DECODE DDDC pilot: An economic diagnosis of the current ‘data economics’ landscape

Executive summary

This document intends to make the reader aware of existing data-driven economic models. Along with the other two ‘diagnosis documents’ (i.e. ‘data governance’ and ‘data regulation’), it falls within the preliminary phase of DECODE’s ‘Digital Democracy and Data Commons’ pilot, whose main goal is that of constituting, through experimenting a technologically-enabled participatory process, a ‘recursive data commons’, namely « *a sociotechnical system of production, appropriation and use of a dataset under conditioned or unconditioned open access, on the basis of a normative framework (be it formal, informal or both) that defines the governance of the dataset, of the community, and of the infrastructure, in a democratic way, and that enables or promotes shared practices that generate goods under a similar model of production, appropriation and use* » (Calleja-López 2018: 33).

This document is mainly based on the research results achieved so far within the DECODE project. After a brief introduction aimed at highlighting the ever growing importance that digital data plays in the nowadays and future global economy, we will schematically describe what can be considered the three major current economic models around data:

1. **Profit-driven models**, namely those of private digital companies that appropriate user-generated data and exploit it for commercial purposes (such as targeted advertising or improvement of the service offered), and those of companies known as ‘data brokers’, that collect (and then cluster and sell) data from other businesses rather than from consumers first hand.
2. **Public value-driven models**, based on a ‘data open for all’ logic in order to strengthen innovation and transparency, pool previously unrelated dataset and get unprecedented insights out of it, or contribute to a cause of public interest.
3. **Common-driven models**, where data is considered as a not merchantable resource that is produced, governed and used by a given community to feed either the commons itself or nurture common-oriented networks and models.

The document’s Annex will concern past and current attempting measures to tax digital businesses. Indeed, a reflection on taxation is necessary in order to discuss the possible economic sustainability related to the three different models.

Introduction

The world is experiencing something similar to a data explosion. The amount of information generated on a daily basis are not only vast, it is also growing exponentially: 90 percent of the data present in the world as of 2017 have been created in the previous two years, and this amount is expected to be 40 times bigger by 2020[1], a forecast consistent with the diffusion of smartphone usage and the expansion of the Internet of Things' ecosystem. The infographic here below, showing the activity taking place globally on various digital platforms in a 60 seconds span, can help providing a clearer idea of such greatness.

2018: This is what happens in an Internet minute
Source: Cumulus Media

Data has been defined in different ways: as the new oil[2], a new factor of production, and the currency of today's digital market. Be that as it may, what is unambiguous is that owning big volumes of data and adequate analysis methods opens up enormous opportunities, as well as equally enormous risks, in particular when personal data processing is involved[3]. Indeed, 'big data' refers to information assets that, due to several varying characteristics (among which their massive volume, their variety in quality, and their high-velocity production) can be properly exploited only using a new generation of analytics technologies (such as Artificial Intelligence, achieved through Machine Learning tools) designed to collect, aggregate and analyse it quickly and cost effectively, in order to extract variably reliable prediction and decision-making patterns for whatever purpose in whatever

1 IBM Marketing Cloud (2017), *10 Key Marketing Trends for 2017 and Ideas for exceeding customer expectations*, https://www-01.ibm.com/common/ssi/cgibin/ssialias?htmlfid=WRL12345USEN

2 The idea that big data represent 'the new oil' in the economic system has been criticized in the DECODE report *Me, my data and I*. Indeed, while oil is a scarce physical resource, data can conversely be replicated and, often, the more it is used, the greater its value.

3 In 2012 the '*Digital Universe Study*' report estimated that the amount of data requiring protection and precise regulation will reach, by 2020, the 40 percent of the total amount of data present in the world. See: Gantz J. and Reinsel D. (2012), *The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Grow the in the Far East*, IDC iView, sponsored by EMC Corporation, https://www.emc.com/collateral/analyst-reports/idc-thedigital- universe-in-2020.pdf

sector, be it public or private. Assessing and optimizing credit risk in the lending industry, implementing citizen-focused public policies (e.g. restructuring bus routes to meet the needs of travellers in specific areas of a given city), deploying more police forces where crimes are presumably most likely to occur, identifying skin cancer at a rate equal or better than a human dermatologist by 'training' a computer with lots of example images, or avoiding overproduction and resource waste in agriculture are just a handful of examples of big data analytics' applications. However, a key issue for organisations using personal data for big data analytics is that data is "processed fairly, lawfully and in a transparent manner in relation to the data subject" (Article 5(1) (a) GDPR). Indeed, personal data may be used to draw conclusions about individuals with unexpected and sometimes unwelcome effects (e.g. users' race-based profiling) or for purposes that fall out of the reasonable expectations of the people concerned (e.g. mass surveillance programmes like PRISM).

It is worth noting that the so-called 'pure players' of the big data market (that is companies obtaining all of their revenues from the sale of big data products and services) account for just 5 percent of the overall market, while the remainder is generated by powerful profit-driven tech-companies (e.g. *Google*, *Facebook*, *Uber*, *Airbnb*, *Amazon*) providing to a great extent online services. Within this first model we find also a multi-billion-dollar industry that operates behind a veil of secrecy without direct regulatory oversights, namely that of so-called 'data brokers'. All these companies usually rely on their own big data analytics tools to optimize their revenues, sometimes offering them as paid services (as in the case of *Google Cloud Platform* and *Amazon Web Services*) to governments, enterprises and professionals, for whom it's cheaper to rent applications than to buy infrastructure and build it out from scratch.

The 'open data' philosophy is the one underlying several initiatives and whose main (but not the only one, as we will see) example is represented by Open Government projects (e.g. *Open Data BCN*), fuelled by public value-driven motivations such as transparency or innovation-fostering. However, the fact that data is 'open to anyone' means that it is frequently actors endowed with economic and knowledge capital (e.g. the just cited profit-driven platforms) who benefit the most of its exploitation.

When a community contributing data in the context of a given platform relies on a democratically conceived (by the stakeholders of the dataset) set of rules that let participants freely decide (individually or collectively) upon the uses they want to grant (or not to grant) to third parties with regard to the further processing of the data they produce, and when these uses are firstly aimed at improving the community's overall needs (e.g. commercial purposes' restrictions) and allowing the spread of its underlying logic through similar initiatives (e.g. adoption of Free-Libre/Open Source Software licenses), we can talk of a common-driven model[4]. The closest project to such a model can be considered *OpenStreetMap*, whose Open Database License (ODbL[5]) does not permit however users' granular control over the data they produce and it allows for private appropriation of the database, provided that an open version free of restrictions is always available.

4 Our approach to the notion of common is based on the idea that it cannot be relegated to specific common goods. The common is the potential of expanding social cooperation which attends the transformation of productive forces and the prominence of new forms of labour, such as the increasingly socialised production of knowledge.

5 https://en.wikipedia.org/wiki/Open_Database_License

1. Profit-driven models

Aim

Generally speaking, this model includes private companies that play the role of intermediaries between users and advertisers (e.g. search engines and social networks, such as *Alphabet*'s *Google Search*[6] and *Facebook*) or between services and products' suppliers and requesters (e.g. on-demand economy platforms such as *Uber* and *Airbnb* + *Amazon Marketplace*). Other companies (i.e. *data brokers*) collect, compile, and package information[7] from sources other than data subjects themselves and sell it to both commercial and institutional clients.

Data governance

These companies track and collect personally identifiable information in a non-transparent way and process it (i.e. extract knowledge) to pursue their commercial purposes. Usually, these companies' 'privacy policies' agreements are not easily understandable. In the case of *data brokers*, consumers are often prevented from accessing, correcting, opting out, or requesting deletion of the information they hold about them.

Business model

Information is either used for targeted advertising[8], which may be deemed as the fee to pay for the 'free' service the users enjoy (as in the case of *Google Search* and *Facebook*), to improve the service offering (as in the case of on-demand platforms such as *Uber* and *Airbnb*, and in the case of *Amazon Marketplace*, which rely instead on the deduction of transaction fees), or sold to customers after being clustered (as in the case of *data brokers*). These companies are normally stock-market listed, oriented on optimizing profit for their investors and minimizing costs and taxation. Their main asset (algorithm) and raw material (data produced by users) are the core of their value creation chain but, being intangible and often not traded in a traditional way, they are difficult to estimate or to monitor, consequently challenging the assessment of potential anti-competitive behaviour[9].

---

6 In 2015, *Google* was reorganized into the biggest subsidiary of a larger public holding company named *Alphabet Inc*. As a consequence of *Alphabet*'s creation, projects that were not part of *Google*'s core business (that is, inter alia, *Ads*, *Android*, *YouTube*, *Cloud*) were split into separate companies, each one with its own CEO. These 'moon-shot' projects, labelled as 'Other Bets' businesses, cover a vast bundle of early-stage, futuristic, but promising pursuits in areas such as smart home technology, healthcare and urban planning.

7 Common data broker products comprise: lists of consumers assembled in 'data segments' based on expected attitudes; predictive 'look-alike models' about a person, built on similarly-targeted people whose data the broker already possesses; marketing/credit/fraud 'scores' to foresee future behaviours; and 'data appends', namely additional information to fill gaps in a given consumer's profile that an organisation already holds.

8 Advertising revenue accounts for almost all of *Alphabet*'s and *Facebook*'s total revenues. In 2017, *Alphabet*'s total revenue amounted to roughly 110 billion US dollars with an advertising revenue of about 95 billion dollars (see: *Alphabet* 2017 Annual Report, https://abc.xyz/investor/pdf/20171231_alphabet_10K.pdf). Similarly, *Facebook*'s 2017 advertising revenue was of more than 39 billion US dollars on a total revenue of almost 41 billion US dollars (see: *Facebook* 2017 Annual Report, http://www.annualreports.com/Company/facebook). *Alphabet* and *Facebook* capture more than a half of all the growth in global online ad spend. They use programmes (i.e. *Google AdWords/AdSense* and *Facebook Ads Manager/ Facebook Audience Network*) which work on the basis of real-time auction mechanisms and allows advertisers to purchase CostPerClick-based advertising spots and to deliver text/image/video-based ads tailored on an incalculable amount of users-generated information.

9 The OECD (2014) indicates, in particular, three critical points that challenge the traditional approach used by competition authorities for assessing potential abuses and harms of market dominance and mergers. (1) *Challenges in defining the relevant market*: since traditional market definitions are based on economic transactions, even if data plays an important role relating to a company's turnover, when it involves no monetary transactions – e.g. a company that uses data to improve its own services but does not sell it –, then the data market is not considered as relevant; moreover, the multi-sidedness of digital platforms means that the market is not constituted by the typical producer/seller/consumer relationship, but by a complex interplay of actors playing differentiated roles and that focussing on one side of market will rarely lead to a proper market definition. (2) *Challenges in assessing the degree of market concentration*: we tend to associate the concept of monopoly or oligopoly with the idea of high prices and control over them; however, a large share of data-driven services is provided for 'free' (in exchange for access to personal data) and the value of data is not

Licenses

Profit-driven models usually rely on proprietary software licenses. Moreover, it happens that free software is at least partially appropriated, as in the case of Google's *Linux*-fork (i.e. *Android*), whose *Google Mobile Services* (as opposed to the *Android Open Source Project* where the core parts of *Android*'s *Linux*-based source code is publicly available) constitutes the proprietary part of the software. Users-generated data are subject to appropriation strategies (i.e. IP-based strategies - such as copyright, patents, sui generis database right - and trade secret) that grant companies ample freedoms[10].

2. Public value-driven models: a) Open Government Data b) Multi-stakeholder data pooling c) Non-commercial standalone open data production

2a) Open Government Data

Aim

Public bodies produce huge quantities of data and information. 'Open Government Data' (OGD) is a philosophy (and increasingly a set of policies) that promote transparency, accountability and value creation by making government data available to all. By encouraging the use, reuse and free distribution of (some disclosed) datasets, governments want to stimulate business creation and innovative citizen-centric services. An example is O*pen Data BCN*[11].

Data governance

Data can be accessed, used and redistributed by anyone and are produced by State bureaucracy or collected from third parties (e.g. transportation companies' information).

Business model

The production, collection, storing and maintenance of the datasets is financed using the State budget.

Licenses

With regard to software, these models may rely on proprietary or FLOSS licenses. Data licences are normally permissive, allowing for commercial uses (e.g. the UK's Open Government License[12]).

---

intrinsic but depends on the context of its use; moreover, data may act as a 'barrier to entry' when new small but promising entrants are not able to match the volume or quality of data processed by incumbent actors. (3) *Challenges in assessing potential consumers' detriments*: the degree to which privacy harms should be considered when assessing anti-competitive behaviour and mergers is an ongoing debate. Arguments in favour of such an approach are, for instance, the fact that a company can cross the limits accorded to data protection in order to extract more value and thus gain a competitive advantage over its rivals; or that a merger between two data-driven companies may allow the resulting entity to lessen the degree of privacy protection accorded to consumers.

10 For instance, *Facebook*'s Terms of Service states: "Specifically, when you share, post, or upload content that is covered by intellectual property rights (like photos or videos) on or in connection with our Products, you grant us a non-exclusive, transferable, sub-licensable, royalty-free, and worldwide license to host, use, distribute, modify, run, copy, publicly perform or display, translate, and create derivative works of your content (consistent with your privacy and application settings). This means, for example, that if you share a photo on Facebook, you give us permission to store, copy, and share it with others (again, consistent with your settings) such as service providers that support our service or other Facebook Products you use. You can end this license any time by deleting your content or account. You should know that, for technical reasons, content you delete may persist for a limited period of time in backup copies (though it will not be visible to other users). In addition, content you delete may continue to appear if you have shared it with others and they have not deleted it".

11 http://opendata-ajuntament.barcelona.cat/en/

12 http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/

2b) Multi-Stakeholder Data Pooling

Aim

Two or more entities of any sort (private firms, governments, NGOs, collectives of citizens, private individuals, etc.) create an enriched and enlarged dataset through pooling their respective datasets. For non-profit entities (e.g. PRIDE[13])) the logic of this data pooling consists in being able to fulfil better their non-commercial mission. For profit-oriented firms, motivations stem mostly from the fact that, each stakeholder providing data and/or expertise on a domain, linkage of previously unrelated data reveals valuable information and business opportunities, as well as good publicity when the pooling also helps to tackle a societal issue (e.g. CIESIN & *Facebook*: Open, Improved Settlement Data[14]).

Data governance

Data can be accessed, used and shared by anyone. Data pooling is followed by the provision of respective expertise to increase data quality and interoperability, the latter being essential for the data pool to be usable and valuable.

Business model

The creation and maintenance of the open dataset is carried out by financial contributions from promoting stakeholders or public funding when the State is involved.

Licenses

With regard to software, these models may rely on proprietary or FLOSS licenses. Data licences are normally permissive, allowing for commercial uses (e.g. Open Database License - ODbL).

2c) Private firm standalone open data

Aim

It refers to cases in which a single firm decides to open a certain dataset it owns. Two objectives motivate this decision. The first one is the enlargement (increasing value through *size*) and/or enrichment (increasing the value through *quality*) of the original opened dataset. The second one is creating business opportunities related to the dataset.

Data governance

The original dataset is produced either by an intra-firm data collection (e.g. crowdsourced data about the location of bus stops) and/or production (e.g. intra-firm data about clinical trials carried on) and, once opened, through the enlargement/enrichment of the data provided by third parties.

Business model and Licenses

- When a ‘freemium model’ is chosen, a core dataset is kept open but an extended one (usually useful for commercial purposes) is sold. In these cases, licenses distinguish commercial from non-commercial use of data.

---

13 PRIDE (Platform for Regional Innovation to enable Data exchange on Energy) aims at implementing a platform for gathering, handling and visualising energy data from very diverse sources in the Brittany and Pays de la Loire regions. See: https://smile-smartgrids.fr/en/projects/en-pride.html

14 Carried on by the Centre for International Earth Science Information Network (CIESIN), *Facebook* and the World Bank, the project’s resulting pooled dataset helps understanding how human settlements are distributed across landscape, allowing for many different applications such as research, humanitarian planning or crisis response. *Facebook* has shared commercially purchased satellite imagery data with CIESIN, which in turn has census data of the places to which the satellite images correspond. In addition, *Facebook* has shared ‘state of the art computer visioning techniques’ (i.e. an increase in the quality of data) with CIESIN to identify buildings.

- Another model consists in developing data-related services around the open data, such as paid data visualization toolkits to analyse the open dataset. An example of this is HERE Open Location platform[15], which provides open cartographical data and offers cartographical licenses for firms that want to use their maps.
- Finally, a third way of monetizing private firm standalone open data is creating future business opportunities by gaining knowledge and expertise through the study of the dataset.

2d) Non-commercial standalone open data production

Aim

Creation - from scratch or over already-existing open and/or crowdsourced datasets - of open datasets by a single organization or individual, usually motivated by contributing to a cause and not for commercial purposes (e.g. *Inside Airbnb*[16] collects publicly available information about cities' *Airbnb*'s listings and provides filters and key metrics that can be used by anyone in order to explore how *Airbnb* is being used to compete with the residential housing market).

Data governance

Data can be accessed, used and shared by anyone. The organization/individual is in charge of data production, storing, and maintenance.

Business model

The production, collection, storing and maintenance of the datasets is financed by donations and voluntary labour.

Licenses

With regard to software, these models may rely on proprietary or FLOSS licenses. Data licences are normally permissive, allowing for commercial uses (e.g. Creative Commons CC0 1.0 Universal (CC0 1.0) 'Public Domain Dedication' license[17]).

3. Common-driven models

Aim

The most relevant feature of this set of models is the attempt to develop a new mode of regulation and production opposed to the principle of property as an exclusive right and aimed at providing an open dataset that benefits the population at large. Therefore, the common establishes the concepts of *horizontal social cooperation* among participants in the common's construction and management, and of *inappropriability* of the means of production (and of the value generated). A good example is *OpenStreetMap*, a 4 million members' project whose geographical database is mainly composed of data crowdsourced by individuals (using handheld GPS devices) and also enriched by other open data from third parties (such as *Yahoo* and *Bing* aerial maps, as well as those produced by municipal administrations). Anyone can contribute, enriching or correcting the data, but different levels of governance guarantee the information reliability.

Data governance

15 https://www.here.com/en/vision/innovation/here-open-location-platform
16 http://insideairbnb.com/index.html
17 https://creativecommons.org/publicdomain/zero/1.0/

Data is open (it can be accessed, used and shared by anyone), produced through crowdsourcing and improved through other open sources.

Business model

The production, collection, storing and maintenance of the datasets is usually financed by donations, voluntary work, public funding, or (less often) the selling of project-related products.

Licenses

With regard to software, these models usually rely on FLOSS licenses (e.g. GNU General Public License[18]). Data licences are normally permissive, allowing for commercial uses (e.g. Open Database License - ODbL).

Annex A - Digital businesses' taxation

Any attempt at implementing an international regime as regards the taxation of digital businesses looks set to stumble upon conflicting political agendas and national legal regimes. As an evidence

18 https://www.gnu.org/licenses/gpl-3.0.html

of this, some EU countries[19] has started to implement unilateral fiscal measures, since the legislative proposals submitted in March 2018 by the European Commission - which will be exposed here below after briefly describing prior attempts to fiscally regulate the digital economy - seem far from reaching the needed unanimous agreement among the 28 EU Member States. Indeed, low-tax countries such as Ireland and Luxembourg, housing many Internet giants' subsidiaries, see this potential change as a threat to their economic models and fears that it would pave the way for the imposition of a minimum and common corporate tax rate[20]. Germany is concerned that the proposed reforms could lead international partners to respond with retaliatory tax measures that will penalize German exports[21], while Sweden started to oppose after Prime Minister was urged to do so by *Spotify*'s founder[22]. Overall, according to 'Politico'[23], most of EU members are on the fence, while supporters and opponents are distributed in a fairly balanced measure.

A first proposition aimed at considering the possibility to fiscally discipline the value generated through - at that time only emerging – the so-called information society (IS) can be traced back to the 1997 report 'Building the European information society for us all' (EC 1997), prepared by a high level group of experts set up by the European Commission and focused on policy suggestions addressing the many social aspects connected to the IS - recognized as being neglected until then. Among these suggestions, was a recommendation by Professor Luc Soete, chairman of the high level expert group, to investigate further whether a 'bit tax', "based on a simple count of bits flowing over telecommunications lines" (*Ibidem*: 50) - namely a taxation based on the intensity of electronic transmission - might be a feasible tool to more equally distribute the benefits arising from the "trade in intangible information services, where notions of value are difficult to estimate or to monitor" (*Ivi*). In a parallel paper (Soete and Kamp 1996) the main arguments in favor of such a tax, described as a 'no-man's research land', are summarized. The paper starts by assuming that « *in the preset, global free-market environment, any suggestion for a new tax is likely to be greeted with skepticism and to be quickly rejected* » (*Ibidem*: 353), as showed by the predominantly negative reaction amongst policy makers (fearing that a similar measure might give the wrong signal to potential investors), technical experts (considering 'bits' as an irrelevant or ineffective measure of transmission intensity) and individual Internet users (perceiving it as an attempt of the state to tax freedom of speech). After the 'bit tax' was popularized by the United Nations Development Program (UNDP 1999), the US Congress threatened to withdraw from the United Nations if global tax proposals were to continue to be put forward (Thorndahl 2003: 203).
Indeed, this first, timid call for intervention has been occurring right in the middle of the 'Get Big Fast' business model's period, shortly before the 2001 bursting of the dot-com bubble, in response to which – given the bankruptcy *domino* it caused - the discourse around fiscal taxation of data-driven digital firms has been silenced, leveraging even more than before on the neoliberal assumption that government interventions are likely to distort decisions in a manner harmful for the

19 For instance, Slovakia's 2018 Tax Reform Law introduced new permanent establishment (PE) rules for digital platforms (See: Financial Administration Slovak Republic, 16/03/18. URL: https://www.financnasprava.sk/sk/pre-media/novinky/archiv-noviniek/detail-novinky/_zdan-dig-platfor) and in 2017 Italy established a 3 percent 'Web Tax' on digital transactions, coming into force as from 1st January 2019 (See: Losito A. (2018), *Web Tax*, Guidafisco.it, 01/10/18. URL: https://www.guidafisco.it/web-tax-italia-cos-e-come-funziona-imposta-sulle-transazioni-digitali-1999).
20 Stratfor (2018), *Digital Tax Proposals Produce New Discord in the EU*, Stratfor, 14/02/18. URL: https://worldview.stratfor.com/article/digital-tax-proposals-produce-new-discord-eu
21 Guarascio F. (2018), *EU digital tax on corporate turnover faces uphill road*, Reuters, 28/04/18. URL: https://www.reuters.com/article/us-eu-ecofin-tax/eu-digital-tax-on-corporate-turnover-faces-uphill-road-idUSKBN1HZ0JS
22 Business Insider (2018), *Spotify founder Daniel Ek gets his way on tech tax – after emailing Swedish PM*, Business Insider, 25/04/18. URL: https://nordic.businessinsider.com/spotify-founder-daniel-ek-gets-his-way-on-tech-tax--after-emailing-swedish-pm--/
23 Plucinska J., Vinocur N., Smith-Meyer B. (2018), *Europe's digital tax map: Where countries stand*, Politico, 12/04/18. URL: https://www.politico.eu/article/europe-digital-tax-map-where-countries-stand-analysis-deep-divisions/

efficient functioning of markets and may slow down investment and innovation. It is precisely this kind of *esprit* that helped some US tech-companies to reach a dominant position, at such a point that a stance with regard to a proper taxation strategy - beyond and separate from other 'adjustment' legal mechanism like antitrust laws and personal data protection regulations - became necessary.

Therefore, the Base Erosion and Profit Shifting Project (BEPS Project[24]), led by the OECD's Committee on Fiscal Affairs in conjunction with OECD and G20 countries, was set up in 2013 in order to collaboratively elaborate a response to "tax avoidance strategies that exploit gaps and mismatches in tax rules to artificially shift profits to low or no-tax locations"[25]. The project consists of 15 actions[26], the first of which is emblematically aimed at addressing the tax challenges of the digital economy. Following an interim report (OECD/G20 BEPS 2018) on the implications of digitalization for taxation, more than 110 out of 116 countries and jurisdictions participating in Inclusive Framework - a platform specifically conceived to tackle BEPS Action Plan – has agreed to work towards a consensus-based solution with regard to 'profit allocation' rules.

The following recent EU legislative proposals - mentioned above - can be considered one of the implementation of the BEPS package.

1. The first proposed directive (COM 147 (2018)) represents a long-term comprehensive reform of EU corporate tax rules, aimed at securing a real link between where digital profits are made and where they are taxed. It would enable Member States to tax profits that are generated in their territory by a digital business that, although not having a physical presence *in loco*, has nevertheless a 'significant' and therefore taxable commercial presence. The latter is reached if one or more of the following criteria are met:

   a) The company exceeds a threshold of €7 million in annual revenues in a Member State during a given taxable year;
   b) The number of users of its digital service(s) in a Member State during a given taxable year exceeds 100.000;
   c) The number of business contracts for the supply of any such digital service that are concluded by users in a Member State during a given taxable year exceeds 3.000.

   The attribution of profit will take into account the market values of the following activities:

   a) The collection, storage, processing, analysis, deployment and sale of user-level data;
   b) The collection, storage, processing and display of user-generated content;
   c) The sale of online advertising space;
   d) The making available of third-party created content on a digital marketplace;
   e) The supply of any digital service not listed in points (a) to (d).

2. The second proposed directive (COM 148 (2018)), embodied by an interim tax (Digital Services Tax - DST) at a rate of 3 percent on digital companies' gross revenue net of VAT and other similar taxes, represents a short term solution aimed at ensuring that those activities which are currently not effectively taxed would begin to generate immediate revenues for Member States, helping this way to avoid the implementation of unilateral measures. The tax, which would only concern businesses with total annual worldwide revenues exceeding €750 million and EU revenues above €50 million, would only apply to those types of activities where users play a primary role in value creation, namely from:

   a) The sale of online advertising space;
   b) The sale of collected user data;
   c) The intermediary activities of those platforms that facilitate the sale of goods and services between users (e.g. *Uber* and *Airbnb*).

---

24 See: http://www.oecd.org/tax/beps/
25 *Ibidem.*
26 See: http://www.oecd.org/tax/beps/beps-actions.htm

Both these directives are envisaged to come into force by January 1st 2020. However, to the problem of a substantial lack of consensus discussed above is added the likely introduction of overseas counter measures, since it has been estimated that half of the around 120 to 150 companies that would be affected by the new rules are located in the United States. Concerns about the second proposal rise from the fact that, being the DST a revenue tax, this means that it would be paid as well when the company is loss making. The both measures share a second major concern, namely that of the potential distortionary effects of taxation: it is indeed probable that highly profitable firms will be able to pass the tax burden on to their consumers. When the product is online advertisement - and the customer is an advertiser rather than a consumer – the advertiser is likely to increase the prices of the goods he sells in order to sustain the tax burden that has (allegedly) shifted upon him. This would mean that a tax that is meant to capture a share of firms' profits may have the consequence of primarily penalizing consumers.

Hence, the question of how such a digital taxation should be organized in political and legal terms is still an open challenge. Should it result in a unilateral or multilateral CSR scheme or should it be a mandatory mechanism introduced at either the national or supranational level? As digital data transcends national borders, the ideal form of implementation may be this last one, provided that a reliable and verifiable method of calculating dues is found – given the difficulty in detecting a technique to measure the effective value of data. However, supranational settlements require, as we have seen, reconciliation of antagonistic interests.

There exist sound economic, social and ethical arguments for justifying recourse to wealth redistribution mechanisms in the digital economy. Lehdonvirta *et al.* (2016) proposes four broad data financing models:

> a) A *global Internet subsidy* would be used to reduce the 'digital divide' between online and offline individuals, at the same time benefiting other users and online service providers by growing the population of Internet users.
> b) A *privacy insurance for personal data processing* may provide compensation for victims of potential data leaks, spent on the development of privacy-enhancing technologies, and help data processors to 'insure' themselves against the reputational damage of a data breach.
> c) An *attention levy for digital marketing* on intrusive advertising as a violation of individuals' sovereignty over their personal or private sphere.
> d) A *shared-knowledge duty for open and public data* which would ensure that digital businesses honour their social contract by sharing more of the benefits they earn through exploitation of public resources (such as publicly-funded open data) or through appropriation of users' *digital labour*.

This last point - that is data generated by users' *free digital labour* are the core of online companies' value creation chain - has been openly recognized in a recent French government report (Collin and Colin 2013) and pinpointed as one of the explanation for the low marginal operating costs and the exponential returns to scale that are specific to the digital economy: the fact that « *the labour factor has been squeezed out by the data generated by the activity of the users of online applications* » (*Ibidem*: 105) has allowed digital firms to avoid hiring employees to create content. Taking into account this fundamental aspect and after acknowledging that international tax law gives the power to tax profits in the country where the company's head office is located rather than in the country where the company does business, authors suggest three sets of proposals:

> 1. A tax law reform aimed at identifying a new definition of a permanent establishment (PE) within the context of the digital economy making sure that such a notion more effectively captures the *free labour* phenomenon.
>
> 2. In the meantime, a tax for businesses that collect data obtained through regular and systematic monitoring of users in a given country will prompt the companies to adopt practices conforming to four public interest objectives:

- Strengthening the protection of individual freedom;
- Promoting innovation in the digital trust industry;
- Encouraging the emergence of new services for users;
- Generating productivity gains and growth.

3. Adaptation of R&D and market financing taxation environment to the realities of the digital economy.

A series of arguments (Symons and Bass 2018) uphold the opinion that city governments and local authorities are the most suitable entities to promote and protect people's digital rights and consequently to manage their taxation: cities are emerging as new battlegrounds over personal data, they are closer to the lives of everyday people, they are often more flexible than regional or national governments and represent the most appropriate focus for entrepreneurial ecosystems. This is all the truer if we consider the process of 'regionalization' and consequent strengthening of the principle of subsidiarity that has interested several European countries in response to the crisis of the nation-state, giving local and regional realities a greater chance to claim strong forms of autonomy. This would lead towards a process of federation able to simplify and facilitate the implementation of a set of primary policy actions (*Ibidem*), on which to build the foundations for a network of *data commons*-based and people-centric cities.

Bibliography

Calleja-López A. (2018), *Technopolitical Democratization and Digital Commoning: the Case of the Digital Democracy and Data Commons (DDDC) pilot*, D2.5 DECODE project.

Collin P., Colin N. (2013). *Mission d'expertise sur la fiscalité de l'économie numérique*. Paris: Ministère de l'économie et des finances.

COM 147 final (2018), *Proposal for a Council Directive laying down rules relating to the corporate taxation of a significant digital presence*, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52018PC0147

COM 148 final (2018), *Proposal for a Council Directive on the common system of a digital services tax on revenues resulting from the provision of certain digital services*, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A148%3AFIN

EC (1997), *Building the European information society for us all - Final policy report of the high-level expert group*, Directorate-General for Employment, Industrial Relations and Social Affairs. Unit 8.4. Manuscript completed in April 1997.

Fuster Morell M., Carballa Smichowski B., Smorto G., Espelt R., Imperatore P., Rebordosa M., Rocas M., Rodríguez N., Senabre E., and Ciurcina M. (2017), *Multidisciplinary Framework on Commons Collaborative Economy*, D 2.1 DECODE project.

Lehdonvirta, V., Mittelstadt, B. D., Taylor, G., Lu, Y. Y., Kadikov, A., Margetts, H. (2016) *Data Financing for Global Good: A Feasibility Study.* Oxford: Oxford Internet Institute.

Lucarelli S., Musolino E., Rocchi G., Ciacciarelli S., Laniado D., Espelt R., Senabre E., Fuster Morell M., Bria F., Ciurcina M., Bassi E., Symons T. (2017), *Economic and regulatory analysis of data platforms and value creation models of the on demand economy*, D 2.2 DECODE project.

OECD (2014), *Data-driven Innovation for Growth and Well-being: Interim Synthesis Report*, https://www.oecd.org/sti/inno/data-driven-innovation-interimsynthesis.pdf

OECD/G20 BEPS (2018), *Tax Challenges Arising from Digitalisation – Interim Report 2018*. URL: http://www.oecd.org/tax/beps/tax-challenges-arising-from-digitalisation-interim-report-9789264293083-en.htm

Soete L., Kamp K. (1996), *The "bit tax": The case for further research*, Science and Public Policy, 23(6), pp. 353–360.

Symons T. and Bass T. (2017), *Me, My Data and I: The Future of the Personal Data Economy*, D 1.7 DECODE project.

Symons T. and Bass T. (2018), *Policy Requirements and Models of Implementation*, D1.10 DECODE project.

Thorndahl M. (2003), *Financements et partage des richesses: des sujets tabous*, Annuaire suisse de politique de développement, Vol. 22 (2), pp. 201-206.

UNDP (1999), *Human Development Report 1999*. New York, Oxford.

Vercellone C, Brancaccio F., Giuliani A., Puletti F., Rocchi G., Vattimo P., *Data driven disruptive commons-based models*, D2.4 DECODE project.