Open Data, Big Data and Crowdsourcing: Emergent Mobile Apps Business Models

Samy GUESMI
IDRAC Business School, Paris

Abstract: The aim of this paper is to discuss the business model of a free social GPS mobile app for public transportation. The app functioning relies on open data and crowdsourcing. The crossing of transporter data (open data) and service users' data (crowdsourcing) feed the app's algorithm, which in the end enhance the public transportation experience. The hinge and tensions between for profit and not for profit sides of the data stemming from open data and crowdsourcing are questioned in the implementation of the company business model as well as the role and function of the business model in the start-up strategy.

Key words: open data, crowdsourcing, big data, business model, n-sided markets.

We have arrived at a turning point in our IT and industrial history: data is potentially the new oil of our digital economies. In 2009, 20 years after inventing the World Wide Web, Tim Berners Lee called on governments to release all forms of raw data to foster innovation, transparency and economic growth [1] (BERNERS LEE, 2012).

The digitisation of the economy and the rise in information and communication technologies, coupled with the development of network infrastructures [2], have all made us produce ever more data. We produce data when travelling (travel card, road traffic flow, …), when shopping (payments by credit card, online purchases, ...), when surfing the internet (cookies, clickstream data) and when manipulating the intelligent objects

[1] This call for raw data initially took place in a Ted Talk: http://www.ted.com/talks/tim_berners_lee_on_the_next_web#t-586111 and was commented later: "Opening up data is fundamentally about more efficient use of resources and improving service delivery for citizens. The effects of this are far reaching: innovation, transparency, accountability, better governance and economic growth".
http://www.wired.co.uk/news/archive/2012-11/09/raw-data

[2] A development which can be seen in the ongoing growth in demand for computer equipment, broadband connections and mobile equipment.

which have entered our daily lives (smartphones and their applications, but also contactless technologies [3]). All these objects produce data, leaving traces of our actions, indications of our behaviour, our habits, etc. to such a point that the internet has become an archive of a virtually exhaustive imprint of human activity (COLIN & VERDIER, 2012).

It was with the development of web 2.0 [4] that several technological trends and movements appeared: open data, big data and crowdsourcing, along with the emergence of various issues associated with the production and exploitation of data in the process of value creation. Open data covers several definitions, and constitutes both a philosophy of data access and a publication method for use and reuse of data by everyone, while big data encompasses a paradigm of commercial exploitation of data as a new resource for organisations. The two can be achieved and/or complement each other with a form of data production: crowdsourcing. This represents the changes associated with the wider participation of internet users in the creation of web 2.0 content.

These concepts are all associated with new technologies and the production of data, information and knowledge, which result from their use. Given that data liberation (open data), exploitation of large bodies of data (big data) and the numerous products resulting from crowdsourcing processes all aim to create value, the concept of the business model (BM) seems relevant to each. The BM frameworks (OSTERWALDER & PIGNEUR, 2010; DEMIL *et al.*, 2013) provide representations of the process of value creation in a systemic way and explain the creation, capture and sharing of value.

This backdrop, and the fact that the interactions described above occur increasingly nowadays via smartphones – these objects which have become central to our daily lives – raises the question of the place and status of data in the development of applications and value creation. How, in the commercial paradigm of big data, can an application combine data resulting from open data and crowdsourcing to formulate a value proposition which is satisfactory for the stakeholders, through the development of a sustainable business model?

---

[3] RFID (Radio Frequency Identification) chips or NFC (Near Field Communication) devices. Referred to as the "internet of things".

[4] Web 2.0 is characterised by a greater participation of internet users in the creation of content, and aspires to be more participative.

The present article attempts to address this question using a case study of an application [5], by illustrating the construction of its BM and detailing the different types of data it uses in the process of value creation and value proposition. We will show that the notion of a BM has evolved from being a static to a dynamic concept and could lead to a distinction between the theoretical BM and the effective one when associated with the operation of start-ups, and in particular those which develop their business in a multi-sided market. The theoretical BM would consist of an ideal multi-period and sequential model describing the path of the start-up from where it begins to where it heads up. Whereas, the effective BM would be an instantiation of the theoretical one at any given time and could diverge from it according to the changes in strategic choices in response to a rapidly changing environment. The BM can also be a useful tool for public authorities and policy makers in understanding the stakes involved in releasing data to promote open innovation.

■ Web 2.0 and the emergence of data

According to O'REILLY (2007), the emergence of web 2.0 meets the requirements of new business models, which have developed to support the recovery of e-commerce after the internet bubble burst. This new web is characterised by a wider participation of internet users. With the advent of web 2.0, internet portals and aggregators have lost their importance and their previous value has migrated from user traffic to the layer of databases and user interfaces (O'REILLY, 2007; GEHL, 2010).

These changes, reflected in the business models of the web 2.0 economy, have been widely studied, exploring both the role of the internet and the place it has given to interactivity [6] in the creation of value (MÖLLER & RAJALA, 2007) and analysing the different ways to capture such value (BEUSCART & MELLET, 2008; CLEMONS, 2009; WIRTZ *et al*., 2010). Web 2.0 revenue is mainly generated from the "work" of internet users. Companies involved take advantage of data and content produced in this

[5] The company wishes to remain anonymous. Throughout the article, they will be named GPS App.

[6] This characteristic provides organisations with two essential elements: access to the creative capacity of internet users, and the advantage in terms of cost, simplicity and rapidity which this method of creation covers in comparison to other methods (internally or by traditional outsourcing).

way as it makes their services more attractive. The data increase the usefulness of their service for other users and represents a competitive advantage for them through the database developed in parallel.

Literature on the subject of BMs developed from the mid-1990s with the explosion of start-ups and the new economy. The managerial interest in these tools, first practical then theoretical, has proved consistent ever since. Several works have attempted to summarise the contributions of this concept in the various sectors exploiting them, such as e-commerce, innovation management and strategy (ZOTT *et al*., 2011). While these works have not resulted in a satisfactory consensus definition relevant to all areas of study, they nevertheless inspired an interesting emergence of competing conceptual frameworks such as the VISOR [7] model proposed by EL SAWY & PEREIRA (2010), the Business Model Canvas proposed by OSTERWALDER & PIGNEUR (2010), and the RCOV [8] model developed by LECOQ *et al*. (2007).

In the following developments, we define the business model as an explanation of the way in which an organisation creates and captures value. Creation is generated via one or several value propositions, which the company presents to their clients, and the capture and distribution of this value are undertaken in such a way as to ensure value production is economically sustainable.

■ Data: a new source of value in the post-industrial era

It is exactly because it works as a platform economy that the web 2.0 is the source of the informational paradigm change (REBILLARD, 2007) which has inspired the emergence of big data and open data. It takes advantage of users' "work" and exploits collective intelligence through the use of crowdsourcing (LEBRATY & LEBRATY-LOBRE, 2013).

Open data represents both a political stance of transparency and empowerment of citizens *vis-à-vis* public data (BOUSTANY, 2013), a philosophy of access to information which considers information as a shared asset, similar to other ‘open' movements (source, standard, innovation, …)

[7] Value proposition, Interface, Service platforms, Organising model, Revenue / cost sharing.

[8] Resource and Capabilities, Organisation and Value.

and a practice of publishing all types of data, which promotes economic and social innovation (LACOMBE *et al.*, 2011). Open data [9] respects the following main characteristics: data is free, complete, interoperable, non-proprietary and copyright-free. Data can be public or private (i.e. belonging to local authorities, States, public service contractors and also businesses). Some data do not fulfil all the conditions to be considered as strictly open, but rather constitute hybrid forms which can be placed on a continuum between closed proprietary data and open public data. Outlining four criteria, Mc Kinsey categorises data along a continuum of "liquidity" of data, ranging from the most open to the most closed, according to whether they are accessible (to everyone or only to certain people), machine-readable (either through an automated process or after reprocessing), reusable (freely or not at all reusable) and according to whether access is free or paid for (DOBBS *et al.*, 2013). For public government data alone, the Sunlight Foundation defines 10 criteria which characterise open government data [10].

In parallel, big data raises questions about ways in which to develop and use the large bodies of data in a way that increasingly integrates them into economic value creation. The managerial leaning of the concept accounts for the commercial vision associated with big data, seen as the main by-product of economic activities, but also as a way to guide and improve the main activities from which the data are extracted (McAFEE & BRYNJOLFSSON, 2012). Organisations are encouraged to exploit the hidden value of their informational assets. This approach also raises, although from a different angle, the essential question of opening and sharing data through the filter of privacy (RALLET & ROCHELANDET, 2011; REY, 2012) by questioning the production of such data by users, and users' control over reuse of this data.

Finally, crowdsourcing is one of the new means of cooperation and production which emerged with the facility offered to internet users to participate in the creation of web content and its categorisation, development, scoring and sharing, etc. The web 2.0 lowered coordination costs and enabled new forms of organization (SHIRKY, 2008) as well as stigmergic collaboration processes (e.g. the ant-like work processes of

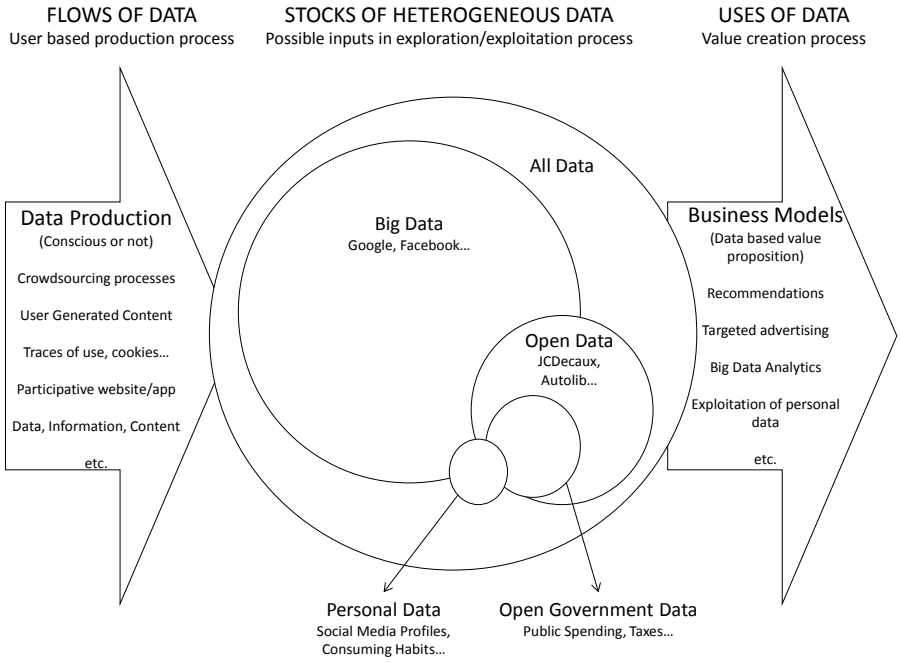
---

[9] Open Knowledge Community defines: "A piece of data or content is open if anyone is free to use, re-use, and redistribute it – subject only, at most, to the requirement to attribute and/or share-alike.". For a more detailed definition refer to: http://opendefinition.org/od/

[10] Completeness, primacy, timeliness, ease of physical and electronic access, machine readability, non-discrimination, use of commonly owned standards, licensing, permanence and usage costs. http://sunlightfoundation.com/policy/documents/ten-open-data-principles/

Wikipedia) as described by ELLIOTT (2007). We can define crowdsourcing as the fact of an organisation externalising a problem or activity to a large number of anonymous individuals, with the idea of finding among them people who are capable of solving the problem or successfully carrying out the task in question either individually or collectively (HOWE, 2006). This usually relies on a certain sense of community spirit while individuals from the crowd consciously participate in this process for free.

Figure 1 - Relationships between production methods, types, and methods of data



Source: Adapted from DOBBS et al. *(2013)*

We can define two broad kinds of crowdsourcing (AFUAH & TUCCI, 2012) according to the objectives of the organisations which are using it: integrative and selective. The first consists of entrusting the crowd with several simple tasks which do not require any special individual abilities. The organisations which use this kind of crowdsourcing are looking to avoid the cost they would have to absorb if these tasks were carried out internally or by traditional subcontractors, or they simply crowdsource tasks which would be otherwise unachievable (large scale collaborative process without having to pay for each and every contribution). The second kind invites the crowd to produce something, and then select the best propositions put forward. These

are complex and/or creative tasks, and the advantage of turning to crowdsourcing lies in the originality of competing choices.

Whatever their method for collecting and processing data, it now represents an essential value: data is central to the BMs of this platform economy which creates databases as strategic assets. In this new informational paradigm, enhancement of heterogeneous data can take many forms, particularly cross-tabulating them. Figure 1 illustrates types of data (digital and analog) and their associated means of production (digital), along with how they are valued by the BMs (digital) which use them.

■ Typology of data and the impact of their status on the value-creation method used

The first body of usable data is comprised of open data, public or private, destined for use freely by all.

The other body of usable data is data from users, produced by crowdsourcing. Crowdsourcing can take several forms, pursuing various goals and using various tools and strategies. There is a broad range of externalised network relationships (LEBRATY & LOBRE, 2010) which have different implications for commercial and non-commercial connections (e.g. relationships within Wikipédia [11], MechanicalTurk [12], InnoCentive [13]). The user may, consciously or unconsciously, participate in the process of crowdsourcing (when conscious, this generally comprises a community of willing users). They can do so for free or against payment. The user's contributions can be limited to browser history or internet use [14], or be the result of a real effort, a creation. Such a wide range of configurations raises questions concerning the commercial or non-commercial nature of the

[11] http://www.wikipedia.org/ is an open and collaborative encyclopaedia.

[12] http://www.mturk.com/mturk/welcome is a service proposed by Amazon which enables the crowdsourcing of low level, underpaid micro tasks which require cognitive capabilities, making automatic functioning impossible.

[13] http://www.innocentive.com offers companies the possibility to crowdsource R&D tasks by placing participants in competition against one another with a high prize money incentive.

[14] These traces of use, along with metadata and geolocated data, do not carry any detectable intelligence when considered separately, but their aggregation and the application of algorithms enable extraction of profiles and consitution of databases.

service, its role in value creation and the place of the consumer's "work" within these free services (DUJARIER, 2008; KLEMMANN *et al*., 2008).

■ N-sided markets theory

A multi-sided market can be defined by its functioning as a platform addressing multiple interdependent sides. Thus it articulates several value propositions in order to initiate and attract the participation of market players on all sides of the market. The participation of one side increases the value of participation for one or several other sides. We can see here that the theory of two-sided markets (ROCHET & TIROLE, 2003) and multi-sided platforms is closely linked to the theories of network externalities and of composite prices. In network economics, externality designates a situation in which the usefulness of a consumer depends on the consumption of the same product or service by other consumers (KATZ & SHAPIRO, 1985).

Heterogeneous data (public and open on the one hand, closed and proprietary on the other) can be aggregated in various ways to form the strategic asset of a technical platform and enable them to address targeted value propositions to each of the market players on the different sides of the market. The economic theory of multi-sided markets enables the development of web 2.0 platform business models (PARENT & CHANAL, 2009).

In order for private actors to invest in the production of an innovative BM, merely having access to open data would not be sufficient, as this would not enable them to make their investments profitable. To acquire this capacity, they must be able to establish competitive advantages. However, if nothing differentiates the data freely available to them from those which any other competitors may access, they will not invest. Consequently, there are two possibilities: either the data is not strictly open (according to the open definition) and the privilege of differentiated access provides a competitive advantage, or data is open, but private actors combine it with closed proprietary data to develop their database, the strategic asset on which their competitive advantage is based.

Below is a proposed table of data categorisation which enables an identification of issues associated with the organisation of heterogeneous data:

Table 1 - Typology of data

Type of data	*Big data*	*Private open data*	*Public open data*	*Crowdsourcing data*
Producer / Transmitter	Organisations which produce large bodies of data (Google, Facebook)	Private services (JCDecaux with Vélib' city rental bikes, Amazon)	States/Local authorities (Rennes, Etalab)	Users (InnoCentive, MecanichalTurk…)
Is the user aware of the data production process (use of an artefact)?	No The majority of bodies are developed through the use of the service in a seamless way	Yes Developed through the use of an artefact or the acceptance of a transaction	No The data already exists (raw) via use of a public service.	Yes Differing degrees according to the system of user involvement in the process
Issues associated with Cost / Value	Cost: variable and dependent on the activity (by-product or primary output). Value: strategic asset upon which the competitive advantage or non-monetary uses are based	Retrieve value upon data release and propose data liberation as an offer of additional services	Collection, processing, provision supported by public actors. Betting on open innovations, empowerment and transparency	Diversified typology of cross-tabulating between sourced data, implication of users and user remuneration
Questions raised	Privacy: Control over personal data, Third-party access to data	Free or paying data access	Balancing the budget, Between open innovation value creation and value capture	Exploitation of users: Ethics of value creation based on User Generated Content
Status of data within the production process	Product of the main activity or by-product of another activity	By-product of a main activity	Product of the main activity	Paid: product of the main activity Free: by-product of an activity, low involvement
Native format of data	Digital	Analog or digital according to the type of data	Analog (except for e-administration initiatives)	Digital

Source: The author

■ A case study of GPS App: a social GPS application for public transportation

The stance adopted in this research falls in line with the interpretive epistemology often employed in information systems research (WALSHAM, 1993). Access to data has led us to adopt a qualitative methodology and carry out a case study (YIN, 2014) on the GPS App application. Sources of

cross-tabulated data resulting from 15 semi-structured interviews conducted with an operating employee of the start-up at regular intervals during a period from end of 2012 to the beginning of 2014, of primary and secondary source documents. A thematic content analysis was then carried out on the interview scripts, using a conceptual coding based on the specificity of multi-sided markets. The three main themes are short and medium-term strategy, relationships with users and relationships with transport operators.

The GPS App is a social and collaborative mobile GPS application for public transportation which aims to improve the experience of public transport users. At the end of the year 2013, GPS App managed to raise several tens of millions of dollars from one of the largest venture capital companies in the world [15]. Around the same time, the application counted 3 million users across more than 100 towns and cities in 25 countries. These numbers rose rapidly, to 400 towns and cities in 40 countries totalling 9 million users by September 2014 [16]. As of writing this paper, GPS App still doesn't monetise its multi-sided market, hence it doesn't generate revenues. The theoretical business model has yet to be fully completed, modified or improved.

GPS App operates using two data sources: transporter data and user data. Each of these sources can provide two types of data. Transporters provide static data resulting from GTFS [17] files which contain static information (location of stops, network line routes, timetables) and real-time data transmitted via APIs [18] which enable the diffusion of information in real-time about the situation of the different means of transport (trains, metro, bus, etc.) as well as any network incidents. On their side, users supply the platform with passive reports collected by geolocation (location within the network according to the route taken) as soon as the application is activated and they have authorised transmission, but also with active reports when users provide information about their location to share with other users

15 This company holds stakes in the capitals of the big names of the digital economy.

16 The pace of user acquisition, retention rate and critical mass constitute performance indicators for audience models often associated with free applications and web services, and compete with traditional notions of turnover, revenues and profitability in the valuation of start-ups.

17 General Transit Feed Specification. This is a file format combining information concerning public transport timetables and associated geographical information.

18 Application Programming Interface. These are interfaces which are offered by internet services in order for the user to be equipped and develop additional services. Data transfer is the result of requests submitted to the API.

(density of transport users and cleanliness of the transport taken, onboard available equipment, network incidents, driver's conduct). These various types of data, crossed one upon the other, supply the algorithm at the heart of the service, which improves over time and with the accumulation of data. This data enables them to make recommendations to users for the most efficient multimode routes possible, to alert them about network incidents in real-time and also to provide transporters with a summary feedback of network flow and users' experiences.

The service GPS App provides to users relies on different types of data: open, closed, public and private. These come both from public or private open data with the opening up of access to transporter data and crowdsourcing, i.e. user contributions. This database can be seen as their shared asset: it is in their interest to contribute by providing indications about traffic density and any incidents to provide all the users with a fine tuned transportation networks situation in real-time. To coordinate the different types of data into a business model, which satisfies all actors present and enables the start-up to potentially generate a revenue, GPS App operates like a multi-sided market making distinct value propositions to the various contributors who make up their marketplace.

Regarding the functioning of the app based on its theoretical business model, in the towns and cities where it is implemented, the application is initially only targeted at two contributors: "transporters" and "users". To transporters, it offers summary information, based on the data generated by users, of the state of the network, its weak points and recurring problems. For them, the incentive is the perspective of an increase in the number of network transport users and their satisfaction once the application is implemented in their city. To the transport users, it offers a tool for improving the experience on public transport and a way of regaining control over such experience by enabling them to prepare and coordinate their journeys independently of the transporter. For them, the incentive lies in saving time on their daily itinerary.

In contrast with traditional two-sided markets such as free press, neither of these two primary sides generate revenues. They must therefore be subsidised until a critical mass is reached [19] which enables the service to

[19] In the context of GPS App, the critical mass represents the number of users from which the production of crowdsourced data is sufficient to connect up the network and produce real-time crowdsourced data. It is a threshold from which absence of access to transporter data is less harmful to the attractiveness of the application and relevance of its results. Evidently, the critical

become profitable once a third advertising side joins, which would form the model of advertising revenue or from offers of service and expertise addressed to the transporter side once the application is able to produce enough crowdsourced network data. Depending on the outcomes of negotiations and partnerships with the transporters, GPS App could also establish other revenues by offering complimentary assets and expertise concerning management of mobile payment infrastructures (m-ticketing). The functioning of the application, data collected and used, value propositions and flows of potential revenue are summarised in Figures 2 and 3 below (representing the theoretical business model).

■ Usefully combine open data and commercial services

The liberation of data leads to a choice of who will absorb the cost, and requires an understanding of the distribution of value made possible by access to such data. The issue is therefore that of finding a balance between facilitating value creation through economic and social innovation, and a fair division of costs.

A major stake for GPS App is to establish themselves as a commercial company within an ecosystem of open innovation favourable to the co-creation of value, through the provision of a quasi-public service. Through the lens of its theoretical BM, the success of the platform depends on the active participation of the two main contributors of the market, which are transporters on the one hand, who have the power to release data in real-time, and the transport users on the other, who can participate in the management of the shared asset, which is the database.

To find a balance between, on the one hand, the political desire to release data and non-commercial exchange driven by the open data movement; and on the other hand, the commercial operation and private property of users' crowdsourced data, GPS App provides transporters with access to summarised information concerning their networks. In this spirit of openness, if GPS App obtains access to data in operator real-time, transport

mass depends on the territories concerned (network size, number of users, etc.). For Paris, an estimated 100,000 users would be sufficient to be able to provide reliable real-time information across the network at rush hour.

users will be more motivated to use the application, whose utility will consequently increase.

Another problem an application like GPS App could encounter is competition with the very creator agencies and potentially the data broadcasters which enabled them to develop this strategic asset. Indeed, it can happen that, in releasing data, a public organisation then decides to exploit such data and produce a value added service. It therefore has a considerable advantage over third-party companies, as it can capture the value upstream in the value chain via the primary service (transport) and is exempt from monetising the secondary service.

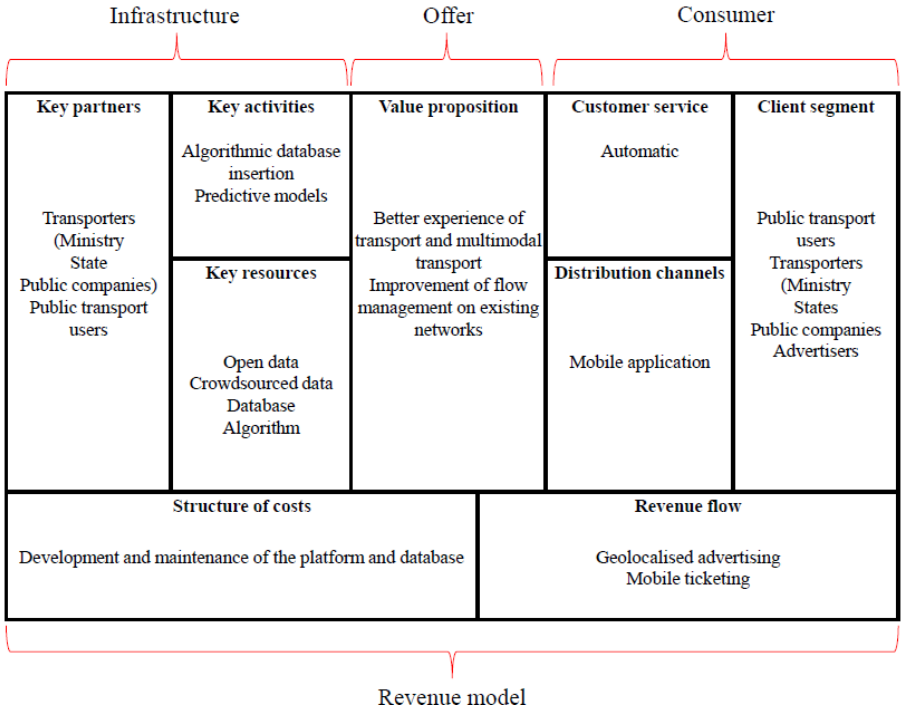
Value creation linked to open data thus presents a differentiation which overlaps the typology presented in Table 1. According to whether the actors are the producers and/or owners of the data used in the service, and whether they are used in a public or private sector, monetisation of data and reticence associated with data dissemination are not the same. This creates differentiated competitive conditions and does not encourage autonomous transporters to liberate their data to be used by third-party companies. The commercial/non-commercial aspects resulting from exploitation of data by third-party companies could also produce tensions in the relationship between value proposition(s) made by the application and the work or collaboration of transport users to the operation of the application.

■ Multi-sided business models: defining what is meant by 'free of charge' and 'payment'

To illustrate the GPS App BM and highlight the problems faced by a multi-sided market using an application which combines data resulting from open data and crowdsourcing within a commercial service, we will use the Business Model Canvas proposed by OSTERWALDER & PIGNEUR (2010). This model describes how an organisation operates according to 9 building blocks, giving a systemic vision of the organisation. These blocks correspond to four main areas of an organisation's business activity: the offer (value proposition explaining why clients come here rather than elsewhere) the consumer (targeted client segment, type of client relationship set up by the organisation and distribution channels to reach clients) the infrastructure (key activities, resources and actors in the production process) and financial viability (the difference between generated revenue and

production costs). Figure 2 presents the GPS app theoretical business model canvas.

Figure 2 - Theoretical business model canvas of the GPS app



Source: The author

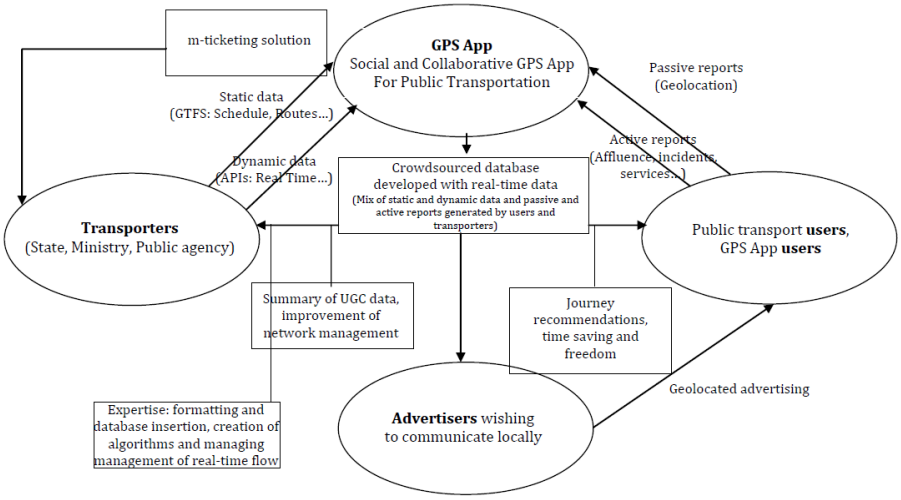
Once the critical mass of users is reached, the platform can focus on its profitability by providing advertisers with new channels of geolocated communication (behavioural targeting, habits, etc.) and their expertise in terms of mobile ticketing infrastructure management, network data production or network database management.

Figure 3 describes the complete theoretical business model with the various possible sides and their interrelationships. This model competes with the models of autonomous transporters who do not need to monetise their applications but it has a leverage effect with a worldwide coverage. The GPS App application is deployed in countries where transporters do not at all have the same motivation to open their data. In many Latin American countries (Colombia, Chile, …), such motivation is greater as their public transportation infrastructures are very recent and management is not

dependent on autonomous operators. Furthermore, due to extremely fast development in some countries (USA, Italy, South Africa, ...), and in large cities (Rio de Janeiro, Santiago, New York, Rome, Los Angeles, Paris, Montreal, Cape Town, ...), GPS App derives a capacity for negotiation with even the most reticent operators.

Theoretically, the multi-sided model of this company is made up of sides which emerge sequentially. The first two sides (transporters and users) are financed by funds raised from venture capitalists. Once the critical mass is reached, the advertising contributor could be introduced into the market to seek profitability. Finally, in the case of developed partnerships, the offer of expertise and provision of mobile ticketing infrastructures may constitute a diversification of the revenue model.

Figure 3 - Representation of the theoretical GPS app BM with all its sides



Source: The author

The BM is no longer a simple static tool used merely to legitimate the business activity to investors, but is rather a dynamic tool which develops concurrently with the strategy and various phases of the life of a start-up. Indeed, the effect of the critical mass on the users' contribution could enable the application to operate without requiring transporter contribution any longer (one less side on the market means no more chicken and egg problem). Monetary value creation no longer dominates strategic choices and investors see, in the value attached to its use (worldwide app coverage and users' involvement metrics) the potential to increase the monetary value

(revenue and financial market value creation) of a start-up business in the longer term. Consequently, there is a dissociation between the traditional key performance indicators, which are the turnover and result, and the measures of use and involvement (number of users, organic downloads, monthly and daily usage). We can observe that a political economy of attention (GEHL, 2010; KESSOUS *et al.*, 2010) develops, and that the value creation of a start-up depends both on its virtual development (value attached to its use) and the credibility of its BM (monetary value). This first point leads us to dissociate a theoretical business model from an effective one. Indeed, the theoretical BM is a model with a general scope which proposes operational functioning and development in the long-term. It is this dynamic sequential model which would be positively sanctioned and legitimated by fundraising. This theoretical model presents the various possible development scenarios of the application and proposes elements of models for possible revenues. However, if funds raised are sufficient and managers and investors decide to do so, the start-up may adopt a different strategy to that reflected in the theoretical model. The effective BM, completing the theoretical BM in a specific period, would not necessarily seek monetisation (according to theoretic sequences) but would seek to increase the presence and dissemination of the application by continuing to expand its user base. As the value attached to use increases, with cross-side network externalities, it is the financial valuation of the start-up which increases. Moreover, monetisation of an app usually comes with a drop in the download trend due to users' aversion to advertising. This divergence between theoretical and effective business models could simply be the result of a mere time lag in implementing the theoretical BM, or it could alternatively be the consequence of a desire to support an increase in the user base enabling the start-up to make a potential initial public offering (e.g. Instagram, WhatsApp, ...) or to be bought out by established actors (Google, Facebook, …)

The free app business models fall within a political economy of attention, and as such, enter a race for audience, whether this is monetised or potentiated. This race for users, downloads and an audience leads start-ups to adopt monopolistic behaviour as they are in a type of model of technological competition, as described by ARTHUR (1988), without, however, having the same technological lock-ins, as these are easily interchangeable applications, but with switching costs formed by the digital enclosures (the users' shared asset) which are the databases they have developed and which contribute to creating "winner takes all" situations.

■ Conclusion

One of the difficulties associated with data freedom is absorbing the cost of such freedom. Organisation with commercial services appears to be a viable solution for creating commercial value by innovating, using open data and personal and/or geolocated data and distributing this value between the various stakeholders. However, the mix of types, between open, public and free data, and private, closed data and proprietary databases remain problematic. On the one hand, users may have concerns in terms of privacy, but also in terms of approving the value which they helped to create. On the other hand, transporters may be tempted to retain control over the data which could expose their errors, or to innovate in the place of private actors and other platform operators.

The issue with multi-sided markets is inevitably that of which came first, the chicken or the egg. But much of the power lies with the policy-makers who decide whether to open access to data or not. The two contributors must exist and, in contrast with traditional models, both must be financed in order to populate the service before it can be monetised. There is consequently a sequence with a two-sided market, in which no contributor is initially profitable and must therefore be financed by investment, before a second phase which divides into monetisation or the path towards profitability or takeover.

Within a logic of value creation, open data cannot be only considered from the angles of transparency and democratic and emancipating virtues of liberating data. In order to produce new private services, paid for or free, sets of public and private data must be situated within platforms. The stakes of these platforms are to place themselves within a context of open innovation by enforcing cohabitation of approaches and antagonistic regimes of public, common and private goods in the distribution of created value and through the value propositions on offer. Policy makers have a role to play in facilitating data liberation and the BM offers cognitive perspectives for their understanding of the creation, capture and distribution of the potential value created by such liberation.

References

AFUAH A. & TUCCI C. (2012): "Crowdsourcing as a solution to distant search", *Academy of Management Review*, Vol. 37, n°3, pp. 355-375.

ARTHUR B. (1988): "Competing technologies: an overview", in DOSI *et al*., *Technical Change and Economic Theory*, Pinter Publishers, London and New-York.

BEUSCART J.-S. & MELLET K. (2008): "Business models of the web 2.0: advertising or the tale of two stories", *Communications & Strategies*, Special Issue "Ultrabroadband: the next stage in communications, November, pp. 165-181.

BERNERS LEE T. (2012): "Sir Tim Berners-Lee: Raw data, now!", *Wired Magazine*, Wired UK, November.

BOUSTANY J. (2013): "Accès et réutilisation des données publiques", *Les Cahiers du numérique*, Vol. 9, pp. 21-37.

CLEMONS E. (2009): "Business Models for Monetizing Internet Applications and Web Sites: Experience, Theory, and Predictions", *Journal of Management Information Systems*, Vol. 26, n°2, pp. 15-41.

COLIN N. & VERDIER H. (2012): *L'âge de la multitude. Entreprendre et gouverner après la révolution numérique*, Armand Colin, 288 pages.

DEMIL B., LECOCQ X. & WARNIER V. (2013): *Stratégie et business models*, Pearson, 242 pages.

DOBBS R., MANYIKA J., WOETZEL J., FARREL D. & VAN KUIKEN S. (2013): "Open data: Unlocking innovation and performance with liquid information", McKinsey Report, October.

DUJARIER M.-A. (2008): *Le travail du consommateur*, La Découverte, Paris, 246 pages.

ELLIOTT M. (2007): "Stigmergic Collaboration. A Theoretical Framework for Mass Collaboration", Ph.D. Dissertation, The University of Melbourne.

EL SAWY O. & PEREIRA F. (2013): *Business Modelling in the Dynamic Digital Space. An Ecosystem Approach*, Editions Springer, Collection Springer Briefs in Digital Spaces, 68 pages.

GEHL R. W. (2010): "A cultural and political economy of Web 2.0", Ph.D. in Cultural Studies, George Mason University.

HOWE J. (2006): "The rise of crowdsourcing", *Wired Magazine*, Issue 14.06, June.

KESSOUS E., MELLET K. & ZOUINAR M. (2010): "L'économie de l'attention : entre protection des ressources cognitives et extraction de la valeur", *Sociologie du travail*, Vol. 52, n° 3, pp. 359-373.

KATZ M. & SHAPIRO C. (1985): "Network Externalities, Competition and Compatibility", *The American Economic Review*, Vol. 75, n°3, pp. 424-440.

KLEMMANN F., VOSS G. G. & RIEDER K. (2008): "Un(der)paid Innovators: The Commercial Utilization of Consumer Work through Crowdsourcing", *Science, Technology & Innovation Studies*, Vol. 4, n°1, pp. 5-26.

LACOMBE R., BERTIN P.-H., VAUGLIN F. & VIEILLEFOSSE A. (2011): "Pour une politique ambitieuse des données publiques. Les données au service de l'innovation et de la transparence", Rapport pour le Ministre de l'Industrie, de l'Energie et de l'Economie numérique, Juillet 2011, 116 pages.

LEBRATY J.-F. & LOBRE-LEBRATY K.:

- (2010) : "Créer de la valeur par le crowdsourcing: la dyade Innovation-Authenticité", *Systèmes d'Information et Management*, Vol. 15, n°3, pp. 9-40.
- (2013): *Crowdsourcing: One Step Beyond*, Willey-ISTE, Willey and Sons, 144 pages.

McAFEE A. & BRYNJOLFSSON E. (2012): "Big Data: The Management Revolution", *Harvard Business Review*, Vol. 90, n°10, octobre, pp. 60-68.

MÖLLER K. & RAJALA A. (2007): "Rise of strategic nets - News modes of value creation", *Industrial Marketing Management*, Vol. 36, n°8, pp. 841-872.

O'REILLY T. (2007): "What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software", *Communications & Strategies*, n°65, pp. 17-37, 1st Q.

OSTERWALDER A. & PIGNEUR Y. (2010): *Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers*, Editions Willey, 288 pages.

PARENT R. & CHANAL V. (2009): "Quels business models pour les plateformes web 2.0: les apports de la théorie des marchés bifaces", *Facef-Pesquisa*, Vol.12, n°3, pp. 379-395.

RALLET A. & ROCHELANDET F. (2011): "La régulation des données personnelles face au web relationnel : une voie sans issue ?", *Réseaux*, Vol.3, n°167, pp. 17-47.

REBILLARD F. (2007): *Le Web 2.0 en perspective Une analyse socio-économique de l'internet*, L'Harmattan, Coll. Questions contemporaines, Série Les industries de la culture et de la communication, Paris, 162 pages.

REY B. (2012): *La vie privée à l'ère du numérique*, Lavoisier, Collection Traitement de l'information, 297 pages.

ROCHET J.-C. & TIROLE J. (2003): "Platform competition in two-sided markets", *Journal of the European Economic Association*, Vol.1, n°3, pp. 990-1029.

SHIRKY C. (2008): *Here comes everybody: The power of organizing without organization*, Penguin Press, New York, 327 pages.

SUROWIECKI J. (2004): *The Wisdom of Crowds*, Doubleday, New York, 308 pages.

TURNER F. (2008): *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*, University of Chicago Press, 354 pages.

WALSHAM G. (1993): *Interpreting information systems in organizations*, John Wiley Series in Information Systems, John Wiley and Sons, 286 pages.

WIRTZ B., SCHILKE O. & ULLRICH B. (2010): "Strategic development of business models. Implications of the web 2.0 for creating value on the Internet", *Long Range Planning*, Vol. 43, n°1-2, pp. 272-290.

YIN R. (2014): *Case Study Research: Design and Methods*, 5th Edition, Applied Social Research Methods Series, Sage Publications, 264 pages.

ZOTT C., AMIT R. & MASSA L. (2011): "The business model: recent developments and future research", *Journal of Management*, Vol. 37, n°1, pp. 1019-1042.