevant for the user). Obviously, a reduction in the value of the personal data would run against those who can provide it, i.e., the individuals, who would get less web services and contents in exchange of it than otherwise.

Summing up, the increased use and valuation of personal data place individuals in a very good position, as each of us is the only possible originator for our personal data. We should also be conscious that when we choose to provide these data, the recipient is acquiring an asset which could become very profitable. So, it is for us to obtain in turn the most profitable conditions for this exchange, not only in terms of web services, but also on privacy conditions.

To date the value of Internet companies has focussed on the amount of data they own (information about users – using subscriber numbers as the proxy) and the quality of the algorithms they put to work on that data. A fuller understanding of property rights also shows that, properly informed, consumers might be willing to originate more data (and hence more value) to firms, if firms were willing to offer higher levels of privacy and digital confidence in exchange.

References

Böhm-Bawerk, E. von (1891). The positive theory of capital. Londres: McMillan & Co (Translated into English by W.Smart, 1891)

Coase, R.H. (1959). The Federal Communication Commission. *Journal of Law and Economics* Vol. 2 (Oct., 1959), pp. 1-40.

Hoppe, H.H. (1989). A Theory of Socialism and Capitalism, p. 138.

Hume, D. (1751). An Inquiry Concerning the Principles of Morals: With a Supplement: A Dialogue. New York: Liberal Arts Press.

Regulatory Services

Telefónica S.A.

Edificio Oeste, 2- Plta. 1ª E28050 Madrid- Spain

For further information, please contact

Fernando.herreragonzalez@telefonica.com