WMG Service Systems Research Group
Working Paper Series

Big Data, Brand Loyalty, and Business Models: Accounting for Imprecision and Noise in Consumer Preferences

Ganna Pogrebna

THE UNIVERSITY OF WARWICK

ISSN: 2049-4297
Issue Number: 06/15

About WMG Service Systems Group

The Service Systems research group at WMG works in collaboration with large organisations such as GlaxoSmithKline, Rolls-Royce, BAE Systems, IBM, Ministry of Defence as well as with SMEs researching into value constellations, new business models and value-creating service systems of people, product, service and technology.

The group conducts research that is capable of solving real problems in practice (ie. how and what do do), while also understanding theoretical abstractions from research (ie. why) so that the knowledge results in high-level publications necessary for its transfer across sector and industry. This approach ensures that the knowledge we create is relevant, impactful and grounded in research.

In particular, we pursue the knowledge of service systems for value co-creation that is replicable, scalable and transferable so that we can address some of the most difficult challenges faced by businesses, markets and society.

Research Streams

The WMG Service Systems research group conducts research that is capable of solving real problems in practice, and also to create theoretical abstractions from or research that is relevant and applicable across sector and industry, so that the impact of our research is substantial.

The group currently conducts research under six broad themes:

- Contextualisation
- Dematerialisation
- Service Design
- Value and Business Models
- Visualisation
- Viable Service Systems and Transformation

WMG Service Systems Research Group Working Paper Series
Issue number: 06/15
ISSN: 2049-4297
April 2015

Big Data, Brand Loyalty, and Business Models: Accounting for Imprecision and Noise in Consumer Preferences

Pogrebna, Ganna
Service Systems Group, Warwick Manufacturing Group
University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK
E-mail: G.Pogrebna@warwick.ac.uk

If you wish to cite this paper, please use the following reference:

Pogrebna G (2015) Big Data, Brand Loyalty, and Business Models: Accounting for Imprecision and Noise in Consumer Preferences, *WMG Service Systems Research Group Working Paper Series*, paper number 06/15, ISSN 2049-4297.

Abstract

This paper considers how context-independent data (content data) and context-dependent data (metadata) about consumer choices can capture brand loyalty and affect the creation of new business models. We find that metadata can provide more precise account of consumer preferences and more accurately predict future user choices by increasing the visibility of user context. This implies that metadata should be preferred to content data to achieve more efficient business model innovation.

Keywords: Big Data, content data, metadata, business models, case study
JEL classification: M10, M31

Big Data, Brand Loyalty, and Business Models: Accounting for Imprecision and Noise in Consumer Preferences

1. Introduction

On a daily basis, consumers (henceforth, users) generate large amounts of individual as well as household consumption data. According to IBM, in 2012 more than 2.5 exabytes (2.5 billion gigabytes) of data was generated daily.[1] By 2015 this number has grown and, according to forecasts, will continue to grow. For example, Gantz and Reinsel (2012) estimate that data created annually will increase from 1,200 exabytes in 2010 to 40,000 exabytes by 2020. Under these circumstances, businesses (henceforth, product and service providers) develop innovative techniques to extract and analyse data "*on the fly*" in order to create quick value propositions for the consumers. The availability of large masses of data catalyses the rise of the domain of data-driven business models (DDBM) which looks at how the data can be used in order to develop new and improve existing business modelling mechanisms (e.g., Hartmann et al. 2014).

Yet, the creation of meaningful analytical tools for DDBM is complicated not only because of the volume of the data but also because of the complexity of human decision processes and the way these processes are reflected in the data. Particularly, household consumption data shows that users tend to make different choices from the same closed set of products and services (e.g., Simon, 2013). For example, when making grocery purchases, users tend to alternate brands of products they choose. This is one of the reasons why current online systems developed by some providers such as, e.g., Amazon, which suggest products and services to users and which are intended to *nudge* users to purchase suggested services and goods, have not gained much popularity (e.g., Thaler, R. H., & Sunstein, 2008). One of the main disadvantages of the currently available purchasing data is that even though it allows analysts to observe consumer choices as well as provides them with useful demographic information about consumers; it is hard to tell whether observed choices are a result of consumer *true* preferences or merely a product of noise in these preferences. Analytics is particularly complicated for cases when users opt for products and services from different brands in different environments. Under these circumstances, it is important to not only pay attention to the models which help us analyse the data generated by consumer choices, but also to the types of data used for the analysis.

Recent literature on servitization and business models makes a distinction between content data and metadata (e.g., Ng, 2013). In application to user choice, content data provides an account of decisions made by the users with regard to purchasing products or services. However, this data does not provide information about the context in which these decisions were made. Content data includes Big Data and Connected Internet-of-Things (henceforth, IoT) Data as we know it. At the same time, metadata refers to the data which contains specific references to the context

[1] *See* **http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html** *for more details.*

and gives an opportunity to understand how decision architecture (features of the decision environment) affects choices made by users (e.g., Hastie and Dawes, 2010). While content data is used to create technology-based servitization mechanisms, metadata places an individual user in the centre of the servitization system. A recent study by Parry et al. (2015) shows how a servitization system which is based on human-oriented approach (e.g., Sundbo and Toivonen, 2011) can function through metadata and finds that visibility of consumer context through metadata can improve reverse supply mechanisms and create more efficient supply chains.

This paper considers whether and to what extent brand loyalty can be analysed and predicted through content data versus metadata and considers implications of obtained predictions for business models. We propose a simple framework which allows us to analyse consumer preferences in the context of daily household consumption and consider predictions which can be generated within this framework using content data and metadata from a case study. Our analysis reveals that user preferences can be better understood and predicted using metadata rather than content data which has significant impact on creating efficient DDBM.

The remainder of this paper is organized as follows. Section 2 describes research methodology which is based on the interdisciplinary approach incorporating literature from decision science, service systems, and marketing. Section 3 summarizes data and results obtained from the household case study. Section 4 analyses how obtained case study results can influence business models. Finally, Section 5 concludes with the general discussion about theoretical and practical implications of this research as well as with an account of possible limitations and future directions for the research on DDBM.

2. Research Methodology

In order to consider how content and metadata can capture brand loyalty, we propose a simple approach which stems from decision-theoretic literature on preferences, marketing literature on brand loyalty, as well as human-based servitization models. Decision science literature makes a conceptual distinction between precise, noisy, and imprecise preferences. Precise (or deterministic) preferences lie at the core of the majority of economic and decision-theoretic models. If an individual is choosing between option A and option B, such models predict which of the two options will be chosen by an individual. This individual would either prefer A to B ($A \succcurlyeq B$) or B to A ($B \succcurlyeq A$) and this precise (deterministic) preference will not change irrespective of the features of the decision environment (e.g., Kahneman and Tversky, 1979).

Experimental literature from decision science reports that when making a decision between A and B on multiple occasions people are likely to choose different options (e.g., Hey and Orme, 1994), i.e., in some cases revealing a preference for A over B ($A \succcurlyeq B$) and in others for B over A ($B \succcurlyeq A$). This finding provides evidence against precise (deterministic) preferences suggesting that preferences have a stochastic component. To explain these choices, two approaches to stochastic preferences are introduced. According to one, individuals have precise preferences over options but these preferences are distorted by noise or errors. For example, an individual prefers A to B (A≽B) but due to noise or by mistake chooses B over A (e.g., Fechner, 1966). A different approach suggests that individuals make different choices in different circumstances because they do not have precise preferences (e.g., Loomes and Sugden, 1995). In other words, if we sometimes observe A≽B and sometimes B≽A, this means that people do not have precise preferences between A and B and in different contexts these imprecise preferences will have different realizations.

Using decision-theoretic topology of preferences, we can apply concepts of precise, noisy and imprecise preferences to brand choice and propose a simple mechanism which establishes the link between the preference type and brand loyalty (see Figure 1). According to this mechanism, various offerings (products and services) can be partitioned into three categories: Green items; Yellow items and Red items. Red items include offerings for which an individual has strong precise preference: if these offering are available, an individual would always prefer these offerings to any other offerings. This means that for these offerings an individual would have high brand loyalty. Yellow items include offerings for which an individual has strong preference but this preference may be in some contexts distorted by noise: an individual would have chosen these offerings over others every time they were available, but, due to fatigue, tremble error or more sophisticated mistakes, this individual may choose other options over the offerings he or she prefers. This means that an individual would often choose the same brand but choice of other brands may also be observed. Finally, Green items are a product of imprecise preferences: an individual will purchase offerings from different suppliers and the brand loyalty will be low.

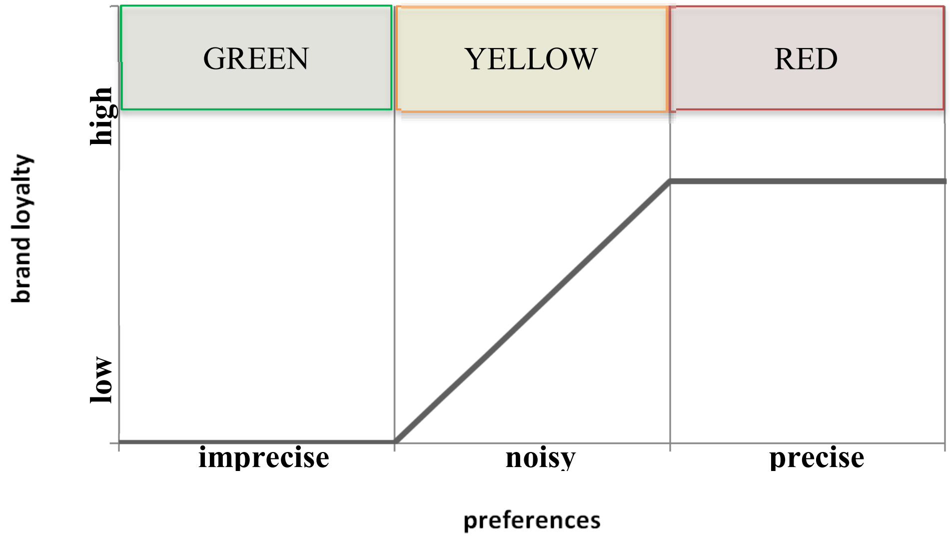

Figure 1 Preference-Brand Loyalty Mechanism

Notice that this simple approach may be very valuable for creating new business models because by accurately predicting the product or service category in user preferences (Green, Yellow, Red) it is possible to anticipate with large degree of precision which brands will be purchased in the future by the users. This mechanism is much different from the existing approaches proposed in the marketing literature. According to one of such approaches, market strength of the brand is a good predictor of brand loyalty for online purchases (e.g., Danaher et al., 2003). Another approach argues that brand loyalty can be established by carefully analysing different types of consumer satisfaction and the driving factors of satisfaction (Oliver, 1999). Villas-Boas (2004) proposes another way of predicting brand loyalty through understanding how users learn more about the product or service through experiencing it after purchase. While some of these approaches start with the brand and others start with the user, they seem to rely heavily on the data which providers have about user understanding of the brand market strength, satisfaction and experience which may be (i) difficult to collect and (ii) difficult to explain to the average user. In contrast, the proposed mechanism starts with the user and employs simple categories to identify characteristics of user preferences. Via a case study, we show how the Preference-Brand Loyalty Mechanism depicted on Figure 1 can be applied in practice using content and metadata and analyse how item categories generated from content and metadata can affect business models.

3. Case Study

The data for this paper was provided by the research project "*Smart Me versus Smart Things: The Development of a Personal Resource Planning (PRP) System through Human Interactions with Data Enabled by the IoT*" which implements the *Hub of All Things* (HAT) technology[2] using real households. A household consisting of two young professional adults (male and female) was asked to monitor consumption of shower products using smart shower sensors and a specially designed

[2] See www.hubofallthings.com for more detail.

consumption monitoring "*Beauty Box*" device (Oliver, 2015). For the purposes of this case study, the following data was recorded:

(1) Duration of the shower activity for each study participant in minutes was captured using the motion sensor placed in the shower.
(2) Duration of water usage as a part of the shower activity for each participant was recorded using the flood sensor placed on the shower floor.
(3) Water temperature used in the shower by each study participant.
(4) Weight of all shower products after each use was recorded using the "*Beauty Box*" device. The "*Beauty Box*" captured the bar code of the product and weight of each product (see Figure 2 for the photograph of the "Beauty Box" prototype). All information from the "Beauty Box" was recorded on the HAT cloud and could be viewed by the study participants at all times.
(5) Study participants were also asked to write a detailed diary recording their purchasing behavior.

Figure 2 "Beauty Box" Prototype developed by Helen Oliver for the HAT project (Oliver, 2015)

While (1)-(4) provide content data, (5) creates an opportunity to combine datasets into metadata. Therefore, in our analysis we will use a combination of (1)-(4) to construct content data and a combination of (1)-(5) to construct metadata. For the purposes of this case study we will concentrate on the consumption of three products: shower gel, toothpaste, and shampoo. These items will allow us to illustrate how content data versus metadata can capture choices of shower gel, toothpaste and shampoo brands. Shower gel is joint use while each member of the household uses his/her own shampoo and toothpaste brands. Since shower activity in the household is clustered around morning and evening hours, we simplify the dataset and after cleansing obtain 296 use points for toothpaste and 169 use points for gel and shampoo each across 74 days of household observation.
Table 1 provides basic statistics summarizing all variables measured in the case study. According to Table 1, person 2 (female member of the household) uses more quantity of the shower gel, toothpaste and shampoo compared with person 1 (male member of the household). Yet, person 1 on average tends to spend more time in the shower and uses more water.

Table 1 Basic Statistics

Person	Shower gel average consump-tion per use (grams)	Toothpaste average consump-tion per use (grams)	Shampoo average consump-tion per use (grams)	Duration of shower activity per use (minutes)	Duration of water usage per use (minutes)	Average water temperature per use (degrees ^{o}C)
Person 1 (♂)	12.5	3.7	10.5	21.9	16.9	37.5
Person 2 (♀)	17.7	7.5	13.9	15.0	8.7	39.7
Total	15.4	5.6	12.4	18.0	12.3	38.8

Indeed, these patterns are confirmed by the results of OLS regressions with robust standard errors (errors are clustered at the level of the day of the week, i.e., Monday-Sunday). Results of these regressions are summarized in Table 2.

Table 2 Results of the OLS Regressions with Robust Standard Errors

Explanatory variable	Dependent variable					
	Shower gel average consumption per use (grams)	Toothpaste average consumption per use (grams)	Shampoo average consumption per use (grams)	Duration of shower activity per use (minutes)	Duration of water usage per use (minutes)	Average water temperature per use (degrees ^{o}C)
Person (1 or 2)	5.8113*** (0.6322)	3.9611*** (0.2974)	3.7642*** (0.6630)	-1.1149 (0.7491)	-8.2061*** (0.1429)	2.2068*** (0.1176)
Duration of shower per use (minutes)	0.0529 (0.0507)	-0.0017 (0.0343)	0.0366 (0.0379)	-	-	-
Duration of water usage per use (minutes)	0.0119 (0.0782)	0.0077 (0.0205)	0.0121 (0.0544)	-	-	-
Average temperature per use (degrees ^{o}C)	-0.0965 (0.0586)	0.0019 (0.0271)	0.0060 (0.0582)	-	-	-
Constant	8.9787* (2.8247)	-0.5463 (1.1894)	5.5075* (2.2894)	13.8514*** (0.7298)	25.1386*** (0.2132)	35.3337*** (0.2020)

N (observations)	169	169	169	296	169	169
R^2	0.7835	0.8223	0.7159	0.0057	0.8835	0.4533

* significant at 0.05 level; *** - significant at 0.001 level

Table 2 shows that person 2 indeed uses more shower gel, toothpaste, and shampoo than person 1. Person 2 also uses hotter water temperature than person 1. One would expect that more product consumption should be positively correlated with the duration of the shower activity as well as with the duration of water usage, i.e., the longer is the shower activity and the longer is the duration of the water usage, the more shower product should be used. However, duration of the shower activity, duration of the water usage and temperature are not statistically significantly affecting any of the product consumption patterns. Furthermore, surprisingly, person 1 uses water for longer time periods than person 2. Based on content data, we cannot explain why we observe this result. However, metadata allows us to answer this question. From the participants' diary we can establish that person 1 combines shower activity with his daily shave which means that person 1, on average, uses water for longer time periods than person 2. Yet, the difference between explanatory and predictive power of content data versus metadata becomes clearer if we consider each product (shower gel, toothpaste, and shampoo) and explore how decisions about purchasing different brands are made in the household.

Let us first consider shower gel consumption. Figure 3 summarizes consumption patterns for 74 days of observation. Since shower gel is a shared item for this household, the data from both person 1 and person 2 is plotted on Figure 3. The vertical axes shows remaining weight of the shower gel while the horizontal axes depicts the day of observation from 0 (first day of observation) to 73 (last day of observation). On the horizontal axes the data is arranged by week, where the first week of the study runs from 0 to 6 (7 days).

In our case study the information about brand of shower gel was derived from the bar codes recorded by the "Beauty Box". Figure 3 shows that the case study household alternated between different brands of shower gel changing 6 brands during 74 day (12 weeks) of the study. Using content data, we can derive 3 sets of conclusions from looking at these data: (i) a new bottle of shower gel is purchased every 12-13 days by the household; (ii) person 2 consumes more shower gel than person 1; and (iii) the household alternates brands of shower gel without repetition which allows us to put shower gel in the Green items category according to the mechanism proposed on Figure 1.
Metadata provides more information about the household choices of shower gel brands. Specifically, analysis of the household purchasing diary reveals that all shower gels were bought at different locations. The participants also recorded that while Brands A, C, D and E was purchased by person 1, Brands B and F were purchased by person 2. Both person 1 and person 2 indicated that they liked to

alternate and try out different brands of shower gel. Person 2 recorded that she had a preference for red or pink colour of the gel but did not mind about the shower gel brand. The analysis of content and metadata on the shower gel consumption from the case study household allow us to place shower gel under Green items category.

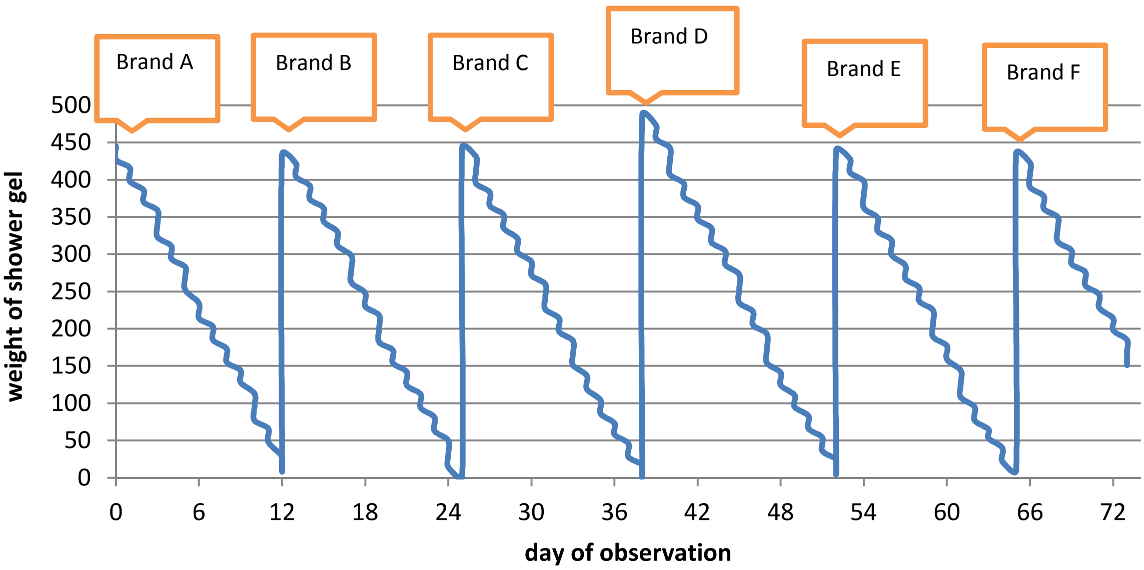

Figure 3 Summary of the Shower Gel Consumption

Let us now turn to the analysis of the toothpaste consumption. Toothpaste is not shared among the study participants. Therefore, we plot toothpaste consumption patterns for each study participant separately (see Figure 4). Brands A, B, C and D shown on Figure 4 are not the same as those on Figure 3. Content data reveals that in 74 days, person 1 bought toothpaste of Brand A three times and of Brand B once. At the same time, person 2 bought Brand C four times and Brand D three times. Using content data we can make the following conclusions from the observed patterns: (i) person 1 needs to replace toothpaste every 2 weeks if he uses Brand A and every 4 weeks if he uses brand B while person 2 needs to replace toothpaste every 2 weeks if she uses Brand C and every week if she uses Brand D; (ii) person 2 uses larger quantities of toothpaste than person 1; and (iii) for both participants, toothpaste can be classified as a Yellow item because both participants mostly use one brand of toothpaste (A for person 1 and C for person 2). However, they both occasionally deviate from their preferred choice in favour of other brands (B for person 1 and D for person 2).

Person 1 (♂) **Person 2 (♀)**

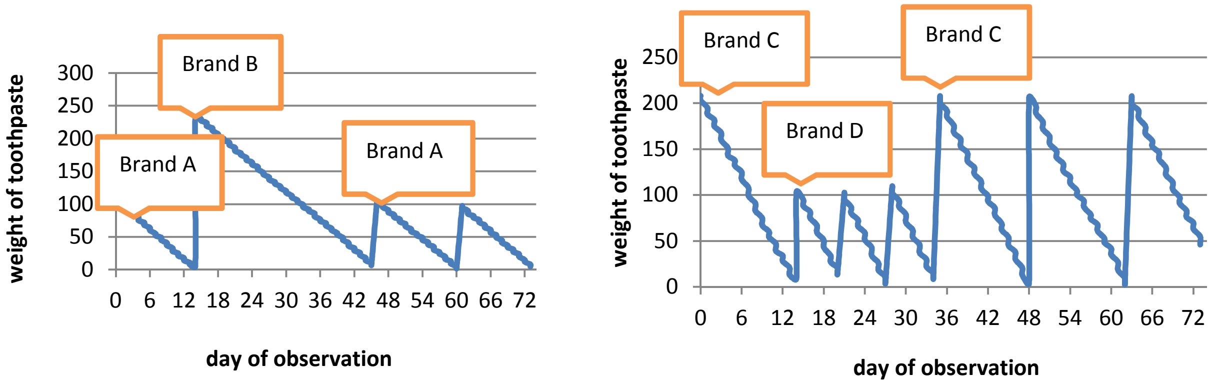

Figure 4 Summary of the Toothpaste Consumption

Metadata, however, reveals that the picture is more complex than that depicted by the content data. Specifically, the purchasing diary of person 1 reveals that this person indeed prefers Brand A to any other brand of toothpaste. Brand B was purchased by person 1 because he was shopping in a pharmacy rather than a chemist and even though Brand A was available picked Brand B. This story is consistent with classifying toothpaste as Yellow item for person 1. However, the picture is different for person 2. Specifically, person 2 revealed that she always buys toothpaste of Brand C. Yet, Brand D was used for 3 weeks during the observation period because it was prescribed to person 2 by the dentist. Therefore, Brand D was not purchased because person 2 really preferred to buy Brand D. Instead, it was purchased on doctors' instructions. Therefore, taking into account metadata, we should classify toothpaste for person 2 as Red item rather than Yellow item.

Finally, let us consider shampoo consumption. Similarly to toothpaste consumption graphs, we provide separate graphs for person 1 and person 2 (see Figure 5). Brands A, B, and C do not coincide with Brands A, B and C on Figures 3 and 4. Content data shows that shampoo consumption pattern for person 1 is very straightforward with person 1 consistently choosing shampoo of Brand A. For person 2, the pattern is more complex with person 2 choosing Brand B twice and Brand C once during the observation period. Content data allows us to make the following conclusions: (i) person 1 needs to replace shampoo every 6 weeks and person 2 every 8 weeks when she uses Brand B and every 2 weeks when she uses Brand C; (ii) person 2 uses large quantities of shampoo than person 1; (iii) shampoo should be classified as Red item for person 1 and as Yellow item for person 2.

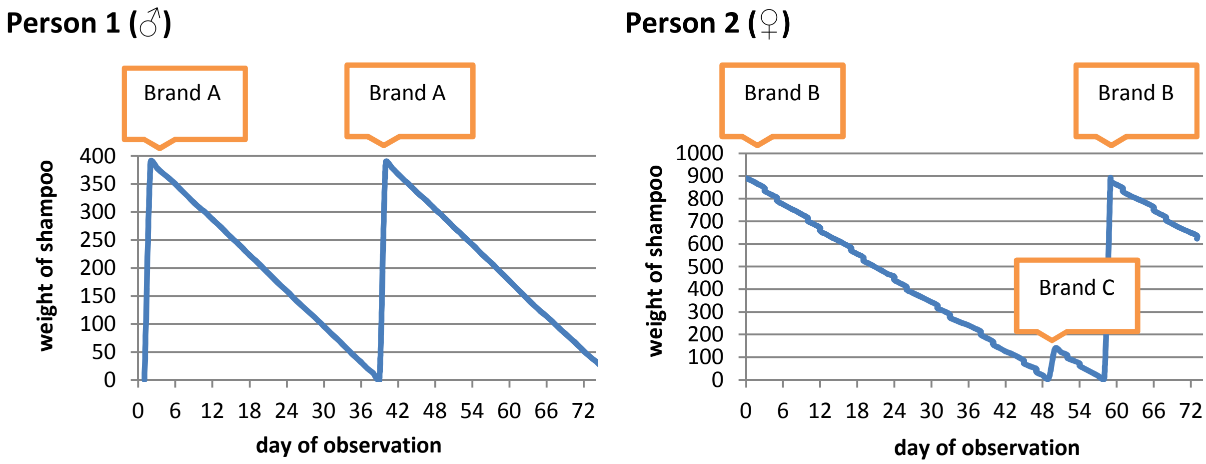

Figure 5 Summary of the Shampoo Consumption

Metadata confirms conclusion (iii) for person 1 but reveals additional information about person 2. Specifically, person 1 indicated in the diary that he indeed preferred shampoo of Brand A and always purchased that brand. At the same time, person 2 stated that she always bought Brand B. Brand C was purchased because the pharmacy where person 2 was shopping for Brand B did not have Brand B in stock. Therefore, a much smaller bottle of shampoo of Brand C was purchased in a hope that it would soon be replaced by Brand B. Therefore, for person 2, shampoo should not be classified as Yellow item. Rather it should be classified as a Red item.
Comparison between content data and metadata is summarized in Table 3. According to this comparison, both content and metadata would predict that person 1 will continue purchasing the same brand of shampoo; will in the majority of cases stick to the same brand of toothpaste occasionally deviating to other brands and is likely alternate among different shower gel brands in the future. At the same time, according to content data, person 2 will alternate among different shower gel brands and will mostly prefer the same brands of shampoo and toothpaste but will occasionally purchase different brands of shampoo and toothpaste. Metadata reveals that person 2 is likely to buy different brands of shower gel but will stick to the same brands of shampoo and toothpaste. This summary reveals that while content data and metadata can lead to the same conclusions in more straightforward cases (e.g., preferences of person 1), in more complex cases (e.g., preferences of person 2) using metadata over content data has obvious advantages because it allows to classify user brand preferences for various offerings more quickly and accurately.

Table 3 Preference-Brand Loyalty Mechanism Categorization Based on Content Data and Metadata

Person	Content Data			Metadata		
	Green	**Yellow**	**Red**	**Green**	**Yellow**	**Red**
Person 1 (♂)	shower gel	toothpaste	shampoo	shower gel	toothpaste	shampoo
Person 2 (♀)	shower gel	shampoo, toothpaste	-	shower gel	-	toothpaste, shampoo

Content data may be able to produce the same results as metadata. However, in order to reach the same conclusions content dataset has to include more data and over a longer time period to establish robust behavioural patterns.

4. Analysis of Business Models from Case Study

In Section 3 we have established that metadata has advantages over content data in classifying offerings in terms of user preferences and brand loyalty. In this section, we will look at how more accurate classification in terms of Preference-Brand Loyalty mechanism can impact business model innovation mechanisms. Content data allows to work within the frame of the data-driven business model shown below on Figure 6(a) (e.g., Hartmann et al., 2014). Businesses tend to use backward induction mechanism in their engagement with the users. Specifically, they start with the development of the product or service, then they collect data, analyse data or attract professional data analysts, create an offering package around the product or service (for example with support of advertisement, apps, other nudging mechanisms). After that, they provide the offering package to the target customer (user) executing a chosen revenue model. In other words, the mechanism is working backwards from the product to the user.

(a) DDBM based on Content Data (Backward Induction Model)

(b) DDBM based on Metadata (Forward-Looking Model)

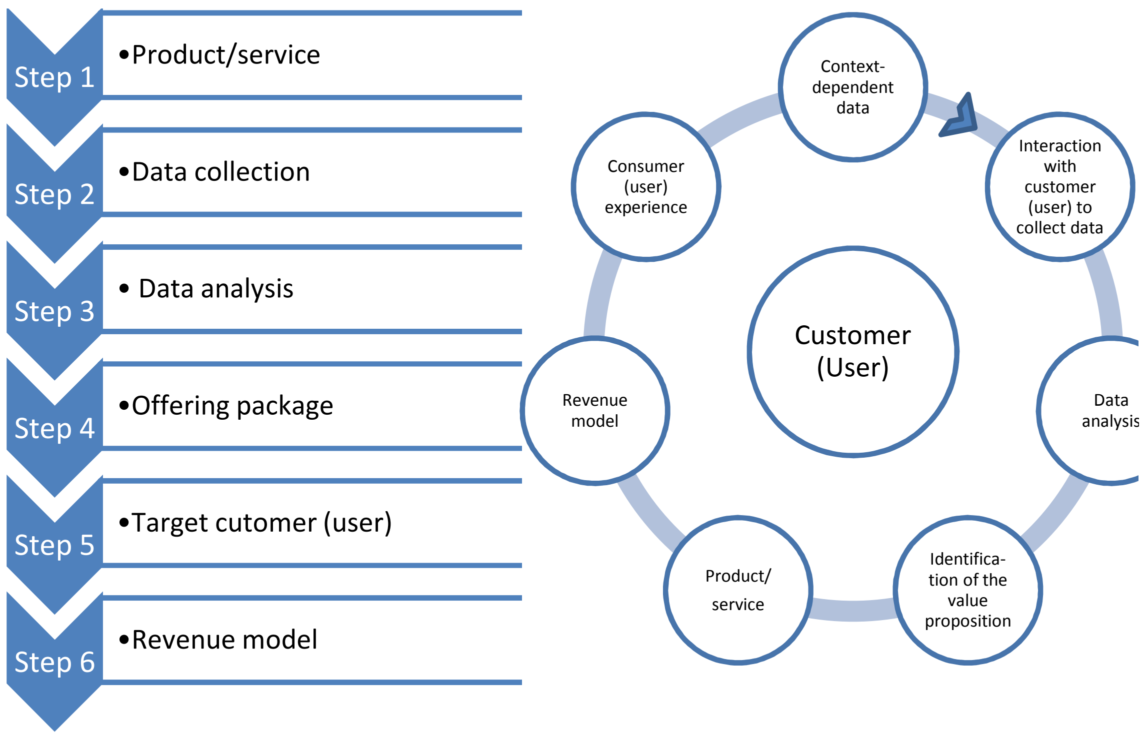

Figure 6 Data-Driven Business Model Mechanisms for Content Data and Metadata

The main advantage of the DDBM which is based on metadata is that backward induction model is replaced by forward looking model – from user to product - as depicted on Figure 6(b). In DDBM, user is at the centre of the system. User generates context-dependent data which then is communicated to the product/service provider. This communication can be proprietary (i.e., users can receive monetary reward for sharing their data) or users can be motivated to share their data in exchange for receiving better (more personalised) product/service propositions.
Context-dependent data is aggregated and analysed generating identifiable value proposition options for the provider. After that, the provider develops customizable and yet cost-effective product or service which then is offered to the user via an appropriate revenue mechanism. Users experience the product or service and generate the new wave of context-dependent data. Such a DDBM is more flexible and dynamic because it allows for constant interaction between users, data and providers.

For both DDBM mechanisms depicted on Figure 6, information about brand loyalty can significantly simplify data analysis and the development of the appropriate revenue model. However, as shown in Section 3, metadata can predict brand loyalty more accurately and requires fewer data points than content data. Brand loyalty predictions can also provide additional benefits in exploring various delivery mechanisms for different types of products and services which can be developed with different revenue streams. For example, instead of concentrating on retail sales of individual items, providers may specialize in the provision of bundles of products or services taking into account consumer preferences and their brand loyalty. Furthermore, instead of distribution through retail chains or home delivery, providers may explore alternative distribution mechanisms such as distributing bundles of products at public transportation hubs, bus stops, etc. via designated safety boxes areas.

5. Conclusion

This paper has considered how content and metadata can predict consumer choices using the concepts of precise, noisy and imprecise preferences and their relation to brand loyalty. We showed that metadata provides more accurate predictions with regard to brand loyalty which have significant implications for data-driven business models of the future.

This research has a number of theoretical implications. First, it extends the work on reverse supply chain mechanisms (Parry et al., 2015) offering a simple and tractable example of how users may be engaged and motivated to interact with their self-generated data in order to become an integral part of the multi-sided market for personal data. Second, resulting DDBMs which are based on metadata are directly related to the work on human-data interactions (HDI). HDI domain studies how individuals can interact with data as a part of a growing data exchange ecosystem (e.g., Mortier et al., 2014).

This paper also has practical relevance: by understanding brand loyalty through user preferences, product and service providers can increase profitability through creating more salient and efficient value propositions through increased user context visibility and better understanding of stochastic components in revealed user preferences. Providers will not only understand which decisions are made by users and when but also why they are made. This would allow providers to better anticipate user wants and needs and create products and services as well as delivery solutions which are more appropriate for their target audiences.

This research has a number of limitations. First, we report results from only one household which recorded context data as a proof of concept. As of now context-recording is not automatized, however, attempts to facilitate the recording of metadata are currently being made within the Hub of All Things (HAT) project. Second, a number of problems will arise with aggregation of metadata since context-dependent datasets are a lot richer than their content data counterparts. Therefore, metadata would require a new (more comprehensive) set of analytical tools which would allow all interested parties to extract, simplify, and work with data more efficiently. Third, current study does not address the question of dramatic preference change when users switch from one brand to a different brand. These questions can be studied using recent advances in social psychology. Finally, collecting metadata would require educating individual users and households to record their self-generated data. This is probably one of the most difficult practical tasks. In this paper we have shown how more accurate brand loyalty predictions can improve value propositions from providers to users. Yet, a lot has yet to be done in order to motivate users and providers to engage in metadata exchange in practice.

References

Danaher, P. J., Wilson, I. W., & Davis, R. A. (2003). A comparison of online and offline consumer brand loyalty. *Marketing Science*, *22*(4), 461-476.

Fechner, G. T. (1966). Elements of Psychophysics. Translated by Helmut E. Adler. *New York: Holt, Rinehart and Winston, Inc*.

Gantz, J., and Reinsel, D. (2012). The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. *IDC iView: IDC Analyze the Future*, 1-16.

Hartmann, Ph. M., Zaki, M., Fendmann, N., & Neely, A. (2014) Big Data for Bug Business? A Taxonomy of Data-Driven Business Models Used by Start-Up Firms, University of Cambridge, Cambridge Service Alliance Working Paper, accessed at http://ssrn.com/abstract=2416869 on March 1, 2015

Hastie, R., & Dawes, R. M. (Eds.). (2010). *Rational choice in an uncertain world: The psychology of judgment and decision making*. Sage.

Hey, J. D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, 1291-1326.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 263-291.

Loomes, G., & Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, *39*(3), 641-648.

Mortier, R., Haddadi, H., HenDerson, T., McAuley, D., & Crowcroft, J. (2014). Human-Data Interaction: The Human Face of the Data-Driven Society, accessed at SSRN 2508051 on March 1, 2015.

Ng, I. (2013). Value and worth: creating new markets in the digital economy. *Innovorsa, Cambridge*.
Oliver, H. (2015) The Automagic Box of Beauty: a prototypical smart device as use case example for user-centered decision support via the Hub-of-all-Things. SENSORNETS, in press.
Oliver, R. L. (1999). Whence consumer loyalty? *Journal of Marketing*, 63, 33-44.
Parry, G., Brax, S. A., Maull, R., & Ng, I. C. (2015). Visibility of consumer context improving reverse supply with Internet-of Things data, University of Warwick Working paper, accessed at http://wrap.warwick.ac.uk/65719/ on March 1, 2015
Simon, P. (2013). *Too Big to Ignore: The Business Case for Big Data*. John Wiley & Sons.
Sundbo, J., & Toivonen, M. (Eds.). (2011). *User-based innovation in services*. Edward Elgar Publishing.
Thaler, R. H., & Sunstein, C. R. (2008). *Nudge*. Yale University Press.
Villas-Boas, J. M. (2004). Consumer learning, brand loyalty, and competition. *Marketing Science*, *23*(1), 134-145.