Data Pricing Models for Data Exchange: Review Analysis

By

Rupsa Majumdar

Anjula Gurtoo

January, 2023

Centre For Society and Policy

Indian Institute of Science
Bengaluru

Table of Contents

Abstract

Data economy has been growing over the years and governments are putting forward data as a tradable community, thereby opening up a new source of revenue. Promotion of data as a tradable community brings forth the arena commonly referred as 'data pricing'. While the issue of data pricing is getting recognized in literature, dealing specifically with the contemporary data pricing strategies remains a gap. Therefore, the report fills the gap through a literature review to identify pricing strategy clusters. By taking into account a wide number of qualitative as well as quantitative literatures, the report classifies the variety of prevailing pricing strategies under five major themes: a) General Pricing b) Quality- based Pricing iii) Query-Based Pricing iv) Privacy-Based Pricing v) Special Cases. Thereby, the report creates a platform for prescribing data pricing formulaes.

Key words: data market, data pricing, public policy, systematic literature review

1.0 INTRODUCTION

Data is intangible, like a service, but can easily be stored and delivered away from place of production, like a good (Wdowin and Diepeveen,2020). In recent times, with rapid digitization of the economy and the evolution of the big data market, the availability and volume of data is growing. The world is producing about 2.5 quintillion bytes of data per day, with ninety percent of all data having been produced in just the last two years (PWC,2019). Exponential growth of data leads to classification of data from various perspectives; representation, source, informational content, usage, method of generation (Statistics Canada Report,2019; PWC Report,2019; Swedish National Board of Trade,2014).

Data possess the potential for generating revenue and hence pricing of data or valuation of data which is the focus of the current literature review has garnered much attention among scholars. Data pricing basically means treating data as an economic good and placing a value on data's worth. Putting formally, data pricing occurs when the data owners give each dataset a reasonable price in order to push those datasets into digital markets (Liang et al.,2017). If a standard model for data pricing existed – one which considered many aspects of value such as the age of the data, the reliability of the sample, and other factors – sellers would be able to price optimally in the market and buyers could make appropriate comparisons across data service providers to get a fair price (Heckman et al.,2015).

The report attempts to capture meaningful information from the bulk of available literatures using a systematic literature searching methodology. The relatively unexplored area of data pricing is explored to provide some examples about how different pricing models are formulated and henceforth paves the way for future research. The analysis is based on an aggregate of 28 papers concerning data pricing. While some papers deal with theoretical background for pricing data, others have portrayed the varied ways for evaluating worth of data. The novelty of the report lies in the attempt to categorize the vast issue of data pricing under some major heads. Therefore, the major research questions the report strives to answer can be listed as follows:

(1) What are the prevailing data pricing options in the economy?
(2) What are the major dimensions guiding data pricing currently?
(3) How can currently used data pricing strategies be applied to real-world data-sets?

Three major factors can be identified which necessitates the need for pricing or valuing data. The first factor is pricing of data would play a pivotal role in building up an effective market or platform for data trading, which in turn is likely to establish data as a tradable commodity. Second, valuation of data is critically dependent on a number of factors like origin, quality, frequency of usage etc. Thus, developing a transparent and rigorous model of pricing is of utmost importance. Finally, targeted marketing strategy often collects data from potential customers to ameliorate their performance and often the data collected is referred to as personal data[1]. With increasing awareness about personal data, data which have been earlier collected free of cost can no longer be acquired in a similar fashion and hence suitable compensation mechanisms need to be designed to compensate the customers for their privacy loss.

According to OECD, the global volume of data is forecasted to increase by 40% from 2015-2020. Thus, arise the need to exploit the growth in the volume of data and hence data pricing becomes a pertinent issue. According to the estimates of the European Data Market study, the value of the data economy in 2016 was worth nearly 2% of the European GDP. By 2020, the data economy is expected to increase to €430 billion with an overall impact of 2.5% on the GDP in the baseline scenario, which is defined by a continuation of the positive but moderate growth trend of the economy (European Data Market Study,2017). Data valuation has become a prominent phenomenon with a wide range of players customising models for valuing data and generating profit. Players range from corporations gathering data on their products and services, organizations collecting data from the target population and also third-party data aggregators. Hence the data markets prevailing nowadays are vertical and restricted.

Data market structures play a vital role in determining data price and just like for any tangible good, the market structures for data are identified as; monopoly, oligopoly and competition. The next stage of data pricing is identifying the appropriate pricing strategy. Muschalle et al., (2012) organizes the different types of data pricing strategies into the following six categories - a) Free Data Pricing Strategy b) Usage-Based Pricing Strategy c) Package Pricing Strategy d) Flat Pricing Strategy e) Two-part Tariff Strategy f) Freemium Strategy. Data market structure along with the pricing strategies determine the type of data pricing model to be employed. Existing data pricing models are classified under two major

---

[1] Personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.(Available at : https://gdpr.eu/eu-gdpr-personal-data/)

heads: (i) Economic-Based Pricing Models and (ii) Game Theory-Based Pricing Model (Liang et al.,2018). Economic Based Pricing Models are guided by economic principles whereas Game theory Based pricing models functions according to the canons of game theory. Again Economic-Based Pricing models are further classified as- Cost Model, Consumer Perceived Value, Supply Model, Demand Model, Differential Pricing and Dynamic Data Pricing. Game Theory- Based Pricing is subdivided as Non-Cooperative Game, Bargaining Game and Stackelberg Game.

Data pricing faces constraints due to intangible, non-rivalrous nature of data and because of the ability to generate positive externalities. In addition, the presence of diverse data sources, complexity of data management, diversity of data act as impediments in pricing data efficiently. The goal of the literature review is to delve deeper into the various pricing strategies adopted over the years to value data.

The report is organized as follows. In Section 1, we present a short introduction about the data market and basics of data pricing. Section 2 takes us through the details of the research methodology adopted and Section 3 concentrates on the analysis of the literatures. Section 4 provides a framework for pricing data via examples and Section 5 concludes the report.

2.0   REVIEW METHODOLOGY

The report identifies the wide variety of approaches employed to price or value data in order to classify them according to themes and to derive the research gaps and avenues for future research. A literature review appears to be a reasonable first step approach, as an essential step in structuring a research field, providing support in identifying the conceptual content of the field and offering guidance towards theory development (Easterby-Smith et al., 2012; Meredith, 1993). In the report, we follow the three-step literature review methodology given by Mangiaracina et al. (2015a): paper collection and selection, analysis of the selected literature and identification of research gaps and potential areas for further investigation. Post literature review, we develop pricing models based on the parameters affecting the demand for data.

For data collection, we focus on databases like IEEE Xplore, ACM digital library, Springer, Science Direct, Research Gate, recognized as the largest bibliographic databases. Websites of Illinois and Washington university are also searched. Since the study centres around data pricing, papers relevant to the issue are selected. The keywords used are 'economic pricing model for data', 'how to price data', 'data pricing models' and 'pricing policies for

data'. After the selection of the first set of papers, the second set of papers are obtained by the application of snowballing approach.

We search for these set of keyword combinations in the title, abstract and keywords of article and conference papers belonging to these databases and websites. The initial search yields 66 publications. We filter them to 28 based on the scope of the article. We exclude articles which focus on electronic business, cloud computing, wireless networks, mobile data, web enabled application services, Internet of Things, e-commerce.

2.1. Descriptive statistics

The concerned section provides various descriptive statistics for the reviewed literature. Fig. 1 depicts the distribution of publications by year starting from 2000 to 2020. The graph shows number of publications remain more or less constant over the period 2000-2012, achieves a peak during 2013 and then records a drop during 2020. A fluctuating trend is present over the period 2014-2020. Given the rate at which the data economy is growing, literatures on the topic are likely to rise.
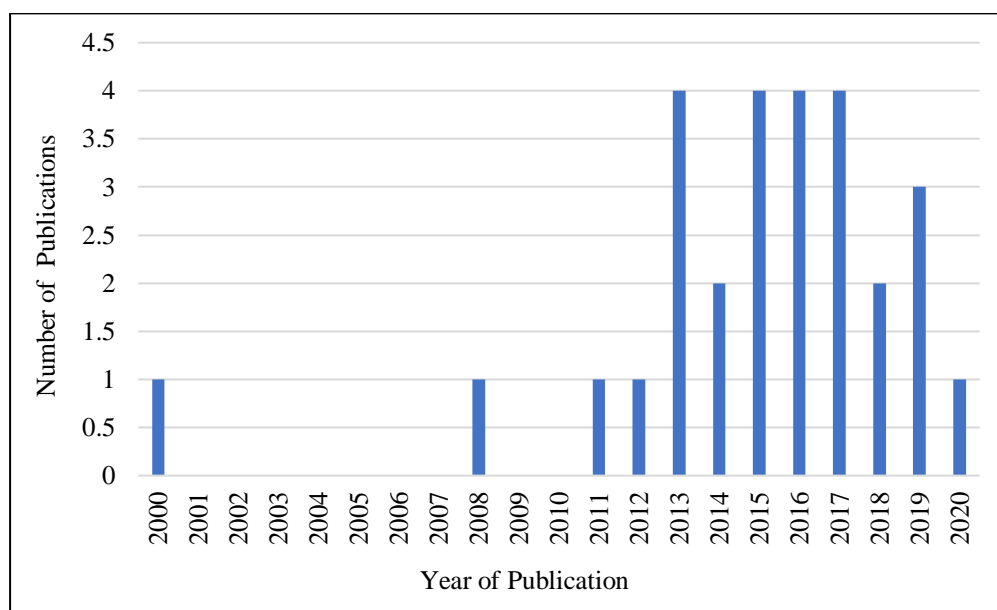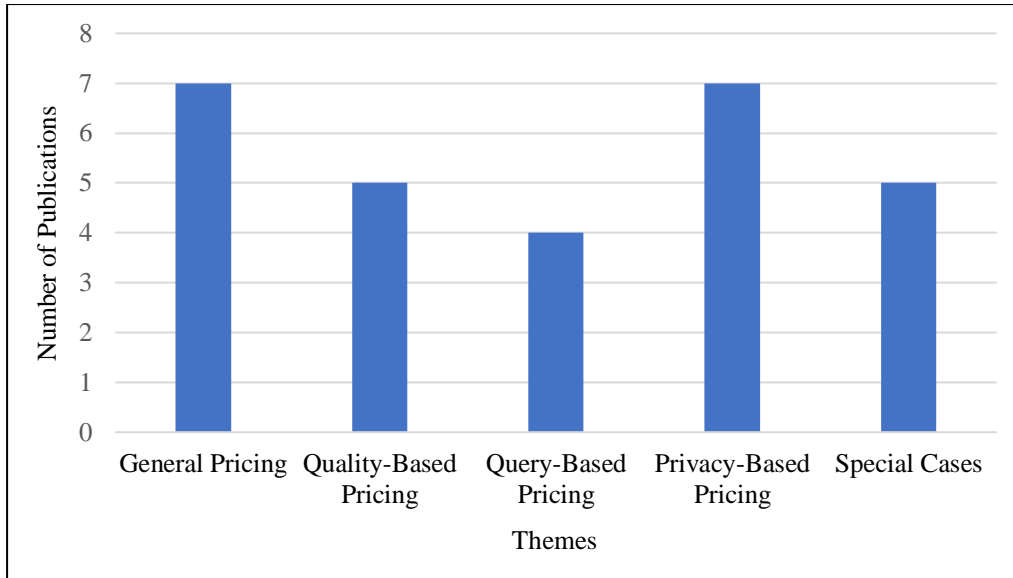

Fig. 1.   Publication by year

Fig. 2.   Publication by themes

Fig. 2 portrays the most widely discussed themes in the reviewed literature. We classify our literature into five major themes based on the dimensions which form the basis of data pricing. General pricing refers to the literatures which provides an overall idea about data pricing. Quality-Based pricing encompasses all the literatures which give importance to the quality of data. Query-Based pricing highlights on the pricing techniques based upon the number of views or queries generated. Privacy-based pricing looks at the pricing techniques for personal data. Special cases dealt with some specific pricing techniques, like the pricing technique adopted by Airbnb

2.2 Publication By Research Objectives

Data measurability plays a pivotal role in data valuation and exchange as economic goods and as a part of data handling on the data infrastructure platforms (Demchenko et al., 2018). Literatures on the particular issue are either review studies examining the already existing pricing strategies or deals with pricing strategies for specific datasets referred to as 'private data' and some even investigate the role of quality in data pricing. We make an attempt to categorize the publications according to the research objectives studied. We again classify the research objectives into subthemes, which centres around whether the valuation strategy is from the side of the data vendors or the data owners or the data market as a whole etc. Table-I shows the two majorly undertaken objectives are valuing private or personal data and examining pricing models for specific type of data, followed by data valuation. The observation is supported, primarily because for personal data, the recently enacted GDPR mandates

multiple measures to protect personally identifiable data, in particular informed consent (Custers et al., 2014) which should build further confidence and trust in digital technologies and build trusted environments for data exchange via the future data market (Osimo et al.,2017).Under the objective of valuing private data, we note majority of the studies focus on the perspective of the data owners for estimating the value of data. The rest of the studies deal with the dimensions of utility, data vendors, data attributes and data market. Under the objective of examining the pricing models for specific data type, studies consider the structure of the data market while assessing data value. Studies revolving around data valuation concentrate on the data market and data vendors, while those dealing with pricing data on the internet, takes into account the number of views, data producers as well as the data market. Literatures reviewing existing data pricing models provide a categorization of the models, while studies which delve into the practice of price discrimination in the data market, point out three major subthemes as discounting, willingness to pay and data market. Finally, for the objective of understanding the impact of quality, the two major ways are either to view from the utility framework or to investigate the importance of versioning.

Table 1: Classification of literature by research objectives studied

Objective Studied	Sub Themes	Publications	Number
Valuing private or personal data	Utility framework	Mehta et.al, (2019)	1
	Data owners	Li et.al, (2014) Lia & Raghunathan et.al, (2013) Gkatzelis et.al, (2015)	3
	Data Vendors	Jaisingh et.al, (2008)	1
	Data Attributes	Shen et.al, (2016)	1
	Data Market	Yang& Xing, (2019)	1
Examine pricing models for specific data type	Earth Observation Data	Harris, (2000)	1
	Dynamic Pricing	Ye et.al, (2018)	1
	Information Based	Li et.al, (2017)	1
	Data Market	Golrezaei&Nazerzadeh, (2014) Zheng et.al, (2017)	2
	Data Vendor	Golrezaei&Nazerzadeh, (2014)	1
Data valuation	Data Market	Heckman et.al, (2015) Schomm et.al, (2013)	2
	Data Vendors	Muschalle et.al, (2013) Kushal et.al, (2011) Schomm et.al, (2013)	3
Framework for pricing data on Internet	View Based	Koutris et.al, (2015)	1
	Data Producers	Li et.al, (2012) Bergemann & Bonatti, (2015)	2
	Data Market	Balazinska et.al, (2013)	1
Review of existing data pricing methods	Classification of models	Zhang & Beltran, (2020) Fricker&Maksimov, (2017) Liang et.al, (2018)	3
Practising price discrimination	Discounting	Tang et.al, 2016	1
	Willingness to Pay	Stahl &Vossen, (2016)	1
	Data Market	Bataineha et.al, (2016)	1
Impact of Quality on big data analysis	Utility Framework	Yang & Xing, (2019)	1
	Versioning	Yu & Zhang, (2017)	1

2.3 Publication by Research Methods

According to Wacker (1998), research methods in operations management are classified into four major categories. Based on the classification, we bucket the reviewed publications based on their research methodology. The majority of the studied literatures employ quantitative and analytical approaches like optimisation (Kushal et al., 2011; Yang et al., 2019) bi-level programming (Yu & Zhang,2017), $O(n^2)$ algorithms (Kushal et al., 2011; multiple choice knapsack problem (Stahl et al., 2016), pseduopolynomial time algorithm (Tang et al.,2016), mechanism design (Mehta et al., 2019), reverse pricing (Shen et al., 2016), regression model (Ye et al.,2018), support vector machine, decision tree and linear discriminative analysis (Li et al., 2017)and gaussian process (Zheng et a., 2017). The rest of the publications employ empirical, review or case-study /survey method. The range of literatures reviewed shows the abundance of contributions from the fields of computer science and economics which can be explained by the enormous use of big data in these fields.

Table 2: Classification of literature by research methodology

Research Methodology	Publication
Analytical, mathematical, theoretical and conceptual [22]	Mehta et al., (2019), Yang et al., (2019), Yang and Xing, (2019) Ye et al., (2018), Li et al., (2017), Yu and Zhang, (2017), Zheng et al., (2017), Bataineha et al., (2016), Stahl and Vossen, (2016), Shen et al., (2016), Tang et al., (2016), Bergemann and Bonatti, (2015) Gkatzelis et al., (2015), Koutris et al., (2015), Golrezaei and Nazerzadeh (,2014), Li et al., (2014), Balazinska et al., (2013), Lia and Raghunathan, (2013), Li and Miklau, (2012), Kushal et al., (2011), Jaisingh et al., (2008), Harris, (2000)
Empirical [1]	Heckman et al., (2015)
Review [3]	Zhang and Beltran, (2020) Liang et al., (2018) Fricker and Maksimov, (2017)
Case Study (Survey-Based) [2]	Muschalle et al., (2013) Schomm et al., (2013)
Note- The numbers in third bracket indicate the number of publications.	

3.0 ANALYSIS OF LITERATURE

Here we conduct content analysis and summarize the issues surrounding data pricing and data market through five major themes. Towards the end of the section, we build connection between these themes.

3.1. *Themes arising from the review*

All the reviewed literatures are classified into five distinct themes depending upon the data pricing strategy. The five major identified themes are i) General Pricing ii) Quality-Based Pricing iii) Query-Based Pricing iv) Privacy- Based Pricing v) Special Cases. The themes are derived based on whether the papers give a general perspective for pricing data or focus on some specific attribute of data or whether it talks about some specific type of data.

Table 3: Classification of literature by themes

Theme	Label	Publication
Theme 1	General Pricing	Heckman et al., (2015), Muschalle et al., (2013), Kushal et al., (2011), Stahl and Vossen, (2013), Zhang& Beltran (2020), Liang et al., (2018), Fricker and Maksimov, (2017)
Theme 2	Quality-Based Pricing	Yang et al., (2019), Yu and Zhang, (2017), Stahl and Vossen, (2016), Tang et al., (2016), Bataineha et al., (2016)
Theme 3	Query-Based Pricing	Koutris et al., (2015), Li and Miklau., (2012),Balazinska et al., (2013),Bergemann and Bonatti, (2015)
Theme 4	Privacy-Based Pricing	Mehta et al., (2019), Li et al., (2014), Lia and Raghunathan, (2013), Shen et al., (2016), Yang and Xing, (2019), Jaising et al., (2008), Gkatzelis et al., (2015)
Theme 5	Special Cases	Harris, (2000), Ye et al., (2018), Li et al., (2017), Golrezaei and Nazerzadeh., (2014), Zheng et al., (2017)

3.2. General Pricing

The concerned theme contains papers where the researchers mainly focus on the taxonomy of the data market and the various prevailing strategies for data valuation. While some of the studies are based on mathematical principles, some are review works, thereby opening up the variety of applicable methodology in the arena. For instance, Heckman et al., (2015) proposes a potential dataset valuation model in order to develop the data market and reduce the issue of 'lemon' market asymmetry. Zhang and Beltran (2020) provide a comprehensive review of the existing data pricing strategies and group the pricing strategies according to the fundamental properties of data to be priced. Work in similar lines, conducted by, Fricker and Maksimov (2017), points out the research gaps in context of data pricing.

3.3. Quality-Based Pricing

The theme mainly deals with papers which investigate the impact of quality on the price of data and the involved studies are primarily quantitative in nature. Yang et al., (2019) introduce the notion of utility of quality in order to derive optimal prices for the data and even highlight upon the impact of quality on big data analysis. Yu and Zhang (2017) provide an insight about the appropriateness of multi-version strategy for the data market where multi- versioning is done with respect to dimensions of data quality. Considering completeness as a quality dimension, Tang et al., (2016) focus on the trade-off between data quality and price. Bataineha et al., (2016) put forward a two-sided market theory for data and opine presence of a broker to be beneficial for the efficient functioning of the data market.

3.4. Query-Based Pricing

Here, researchers highlight how information or data in the internet are priced based on the number of queries generated in relation to the specified information. Most of the literatures of the cluster, formulate] pricing functions for deriving prices. Li and Miklau (2012) propose an interactive pricing scheme for the data market where the price function ought to be non-disclosive, arbitrage free and regret free. Koutris et al., (2013) put forward stipulated prices of some queries and the buyer here purchases the data by expressing their queries. Bergemann and Bonatti (2015) proves cookie-based pricing can serve as an optimal mechanism while selling information.

3.5. Privacy-Based Pricing

The theme deals with the various mechanisms to price personal/ private data, which mainly employs modelling frameworks to obtain prices. The theme deals with how private data are valued and also talks about the compensation to be rewarded to the data owners due to their privacy loss. Mehta et al., (2019) in their work introduces the notion of private valuation of a buyer and derive an optimal price-quantity schedule for datasets. Lia and Ragunathan (2013) studies the trade-off between privacy and data utility from the perspective of the data owner. Shen et al., (2016) looks at big personal data and examine how value of a data tuple is dependent on information entropy, citations etc. Gkatzelis et al., (2015) attempt to set a price for data by maximizing the data buyer's interest in the market and the amount of data traded. Thus, the studies from the cluster imply data priced by the above strategy is likely to be slightly higher due to the presence of the compensation component.

3.6. Special Cases

The above-mentioned theme basically aggregates the pricing issues of specific datasets like earth observation data (Harris,2000), traffic data (Golrezaei and Nazerzadeh,2014). Ye et al., (2018) throws some light upon the dynamic pricing strategy of Airbnb. The paper talks about how Airbnb uses regression model to help hosts to set an optimal price for the entire selling period. Li et al., (2017) puts forward a data pricing metric in order to value data based on the amount of information they possess. Finally, Zheng (2017) concentrate on mobile crowd sensed data market to design a query-based data pricing mechanism called ARETE. However, for datasets like agricultural data or health data specific pricing mechanisms has not been prescribed, thereby opening up scopes for further research.

4.0   CONCLUSION

Given the plethora of discussions, there exists various motivating factors for carrying out research on the issue. Primarily there has been a growing need to value data as a 'corporate asset' and also to address the issue of data transactions (Heckman et al.,2015). Second, data valuation being a subjective issue, quality is likely to have an impact on the price. Yang et al., (2019), proposes a pricing model based on utility of quality to derive optimal prices. Zheng et al., (2017) deals with the data pricing problem considering two dimensions of data quality;

multidimensionality and interaction between dimensions. Third, there are attempts in literatures to price data based upon the number of 'queries' or 'views' generated. Koutris et al., (2015) proposes a framework for pricing data on Internet, which given the price of a few views, allows the price of any query to be derived automatically. Finally, a cluster of literatures tries to link the dimension of privacy with the price of data. Mehta et al., (2019) in their work introduces the notion of privacy by considering valuation of the buyer for an ideal data record to be private.

In contrast to other recent literature reviews, we focus on the variety of prevailing pricing strategies by classifying them under five major themes: a) General Pricing b) Quality-based Pricing iii) Query-Based Pricing iv) Privacy-Based Pricing v) Special Cases. The work takes into account all literatures (qualitative as well as quantitative), discussion papers and literature reviews concentrating on estimating the value of data. The report attempts to capture meaningful information from the bulk of available literatures using a systematic literature searching methodology. The report also tries to provide some examples about how different pricing models can be applied to real-world datasets and paves the way for future research.

REFERENCES

Balazinska, Magdalena., Howe, Bill., Koutris, Paraschos., Suciu, Dan., and Upadhayaya, Prasang., (2013). *A discussion on Pricing Relational Data*. University of Therehington

Bataineha, Ahmed Saleh., Mizounib, Rabeb., Barachic, May El., and Bentahar, Jamal., (2016). Monetizing Personal Data: A Two-Sided Market Approach. *Procedia Computer Science*, Volume.83: 472-479.

Bergemann, Dirk., and Bonatti, Alessandro., (2015). Selling Cookies. *Cowles Foundation Discussion Paper* No. 1920R.

Fricker, Samuel. A., and Maksimov, Yuliyan V., (2017). Pricing of Data Products in Data Marketplaces. *International Conference on Software Business*.

Gkatzelis, Vasilis., Aperjis, Christina., and Huberman, Bernardo A. (2015). Pricing Private Data. New York University.

Golrezaei, Negin., and Nazerzadeh, Hamid., (2014). Pricing Scheme for Metropolitan Traffic Data Markets. *In Proceedings of 3rd International Conference on Data Management Technologies and Applications*: 266-271

Haifei, Yu., and Mengxiao, Zhang., (2017). Data pricing strategy based on data quality,

Computers & Industrial Engineering, Volume 112:1-10

Harris, Ray., (2000). Data Pricing Policy. *International Archives of Photogrammetry and Remote Sensing.* Vol. XXXIII, Part B6:341-350

Heckman, J.R., Boehmer, E.L., Peters, E.H., Davaloo, M., Kurup, N.G. (2015). *A Pricing Model for Data Markets*. In iConference 2015 Proceedings.

Jaisingh, Jeevan., Barron, Jack., Mehta, Shailendra and Chaturvedi, Alok., (2008). Privacy and Pricing Personal Information. *European Journal of Operational Research*.187(3):857-870

Koutris, Paraschos., Upadhayaya, Prasang., Balazinska, Magadalena., Howe, Bill., and Suciu, Dan., (2015). Query-Based Data Pricing. *Journal of the ACM*. Volume.62, Issue.5:1-44

Kushal, Avanish., Moorthy, Sharmadha., and Kumar, Vikash. (2011). *Pricing for Data Markets*. University of Washington.

Lia Bai, Xiao., and Raghunathan, Srinivasan., (2013). Pricing and Disseminating Customer Data with Privacy Awareness. *Decis Support Syst*.59: 63–73.

Li, Chao., and Miklau, Gerome., (2012). Pricing Aggregate Queries in a Data Marketplace. *Journal of the ACM*.

Li, Chao., Li, Yang Daniel., Miklau, Gerome., and Suciu, Dan., (2014). A Theory of Pricing Private Data. *ACM Transactions on Database Systems*. Volume.39, Issue. 4:1-28.

Li, Xijun., Yao, Jianguo., Liu, Xue., and Guan, Haibing., (2017). A First Look at Information Entropy- Based Data Pricing. *IEEE 37th International Conference on Distributed Computing Systems*.

Liang, Fan., Yu, Wei., An, Dou., Qingyu, Yang., Fu, Xinwen., and Zhao, Wei., (2018). A Survey on Big Data Market- Pricing, Trading and Protection. IEEE; Special Section on Privacy Preservation for Large-Scale User Data in Social Networks, Vol.6:15132-15154

Mehta, Sameer and Dawande, Milind and Janakiraman, Ganesh and Mookerjee, Vijay. (2019). *How to Sell a Dataset? Pricing Policies for Data Monetization*. Information Systems Research, Forthcoming

Muschalle, Alexander., Stahl, Florian., and Vossen, Gottfried. (2013.) *Pricing Approaches for Data Markets*. M. Castellanos, U.Dayal, and E.A. Rundensteiner (Eds.): BIRTE 2012, LNBIP 154: 129–144.

Niyato, Dusit., Mohammad, Alsheikh Abu., Wang, Ping., Kim, Dong In., and Han, Zhu.,(2016). Market Model and Optimal Pricing Scheme of Big Data and Internet of Things (IoT). *IEEE International Conference on Communications (ICC)*: 1-6.

PWC Report. (2019). *Putting a value on data.* Available at: https://www.pwc.co.uk/data-analytics/documents/putting-value-on-data.pdf.

Schomm, Fabian., Stahl, Florian., and Vossen, Gottfried. (2013).Market Places for Data: An Initial Survey. *SIGMOD Record*, Vol. 42, No. 1:15-26

Shen, Yuncheng., Guo, Bing., Shen, Yan., Duan, Xuliang., Dong, Xiangquian., and Hong, Zhang., (2016).A Pricing Model for Big Personal Data. *Tsinghua Science and Technology*.Vol,21; No 5 :482-490.

Stahl, Florian., and Vossen, Gottfried.,(2016). Fair Knapsack Pricing for Data Marketplaces. J. *Pokorn´ et al. (Eds.): ADBIS 2016, LNCS 9809*: 46–59,

Tang, Ruiming., Amarilli, Antoine., Senellart, Pierre., and Bressan, Stephane., (2016).A Framework for Sampling-Based XML Data Pricing. *A. Hameurlain et al. (Eds.): TLDKS XXIV, LNCS 9510:* 116–138,

The economic value of data: discussion paper. Available at: https://www.gov.uk/government/publications/the-economic-value-of-data-discussion-paper.

Wdowin and Diepeveen,(2020). The value of data. Bennett Institute for Public Policy.

Yang, Jian., Zhao, Chongchong., and Xing, Chunxiao. (2019). Big Data Market Optimization Pricing Model Based on Data Quality. Wiley

Yang, Jian., and Xing, Chunxiao., (2019). Personal Data Market Optimization Pricing Model Based on Privacy Level. *Information*.10(4):123.

Ye, Peng., Qian, Julian., Chen, Jieying., Wu, Chen-Hung., Zhou, Yitong., Mars, Spencers De., (2018). Customized Regression Model for Airbnb Dynamic Pricing. Applied Data Science Track Paper.

Zhang, Mengxiao., and Beltran, Fernando., (2020). *A survey of Data Pricing Methods*.

Zheng, Zhenzhe., Peng, Yanquing., Wu, Fan., Tang, Shaojie., and Chen, Guihai. (2017). An Online Pricing Mechanism for Mobile Crowdsensing Data Markets. Proceedings of the 18th

ACM International Symposium on Mobile Ad Hoc Networking and Computing. Article.6(July):1-10