Volume XV, Issue (1) (2019) 68-75

International May Conference on Strategic Management

COMMERCIAL OPEN DATA APPLICATIONS: DATA PROCESSING METHODS AND PRICING STRATEGIES

Minna Jukka [1], Janne Saarikoski [2], Mikko Lampi [1]

[1]South-Eastern Finland University of Applied Sciences, Finland;
[2]Kotka Maritime Research Centre, Finland

Abstract: Open data released by governments and public administrations to promote democracy could also be a source for new business opportunities. Open data have a specific mission to create wide benefits for the whole society. Despite the increasing economic impact of open data applications there remains a lack of evidence on the ways these applications actually process data and create revenue. This research investigates whether the data processing method and the business models have any connection in the current commercial open data applications (both mobile and web) using a sample (n=64) of Finnish applications. The applications were studied in detail using inductive qualitative analysis. The data processing methods were classified into six inductively formed themes of visualization, data integration, personalization, traditional analytics, semantic search, and predictive analytics and artificial intelligence. According to the results, most applications used combinations of data personalization, data integration, and visualization as data processing methods. The revenue in these low computing applications was based on ads or price differentiation. Traditional analytics, semantic search, predictive analytics and artificial intelligence were less common in the sample. The results showed that advanced computing methods are related to direct product pricing. The evidence from this study suggests that the use of open data in applications gives more added-value for companies with data analytics.

Keywords: Applications, Open Data, Data processing methods, Business models, Pricing strategies

1. INTRODUCTION

Organizational knowledge management processes include acquisition, organization, transferring, sharing and utilization of knowledge [1]. Thus, the exploration and exploitation of any external information can be used as a data resource in knowledge management systems [2]. This also includes external datasets like open data. However, recent evidence has shown that open data is not considered as a business knowledge resource in software development companies [3].

The business potential of open data is huge. Annual potential economic value of open data is estimated to be around $3 trillion to $5 trillion including industries like education, transportation, consumer products, electricity, oil and gas, health care, and consumer finance [4].

Previous research has established that big data has created new business models, energy and cost savings, open innovations, and new understandings that improve decision-making [5]. Open data even amplifies the positive impact of big data by generating additional economic benefits such as innovative new products and services [4].

Managing big data assets as open data can have a positive impact on their use and value. For instance, open licensing and availability via standard web-based APIs can achieve the above-mentioned benefits.

According to the knowledge-based view of the firm, knowledge is the most important resource in gaining a competitive advantage [6]. In this light, it is surprising that research has ignored to study open data as a source of knowledge and competitive advantage in business. Meanwhile, the business models of big data have received some attention [7].

In general, only few studies [5, 8, 9, 10] have been carried out on data business models. Consequently, there remains a lack of evidence on the ways these applications actually process data. In addition, research on applications' revenue creation has received little attention. For that reason, this research aims to answer the following research question: What is the connection between data processing methods and product pricing in the current commercial open data applications?

This paper is structured as follows: the next section introduces the theoretical background of this study and the concept of open data. The third section sets out the research methodology followed by the findings. Finally, the conclusions are drawn in the fifth section.

2. OPEN DATA

Open data is free access data released by a public sector e.g. governments and municipalities [9, 10]. According to the definition of open data open data is based on six principles:

1) The data must be public domain
2) it must be technically accessible and usable
3) access to the data must be free of charge
4) there must be a license that permits reuse of the data
5) the data should be publicly findable
6) the data must be understandable.

For instance, a PDF file is widely used but not technically very usable. Machine-readability is often emphasized in technical accessibility. Simply making a data set available with an open license does not make it open data by default. There needs to be documentation or metadata to describe the data set and its meaning and structure.

In practice, the level of openness and how well the data meets these principles can vary. Some sources even claim that open data should be the unmodified raw data [11]. Typically, open datasets are provided by public sector, and include for example public transport timetables, weather, epidemic, statistics, budgeting, and tourist information data [9, 11].

Open data is public and purposeful data. It is often released with a specific purpose in mind, and it has a specific mission. This mission reflects the mission of the open data movement to make available free, open, transparent data for a variety of uses and for a large group of users [12]. Fostering service and product innovation and economic growth are some significant motivations for opening data [11, 13].

Multiple government open data programs emphasize the potential for new businesses, especially for IT companies [13]. However, it should be noted that open data does not transform into business value automatically. It requires participation by third parties, such as businesses and software developers [11]. Open data value networks are linked data value chains. Linked data value chains contain raw data providers, linked data providers, linked data application providers and end users [14].

Open data enable the use of multiple business models (Table I). The main revenue sources of open data are consulting, raw data transformation into other kinds of structured data, and open data-based application development [14].

Table 1. Open data business models [8]

1 Freemium (free product)
2 Premium (top class product, chargeable)
3 Dual licensing
4 Support and service
5 Charges for changes (billing to convert data to open format)
6 Increases quality through participation (involvement improves clients' income or reduces costs)
7 Supports primary business (open data naturally supports the company's own business)
8 Demand-oriented platform (a paid platform that provides open data)
9 Supply-oriented platform (open data users can bill through the platform)
10 Open source (open source in co-operation)
11 Sponsorship (a free product, funding from sponsors)
12 Infrastructural razor & blades (the service is sold in-expensively and earnings are made by additional products)
13 Cost avoidance (reducing the cost of publishing data)
14 Free, as branded advertising (a free product for pro-motion)
15 White-label development (the company develops a product or service that other companies can sell on their own).

The focus of this research was on application developers. These developers are typically either software companies or private persons aiming for additional revenues from application development.

In general, the software revenue models comprise of revenue sources, pricing policies, cost structures, and revenue velocities [15]. The use of traditional pricing policies (e.g. marginal cost based) is not recommended for pricing of information goods. Instead, price differentiation based on different user needs and consumers' willingness to pay are better alternatives for information goods [16]. This demands personalized products and pricing, and product versioning based on customer needs and customers' willingness to pay [16].

Information goods are experience goods, meaning the users need to have an experience of the utility of the goods before they can assess how much they are willing to pay for it. Thus, versioning of the goods should also include free versions as sales promotion for the potential customer groups with zero willingness to pay at the current time [16].

3. DATA AND RESEARCH METHODOLOGY

The sample of 93 Finnish open data-based web and mobile phone applications was collected from the Finnish open data Databusiness.fi website [17] in March 2018. This official website, provided by the six largest Finnish cities, presents open data showcases and collects together examples of using open data for applications, services, and other purposes. It is created to accelerate data-based business in Finland and especially promote the use of open data in business.

The sample was a comprehensive sample of Finnish open data applications publicly available on the Internet. The inclusion criteria were that the applications were created in Finland or owned by Finnish firms. The demo versions and non-commercial applications were excluded from the sample. As a result, 64 applications created for business were included in the analysis (Fig.1).

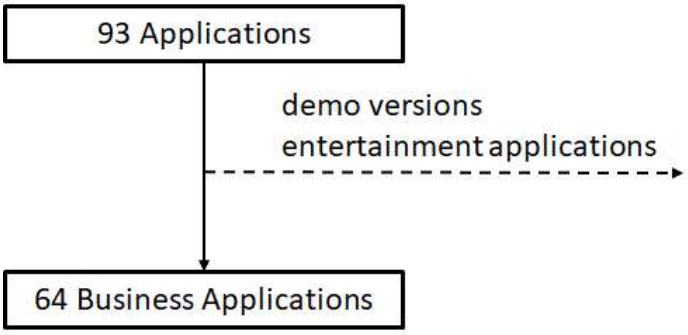

Figure 1. Sample collection procedure

The applications and the related websites were studied in detail using thematic qualitative analysis. This analysis method is widely used for identifying, analysing and reporting patterns as themes and is very useful for summarizing key features of datasets [18]. This study used analytic induction [19] as coding strategy.

Two researchers classified the applications based on the role of data in processing into six inductively formed themes: 1) visualization, 2) data integration, 3) personalization, 4) traditional analytics, 5) semantic search, and 6) predictive analytics and artificial intelligence. The latter three were defined as more advanced computing.

It should be noted that the themes are not mutually exclusive. For instance, an AI-based personalization application would be classified into the themes three and six.

The revenue models were first classified using open coding. The following 2nd level of coding categorized the revenue potential into three classes: 0 (non-profit), 1 (moderate), and 2 (high). The applications categorized as non-profit were free applications developed to demonstrate the company's skills for sales promotion or brand building.

The earning potential of ad-funded applications is still modest in Finland, so applications using only ads for revenue were classified as moderate. For paid products and

services, where the consumer's willingness to pay is higher, revenue is based directly on transactions. These were classified having a high revenue potential per product.

4. FINDINGS

A variety of different data processing combinations were used in the applications, see Fig.2. Most were based on combinations of personalization (46 out of 64), data integration (43 out of 64), or visualization (37 out of 64).

Personalization means e.g. combining your own location information with open data weather information, location of available services nearby, or providing open data -based price comparisons for your personal assets (e.g. house, car).

Pricing in these low-computing intensity applications was based on ads or price differentiation, e.g. freemium as the free version and premium as the paid product with additional features [8]. In these applications, the free version (freemium) as sales promotion was widespread.

The more advanced computing was less common in the sample (traditional analytics 24 out of 64; semantic search 2 out of 64; predictive analytics and artificial intelligence 5 out of 64).

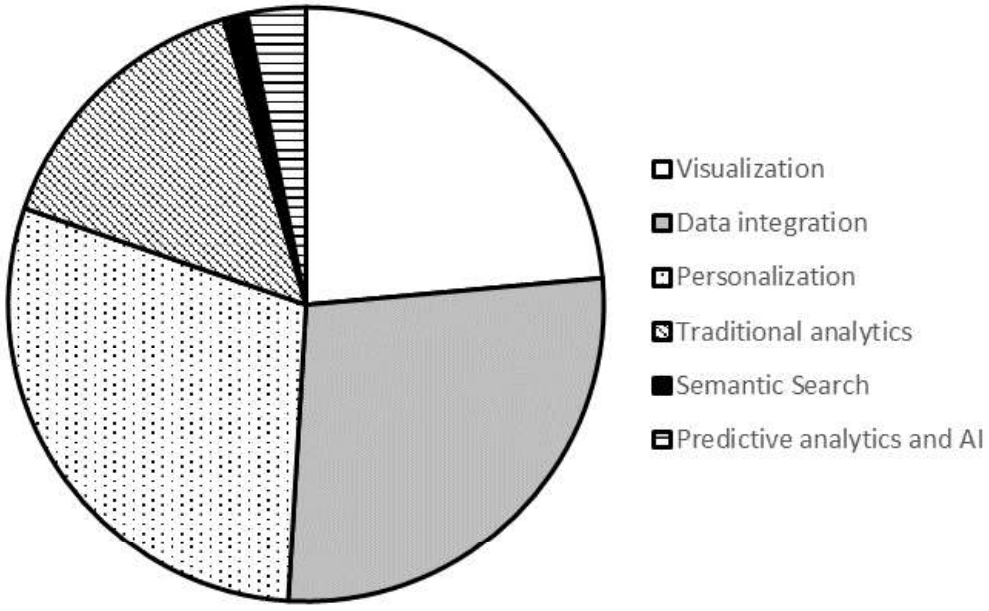

Figure 2. The role of data processing in applications

Based on the results, it seems that the increased level of data processing enabled pricing the product higher, thus assuming better earning potential (Fig.3). In such cases, open data was often integrated into other products of the company. Typically, these products were open data containing high-tech service products.

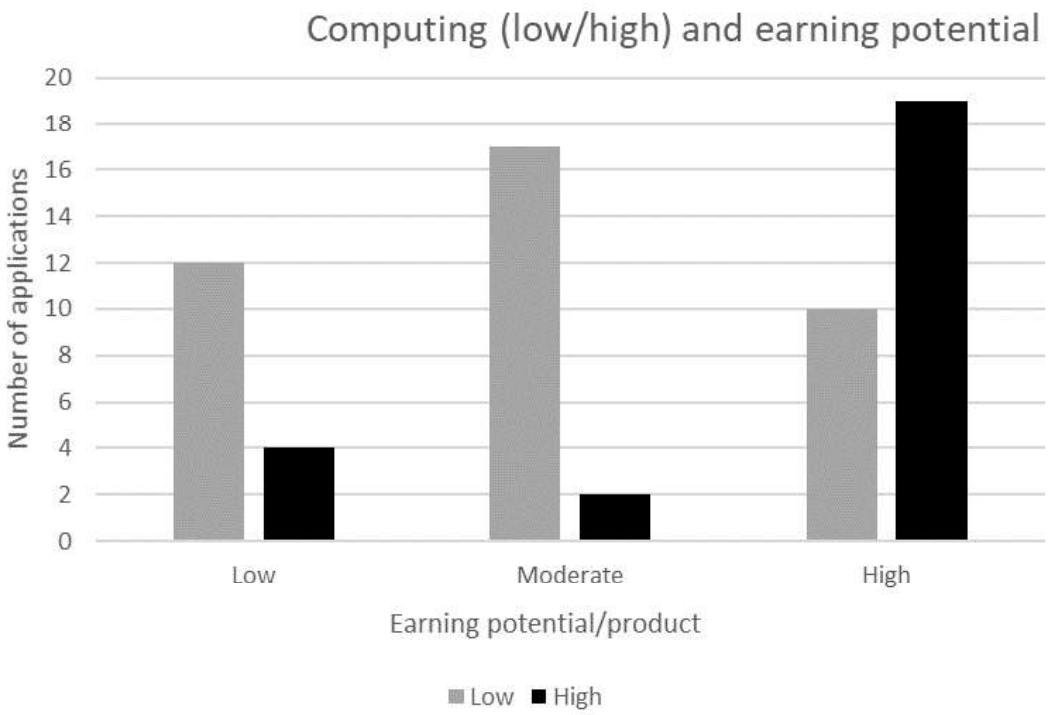

Figure 3. Advanced computing enables direct pricing

5. DISCUSSION AND CONCLUSIONS

This research investigated what is the connection between data processing and pricing in the current commercial open data applications. In response to the research question there seems to be a connection between data processing and the product pricing strategies.

It can be argued that the different roles of data processing can be combined with certain pricing models. It seems that raising the degree of data processing (e.g. data analytics) increases the possibilities for direct product pricing. Therefore, it can be proposed that the use of open data is most valuable as a business resource with advanced computing.

Although, the findings should be interpreted with caution, this study has strengths. First, it has demonstrated the common use of data personalization, data integration and data visualization in the commercial open data applications. These results indicate that added value to customers and new business can be created with fairly simple personalization, integration, visualization operations.

Second, it has demonstrated how the role of data processing and pricing are related. Hopefully, these results will encourage companies to develop more advanced computing containing applications including data analytics. These findings broadly support the evidence of earlier studies, which has shown that most attractive open data innovations for business demands data combinability with other data [20].

According to prior knowledge [3, 20, 21] the use of open data in companies is not yet primarily aimed at gaining economic benefit. As managerial implications for companies, this study has demonstrated ways to gain these benefits and new business opportunities using open data.

Utilizing open data for business helps to develop new innovative service applications and start-ups, particularly with close collaboration with open data suppliers having an ability to share high volumes of information [22]. This also demands, in addition to a good idea, business knowledge on markets and knowledge on end-users' needs [23]. Hopefully, the findings of this study inspire and encourage entrepreneurs to create novel service innovations using open data as a resource.

These findings may be limited by the small sample of Finnish commercial applications. Therefore, more research is required to develop a deeper understanding of the commercial potential of open data applications. It should also be noted that this study was limited to application providers of open data. The open data value networks, like other kinds of linked data value chains, include also raw data providers, linked data providers, and end-users [13]. These were not included in the material of this study.

In the future, it would be important to explore the whole open data business ecosystems including all the above-mentioned actors in these systems.

REFERENCES

1. Bhatt G.D., Knowledge management in organizations: examining the interaction between technologies, techniques, and people, Journal of Knowledge Management, 5 (1), (2001), 68-75.
2. Alavi M., Leidner D.E., Knowledge management and knowledge management systems: conceptual foundations and research issues, MIS Quarterly, 25 (1), (2001), 107-136.
3. Herala A., Kasurinen J., Vanhala E., Views on open data business from software development companies, Journal of Theoretical and Applied Electronic Commerce Research, 13(1), (2018), 91-105.
4. Manyika J., Chui M., Farrell D., van Kuiken S., Groves P., Doshi E.A., Open data: Unlocking innovation and performance with liquid information. McKinsey Global institute, (October 2013). Retrieved Apr 10, 2018 from https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information
5. Ylijoki O., Porras J., Conceptualizing big data: analysis of case studies, Intelligent Systems in Accounting Finance & Management, 23(4), (2016), 295-310.
6. Foss N.J., Knowledge-based approaches to the theory of the firm: Some critical comments, Organization Science, 7(5), (1996), 470-476.
7. Tian X, Big data and knowledge management: a case of déjà vu or back to the future? Journal of Knowledge Management, 21(1), (2017), 113-131.
8. Zeleti F.A., Ojo A., Curry E., Business models for the open data industry: characterization and analysis of emerging models, In Proceedings of the 15th Annual International Conference on Digital Government Research. ACM New York, NY, USA, 2014.
9. Lindman J., Kinnari T., Rossi M., Industrial open data: case studies of early open data entrepreneurs, In Conference proceedings of the 47th Hawaii International Conference on System Science HICSS '14. IEEE Computer Society Washington, DC, USA, 739-748.

10. Danneels L., Viaene S., van den Bergh J., Open data platforms: Discussing alternative knowledge epistemologies, Government Information Quarterly, 34, (2017) 365-378.
11. Janssen M., Charalabidis Y., Zuiderwijk A., Benefits, Adoption Barriers and Myths of Open Data and Open Government, Information Systems Management, 29(4), (2012), 258-268.
12. Gurin J., Open Data Now: The secret to hot startups, smart investing, savvy marketing, and fast innovation, McGraw-Hill Education, New York, NY, 2014.
13. Huijboom N., Van den Broek T., Open data: an international comparison of strategies, European Journal of ePractice 2011, 12(1), (2011), 1-13.
14. Latif A., Us Saeed A., Hoefler P., Stocker A., Wagner C., The linked data value chain: a lightweight model for business engineers, In Proceedings of I-KNOW '09 and I-SEMANTICS '09. 5th International Conference on Semantic Systems, Graz, Austria, 2009, 568-577.
15. Rajala R., Determinants of business model performance in software firms, Helsinki School of Economics. Acta Universitatis Oeconomicae Helsingiensis, (2009) A-357.
16. Shapiro, C., Varian, H.R., Information rules: a strategic guide to the network economy, Harvard Business School Press, Boston, Massachusetts, 1998.
17. Open data showcases. Databusiness.fi, Making our open data your business. Retrieved March 15, 2018 from https://www.databusiness.fi/en/front-page/
18. Braun, V., Clarke, V. Using thematic analysis in psychology, Qualitative Research in Psychology, 3, (2006), 77-101.
19. Ryan, G.W., Bernard, H.R., Data management and analysis methods, in The Handbook of Qualitative Research (Denzin N.K. and Lincoln Y.S.), Sage Publications, Thousand Oaks, CA, 200, 769-802.
20. Susha, I., Grönlund, Å, Janssen, M., Driving factors of service innovation using open government data: An exploratory study of entrepreneurs in two countries, Information Polity, 20, (2015), 19-34.
21. Kitsios, F., Kamariotou, M., Open data hackathons: an innovative strategy to enhance entrepreneurial intention, International Journal of Innovation, 10, 4 (2018), 1757-2223.
22. Stone, M., Aravopoulou, E., Improving journeys by opening data: the case of Transport for London (TfL), The Bottom Line, 31, 1, (2018), 2-15.
23. Donker, F.W., van Loenen, B., How to assess the success of the open data ecosystem? International Journal of Digital Earth, 10, 3, (2017), 284-306.