

**Inter-Organizational
Information Sharing of Customer Data in Retail**

By

John C.F. Tengberg

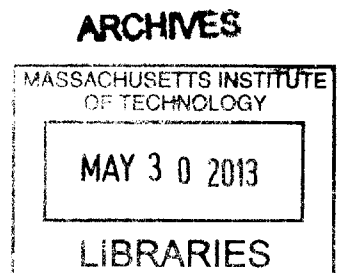
B.Sc. Financial Economics
Lund University, 2010

SUBMITTED TO THE MIT SLOAN SCHOOL OF MANAGEMENT IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

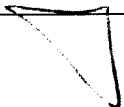
MASTER OF SCIENCE IN MANAGEMENT STUDIES
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

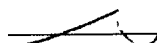
JUNE 2013


@2013 JOHN C.F. TENGBERG, All rights reserved.



The author hereby grants to MIT permission to reproduce
and to distribute publicly paper and electronic
copies of this thesis document in whole or in part
in any medium now known or hereafter created

Signature of Author:  MIT Sloan School of Management
May 10, 2013

Certified by:  Stuart E. Madnick
The John Norris Maguire Professor of Information Technologies
and Professor of Engineering Systems
Thesis supervisor

Accepted by:  Michael A. Cusumano
Sloan Management Review Distinguished Professor of Management
Program Director, M.S. in Management Studies
MIT Sloan School of Management

**Inter-Organizational
Information Sharing of Customer Data in Retail**

By

John C.F. Tengberg

Submitted to MIT Sloan School of Management
On May 10, 2013 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in
Management Studies

ABSTRACT

As massive online retailers are putting increasing pressure on the traditional brick-and-mortar retailers, new ways to compete for customers is needed. Identifying customers' behavior and understanding their needs could be a crucial way for firms to improve their competitive advantage. One way to better understand customers could be for retailers to collaborate and aggregate their customer data in order to gain more holistic understandings of customers.

This thesis examines on a high-level how customer data is used in retail today and what new value can be generated by collaborating and aggregating customer data, both for businesses as well as for consumers. It then goes on to identify structures of customer data sharing that we observe today and what the underlying factors are for each structure. By projecting how these factors will evolve over time, different scenarios are drawn up of what we might expect to see in the retail industry in the future in terms of customer data sharing. The implications of those projections could possibly open up for new business opportunities, both for entrepreneurs as well as incumbent retailers.

Thesis Supervisor: Stuart E. Madnick

Title: The John Norris Maguire Professor of Information Technologies and Professor of Engineering Systems

Table of contents:

I: Introduction	4
<i>Background</i>	4
<i>Methodology</i>	6
<i>Structure</i>	6
II: Current practices and the need for improvement	7
<i>How and what information retailers collect</i>	7
<i>Other firms aggregating consumer retail data</i>	9
<i>How retailers analyze information</i>	9
<i>The need for improved customer relationship</i>	11
III: The value and motives of collaboration	14
<i>Collaborative sharing occur in many places</i>	14
<i>Unique value creation through collaboration and aggregation</i>	15
<i>Network externality boosts value</i>	17
Include customers	19
<i>Motives that influence collaboration</i>	19
Legal and privacy policy concerns.....	20
Technical concerns.....	21
Knowledge gains.....	21
Strategic concerns.....	22
IV: 5 generic structures of information aggregation	24
<i>Levels of information granularity</i>	24
<i>Models of information sharing structures</i>	24
<i>The 5 structures</i>	27
No information sharing.....	27
Direct sharing	29
Joint collaboration program.....	31
Third party aggregator of information	34
Personal data wallet.....	37
<i>Three major underlying factors are shifting</i>	40
V: Value from the customer’s perspective	41
<i>Advantages</i>	41
<i>Privacy concerns</i>	42
VI: Macro scenarios of customer data aggregation in the future	43
<i>From micro to macro</i>	43
<i>4 macro scenarios</i>	43
Current state.....	44
Multiple alliances.....	44
Universal repository	45
Personal data wallets.....	46
<i>The road ahead</i>	46
VII: Conclusions and opportunities	49
<i>Future research</i>	50
Bibliography	51

I: Introduction

This part will set the context for the topic of information aggregation of customer data in retail. It will also give a brief overview of the methodology used as well as give an introduction to each chapter.

Background

One-to-one marketing¹ has since Don Peppers and Martha Rogers first introduced it in their book “The one to one future” (Peppers and Rogers 1993) been the ultimate goal for marketing departments in companies when trying to communicate with their customers as well as their potential future customers. A fundamental key to successful one-to-one marketing is understanding your customer and her needs. This understanding requires information about them and their behavior. Information has traditionally been hard as well as costly to get about customers, collected through surveys, focus panels, sales and purchased reports from marketing experts. So up until recently, truly personalized marketing activities had yet to be realized. However, as digitization and IT solutions have improved and increased, more data is being recorded, stored and becoming available for analysis for companies, making this holy grail of marketing more and more a realistic goal. Arguably the most IT-savvy firm in retail, Amazon.com, even markets themselves as “Your Amazon”, marking how they are thinking of their relationship to customers as a unique and personalized one between Amazon and each individual customer.

Massive amounts of customer information are being created everyday and the data mining tools are advancing to more sophisticated levels that can better extract valuable conclusions and help managers to make better decisions (Strehl and Ghosh 2000) (Davenport, Barth and Bean, How 'Big Data' is different 2012). In his article, “Competing on Analytics”, Thomas Davenport (2005) argues that firms that use advanced analytics of collected data in their operations perform better and earn a higher return (Brijs et al). The casino Harrah’s is a, in academic circles, well-known classic case of how better understanding of your customers through analysis of data can lead to superior performance. There are many other cases of valuable conclusions drawn from data, like the notorious example of when Target sent out personalized advertisements of baby items to a pregnant woman before her family even knew about it (Hill 2012). This was done by comparing the spending patterns of women who had signed up for Target’s baby registry. By looking at purchases of around 25 products including lotions, health supplements and other articles, together with demographic information collected or bought, they are able to give a pregnancy prediction score and even estimated delivery date. This also shows how discerning the increased knowledge of companies about customers can be, something the Wall Street Journal picked up on in their article series “What they know” (Wall Street Journal 2010) as well as New York Times in their “How companies learn your secrets” (Duhigg 2012). The balance between improving relationships to customers and becoming too intrusive or ‘creepy’ is one that companies are becoming increasingly aware of and something they will need to manage in the future. Alexander Pentland of MIT also describes how it is crucial to use network data “in tandem with an understanding how to create value for producers and owners of data, while at the same time protecting the public good” (Pentland 2012).

A huge driver in the massive increase of consumer information available today has come through the Internet and the increasing role of major e-tailers such as Amazon. These firms consolidate the offerings

¹ Also referred to as personalized marketing and relationship marketing

of many external retailers (through for example Amazon's marketplace) and can in this way gain access to not only their own store's specific sales, but to a more holistic view of consumers purchasing patterns across other segments in retail. This has partly allowed them to recently start moving in to the advertising business as they can offer very specialized and valuable segmentation and personalization of advertisements based on actual shopping and browsing history, charging a premium fee for advertisements (Leber 2013) (Hughes 2011). Furthermore, even though Amazon itself does not publicize sales generated by their recommendation engine, some sources estimated that up to 35% of Amazon sales in 2006 were generated through this function. The efficiency of this recommendation engine is thus creating significant value in terms of increased cross sale. This engine is in turn built on the information Amazon has collected about customers over the years. It's thus the massive information databases that has enabled this improvement in sales and is just one example of how customer data can improve operations.

For offline retailers, or online retailers that just have their own offerings for sale, the ability to collect consumer data and behavior and then personalize advertisings, recommendations and improve operations accordingly has been rather limited when compared to Amazon, but is gaining increasing attention (Linden, Smith and York 2003). Offline retailers can and do track their own customers in various ways – loyalty cards, backwards tracking through credit cards, phone numbers, simple registration online or even IP-address backtracking for online purchases (Yan 2006) (Pridmore 2010). However, their data of consumers is still limited to, first of all, existing consumers and only to certain easily recorded behaviors in *their* stores. Customers' behavior in other places cannot be seen and leads to a limited understanding of their customers. Companies are trying to get a more holistic view of customers by adding information from third party data sources such as Acxiom and Experian (Dyché) (Altair; Loyalty360 2012). But as former principal engineer at Amazon and now chief scientist at RichRelevance (a data mining company) Darren Vengroff says, Amazon's knowledge of customers is "laser sharp" in comparison to most offline retailers (Thomson 2012).

An interesting question to consider with this context in mind is therefore if there is a way for offline retailers to create a better understanding of their customers. Can isolated retailers in any way *collaborate* to create a better understanding of their customers that is at par with or closer to what Amazon and other retailers with massive and granular data on customers. One way would be for retailers to share data about their customers' transaction history and synchronize the various sources to form more comprehensive profiles of their customers. New forms of collaboration are occurring and being investigated in many industries (Hirshorn 2013) (Miller and Tucker 2011) (Eriksson and Wallman 2011), but there are also a few cases where sharing of customer information seems to be happening in new ways in retail with novel collaborative approaches taking place. Examples include coalition programs like Payback (Germany, Poland, India and Mexico), Nectar (UK, Chile and Italy), Kooperativ Förbundet KF (Cooperative Federation) in Sweden, FlyBuys in Australia and New Zealand and perhaps most recently Facebook working together with companies like Datalogix (Winkler, Facebook Adds to Google Battle 2013). However, these collaborations vary both in structures and underlying motives and the dynamics of such collaborations has not been studied to any great extent to the knowledge of this author. This paper therefore aims to be a starting point for discussions around this topic. It will aim to clarify the advantages and drawbacks of such collaborations from both the business point of view as well as consumers and identify the forces at play in each different structure that can be observed. By identifying the underlying factors for each structure, the future evolution of these factors should provide a guide to what we might

observe in the future. Based on this, this paper will try to identify various future macro landscapes of the offline retail sector and assess the most likely road ahead. It is not to be regarded as an exhausting or complete map out of the dynamics however, but rather to be seen as an initial start for future discussions about collaborative practices in sharing of information about customers and their behavior and its implications for the retail sector.

Methodology

The methodology for this paper is an extensive review of literature, articles and observed examples in mainly Europe and to some extent the US. Due to limited writings found on the this particular topic, analogous cases from various industries will be used to hypothesize on benefits for the retail industry in particular. From this research, a new framework of information sharing structures will be created along with the determining factors for those structures.

The hypothesized structures and factors will then be substantiated through examples and interviews with current players in Sweden and the UK involved in the various forms of structures. Interviews will take place with active members relevant to the retail and consumer analytics sector. Interviewed companies cover in total 20-25 retail chains in Sweden and the UK, 1 bank issuing loyalty cards, 1 vendor of Point-of-Sale-systems for retailers and 1 so called financial aggregator (a company that among other things consolidates consumers' different bank statements and credit card transactions, similar to Mint in the US).

By using the suggested framework developed in this paper, some potential roads ahead and the underlying factors for each scenario will be identified for the retail industry. The various future scenarios will also give room for possible new opportunities that may arise due to new practices of information sharing.

Structure

The outline of the thesis will be as follows. In order to understand where the retail industry is today, current practices of information gathering and analysis will be the subject of the next chapter. Chapter 3 will move on and spell out what additional value sharing and aggregating customer information could bring. It will also discuss and identify forces and dimensions that affect the decision to share or how to share customer information. Using these dimensions as well as observed examples, chapter 4 then tries to give a new framework of information sharing structures and discuss why we see each particular form. Chapter 5 will look at the matter through the eyes of consumers as opposed to the businesses and discuss the possible concerns of consumers. The next chapter will then translate these microstructures into different macro scenarios for the retail industry and suggest what might be expected in the future. The last part will summarize the new framework, its implications and the important factors for the future in order to identify possible directions ahead. It will end on a note of what new business opportunities that might arise from viewing information sharing practices through this new suggested framework.

II: Current practices and the need for improvement

The first part of this chapter will discuss how retailers currently both gather as well as analyze customer data in order to gain insights and generate business value. It will also briefly mention other firms that gather information about customers. The second part will address the need for improved understanding of and relationship with customers and how this improvement could be achieved through data collection and mining.

How and what information retailers collect

The way retailers collect data about their own customers vary. The most significant difference lies between the online and the offline retailers. Online retailers can track every move of customers in their virtual store and can for this reason do far more detailed, but also more complex and process consuming, analysis. Another dimension that is only applicable to the large online retailers such as Amazon.com is that the product offering they provide is vastly greater than what brick-and-mortar store provide. Since Amazon launched “marketplace” in 2002, independent retailers became able to sell their products through Amazon’s platform. This gives Amazon the opportunity to collect data from a far larger set of items too, including other external retailers, as Amazon then controls a kind of choke point from where they can gather all information on every transaction. As opposed to single-store retailers, Amazon can in this way see purchasing behavior across many different types of products. This means that instead of only seeing what kind of books a person buys, they see everything from purchases (as well as viewings) of electronics, health products, clothes, pet status, office supplies to jewelry and musical instruments and all this on an SKU-level (stock keeping unit). The profiling of customers that this kind of comprehensive data allows for is far superior to what individual retailers can access. Amazons massive database of this kind of data has been used to provide recommendations for other products that might be attractive to users, and an impressive 35% of their sales have been estimated to have been generated through this recommendation function. Recently, Amazon has started to monetize this database and knowledge even more by moving into the advertising business, providing highly sophisticated profiles and segments (Habegger 2011) where advertisers can reach out to very specific profiles instead of the relatively crude methods used today. Amazon only reveals parts of their data on a very aggregated and segmented level, knowing that their rich database is perhaps one of their greatest assets (Leber 2013). They for example provide two main buckets of targets for advertisers, one is based on lifestyle and the other on people who are in-market for particular items.

Online retailers track their customers in many ways, from IP-adresses and cookies to registration and other forms of direct authorization. They can also use credit cards to tie purchases together, without the customer necessarily having created an account and logged in (Yan 2006) (Pridmore 2010). More than just tracking transactions, online retailers can track very detailed behavior of customers. Amazon for example collects information such as (Kohavi et al) (Tsiptsis and Chorianopoulos 2009):

- time spent on each page
- how navigation occurs
- successful and unsuccessful recommendations
- search history and trend spotting
- shopping basket events (additions, changes, deletions of items in cart)

- conversion rates from viewing and buying

These metrics are just a few out of a vast amount available for collecting in the online and virtual store and they occur on a SKU-level. To the offline world however, most of these metrics remain costly or impossible to track.

Despite not having all the tools that online retailers have, brick-and-mortar retailers still have become and are becoming increasingly sophisticated in collecting data about their customers too. They can for example use credit cards to link different transactions together and recent development in surveillance can allow for automatic video monitoring of customers in stores (Kubat 2012) (though crude today, it is not unlikely that it will evolve rapidly over the next couple of years). However, the perhaps most common and convenient way for retailers to collect information about their customers is through membership and loyalty programs. These programs allow companies to link each transaction and the specific details of those transactions (basket items, time of purchase, volume, value etc) in their POS-data to specific individuals. This data provides a huge insight into customers and data mining tools can extract value from this data in order to monetize this knowledge.

In the future, possible additional ways of gathering information could be to track your phone to see how you behave. Geo-fencing, a virtual parameter for real world geographies, with the help of smart phones location is one method that is becoming increasingly common. The exact potential of this method is yet to be explored. Phone carriers are however sitting on massive databases on geolocation data, and experiments and studies have shown tremendous potential in this dataset. Currently, restrictions in legislature hinder them from utilizing this information as they wish, but in the not too distant future, this might very well change as there is a huge amount of untapped value in this information (Hughes 2011). Another more sophisticated method that is gaining attention is the use of RFID and NFC (near field communication) technologies, both on each SKU as well as in your membership card or cell phone. RFID (and video surveillance possibly) could allow companies in the future to track their customers specific movements and behavior in the store as well as what items get picked up and looked at, but not bought (similar to viewings on the web) (Bashir 2011) (the Guardian 2012).

Apart from collecting their own data, firms also complement their data with data from third party firms specializing in consumer data, such as Axiom, Experian and DataLogix. Loyalty 360, an association in the US for marketers, stresses the importance of complementing your own customer data with additional outside sources in order to be able to create a more holistic view of their customers (Altair; Loyalty360 2012). The information held by these companies is typically classic demographic and socioeconomic characteristics, such as age, address, income, family status etc, that add increased granularity and dimensions to companies data about customers. More specific data sources like owned car model and even retail transactions are sometimes also available (Dyché). They draw this information from a wide array of public sources and participating partners and consolidate it to their own database, which is then made available at a fee for marketing departments of companies, retailers just as well as other organizations, even including political parties². The cost of this data can however often be very high, but can help companies reaching a more complete picture of their customers.

² See Aristotle as an example, www.aristotle.com

The contrast between e-tailers and offline retailers is as discussed quite significant. Large e-tailers have much more sophisticated tools and opportunities to track and analyze customers' behavior as well as their actual purchasing patterns in the virtual store. This enables them to reach a greater holistic understanding. The offline world however is trying to catch up as awareness, capabilities and technologies are improving.

A dimension not addressed here, but that is very relevant, is the need for intra-organizational aggregation of consumer data. Many companies have still today not a single and common repository of all their customer data, but store the data in different silos spread out across departments and locations. Sharing and aggregating data throughout the organization and using technologies to leverage this can significantly improve operations and the relationship to customers. This paper will assume that a complete intra-organizational aggregation of customer data exists for each organization, even though this is not necessarily true for today's organizations (Chen and Popovich 2003). According to The Data Warehousing Institute, only 11% of companies have integrated their customer data completely within the organization (Eckerson 2004).

Other firms aggregating consumer retail data

Retailers have their own POS software and system that records granular transaction data. Membership programs can then tie those transactions to individuals and allow for identification of users and their unique behavior. That granular data is strictly owned and collected by each individual retailer. However, some external firms, like payment companies such as VISA, will automatically collect transactions on a lump sum level. This information allows these kinds of firms to analyze transactions on a less granular level, but they too provide useful analysis.

VISA and MasterCard collect lump sum transaction information from their cards (despite being issued by different banks). This allows them to do market analysis, and have traditionally performed certain kinds of very high level market analysis and they publish industry reports, like for example VISA's Industry Insights. In their 2011 retail report however, VISA state they are starting to look into mining and segmenting their payment data to a greater extent (Ayliffe 2011). Interestingly, issuing banks can also collect and analyze behavior from their customers. A, by the author, interviewed smaller bank however found little value in that data given their limited customer base, but larger retail banks could possibly engage in this kind of analysis of sufficiently large customer base.

American Express has seemingly come farther than VISA in this regard with their American Express Business Insights. Retailers can buy more detailed reports and analysis of relevant consumers in order to gain better insights about their customers such as share of wallet (or more precisely share of spending on their American Express card). American Express however only reveals their data on an aggregated level (American Express) (Bulik 2009).

How retailers analyze information

Once companies have collected and aggregated information, the next step is to analyze that information in order to generate value in terms of better decisions or new insights. POS (point-of-sale) systems can together with membership programs track your purchases on SKU-level and link these transactions to your specific profile. There are many techniques to analyze that data, each focusing on different insights and measures. Using data mining tools, analysis can reveal patterns in behavior among different segments

and create personalized offers in order to increase your spending in their store. They also see the profitability of different types of customers, which can be very important. A survey conducted by Bain, a top consulting firm, revealed for example that for retailers, as much as about 30% of customers are typically non-profitable customers (Fournier 2012) and thus focus should be on maintaining and increasing sales from the other customers, while discouraging these lesser customers. The Boston Consulting Group provide similar advice in their 2012 Retail report (Mercier, Jacobsen and Veitch 2012). For these actions to be useful however, customers must first be classified as high or low value correctly, and Fournier et al (2012) argues that many companies today “close the door” to customers too early in belief that they are unprofitable.

As stated earlier, many types of analysis can typically be done on retail data. Each focus on different insights and lead to different actions and decisions. Some of the most important types of analysis in retail include (Han, Kamber and Pei 2012) (Chiu and Tavella 2008):

- Analysis of effectiveness of sales campaigns
 - o What is the response in sales after an advertising campaign, coupons or other promotions?
Thorough data mining can improve profits and also help identify products that are likely to be sold in association with the promoted products.
- Multidimensional analysis of sales, customers, products, time and region
 - o This includes many types of analysis of product trends, fashions, customer needs, sales as well as financial measures of cost and profit of items. Types are customer category mix analysis (typical items among different types of customers) and item basket analysis (how characteristics of the item affects its sales).
- Customer retention – analysis of customer loyalty
 - o Returning customers can be analyzed with help of their loyalty card to see patterns in their recurring purchases as well as changes in order to improve retention as well as attract new customers.
- Product recommendation and cross-referencing items
 - o These tools allow companies to increase up- and cross-selling from existing customers as new associations rules can be identified. Customers who buy item A is likely to buy item B and C as well. This kind of collaborative recommender systems recommends items to customers based on the behavior of other similar customers (Linden, Smith and York 2003).
- Fraudulent analysis and identification of unusual patterns
 - o Millions of dollars a year is lost to fraud in retail. There are multiple tools and methods that help identify and stop this kind of activities.
- Customer segmentation
 - o Can be done along several dimensions such as need-based, demographics-based, value-based, product purchase-based and profitability-based. This is an important step to understanding your customers.
- Customer profiling
 - o Creates detailed descriptions of segments. A segment is profiled with their distinguished behavior, characteristics and attributes.

All of these different insights are then used to personalize communications and offerings to customers, in order to improve operations in terms of both sales and relationships.

Analysis of data can, not only be done on transactions, baskets and customers, but also on sales associates performances as well. Oracle's retail data mining tools for example include tools to identify the characteristics of each sales associate in terms of sales, their customer's baskets, losses during their shifts and more (Oracle).

The previously mentioned case of Harrah's serve to show the potential benefit of such customer analysis (Thelen, Mottner and Berman 2004). Harrah's is a casino that was able to dramatically improve their performance in many dimensions, not just financial but also in terms of customer satisfaction. The way they were able to do this was by analyzing the massive amount of information that was being generated everyday by their customers. Through membership programs, Harrah's was able to distinguish between different customers as high value and low value, even if they weren't spending much money in Harrah's. Analysis revealed certain observed low spenders to truly be high spenders (but possibly in other casinos) based on *the way* they played. Collected data was analyzed to see how much they spent on each 'round', time between bets, age, frequency of visits and many other metrics. From this, Harrah's was able to infer who belonged to their high-value segments (regardless of spending levels). Through this improved understanding of their customers, they were able to through direct marketing, reach out to their 'truly' most important and profitable customers and they changed their operations to attract more of them. Harrah's shows how it's possible to extract value from information and dramatically improve the bottom line.

The need for improved customer relationship

Today, the level of relationship retailers have with customers vary from simplistic to rather sophisticated. CRM tools have in recent years become more and more advanced at the same time as data collection has increased, leading to better analysis and an improved relationship with, or at least understanding of, customers for those retailers who deploy and use these tools in the right way. This topic is increasingly becoming an area for further exploration by companies and it is estimated that just expenditure on analyzing past purchases of customers reached around \$9 billion in 2008 and continues to grow (Fournier 2012). A schematic view of current practices for retailers can be viewed in figure 1.

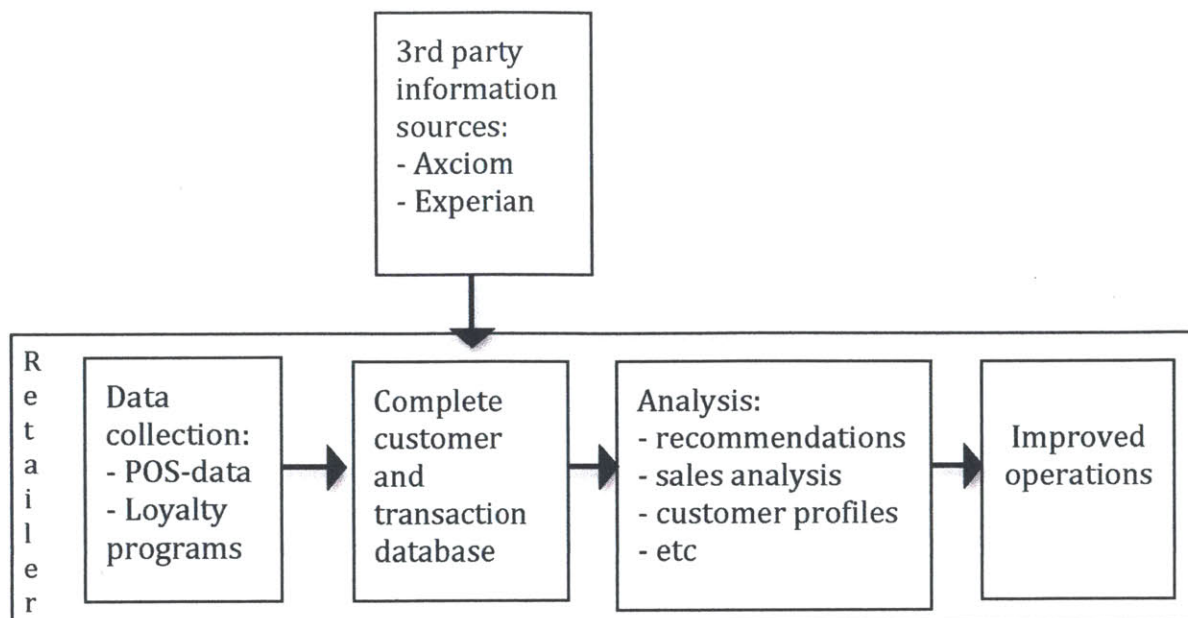


Figure 1: Schematic view of retailers' customer data process

In the Boston Consulting Group's retail report *Retail 2020: competing in a changing industry* (Mercier, Jacobsen and Veitch 2012), they state how established brick-and-mortar retailers must find ways to compete on more than just price, since they tend to be at a cost disadvantage against pure e-tailers in many product categories. They stress that consumers generally look for best value, best service or best experience. If brick-and-mortar retailers will have a harder time to compete on value, customer service and experience will increasingly become important. They call for a customer focus "based on their shopping habits and spending levels" and to focus on the high value customers. In order to move towards such a state, retailers must move beyond current practices and become more sophisticated in how they analyze and classify customers. As retailers are facing increasing pressure from e-tailers to improve operations, in order to compete effectively, one way to do so is to improve the way they both collect data as well as analyze that data, as seen in figure 1.

The pioneering book *The one to one future* (Peppers and Rogers 1993), stresses the importance of understanding your customer in a broader sense than just what she buys of your company and their basic demographics. In order to really build loyalty and increase sales, you need to reach the customer by being close to her, not only physically but also mentally. This requires a more holistic view of the customer and to know as much as possible about them in all aspects of their life, sometimes seemingly irrelevant to your business. Amazon's global head of advertising, Lisa Utzschneider, have for example had contact with Chrysler's Town&Country (a car model) team, who wanted to improve their understanding and connect with moms, a key segments. To better address and understand this target segment, Chrysler did not seek out information about their behavior explicitly related to cars, but rather considered all other aspects of their observable behavior to better understand and adapt to this segment. Utzschneider says "We can give them insights into how moms are spending time on Amazon. It can be totally unrelated to cars" (Learmonth 2012). This shows real life traction for companies striving towards understanding their customers beyond directly related purchases and behavior.

Peppers and Rogers (1993) also stress the importance of personalization and customization to each individual based on their characteristics and behavior. According to Fournier et al (2012), the limited view

of your customer often leads to businesses classifying their customers wrongly, or “closing the door to them too early” as they view some potentially profitable customers as low-value customers. A holistic view with more complete information would more likely show them to be high-value customers. This level of sophistication is made easier with knowledge of the customer on a deeper level rather than their behavior in a single retailer. This could be even more true for retailers with low frequency transactions per customer, i.e. customers buy on few occasions in their store (as opposed to grocery stores that have an immensely rich transaction history). This is because rich transaction histories (high frequency sales) have many internally generated data points on which to do analysis. Retailers with lower frequency, for example durables, have fewer data points and thus are limited in what analysis they can do and what insights about their customers they may draw.

Faced with these challenges in reaching a better understanding of customers while at the same time facing an increasingly competitive environment, brick-and-mortar retailers might consider sharing, aggregating and analyzing customer data between organizations. This could potentially provide very powerful tools to improve how retailers understand, connect and target customers in the future.

III: The value and motives of collaboration

By scaling the data volume about customers through collaboration, this chapter will discuss what new kinds of analysis and insights that could be gained. Furthermore, a closer look will be taken at the dimension of network externalities in this kind of collaboration and how it can impact the dynamic of sharing initiatives. Lastly, the chapter will explore factors and dimensions that affect firms' decision to share information.

Collaborative sharing occur in many places

It is clear that companies can gain valuable and actionable insights from customer information that allow for improved operations, offerings and relationships. Even with information limited to only their own customers' behavior in their own stores, this is evidently true. But by scaling the information base to include a much more comprehensive dataset of customers as well as non-customers' behavior through collaborating and aggregating multiple sources, resources could be created and accessed which are impossible to generate internally. As Pillar and Schaller argues for collaborating vertically in retail, "collaboration may itself become a dimension of competition" as it drives competitive advantage "through building up and extending knowledge about the individual customer needs (Hamel et al, 1989)" (Piller and Schaller 2002). Given that it would be possible to aggregate retail data between organization, the question then arises: what additional value could be generated by collaborating to aggregate information?

Today there are more and more examples emerging where sharing of information is occurring between actors in the same line of business. Peer groups have existed for a long time, where sharing information with relevant parties have proved to be beneficial for all participants. Peer groups can be both industry specific as well as cross-industry and share information that can be seen as proprietary and a source of competitive advantage (Hirshorn 2013). Another example is the PIMS (Profit Impact of Market Strategy) network by the Strategic Planning Institute. They collect detailed information about companies and their performance in order to provide a way for companies to benchmark themselves to relevant peers. The individual details of each company are never shared and only aggregated results and benchmarks are revealed for comparison to individual companies (Strategic Planning Institute).

Frequent flyer programs is another type of collaboration. These programs all work in alliances of airlines that share information, including even hotels and rental cars. Las Vegas casinos all share information about their customers behavior (mostly for anti fraud analysis). The FBI has since 9/11 taken initiatives to create ISE, Information Sharing Environment, to share information between agencies in order to fully extract the value of their data, previously isolated in different departments and agencies. Vertical information sharing in retail, i.e. between suppliers and retailers, is occurring on a broad basis for operational improvements, such as stock replenishments, by many businesses (Barret and Konsynski 1982) (Kulp, Lee and Ofek 2004). But Piller and Schaller (2002) also discuss collaboration of customer information in a vertical way in order to allow for individualization of products and services.

A few examples that are very relevant for this paper include Nectar (UK, Chile, Italy), Payback (Germany, Poland, India, Mexico), KF in Sweden and FlyBuys in Australia have created a combined loyalty program, a so-called coalition program. These programs allow customers to collect points through one

loyalty program (and card/phone app) in member retailers, instead of a separate program for each individual retailer. Furthermore, these programs consolidate offerings and advertising in emails and catalogues. The programs then not only analyze data on an individual partner level, but also perform analysis on aggregated data, much like PIMS, to see broader patterns. The full benefits of these programs will be discussed later on.

There are a large number of instances where sharing and collaboration of information occurs and gives rise to positive net benefits. This paper will address and focus on types of information sharing of customer data in retail, meaning not for stock-outs and similar operational concerns, but between retailers themselves as well as other marketing channels for the purpose of customer analytics and marketing. The next part will look closer at how new and additional value and insights could be generated through information aggregation, value that would be impossible to generate internally.

Unique value creation through collaboration and aggregation

As discussed previously, key to better understanding of customers and moving towards personalization is analysis of customer information. Yan (2006) mentions two fundamental ways to improve the level of analysis possible: 1) more, or better, data and 2) improved methods to analyze that data. In the digital era, information available has skyrocketed without necessarily having the capabilities to extract value from that information growing at the same pace. It is only recently that the analytical competence among firms has been receiving widespread attention and is increasingly gaining ground, rendering statisticians and data scientists, especially with a business sense, highly desirable and valuable but hard to come by (Davenport, Barth and Bean, How 'Big Data' is different 2012). This paper does not aim to look at value generation through suggesting new ways to analyze data, it will though give an overview of techniques and types of analysis that could be deployed to achieve new insights. The focus of this paper is instead to use the other lever of the value generation potential of information – improving and enlarging the dataset. By combining individual datasets of different retailers into a common repository, the dimensions of the datasets increase as well as number of data points, allowing for new patterns and behaviors to become visible and uncovered. These patterns are connected to individuals and can thus give retailers an improved understanding of their customers and even see non-customers that are more likely to become future customers. The new and enlarged datasets can provide new insights in many different ways. As Nectar's, a firm currently aggregating data this way, CEO says, the “combined effect is without question more powerful” (the Guardian 2012).

Apart from applying the same analysis as done currently by individual stores mentioned in the previous chapter, many more types of analysis could become possible and benefits generated. Previous papers on information sharing and aggregation state various kinds of benefits. Barrett and Konsynski (1983) state three potential benefits in (1) cost reductions, (2) productivity improvements and (3) product/market strategy. Dawes (1996) mentions three categories of benefits and barriers in another paper about interagency information sharing; (1) technical, (2) organizational and (3) political. However, these papers have a stronger operational focus than what the discussed information sharing here has, which focuses more on analytical insights rather than operational day-to-day activities. A perhaps more relevant paper is Madnick and Siegel's “Seizing the opportunity: Exploiting web aggregation”(2001) that discusses the value of web aggregators, which have a much more informational focus. These papers as well as findings

from interviews suggest some potential improvements that could be generated from customer data aggregation from multiple sources. These include:

- **Lead generation:**
 - With datasets containing current as well as non-customers, it could become possible to identify potential new customers based on similarities between them and the existent customers (so called look-a-likes or twins) and collaborative filtering (Linden, Smith and York 2003)
- **Improved cross and up-selling:**
 - Companies' success with cross-selling and up-selling seems to be directly proportional to the variety of detailed data it keeps about its customers. Enlarging the dataset to include more variety could potentially unlock further cross and up-selling (J. Dyché 2002).
- **New classification of high versus low value customers:**
 - Many companies close their door to profitable customers too early, believing they're unprofitable, resulting in their current customers not being maximized. A holistic understanding of customers can improve companies' ability to more correctly identify low and high value customers and thus extract more value form customers.
- **Improved segmentation of customers:**
 - With comprehensive datasets of customers, segmentation can be done based on actual behavior across many segments, allowing for improved segmentation of different types of customers. Amazon for example has segments based on lifestyles, which could be emulated. This kind of segmentation is broader than what segmentation on store level data is.
- **Improved market and segment understanding:**
 - Depending on how extensive the aggregation is, a higher-level understanding of segments and the market in general and how your company relates to it could be enabled as much broader and holistic views can be achieved.
- **Share-of-wallet analysis:**
 - A common measure of interviewed companies' CRM managers is share-of-wallet. An aggregated database has a much better chance of accurately measuring this rather than estimates based on assumptions.
- **Improved measures of advertisement efficiency:**
 - If partnering occurs between online advertisers such as Facebook and offline retailers, a link can be established between online advertisement exposure and offline (as well as online) purchases if all participants are in the same collaborative network.
 - Brands and suppliers could also possibly benefit from this ability to better see efficiency and results from advertisements.
- **Pooled resources for analysis:**
 - Many retailers today lack the scale or capabilities to efficiently do analysis of transaction data to its full potential. By pooling data, sharing resources could make it economically possible to build that kind of competence and improve their analytical capability.
- **Agent as consolidated channel to interact with customers:**
 - Depending on the type of collaboration structure, some structures could provide a consolidated way for companies to interact with customers. This could potentially

increase customer engagement and generate improved relationship and understanding of customers.

- Customized and personalized advertisements:
 - Advertisements could be directly targeted to only those that might be relevant. Advertisements to not only current customers, but also non-customers, could be improved and lead to improved advertisements, both to the benefit of customers as well as businesses. Personalized advertisements are furthermore a source of loyalty according to various studies (Svadkooho 2012).

These are some of the aspects that could be improved with an enlarged dataset. The true value of a comprehensive database of customers and their purchasing behavior can only be seen when actually realized, but Nectar for example are increasingly discovering new value from this enriched database and Amazon's database is perhaps regarded as the most valuable in the world because of its richness.

Network externality boosts value

In the framework of information sharing and aggregation, there are significant network effects that could play an important role in the dynamics of some structures. Network externality is defined as the value for a person to join the network depends on the number of others using that network (Shapiro). For example, if two firms start sharing, both give away their data but they also receive more data. As another participant joins in, the value of that database increases to all participants, not just the joining member, illustrated in figure 2.

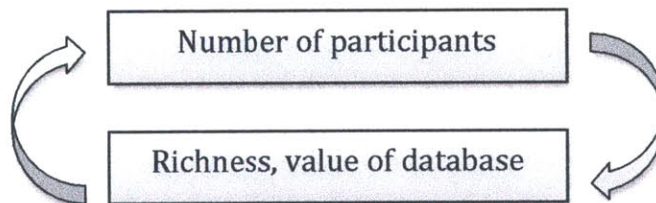


Figure 2: Network dynamic

The network effect will spiral this dynamic in either a positive or negative loop. This structure inevitably means that incentives to join will only grow as the number of participants increase, making the data richer and more valuable, attracting further organizations to join. This dynamic could play an important role in the dynamics of the framework suggested later on. The mechanism actually allows a large number of small retailers to attract a bigger retailer that previously didn't see any value in engaging in any sharing practices with a small individual retailer, as the small retailers' data combined can offer something that is attractive enough for the larger retailer to join. An analogous case could be seen during the emergence of ATMs in the US and UK in the 60's-90's (Batiz-Lazo 2007). As leading banks and institutions were getting a competitive advantage with their proprietary ATM networks, smaller institutions and banks joined together in 'shared Networks' that allowed them to compete at par with the larger institutions. Again, in this case, network externalities played a big part in forcing the shared networks among smaller players to emerge in order to compete with the bigger ones. Interestingly, the service of ATM's was first intended to be a source of competitive advantage but as it became ubiquitous and full interoperability between ATM networks emerged, it later became a necessity rather than an advantage.

The richness of the data is not just increased by having more transactions, but the sheer diversity of possible participants in the database can add completely new dimensions to the value of the repository. For example, online advertisers such as facebook have recently started sharing customer data with companies like DataLogix and Axcion (Delo 2013). The idea behind this is that as they display an advertisement online, they will hopefully be able to connect this online exposure to real life actions that can be seen in DataLogix database of retail transactions. They connect the consumer by matching the email listed in the loyalty programs to the one used for the Facebook account. This can provide a crucial link that has been troubling online advertisers since its emergence, the ability to identify the actual conversion rate of the advertisements. The lack of control of investing in online advertisements has led to more careful advertisers and the cost-per-click model that Google became known for is increasingly questioned as far as its actual efficiency. Complex data analytics companies are become more sophisticated in judging the value and conversion of these advertisements, questioning this old model. However, a direct link between online advertisements and offline purchases would be a giant leap for advertising agents such as Facebook if properly established. Google is trying to close this gap as well to some extent through its Google wallet (Close, Kukar-Kinney and Benusa 2012), and in practice all kinds of payments systems are in one way or another a potential choke point that could provide invaluable information and possibly link advertisement exposure to actual transactions, illustrated in figure 3. Credit cards are restricted however as they don't collect information on an SKU-level, but rather on an aggregated transaction amount level. They could therefore in the current form never provide a very granular link between individual advertisements and specific products. This is because of technical reasons where, for now at least, the POS-system tracking SKU-level transactions is separated from the payment system that credit cards use.

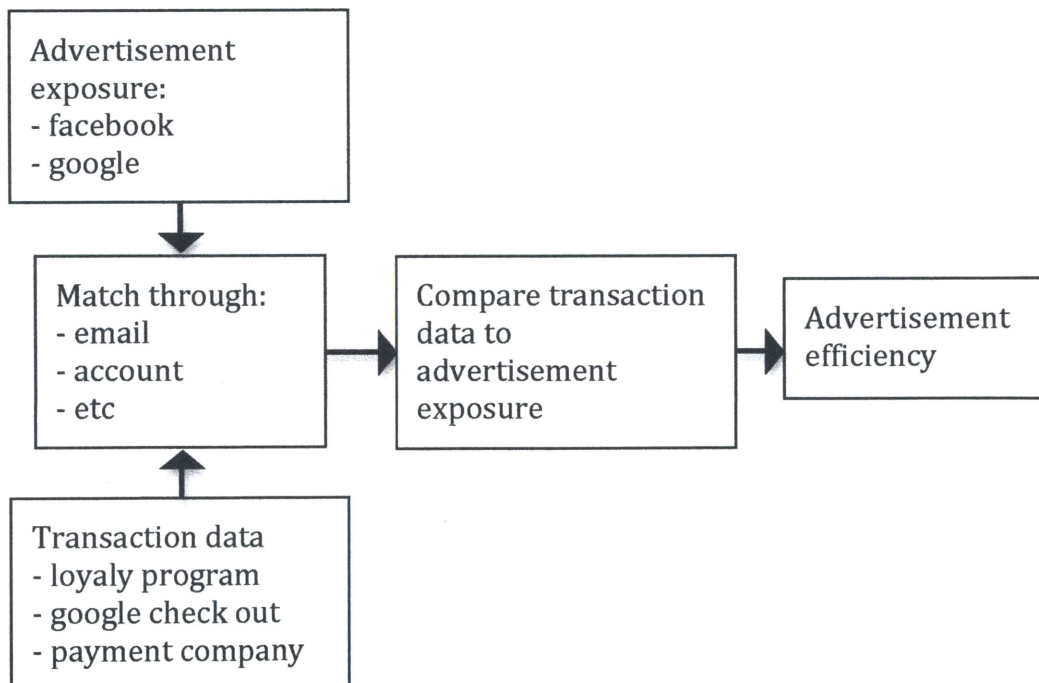


Figure 3: Advertisement exposure-Transaction data measurement process

Apart from retailers and online advertisers, other players could possibly also be part of the program, each providing even further benefits. As mentioned, suppliers and manufacturers could also join. This would allow for new kinds of improvements in the actual products and services, such as individualization, or improve the brand specific advertisements (as opposed to retailers' advertisements). Furthermore, participants could also be firms that specialize in just collecting customer information, such as mentioned Acxiom, Expedia and DataLogix. This way, the more participants such a retail network has, the greater the value for all members.

Include customers

Another dimension of the network externality that is possible, depending on how the information sharing is done, is that it could also affect *customers* willingness to participate in various membership programs. If customers see additional value in becoming part of loyalty programs because the advertisements, deals and offerings generated in return for them are significantly improved and the spam-dimension is reduced, more customers might join these programs. This could be true especially if these programs consolidate to fewer programs (so called coalition programs mentioned earlier), making it easier for customers to manage all their accounts and create a consolidated communication channel. As more customers join the program, this will in turn also increase the value of the database. Thus we can expand the model even further to include customers, as in figure 4. The network externality at play can have a very important part in the incentive companies see to join a sharing initiative, even though there may naturally be other forces that outweigh this aspect, which will be discussed later on. However, this is a situation where the result potentially could be a win-win-win situation.

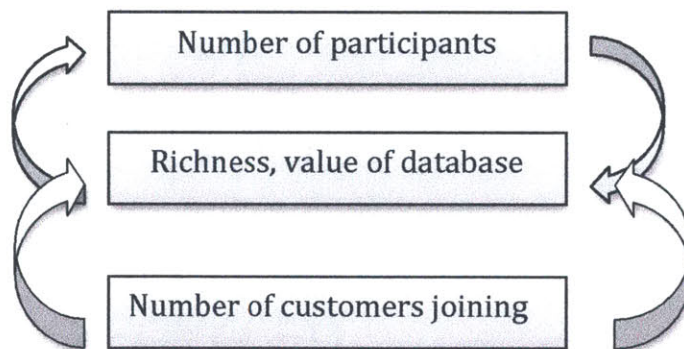


Figure 4: Network dynamic including customers

Motives that influence collaboration

The choice to engage in sharing of customer data between organizations is not a trivial decision. It goes to a large extent against the very natural instinct of competing firms, where anything proprietary and valuable should intuitively not be shared. Issues of trust are very important and privacy concerns of customers and many more issues arise that affect the decision. The true factors for each individual case is impossible to determine, but it is possible to gather clues from literature and interviews to identify some major dimensions and factors that should play a part in deciding whether to join in sharing information.

In this paper, the identified factors have been put into four categories:

- Legal and privacy concerns
- Technical concerns
- Knowledge gains
- Strategic concerns

These four categories each contain a subset of factors, and aim to together capture reasons and factors companies will consider before engaging in any collaborative activity. The proposed models in the next chapter will use these categories in order to evaluate how the different models differ and why some models might arise and some don't. Before that however, more specific factors of each categories will be presented.

Legal and privacy policy concerns

This category focuses on trying to capture both legal aspects of sharing as well as privacy issues and concerns that companies have. The interviewed companies that deal with customer data often consider customers' reactions to perceived privacy intrusion and so are fairly considerate to their concerns.

- Strictly legal regulation:
 - Sharing data about customers is not only a choice of businesses. Depending on what the agreements are about information disclosure to third parties between the organization and consumer, this could limit its possibility to share information.
 - Regulation could also affect the decision to share. The EU is for example currently discussing an overhaul of the use of personal consumer data by companies. The discussions so far reveal that it could forbid companies to use personal data for advertising purposes unless explicitly given consent by each individual consumer. Such a regulation could disrupt current practices and serve as an impetus to new models of handling consumer data but also make sharing practices even more difficult.
- National differences for multinational companies:
 - Regulation concerning how companies may use consumer information differs across countries, and perhaps more significantly between the US and the EU and the rest of the world. Setting operational processes for a multinational company might therefore involve taking many different regulations into consideration, making it more difficult.
- Sensitivity of data:
 - Different industries can have different levels of sensitivity in their data. Retail data can also be more or less sensitive. Pharmaceutical products might be more sensitive than groceries for example and could affect whether organization choose to share or not.
- Customers perception of privacy:
 - Public opinion about consumer privacy can shift over time, and depending on the pressure from the public, sharing information could be less or more attractive. An erosion of trust from customers can have serious long-term consequences (US-CERT, 2005).
 - Expressed consent could also affect how intruding customer perceive companies. If signed up for a reward program, usage of personal data might be more accepted by the customer for that particular company than usage by more unknown sources.

Technical concerns

The decision to share or not to share also take into consideration the technical aspect and difficulties implied in a decision to share. Companies might store data in different systems with different ontologies. Even though UPC and EAN bar codes are fairly consistent for many product categories in Europe and the US, it certainly is not for all products³. So identifying the same SKUs across different databases could pose a huge obstacle that is of concern to companies.

- Compatibility and integration issues:
 - o Because of multiple touch points and channels, companies can often struggle with having a common and consolidated data warehouse of their customer information. Standards in formats vary as well, both internally and between firms⁴. This can create large difficulties when trying to merge the databases into something that can fruitfully and readily be mined to extract value.
- Technical changes over time:
 - o As companies update and change their IT-systems, inter-operability issues could arise in the future. This could possibly mean that a collaboration practices could restrict future changes or lead to problems when introducing changes.
- Ability to exclude aggregation from aggregator:
 - o In some cases, it is impossible for a participant to hinder aggregation and thus the decision to contribute with their data is not an option. Retailers for example can't stop payment companies like VISA from aggregating their customers' purchases.

Knowledge gains

The main driver for engaging in collaboration and aggregation activities is that retailers could gain valuable insights and understanding about their market and consumers. The estimated value in terms of knowledge potential of sharing is therefore an important dimension and needs to be evaluated. Factors could include:

- Relative size of datasets
 - o Size could play an important role in as far as what the additional value of sharing is. A big company has less to gain from sharing with a small organization. Small companies could possibly influence this through collaboration networks, similar to the ATM evolution discussed earlier.
- Skills in analysis
 - o How skilled a company is at extracting value from the increased data will also affect how willing they are to share data. If they can't leverage the richer data, then they will only be helping the other companies to attract customers, possibly even competitors. Thus the less skilled at generating value from data a company is, the more likely to be hesitant about sharing it might be.
- Combined marketing channels
 - o Included in knowledge potential is the factor of potential ability to create a combined communication channel with customers. Interviews revealed that sharing retailers saw

³ For example, there are unique store products in the apparel industry.

⁴ Even though some standards are existant, like for example the NRF-ARTS standard.

great value in being able to provide combined offers through a combined channel that is more convenient to customers and thus leads to greater customer satisfaction and engagement.

- Furthermore, with so-called loyalty coalition programs, consumers subscribe to fewer programs and do not have to deal with multiple and separate reward collections, redemption and loyalty cards, making such a program more attractive to customers.
- Relevance of datasets
 - The knowledge potential will likely vary with the relevance of a dataset. For example, a book store might not gain much insights by sharing with a car rental service, but could get more value from sharing with an apparel retailer.
 - Relevance of datasets could also be affected by what geographic coverage the databases has. Naturally, there is very little point in sharing customer data (but could be some, though not as comprehensive) if your partner is only operating in an area far away from where you are operating.
- Frequency of purchase
 - Some interviews explained how retailers with high transaction frequency, like grocery stores have such rich transaction histories that they can become very sophisticated in their analysis and reach conclusive findings. However, retailers with lower frequencies of transactions, like durables and more unique items, will have very few observations for each individual and so are very limited in what conclusion they can draw from their transaction history. These retailers can benefit more from the richer data as the data becomes enriched with more observations, variables and so better analysis and insights can be generated.

Strategic concerns

The last category aims at considering more strategic concerns for retailers. Even though legal and technical concerns are limited and the knowledge potential is huge, it is not obvious that companies should start engaging in sharing. Some factors that will have an impact are:

- Ownership structure
 - If different retailers are owned by the same parent company, sharing is probably to be much more likely for obvious reasons. Retailer with no ownership connection will face many more concerns of trust and proprietary issues of customer data.
- Competitive relation
 - A company's relation to the other company can also be an important concern. Direct competitors could gain very relevant information about relevant customers, but they at the same time give the same advantage to its competitor, whereas actors that don't perceive themselves as competitors would find it easier to share.
 - Even non-competing retailers can have issues with sharing, as it is possible that the other firm expands and moves into the previous non-competitor's space. The access to data that a sharing initiative would give might facilitate or even initiate such a migration.
- Filtering of shared data
 - How the shared data is filtered before shared could play a major role. Unfiltered sharing of all data about customers might be a hard sell for companies, but filtering of various kinds could make sharing less controversial. One way is to only share behavior, but keeping

- identities hidden. Granularity of data could be limited to only level 1 data instead of level 2. A common practice is also to aggregate individual pieces of information and only share the aggregated results.
- Pooling of resources and capabilities
 - o Many companies participating in some structures found additional value in ‘outsourcing’ their analytics of transaction history to specialized operators and so benefit from expertise and economics of scale. Small companies can not always afford or know how to create analytics departments, and so turn to outsource this. By sharing and aggregating customer data in retail, the scale and cost sharing needed for these kinds of operations to become economically viable could be achieved.
- Lock-in and loyalty concerns
 - o Some firms might view a shared system as something that undermines their ability to create loyalty through membership programs as well as diluting their brand if a strong brand is associated with weaker retail brands.
 - o As many people have multiple membership programs, simplifying this process could allow for customers to increase their participation in more membership programs and thus become less loyal to the initial retailers. This is perhaps especially true if exchanges of points and rewards are made possible, something that seems to be increasingly happening (Postrel and Hlavinka 2012).
- Psychological loss of control and mental model of proprietary of data
 - o Because data has in recent been years regarded as a source of competitive advantage and thus valuable, managers might feel its not good praxis and rather counter-intuitive to share something proprietary and valuable. This point could have a greater impact than suggested by its space in this paper, as mental models can be notoriously hard to change.

These factors are not an exhaustive list of all elements that affect the decision process, but because of the relative novelty of this practice, it is a starting point from which further discussions and research can be made. This paper will now attempt to look at examples of each structure and see which of and how these different dimensions might help to explain the observed behavior and what future scenarios can be expected.

IV: 5 generic structures of information aggregation

This chapter will first introduce two dimensions of information aggregation: (1) the levels of information granularity and (2) structures between participating organizations. The following part will explore each structure in more depth and give examples of cases today as well as look at why each particular structure has emerged and its dynamics. Each structure will be evaluated by using the decision dimensions established in chapter III.

Levels of information granularity

An aspect that is important when looking at how information is aggregated is at what level of detail the information is. As discussed previously, the transaction data can be a total amount or it can be detailed and encompass transaction data on an SKU-level. In this paper, we will define total amount data as a level 1 transaction data and SKU data as level 2 transaction data. This dimension will help us separate between different aggregation types. The level of granularity will matter for how valuable the aggregated dataset will be and what kind of benefits it could generate, as well as affect the underlying decision factors for companies.

Models of information sharing structures

Given that some organizations today currently already share or aggregate information, it is possible to observe different structures of how that sharing occurs and what the relations between players are. These structures will naturally depend on multiple factors. This paper aims to use the four category factors from the previous chapter to see how the identified structures differ from one another. Where possible, examples of those structures will be looked at in more detail.

To introduce the possible structures that will be addressed, each organization will be represented as a data silo. This will allow for easy visualization of how participating members relate to each other. There can be many different ways and structures in which these entities could interact with each other. Information sharing structures have been examined previously by for example Barret and Konsyski (1982) in their paper “Inter-Organization Information Sharing Systems”. However, their paper has a strong operational and system-compatibility focus. Their framework is parallel, but perhaps not directly applicable to the more operation-independent customer data type of sharing. Madnick and Siegel discuss relationships between aggregator (like kaya.com) and aggregatees in their paper “Seizing the opportunity” (Madnick and Siegel 2001) in various industries on the web and provide another way of looking at structures of aggregation between organization, both involuntarily as voluntarily. In their paper, aggregators are entities that collect and aggregate information from different sources and organization whereas aggregatees are the entities that provide the partial information that together make up the new database of the aggregator. Depending on the relationship between the aggregator and the aggregate they identify 4 different kinds of relationships that can occur during aggregation;

- (1) no aggregation,
- (2) aggregation with partnerships,
- (3) aggregation with no partnerships and
- (4) aggregation with ownership.

Similar to those structures, actors in our setting can establish various forms of relationships that are similar to the ones identified in their paper. Therefore, this paper will draw upon that framework, but a modified and extended version will be used to better suit the needs and purpose of this paper. In order to introduce the structures used in this chapter, we will start with differences between Madnick and Siegel's structure and the one used in this paper.

One difference is that this paper also discusses direct relationships, where neither participant is only an aggregator or aggregate. In the many different ways retailers can interact with one another, one way is that two retailers can choose to *directly* share information with one another. In this scenario, they will themselves handle all interactions between the companies and deal with any kind of filtering or trust issues themselves. This dualistic and horizontal relationship can further be expanded to include more players, leading to possibly more complex logistics when sharing information. To handle this increased complexity, one of the retailers (or possibly all of them together) can create a new organization, consortium or a joint venture that deals independently with organizing, gathering and distributing information in a way that all parties are satisfied with. This new entity would thus act as an aggregator in the sense of what Madnick and Siegel call an aggregator with ownership.

Similar to creating a shared consortium with a common repository, it's also possible that a third party actor takes on the role as coordinator and guarantor of data dissemination. In Madnick and Siegel's terms, it could be viewed as both an aggregator with partnerships as well as without. In addition to these four structures, a fifth is also possible that does not exist in the framework of Madnick and Siegel. Alexander Pentland, a prominent data scientist of MIT, talks of how a personal information repository could be created, where information companies and government agencies have about an individual is stored in a personal repository owned by each individual and where the information is free to be used or shared as she wishes, something also supported by the World Economic Forum (Pentland 2012) (World Economic Forum).

Madnick and Siegel's framework for the relationship of aggregators and aggregatees has thus some strong similarities to the structures we've described above. For the purpose of this paper and focus, a modified version of their framework will be used and the various structures described will be classified into five different types closely resembling that framework. The structures identified for this paper and purpose are described in table 1.

As opposed to Madnick and Siegel's categories, which have a greater focus on what kind of relationships there are between the aggregator and the aggregatees, this paper has a greater focus on the actual structures of sharing, with relations between organizations as a secondary dimension. Figure 5 tries to show how Madnick and Siegel's framework maps onto this new modified version.

Information structure	Description
- No information sharing	In the case an individual retailers solely uses its internally generated information and does not share with any other actor
- Direct information sharing	In the case of a individual retailers collaborating directly with each other.
- Joint consortium of information sharing	In the case of the retailers creating a separate unit for managing the information sharing. Can be owned by one member or a joint venture by some/all participants
- Third party aggregator of information	In the case of a third party entity who deals solely with managing the information and is not involved in any retail activities themselves or linked to any retailers through ownership.
- Personal data wallet	All personal data is collected by the individual from each entity. She then controls with whom to share what kind of information.

Table 1: Identified information sharing structures, modified from Madnick and Siegel (2001)

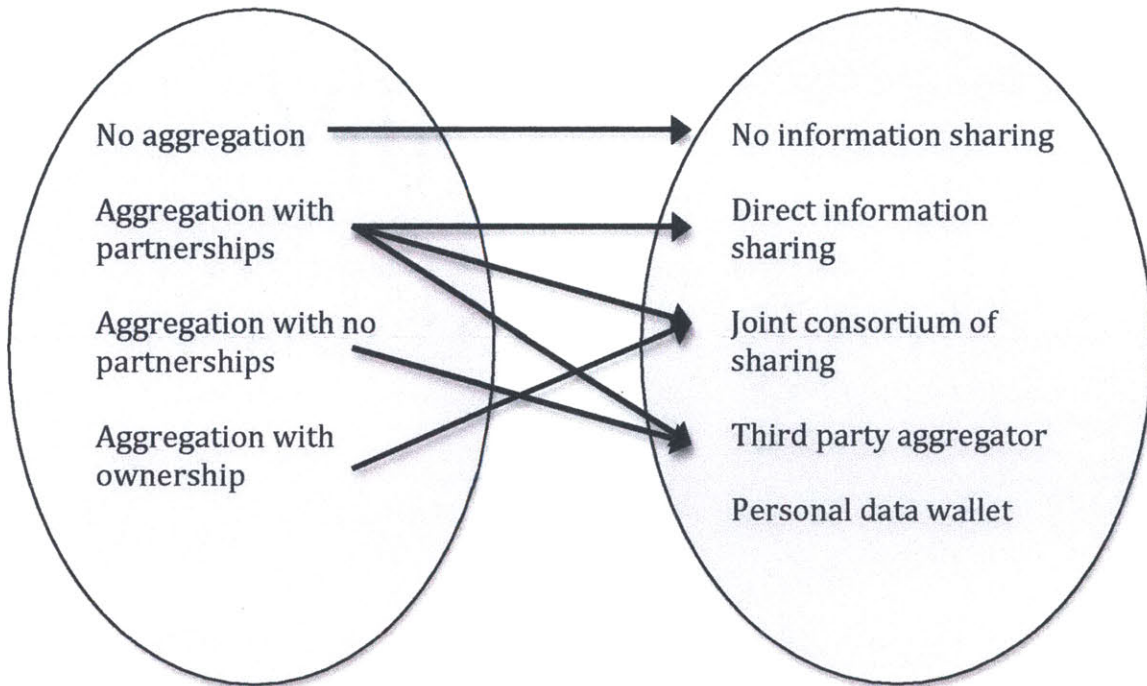


Figure 5: Mapping of Madnick and Siegel's framework onto the new modified framework

All of these different structures have examples in the world today. However, to the best of the author, it has not been examined to any great extent as to why some structures occur in certain circumstances and why some occur or do not occur in other. Each of the different structures will be evaluated along the dimensions we established in the previous chapter in order to see how they differ and what the factors behind each structure are. The four categories of dimensions evaluated in this paper are once again:

- Legal and privacy concerns
- Technical concerns
- Knowledge gains
- Strategic concerns

The 5 structures

This part will go into depth about the examples of each structure and try to explain the dynamics of that particular structure. As mentioned previously, the structures will be evaluated along the dimensions identified in chapter III. Important to note however is that these dimensions will be scored on a *perceptual basis*, i.e. what companies of that category perceive to be the case, regardless of what the ‘true’ nature of each factor may be. By using a perceived score, the differences between each structure will be explained in more reality-anchored terms.

No information sharing

No information sharing is the structure that seems descriptive for most companies and involves no interaction between retailers (figure 6). This is thus the default mode for most companies right now.

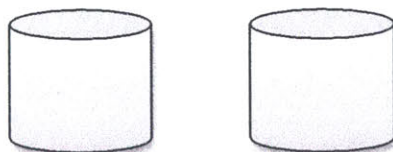


Figure 6: No sharing between organizations

Legal and privacy concerns

Non-sharing retailers are concerned with privacy issues and tend to respect customers’ privacy. For this reason, they tend to fear sharing data with external parties. Furthermore, stated policies in membership programs sometimes exclude sharing to third parties. Sharing on a level 2 basis would be even more difficult as it becomes even more personal.

Technical concerns

Retailers in this category also see major technical difficulties in combining databases from different sources. Compatibility issues as well as cumbersome managerial issues hinder collaboration practices. Level 2 data is more cumbersome to integrate than just level 1.

Knowledge gains

As many of these companies are realizing the value of data and exploring how they can gather and analyze this data, it could be distant for these companies to start sharing it even before they have understood how to use it themselves. New to data, these companies are probably less far down the experience curve of analyzing and using this kind of data, meaning that their skills at analyzing is still developing from basic levels. Helping other firms that are better at analyzing (by sharing their data) puts them at an even greater disadvantage without realizing any value themselves. For this reason, it can make sense to wait, even if the potential value of sharing information is understood.

Another reason in this category is also often size. Small companies appear from interviews done by this author, not to have the capabilities or resources to create sophisticated analytics initiatives or even data collection processes. This naturally renders it hard to start sharing data.

Strategic concerns

A major concern with sharing data is that it doesn't fit with the mental model of managers on how to use data. Customer data is viewed as something valuable and proprietary, and thus should not be handed away. Non-sharing retailers also typically seem to have competitive concerns when sharing, as the most interesting data is perceived to be that of direct competitors.

Little value and many obstacles

Companies in this category seem to view sharing of customer data as an initiative that has large legal and technical concerns yet only limited, or at least not large enough, knowledge value. Strategic concerns are perceived as fundamental and do not fit with existing mental models of how to handle data. The influence each of these dimensions has on the companies' likelihood to start sharing customer data is visually scored in table 2. It's worth noting however that this structure does not necessarily have to be an explicit choice of companies. Rather it can be the default mode from which all companies come, thus they will fall into this category even if they're not even aware of the choice or possibility of doing so.

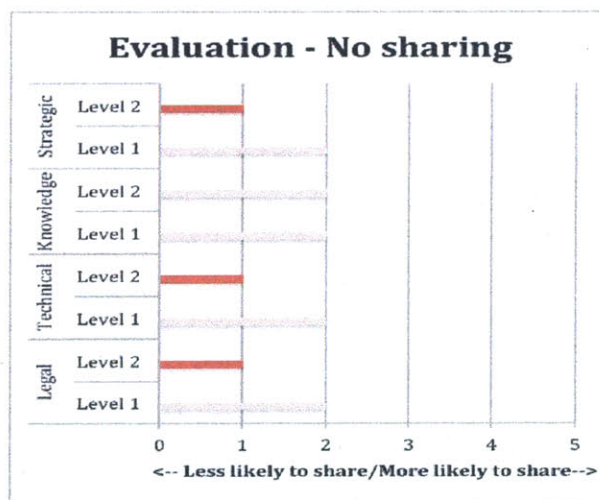


Table 2: Perceived dimension scores for companies not sharing

Direct sharing

Direct sharing aims to describe when companies share information directly with each other, without any intermediary (figure 7). This can happen when there is a specific value to be gained through a particular relationship. This kind of sharing occurs rather frequently for retailers with more sophisticated relations to their suppliers. P&G for example collaborates with their retailers when it comes to transaction data, which allows them to do more sophisticated analysis as well as improve their day-to-day operations. This kind of vertical information sharing is fairly common practice today and has led to improved operations in many ways. However, a form that is more relevant for this paper and less related to stock-outs and replenishment automation is how sharing *customer* data to suppliers and manufacturers could allow for individualization of products and services. Even though it is not focus of this paper, Piller and Schaller (2002) discuss the benefits and difficulties as well as structures of this phenomena. They show how Adidas for example has strong collaborations with retailers about customers and their data in order to allow individualization of sneakers.

One other possible example of direct sharing is the example of Facebook and third party aggregators such as DataLogix and Acxiom. Even though these companies themselves are third party aggregators, the sharing structure between Facebook and each company is direct and very limited). By combining information about online advertisement on Facebook and level 2 type transaction data, advertisement efficiency can be measured to a much greater extent, as illustrated in figure 1 in chapter II. For example, if there is an advertisement for a particular snack on your Facebook page, it is possible to cross analyze that with your transaction data and thus see if it led to a sale. This is done in this case by matching the customer's stated email address in the loyalty program to the email/s associated with a Facebook account. Apart from advertisement efficiency and individualization of products, the advertisements themselves could now become personalized not only based on your browsing history and likes captured by Facebook, but also by your offline purchasing behavior. This relationship is however purely for advertising purposes and claims to be anonymous (Wasserman 2012), but it could possibly evolve to more comprehensive analysis in the future.

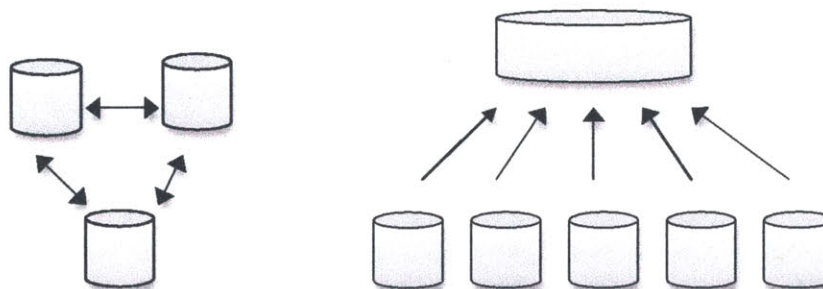


Figure 7: Direct information sharing structures

Legal and privacy concerns

Sharing of customer data does not seem for these companies to have a hindering effect. These collaborations tend to be very specific and rather limited, and has a very specific targets in mind, which may lead to them getting little attention. Facebook's collaboration with retail transaction has gotten some attention for privacy concerns, but nevertheless has not yet resulted in any major complaints. It is a rather

new initiative as well, so it might be that it's effect has not yet fully materialized. However, these concerns seem to be regarded as fairly limited when sharing occurs in this way.

Technical concerns

Due to the very specific nature of these kind of collaborations, technical difficulties can to some extent be limited. If there is enough value, it is worth investing in solving any compatibility and standards issues and it only needs to be done for that specific investment in mind. Despite this, they can be somehow difficult to manage for companies like P&G who deal with many direct sharing partnerships. For the large and meaningful partnerships (like major grocery and convenience stores), they find ways to overcome these concerns.

Knowledge gains

These partnerships are usually very specific and so has a very high relevance for the partnering companies, or at least one of them. The targeted approach allows for analysis with immediate returns, whether it is advertisement efficiency or individualization of products, and is thus perceived as valuable. However, these partnerships seem to require detailed data to be valuable, meaning level 1 data might not be so relevant.

Strategic concerns

The specific sharing of this nature does not seem to happen between directly competing rivals. An interesting phenomena however is that big and powerful companies like Adidas, Puma, P&G and Facebook seem to be pervasive in this structure, which could be an indication that power balance in relationships could have an important. A certain size and value could be needed in order for the cost of creating a functioning partnerships to be justified, limiting these very specific and targeted practices to large companies.

Specific investments make investments worthwhile

This collaboration structure seem to happen between large players with very specific targets in mind. With particular large benefits in mind, narrow investments can be made with clear payoffs. At the same time, privacy concerns and strategic concerns are avoided through the very narrow scope of these investments, rendering a perceived evaluation score as seen in table 3.

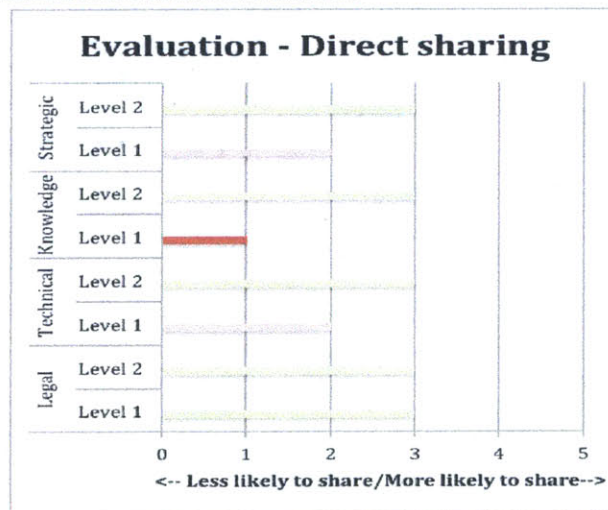


Table 3: Evaluation score of direct sharing

Joint collaboration program

A step further down the collaboration scheme is to create an even stronger and more formal structure in which to share information, owned by one, several or all participating firms. Examples of this in the world today seem to happen in what's called coalition loyalty programs, where there is one loyalty program for several different organizations, similar to many loyalty programs within airlines. StarAlliance is a joint organization between 27 members and allows for flexible miles or points collection and redemption, using the airlines' individual loyalty program for all partners. That means that information across many airlines is consolidated between partner companies and the combined information they have of you and your behavior is then shared, as illustrated in figure 8.

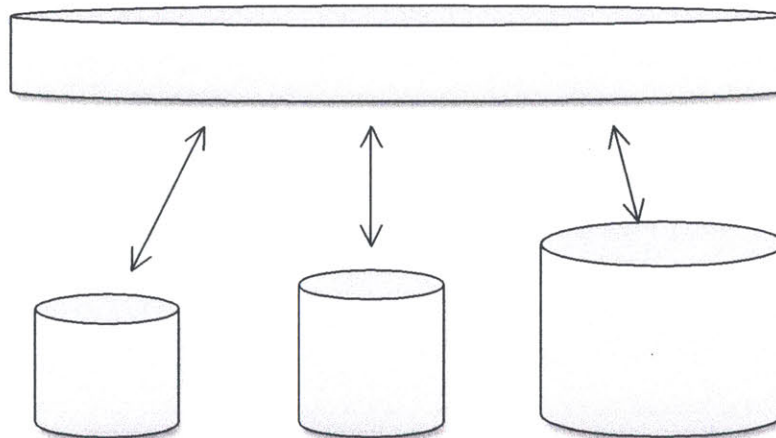


Figure 8: Joint consortium where the aggregation firm is owned by one or several participants

Examples more related to retail include Flybuys in Australia and New Zealand (owned by one member in Australia, but by a consortium in NZ) and Kooperativa Förbundet (*the Cooperative Federation*), KF, in Sweden. KF has one official program called “Medmera-kortet”, a reward program that is useable in many different retailers. It started with the several retail chains owned by KF, such as a bookstore chain and grocery store chain. With the same parent company, sharing became less of a controversial topic and was initiated and has allowed these retail chains in the KF-family to quickly realize both the business value and the value for the customer in sharing a reward system that allows them to consolidate their customer data across retailers on a level 1 and 2 transaction data (currently varies). However, they have also extended this reward-program to incorporate other retailers and organizations, to some extent outside of their ownership sphere for example Hertz, the car rental service. Recently, they also include customer data from an apparel retail company (with multiple large retail chain companies in Sweden) that are now part of the program (at the same time as having their own programs). That only happened however after KF acquired 25% of the company and forced the apparel company to join the program (something similar to what Madnick and Siegel called ‘dependent’ firm). So in this case, one member owns the formal organization collecting information, but information is shared between all participants, with a seemingly unfavorable bias towards the partly owned company.

Legal and privacy concerns

Even though this kind of collaboration inherently produce more sharing of customer data due to their larger scope of partners, organizations like KF and StarAlliance seem to be able to mitigate many of these concerns. Through coalition programs, these firms allow consumers to also benefit from their shared data through programs that span multiple organizations. Studies have shown that consumers seem to be willing to give up some concerns for privacy for benefits (Pridmore 2010). This helps to reduce privacy concerns and allow for increased sharing.

Technical concerns

Larger aggregation projects of customer data like this naturally may create large technical issues. Even though KF for example have not yet fully integrated all their data at level 2 for all participating firms (most notably for the newly joined apparel consortium), these concerns are still overcome in favor of the value potential in doing so.

Knowledge gain

This structure, depending on participating firms, can allow for large aggregation of customer data and so allow for much richer analysis to be done. StarAlliance reveals a better understanding of their customers' travel behavior, including car rental and hotel stays which is sometimes included into the loyalty program (Eriksson and Wallman 2011). This provides a much more holistic travel behavior that can generate improved analysis (similar to what Orbitz, the travel comparison site, can do. Orbitz is in itself a joint collaboration between travel companies). For KF, they also gain better holistic understanding of their customers. Shortly after joining the KF loyalty program, the apparel company uncovered a completely new customer segment, which was only discovered after the customer data from the organizations was combined and analyzed. The apparel store previously regarded families to be their most prominent target group, based on their current loyalty program and analysis. After joining databases and introducing loyalty program compatibility, this still held true but it also turned out that a big customer segment was a certain type of non-families that seem to be buying presents and gifts for other families. This new customer segment represented significant sales and has now impacted the way they market their goods. They also found great value in combining their marketing channel and pooling resources together, generating greater outreach and cost savings.

Strategic concerns

An important element in these partnerships seems to be competitive relations between participating firms. For the airline industry, and the example of StarAlliance, they are at first sight providing the same product, but in fact they are not directly competing. The program intentionally excludes direct competitors and instead allows the program to cover larger geographic regions with their partnerships (Eriksson and Wallman 2011). This way, existing customers are more inclined to use their network of airlines in order to get the most out of the reward programs. An important note here made by Brandenburger and Nalebuff (1996) is that it is very important to think through how you create the rewards program to actually incentivize loyalty and not gaming behavior i.e. customers joining many other programs to extract value but not necessarily be a loyal customer.

Another aspect that is important in this kind of collaboration structure is the ownership situation. KF for example was at first only within their own conglomerate of retail chains and was only able to provide new

meaningful partners after buying 25% of those companies (except for smaller partners like Hertz and similar). For StarAlliance, all participating firms own the loyalty partnership, reducing ownership-bias issues. Bias in ownership is discussed in Madnick and Siegel's paper in the situation of Intershipper and iShip, two price comparison aggregators for parcel delivery services. One was part of UPS, the other independent. Madnick and Siegel argued that there was a, due to the information bias (in UPS' favor), very strong incentive for competitors of UPS to either create a shared aggregator or support the independent aggregator, given that it shared information equally to all players. Fully joint ownership in this structure could therefore be a prerequisite for a stable partnership.

Perceived value from customers is also an important aspect. By creating coalition programs, customers are happier as loyalty programs are simplified and consolidated, making it easier to collect and redeem points. This is particularly true if greater reward currency exchange happens. In a study by Colloquy and Swift Exchange, airlines and hotels is showed to be on the forefront in terms of collaboration of rewards. Almost all of the major airlines also allow reward currency exchange, both into and out of their program to other companies both in the same as well as different industries. Hotels are more restrictive but still allow an increased level of transferability of reward currency to other industries. These collaborations allow for an improved outreach of promotions and aims to grow the pie for all parties by collaborating (Postrel and Hlavinka).

Large potential with ownership concerns

Joint consortiums have the ability to extract significant value from their combined resources. Coalition programs that allow for customers to benefit from the collaboration as well mitigate privacy concerns. The largest source of potential conflict between firms is the ownership structure favoring the owner of the network. However, this can be avoided through joint ownership or successfully building trust between partners. Building that trust for larger networks however seems to be hard and may be the difficult road to take. Figure 4 shows the perceived scores for the joint consortium.

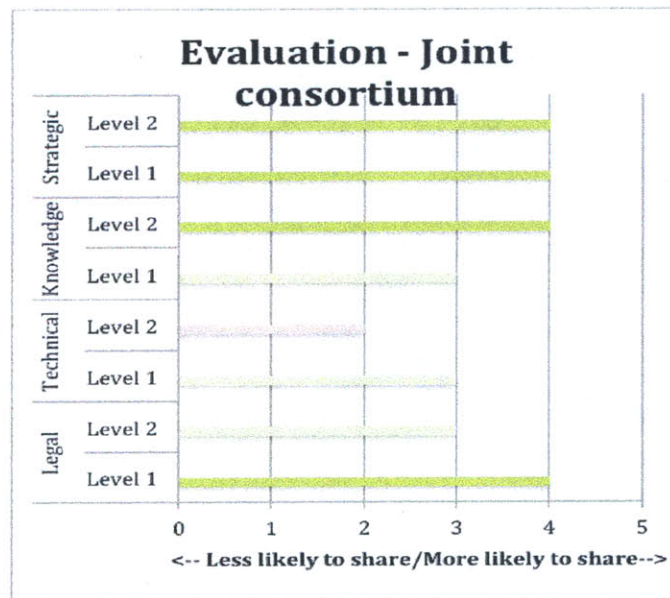


Table 4: Evaluation score for joint consortiums

Third party aggregator of information

Another form of a more comprehensive way to aggregate customer data is to use a third party organization that does not itself participate as a retailer, thus avoiding some of the concerns that the joint collaboration could have. It turns out that there are two major branches of this kind of aggregator: coalition programs and payment companies. Since they function in different ways, each will be analyzed separately here.

Coalition programs

This structure has a, often specialized, organization that collects information from different retailers and then distributes it to participants in an agreed way. An example that could potentially do this, but that does it in a different way is Payback (owned by American Express⁵) in Germany, Poland, India and Mexico. They have created one loyalty program that is valid for a vast number of retailer, both online as well as offline, in Germany and can thus consolidate all the information gathered. This card is held of around 20 million Germans and includes a large number of small as well as large retailers. In the UK, a similar coalition program called Nectar exists with around 50% of UK households signed up. This is a partnership between more than 14 retailers including Sainsbury's, Barclaycard, BP, Homebase, and Ford. Nectar aggregates this data and performs level 1 type of cross-analysis between retailers in order to get new insights, allowing partnering firms to benefit from expertise and pooled capabilities as well as combined marketing channels and increased perceived value from the customers point of view. They do also provide help with analyzing individual partner data on a level 2 basis.

The structure of these collaborations are illustrated figure 9, where the major difference to joint consortiums is of course that the aggregator is a third party outside of participating retail firms. Otherwise, the two different structures are very similar in many regards.

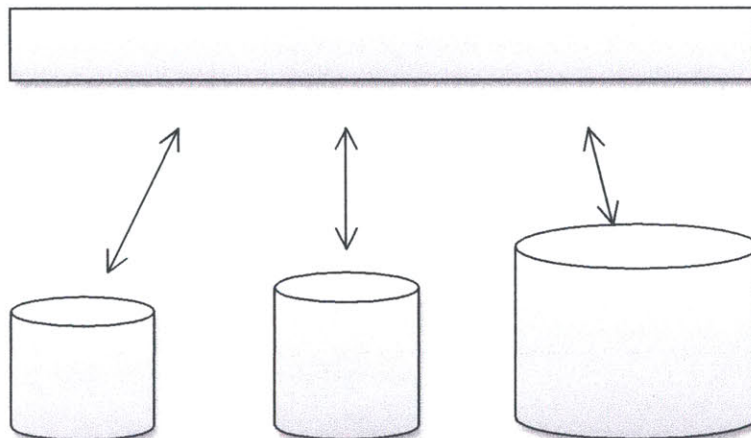


Figure 9: Similar to Joint consortium, but here the aggregator is not a participating retailer but a third party operator

Legal and privacy concerns

The concerns for this kind of coalition programs are somewhat different to joint consortiums, although similar. The customer data is now handled by a third party, something consumers may find both less or

⁵ Payback was interestingly founded by Metro, the 4th largest retailer in the world. It thus started off as a joint collaboration but was in 2011 sold off to American Express.

more comforting. However, by signing up to a specific coalition program, consumers have still to a great extent agreed to have their data being used and shared in return for rewards and deals, mitigating these concerns.

Technical concerns and Knowledge gains

These concerns remain the same as for joint collaborations discussed earlier, even though there are idiosyncratic differences between each example. Payback for example has over 600 participating retailers (some bound to compete) but only publishes large industry reports and helps individual partners with their analytics. Nectar on the other hand do perform aggregated analysis to a greater extent to see how specific customers behave across the network, but then again is limited to a handful of sizeable retail chains, each with exclusivity within their segment as to avoid direct competition.

Strategic concerns

The difference to a joint consortium is that a third party operator could be perceived by managers as a greater psychological loss of control of valuable data, even though this could be regulated in a contract. It could however be preferable with a third party operator if the alternative is a program owned by one or several other retailers, leading to trust issues and information bias. Just like in the case of Intershipper and iShip, an independent operator could lead to greater levels of trust and lessen management and agreement issues that could arise in a joint operation. Third party operators could filter data in an appropriate way as to not share sensitive data, while still achieving value through high-level aggregated analysis. Some interviewed retailers furthermore do not seem to view customer analytics as core and thus regard ‘outsourcing’ to experts as logical and allows them to focus on their core operations.

Strategically easier to create larger networks

Very similar to joint consortiums, the large difference for this structure is in its strategic implications of being an independent operator. This may facilitate for cooperation, but it can also make the start significantly more difficult, unless a large marquee partner can be found from the start. Table 5 indicates a slight advantage over the joint consortiums.

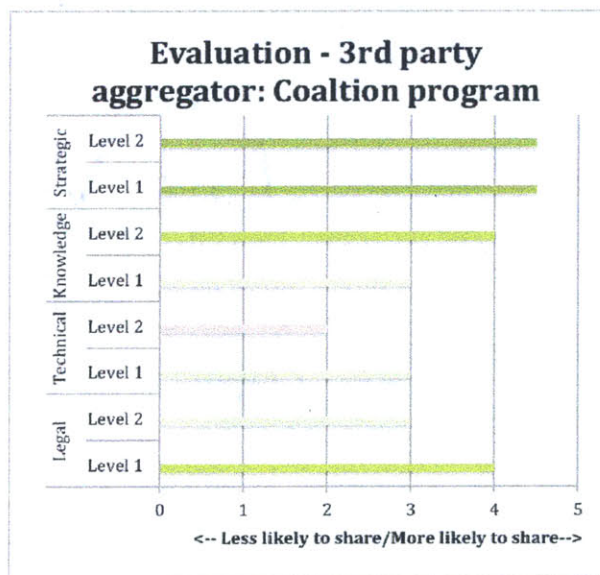


Table 5: Evaluation of third party coalition program

Payment companies

Some other interesting aggregators with this structure are of course the already mentioned payment companies. VISA, MasterCard, American Express, Discover and similar companies can, and do, aggregate information without the explicit partnership of any of the retailers, and thus behave differently from explicit partnerships. In the online world, similar check-out gatekeepers could provide similar possibilities, one of the reasons it is a very competitive space with newer payment companies emerging. The traditional payment companies like VISA and American Express have in recent years looked more into the opportunity to analyze and mine the transaction data they have. However, from a customer data aggregation perspective, the dynamics of such possibilities are somehow different from the two previous structures.

Legal and privacy concerns

Payment companies naturally already have the transaction history of their card holders and thus would be in a very favorable position to exploit their data. However, there could be major concerns with this as customers expect a payment solution when they sign-up for a credit card and today do not expect them, as opposed to loyalty programs, to track your behavior and sell it on to others (which they inherently must to make full use of the insights of such analysis). Payment companies are furthermore restricted to a greater extent as to how they can use consumers' financial transactions. This naturally poses substantial legal and privacy difficulties for payment companies.

Technical concerns

Level 1 aggregation is no issue for payment companies as they already collect that kind of transaction data. However, level 2 data is not something that is possible with the technical architecture of today, as the POS-system (keeping track of detailed transaction records) is separated from the payment system. Even though it could be possible to change this in the future with new technology or integration, it is today not possible.

Knowledge gains

Because of payment companies restricted access to level 1 data combined with legal restrictions, they can only provide high-level patterns in spending of customers. American Express does this in their Business Insight department, and VISA for example publishes "Industry Insights" and is now also becoming more aware of deeper analysis of their data. In VISA's 2011 Retail report for Europe, they report how they see this as an emerging opportunity to segment and mine their payment data for better understanding of customers. It is possible therefore that we will see more of these activities from their side in the future. They do not however provide any combined marketing channel. The greatest advantage is perhaps the coverage of consumers' total transactions. However, issuers of cards seem to some extent starting to engage in bonus programs through their credit cards (for example American Express' membership rewards), but still lack a combined marketing channel that coalition programs provide.

Strategic concerns

If payment companies start to fully aggregate and sell insights and data to retailers, they cannot discriminate who's data is shared to whom. There is difficulty thus in filtering the data, which may make retailers look less favorably at such a program in comparison to a walled coalition program for example.

Furthermore, retailers may still have to engage in normal loyalty programs as level 2 data can provide much other analysis than an aggregated level 1 type analysis by payment companies could. This could lead to coalition programs being attractive despite payment companies' aggregation efforts, which ultimately could lessen the demand for payment companies' insights as retailers now can provide those insights from 'internal' coalition programs.

Already occurring but has significant limitations

Payment companies have a huge advantage in the sense that they already have the aggregated data on a level 1 basis and so possess a huge database of valuable customer data. However, the technical limitations (to only level 1 data) as well as legal and privacy issues are substantial and so it is unclear exactly how much they would be able to leverage the data for knowledge discovery. Table 6 illustrates the scores these factors generate.

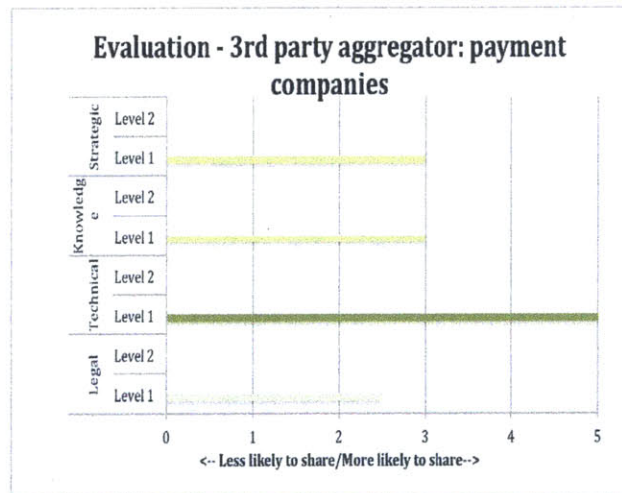


Table 6: Evaluation of payment companies' sharing

Personal data wallet

In the World Economic Forum meeting at Davos in 2012, a new report was released concerning personal data in our new digital age (World Economic Forum). The report explains how personal data is the "new oil of our age", possibly metaphorically hinting at several dimensions of it. First of all, it is regarded as a new asset class, an asset class that is extremely valuable and needed for businesses today. Second meaning is a metaphor in the traditional sense. Just like oil in an engine, personal data can make operations more efficient.

One of the major topics during the Davos report was the ownership of personal data. Both corporations as well as government officials from all around the world agreed that we must find a balance between company as well as government interests and individual privacy. Alexander Pentland, one of the key people in the initiative on Personal Data in the Davos, suggests in his New Deal on Data, that individuals should have the right to "possess, control, and dispose of copies of these data as well as all other incidental data collected about you – for instance, location and similar context – that companies and the government

must not retain” (Pentland 2012). This deal, if implemented, has numerous implications that will affect how businesses can acquire and use personal data. If individuals, can own copies and control non-necessary data (whatever that would be defined as) of companies, they become the ones in power of that important asset. Another important finding was the element of focusing on interoperability and open standards. This means that whatever data a company has of you, you can ask for an export of that data to yourself. And as an owner, you can share it with whomever you like and formats compatibility should not be an obstacle.

Aside from the very important social implications this could have, it also creates another way, perhaps the most powerful way, of sharing and distributing personal data. Theoretically, such a system would allow individuals to consolidate all their data in all kinds of information sources (retailers, hospitals, utilities, education, social networks etc) and then decide what, how and with whom to share that data, illustrated in figure 10. The result is therefor in the hands of the consumer and the outcome can range from companies knowing less than today to rich and comprehensive personal data becoming widely available to all kinds of companies, not only the big ones that gathered it, but also the small ones. Personal data could become ubiquitous and behave like a public good, at the discretion of consumers.

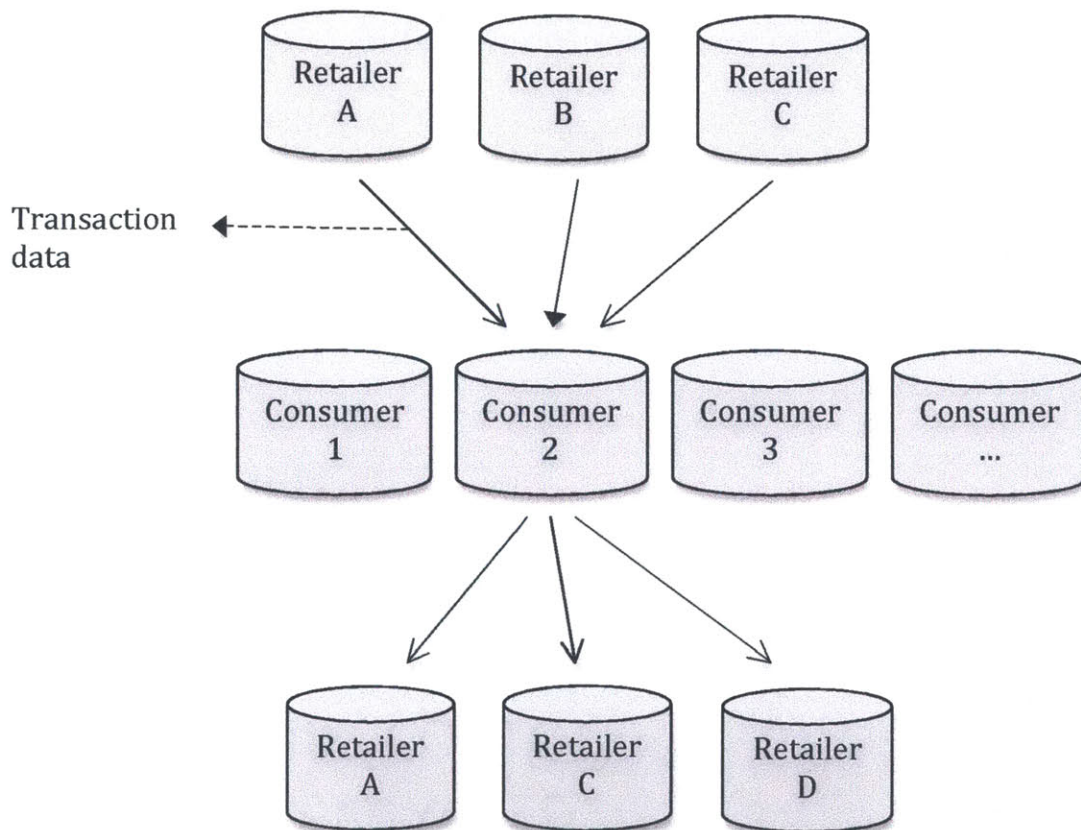


Figure 10: Personal data wallets. The aggregator is the individual consumer and he/she can share information as desired (for example here, not sharing with C, but with D)

The MIT Media Lab and the ID3 initiative, partly led by Alexander Pentland and John Clippinger, are working on the creation of what they call a Personal Data Ecosystem that encompasses personal data in many other areas, not just retail, but also financial data, medical records and location data. It's one example of an attempt to create a personal silo for individuals to store all kinds of personal data in order to more efficiently and in a trusted and legal way disseminate this data for the public good while retaining some control to the individual. Regardless of the many issues and obstacles to such a project, a personal data wallet could be both of level 1 as well as level 2 as far as retail is concerned. For this to be realized however, regulation could be an important factor and influence as well as public opinion about data privacy.

The coming discussion of dimension of evaluation will be purely hypothetical, since this does not currently exist. Therefore, there will neither be any evaluation score table as perceived scores do not exist.

Legal and privacy concerns

Due to the very nature of this structure, the choice to share will be completely at the customers' discretion and so should not provide any challenge from customers' perspective. However, recent discussion for new legislation in the EU about what and how firms can use customer data has led to complaints from corporations' side, stating that some legislations interfere with their right to conduct business, a right under EU charter (Out-Law 2013). The legislative action required however for such a data-wallet scenario to happen is substantial and somewhat unlikely in the near future.

Technical concerns

Even though this kind of aggregation is similar to the previous ones, significant technical solutions must still be found and implemented to generate a functioning and stable ecosystem under which a data-wallet structure could function. This could therefore provide a major hindrance.

Knowledge gain

The potential data available to companies however would be tremendous, especially to smaller companies who would now gain access to otherwise hard to get data. Depending on how much data customers want to share, significant coverage, including competitors, of purchasing behavior could also be achieved. It is also possible that the same platform that allows consumers to share or not share data can be used as a combined marketing channel, albeit less coordinated than a formal single aggregator. However, many companies could see a surge in accessibility to customer data.

Strategic concerns

A structure like this would inherently mean that source of advantage between firms would shift more distinctly to how well they analyze customers' data and are able to produce goods and services that address the needs of the consumers, rather than ability to collect proprietary data, something that could mean a weakening of the positions of the large corporations with lots of proprietary customer data. In this sense, it truly would work as oil in the machinery, making markets more efficient in the sense that more of the right goods should be produced and the right customers can be reached in a more efficient way. It also helps smaller companies to compete with bigger companies by giving them access to resources impossible to generate internally. It does not however provide a pooling of analytics resources by itself, but this could instead be provided by pure analytics services.

Three major underlying factors are shifting

Reviewing all the various structures and the corresponding perceived scores within each evaluation dimension, there seems to be three underlying reasons for the different structures.

1. First of all, the perceptions of organizations and managers clearly vary across the different structures; even though the underlying fundamentals should not vary too much. The reasons for this could be purely different conclusions of these factors given the same fundamentals. This could be due to fixed mental models about best practices and how valuable data is proprietary as well as knowledge about the value of data analysis and what the value of aggregation analysis has to provide.
2. Secondly, there have also naturally been fundamental practical challenges of data aggregation previously (and still is) in terms of compatibility, ease of data transfer and knowledge discovery. Furthermore, the tools and techniques for knowledge discovery of such vast databases have been limited in the past and are still improving. The value and potential of transaction data analysis was (and still is not fully clear), which may lead to managers having different ideas about its potential.
3. A third important point that stands out is that the collaborating structures have dealt with some of the major concerns through various mitigating mechanisms. For example, the coalition program reduces consumers' privacy concerns as they receive benefits in return, consistent with previous studies. There also seems to be some constraints within these programs as to not include direct competitors, reducing many strategic concerns.

Given how the perceptions of the four factor categories have shaped each structure, it's important to note that these perceptions are not static. The three underlying reasons - perceptions, fundamentals and mitigating mechanisms - are likely to change over time and make collaboration easier. Especially as online commerce is expected to face rapid growth in the future and could put increased pressure on the brick-and-mortar world to improve their operations and value propositions. Such pressure could force managers of brick-and-mortar to rethink their perceptions and mental models of sharing and customer data, changing the first major factor. Furthermore, as data analysis techniques become better and more important for operations and so more pervasive in management practices, the value of such sharing initiatives and knowledge discovery could become better appreciated in the future. Even sharing and data integration technologies such as web services, the semantic web, standardized formats of retail POS-data (as the NRF-ARTS⁶) may reduce the fundamental practical difficulties of collaborating. And just as Peppers and Rogers (1993) proposed in 1993, new companies may come in as intermediaries to resolve other sharing initiatives. Nectar manages for example to solve many strategic concerns by being an objective operator that gives exclusivity to each partner in their segment. It is thus possible that as the scores of these underlying factors of structures evolve over time, we might see a shift in management practices when it comes to retail customer data.

⁶ ARTS stands for Association for Retail Technology Standards and aims to introduce all kinds standards in the retail sector, from POS-systems to an XML schema.

V: Value from the customer's perspective

Previous chapters have dealt with factors and structures of aggregating customer data from businesses' perspectives. Even though consumers' privacy concerns are to some extent implicitly included in businesses' considerations, this chapter aims to explicitly consider aggregation of their behavioral data from consumers' perspective.

For consumers, there are both advantages as well as downsides to new forms of sharing of, what to many people is, personal information. The previously mentioned notorious case of how Target managed to figure out and personalize advertisements to a teen that had just recently become pregnant, before her family knew, is to many people a reminder or wake-up call to how much information and how much companies actually know about you. This intrusion or perceived violation of privacy is a factor that must be considered and then contrasted towards to the benefits generated to the customer by accessing such information.

Advantages

The idea behind collecting information about your customers is in general, as companies do today, naturally to use the data in order to provide better products and services. Thus in theory, increased information sharing practices in retail should translate into for example better advertisements that actually are appreciated by customers and reducing spam and making purchase decisions easier. Just like Amazon's or Netflix's recommendation engine, retailers too can suggest products and services that are relevant to you and give deals that you find attractive, rather than feeling spammed.

Furthermore, the increased advantage does not only come in terms of improved advertising. It should also lead to companies understanding their business better and adjusting their offering or value proposition to better meet the needs and preferences of their target customers. Thus there is a potential fundamental improvement in terms of offered products, not just better advertisements and guiding systems to help you cut through the jungle of the multitude of products and services.

By creating collaborative environments of customer data, this could possibly also translate into greater access for smaller players to valuable data. Alexander Pentland mentions this in his article and discuss how this data should not be kept in secrecy by a few and powerful companies, but rather be shared for the public good. By allowing greater transparency and accessibility of customer data, competition should be promoted, generating greater efficiency and improvements for customers.

Depending on what structure is created, there can also be increased value in terms of facilitating for customers. For example, KF's loyalty card extends over a large number of retailers, making it easier for customers as they only need one card and can collect rewards in one place, with possibly a larger set of reward offerings, similar to the airlines' bonus programs. The options and possibility to redeem these programs could also expand, generating even greater bonus (Postrel and Hlavinka). A consolidated reward program could also provide an improved and consolidated communication channel for customers and businesses. From this, improved relationship and customer engagement could follow as seems to be the case for Nectar and KF. Thus there can be significant benefits to be gained from customers point of view,

but for some of these benefits, a common interface or communication channel is required, something that might not be easily established between companies as it dilutes each individual message and is also dependent on structure.

Privacy concerns

The new massive amount of information that is being stored and analyzed about customers has created a situation that is new to humans. How comfortable are we with anyone, corporations or other people, knowing every detail of our behavior at a much larger scale than before, be it shopping behavior, eating habits or browsing history? Viktor Mayer-Schönberger, professor of Internet governance and regulation and the Oxford Internet Institute, discusses one of many issues with this, the issue of companies retaining personal data long into the future. He discusses the social importance of forgetting in society and how that changes in the new digital age where everything is stored for decades. In his book *Delete* (Mayer-Schönberger 2009), he calls for expiration dates of personal data as one way forward in order for our past not to haunt us to an unproportional extent. Regardless of the outcome of that debate or what relationship we establish with personal data in companies and government, the suggested information sharing practice will obviously increase companies' knowledge of customers and their behavior; it is in fact the whole point. Pepper and Rogers (1993) foresaw this in their book and argued that this could potentially spur entrepreneurial activity in order to create intermediaries that are reliable and that consumers trust to handle their information.

According to a study in Canada and the US about loyalty cards and privacy concerns, the majority of customers are not really aware of how businesses are using their loyalty cards in order to track and analyze your behavior. The reaction of consumers when having this explained to them was that they perceived it as very discerning and in the future, as people become more and more aware of this tracking and analyzing, it is possible that customers become more resistant to this kind of behavior and tracking tools (Pridmore 2010) (Dyché). The Payback program in Germany even received the BigBrother award within their segment in 2000, an award given by Privacy International, the oldest international privacy organization in the world, to "governments and private organizations that have excelled in violation of our privacy". Some professionals as well as in the discussions in Davos in 2012, talk about how the future needs to incorporate the customers to a greater extent in what information they have and how it is used (World Economic Forum). By creating a dialogue, customers might become less antagonistic towards corporations gathering information about you, as long as they get something back in return. Some studies mentioned in Pridmore's overview of loyalty concerns find that customers often become satisfied with relatively small compensation schemes in order to give up private information like their retail behavior. Many seem for example perfectly happy to give up this information today in order to get better prices on selected products as well as points they can redeem for cash or products. It is possible though that customers in the future perceive increased discomfort as more companies do this, become better at it and it becomes more obvious. A crucial element in one-to-one and relationship marketing is of course that customers should not perceive they're relationship as too intrusive, that would undermine the relationship. Companies should strike a balance with customers feeling comfortable that a particular advertisement just happens to be very timely for them, whether through naivety or consent will be interesting to see. This could also though be a transitional issue and in the future, awareness of targeted advertisements using your past behavior will be seen as natural or possibly even required by customers, and thus not seen as a privacy intrusion.

VI: Macro scenarios of customer data aggregation in the future

The identified structures and dimensions of information sharing structures in this paper has thus far been on a micro level, between individual organizations. However, in order to look ahead on a broader level, this chapter will consider what the macro implications of these findings are. What are the big industry scenarios we could possibly expect in the future?

From micro to macro

The different kinds of information sharing structures observed and described in earlier chapters were mainly derived from a set of three underlying factors identified in the end of chapter IV. Those factors mainly explain each microstructure at a firm- or network level. However, looking at the retail industry as a whole, it is possible to take a step back and consider something resembling the ‘average’ or ‘distribution’ of information sharing structures across the industry. This paper will refer to this macro level as the ‘landscape of retail’ and entails four different macro scenarios, given the microstructures.

4 macro scenarios

Looking ahead at what the future might hold, it’s useful to see what kind of ‘landscapes’ that the identified structures could possibly generate. If the common structure observed in the future will be similar to what we see today, with no information or limited information sharing instances, then the retail landscape from a customer data sharing perspective will look very similar to what we see today. But if we see increased collaborations between retailers, it is possible that we will observe what we can see in the airline industry today, with multiple alliances that have limited internal competition but greater inter-alliance competition. In the retail industry, this could mean multiple alliances between retailers offering different, non-competing, products and services. These could either be operated by third parties or by joint collaborations. However, if what was partly discussed during the Davos summit in 2012 about end-user focused data silos are pushed ahead by governments and regulation, individual data silos might be the future landscape of consumer data (alongside with organization’s individual retail transaction data as they still can keep their own collected data). The opposite to personal data wallets is perhaps a ‘universal’ aggregated data silo, which is the extreme scenario in the other way, with one universal silo for ‘all’ (or a significant majority of) customer data.

These scenarios can be categorized into four macro scenarios for the future:

- Current fragmented state
- Multiple alliances
- ‘Universal’ data silos
- Personal data wallets

Since each of these scenarios are influenced, but not directly determined, by the five kinds of microstructures identified earlier, a more thorough discussion of each scenario will follow. A brief and simple overview of the relation between microstructures and macrostructures is given in table 7. However, each microstructure in turn is determined by further underlying factors, mainly the three identified factors identified in chapter IV; perceptions, fundamentals and mitigating mechanisms. In order to provide a more qualified projection of the future, each scenario will discuss how the three underlying reasons must evolve in order for that retail landscape to be realized.

	Current state	Multiple alliances	Universal repository	Personal data wallet
No sharing	X			
Direct sharing	X	X		
Joint consortium	(x)	X		
3 rd party operator	(x)	X	X	
Personal data wallet				X

Table 7: Overview of which micro structures apply to each macro scenario. (x) denotes that these happen to some extent already.

Current state

The current state, with mostly isolated retailers and some instances of collaborative information sharing structures can continue given that the underlying factors do not change over time. However, with increasing competition from e-tailers as well as awareness of the value of analytics and customer data, this seems unlikely. Interviewed companies often acknowledged the need for improved analysis and data, but simply lacked the time, scale or skill to do so as well as seeing integration issues and legal or privacy issues. If these factors do not improve over time, with limiting privacy concerns and regulation and data integration issues remaining it is possible that companies do not see enough value in engaging in sharing activities and thus remain in the currently fragmented state.

It is worth noting however that for example some external parties like payment companies can still engage in deeper analysis of their transaction data and VISA have as mentioned explicitly stated that they will explore this further in the future, already moving us away to some extent from this state.

Multiple alliances

Just as the airline industry has several alliances of loyalty programs, each one designed not to compete too much internally with each other, so could the retail industry evolve. In Sweden there is already an alliance that has emerged, KF, with no directly competing partners and in New Zealand and Australia, the coalition program Flybuys exists. The coalition program Nectar in the UK similarly, even though independent from retailers as opposed to KF, restrict their partnerships in order to avoid directly competing partners.

These programs have evolved partly by limiting who can become a partner (not direct competitors), and by creating a coalition program, which seems to consumers' and regulatory concerns. This way major strategic concerns as well as privacy concerns can thus be overcome. For this to happen at a larger scale than presently however, perceptions of managers and organizations must evolve to viewing the creation of such programs beneficial. The value of data mining and data aggregation must be proven to these

managers for them to change their mental models. However, as awareness and understanding of data analytics is improving this is not unlikely to happen.

The idea of multiple alliances resembles how the airlines industry's alliance programs of multiple alliances look, with limited internal competition and more competition occurring at the inter-alliance level. The structure of each alliance could either be as a joint collaboration, like KF, or third-party aggregator, like Nectar. A third party operator for this kind is however perhaps more likely in order to fully limit strategic concerns of companies. If this happens, it also furthermore possible that payment companies like VISA continue their set course of increasing their use of transaction data, becoming their own player alongside the alliances.

Universal repository

An universal repository of customers and their personal data would naturally be a very powerful database. However, it is not clear whether businesses would agree to this as it is not only how valuable a database is that matters, but also how valuable your resources are in comparison to competitors' resources and what you share with your direct competitors. The inherent relative aspect of competitive advantage and competition might make firms less willing to engage in such an extensive level of sharing. A universal repository might not help to reduce the strategic concerns in the same way limited alliances do and would require substantial shifts in perceptions and mental models of managers. It's also questionable whether customers are comfortable with such an holistic data repository of their behavior that is out of their control, again failing to mitigate privacy concerns to the same extent as alliances and coalition programs could.

One way to overcome some of these challenges could be to create a third-party operator that can filter the redistribution of data and analysis to make sure that direct competitors do not share too much information. This will still entail difficulties of trust towards the operator. There is also an issue of where to draw the line between competitor and non-competitor and managing a huge number of different retailers with their own preferences might not be trivial. Interviewed companies expressed concerns of the possibility of non-competing firms moving into their industry (or the opposite, they moving into others') segments with the help of the shared information, gained only through the status of a non-competitor.

Another possible way of creating universal repositories to some extent is through using the choke point at the payment process. Level 1 types of repositories are already being created by credit card companies. These companies only capture items bought on their cards however, but so-called financial aggregators such as Mint in the US and Qapital or Tint in Sweden can import data from all your cards into a consolidated account. This could possibly include all transactions from different cards (except for cash purchases, unless swiped for points in credit cards loyalty programs), closely resembling what could be called a universal level 1 repository. The same issue of resolving how to decide what detailed data to reveal to whom (to get deeper insights than just high-level analysis) still remains for this solution.

Universal repositories face some of the largest obstacles in the future, but also contain the largest value. Strategic concerns are significant as well as privacy issues. Payment companies may have an opportunity, but they are to some extent also faced with greater privacy issues, as it is questionable whether consumers

agree with such usage of their data when they sign up for a payment solution, rather than a loyalty program where they know it will be used for marketing purposes and analytics.

The payment company scenario is furthermore going to be influenced by the increasingly fierce battle going by established and new payment systems. Examples include Google Wallet and Google Checkout, Merchant Customer Exchange (a cooperation between multiple retailers including Walmart and Target (Sidel 2012)), PayPal, Visa, MC, American Express and many more companies. If one company successfully can establish a majority of customers using their method of payment, they could potentially collect information through this process. As opposed to what current credit cards do, a new technology could possibly be designed to collect more information about the purchase, to level 2 granularity, allowing for much more analysis to be done. Regulation could play an important role to see if payment companies can use information in that way and whether businesses as well as consumers' are comfortable with being forced to giving up their transaction data just to be able to pay (for consumers) or get paid (for businesses). It's also possible that the payment industry will become fragmented, hindering a universal silo unless financial aggregators as mentioned above can help avoid this.

Universal repositories face some of the largest obstacles in the future, but also contain the largest value. Strategic concerns are significant as well as privacy issues. Payment companies may have an opportunity, but they are to some extent also faced with greater privacy issues as it is questionable whether consumers agree with such usage of their data when they sign up for a payment solution, rather than a loyalty program where they know it will be used for marketing purposes and analytics.

Personal data wallets

The evolution of personal data wallets as a solution in the future is not perhaps a natural way forward for existing companies. Rather, it seems to depend more on what regulators and consumers push for, as well as the aggressiveness of entrepreneurs. It is certainly possible, but given the power of incumbents, not very likely in the near future without dramatic regulatory changes. Businesses are unlikely to give away the control to the consumer who may pass it on to competitors for example. There are furthermore significant technical challenges with building such architecture, even though it is being worked on currently by for example ID3 as mentioned earlier. But it is possibly the scenario that would be preferred by consumers as they regain a sense of control of their data.

The road ahead

Given the various structures observable today and the underlying factors and dynamics that pertain to each structure, an interesting question is what we may see in the future. Some of the collaborations and examples we've discussed are, just like mining of transaction data in general, ideas and phenomena that have emerged fairly recently. New technologies and methods to capture data as well as analytical tools to extract value from that data have rendered a new kind of asset class. Perceived as valuable and proprietary, sharing has not been the first action companies undertake and understandably so. But given that this kind of customer data exhibit significant network externalities and where aggregation disproportionately could increase the value of the data, controlled and limited collaboration has started to emerge. Will this trend continue, or have we seen all that will happen in terms of collaboration?

This author argues that as the underlying factors for the microstructures evolve, so will the structures they generate and we will see a shift in the retail landscape. Firstly, data integration technologies like web services, the use of semantic data and other format standards are becoming more prominent, making collaboration easier from a technical standpoint. Furthermore, as awareness and capabilities of performing analytics and data mining of transaction and customer data in general gain traction, more data will become more desirable. Competition from massive e-tailers with rich databases of customer behavior and preferences and even moving into brick-and-mortar lockers or providing one-day delivery (Berman 2012) (Matthews 2012) can put increased pressure on the brick-and-mortar retailers, forcing them to in some cases to focus on customer experience and service rather than price. For this, a deep understanding of customers is needed and rich databases could potentially provide this.

Many of the collaborations we've seen have mechanisms that function to mitigate many concerns. Partnership networks are limited so as to not have direct competitors in the same network. This is one way that, just as the airlines industry, companies can avoid some of the larger strategic concerns they have about sharing valuable data, whilst still benefitting from increased insights about customers as well as pooling resources and marketing channels to customers. These collaborations could potentially be done both as joint collaborations like StarAlliance, but given the heterogeneous group of companies (as opposed to the airlines), a third party operator similar to Nectar might be more realistic. Just as StarAlliance face many other competing alliances in the airline industry, this author finds it feasible to see similar dynamics in the retail industry. Multiple alliances in the form of coalition programs, whether in joint collaborations or as third party operators, with each alliance spanning many different segments, but avoiding direct competition, can provide many benefits for both the consumers as well as companies without facing some of the larger strategic issues. Of course, there is always a risk of migration of companies into new segments, something that is just as valid for current collaboration networks as well as future. However, given that they tend to manage it today, it's reasonable that there will be ways to deal with this in the future, or simply worth the risk.

Given that we see these practices in many countries, with examples taken from UK, Germany, Australia, Sweden and many more countries, why haven't such a program been seen in for example in the US. A possible reason for this is because of the sheer size and diversity of the American market. These coalition programs tend to limit their partnerships as to avoid direct competitors. A national program in the US might make a single network of fixed partners a one size fits few, meaning that the optimal constellation of partners within each segment is different across regions. 'Normal size' countries may have greater homogeneity in terms of national retailers and thus are better suited for a national network. This may explain why there are some smaller coalition programs in certain regions of the US like Kick Back Rewards in Western US. Furthermore, the big retailers in the US may have much stronger perceptions about the value and proprietary of their customer data, especially since their datasets are so much more valuable than smaller retailers. The example coalition programs discussed earlier all have had significant marquee partners in terms of a major retailer championing the program. Smaller programs may have difficulty getting traction, as the value really is generated at large volumes of data. However, these factors can still change and it is possible that we may see more coalition programs in the future in the US. Attempts have been made previously. LoyaltyOne, a global loyalty consulting firm, estimated in the end of 2009 that the US would be the next large battleground for coalition reward programs.

One way to visualize the future landscape of retail could be to represent each scenario on a two dimensional space, where one axis has consolidation level of customer data repositories and the other the level of transaction data (1 or 2), illustrated in figure 11.

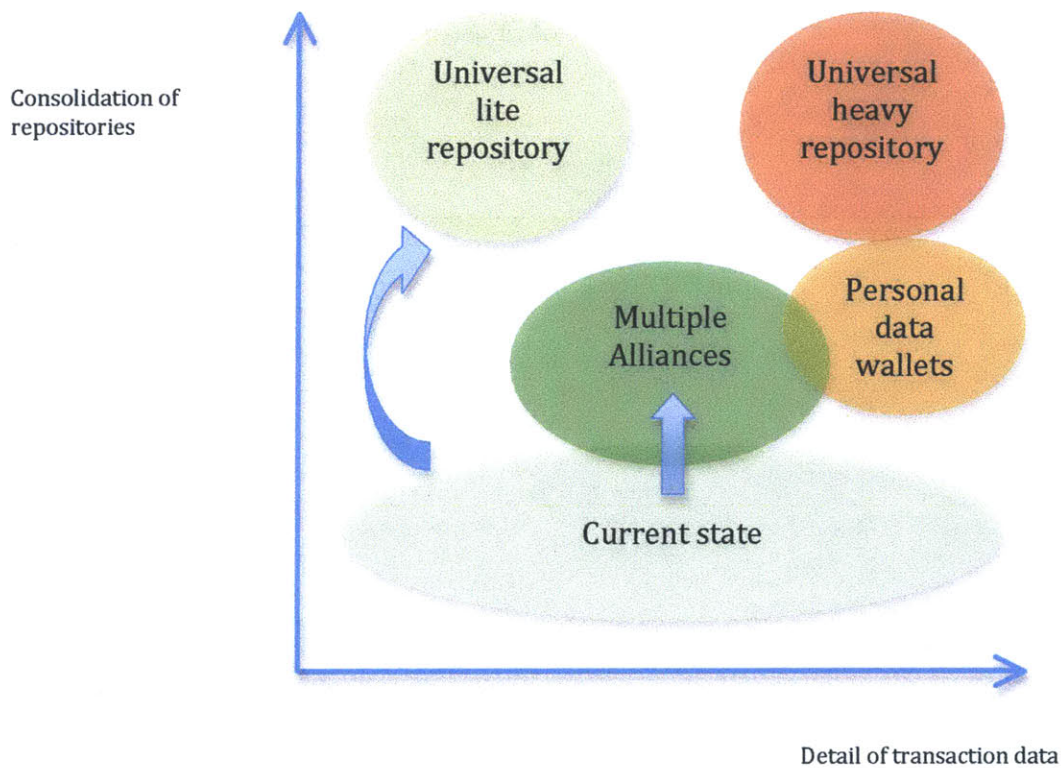


Figure 11: Two dimensional representation of macro scenarios. Green, more likely - Red, less likely

As payment companies seem to be increasing their willingness to participate in this field, as described previously, we might see two trends in this space. Current state of isolated and very limited repositories of individual retailers will move towards ‘Multiple alliances’ that provide richer analysis and better customer understanding, pooled resources, combined marketing channel that simplifies for costumers as well as increases costumer engagement. At the same time, payment companies and financial aggregators could become more active in mining their payment data, which is currently restricted to level 1 data (‘Universal’ lite repository), in order to compete with (or perhaps complement) the alliances in what insights they can provide. Payment companies however lack the active partnership that alliances may have and so may find it harder to be able to generate all the benefits outside of the pure customer data analysis. These two trends do both however seem likely in the future given the current dynamics. Personal data wallets and a universal level 2 repository are possible, but require overcoming many more strategic and regulatory difficulties, sometime given current outlooks less likely.

VII: Conclusions and opportunities

Analysis of collected data is increasingly becoming an important element of competition. Firms are coming up with improved ways to both collect and analyze data. However, the boundaries of retail firms operations limit the information available to collect and analyze, especially for offline firms. Big online players like Amazon are able to collect a huge diverse amount of information about customers, whereas offline retailers do not have the same possibility to do so. One way to improve their information database in order to improve operations, advertisements and value propositions could be however to start collaborate to build larger repositories.

By sharing, new valuable insights can be gained. Many other industries have realized the value of sharing and a few instances in retail show that this may be true for retail too. The decision to share by firms can be traced to four categories of factors:

- legal and privacy concerns
- technical concerns
- knowledge gains
- strategic concerns

The observed examples all exhibits different structures and has allowed us to identify five different structures of relationships between firms involved in sharing information. These structures are:

- No collaboration
- Direct sharing
- Joint collaboration
- Third party aggregator
- Personal data wallets

Each structure is determined by the underlying factors, where they main differentiating factors are managers perceptions, fundamental technical challenges and mitigating mechanisms. The distribution of sharing structures on a micro level across retail in general could be said to generate different macro structures, what this author calls, 'retail landscapes'. In the future, it is possible to see these micro level structures constitute four different macro scenarios of the retail business.

- Current fragmented state
- Multiple alliances
- Universal repository
- Personal data wallets

What macro scenarios that will happen in the future depend on how the underlying factors responsible for each microstructure evolve. As the importance of analytics is becoming increasingly important and technical obstacles becoming less insurmountable, the two most important factors will be perceptions and mental models of managers as well as how these collaborations structures are executed with respect to mitigating concerns. This author argues that third party aggregators and multiple alliances in the form of coalition programs can allow for increased value in terms of pooling of analytics resources as well as insights and customer understanding without compromising too much of strategic concerns or legal and

privacy concerns of consumers. Therefore it is likely that coalition programs as well as payment companies will become more prominent in the future, with multiple alliances being created. If right, this should mean significant opportunities for both incumbent retailers as well entrepreneurs to establish such practices where it is missing.

Future research

This paper aimed to provide an introductory framework for these collaboration practices of sharing customer data. However, it is incomplete in many ways, and many touched topics could by themselves be subject to their own papers and investigations. Because of the limited earlier writings on this particular topic, the idea of this paper is not to necessarily fully explore all dimensions of horizontal information sharing in retail, but rather to give an introductory starting point for future discussions. Further research is therefor needed to validate many of the findings as well as dig deeper into particularities and individual topics of this paper. For example, deeper understanding of current and future practices of some database companies like Acxiom and Datalogix could be helpful to see how they could impact this dynamic.

It could also be valuable to look closer at what the value for manufacturers could be from more comprehensive databases in order to develop both their marketing as well as proposition development. It's not evident however that other players who might want to tap into such a database should be able to. An obvious example is perhaps insurance companies in order to detect (and reject or charge higher premiums to) high-risk profiles and even governments and authorities.

More technical research will be needed as to where the actual data mining tools are today and what value they currently as well as potentially in the future actually can be able to extract from such vast databases. This paper assumes these processes and tools in order to simplify the reasoning, but the actual capabilities of current tools are not necessarily at par with what is suggested. The use of other new technologies such as RFID and video surveillance in combination with these databases could allow for instant identification of customers upon entering stores, and the implications of such activities are far vast and evolving that discussed in this paper. This could also have social implications where retailers screen out 'low value' customers and focus on the big spenders.

In general, 'Big Data' has huge implications for businesses today and this is just as valid for Big Data in retail. Hopefully this paper will have helped to some extent to think about one dimension of how to improve operations today by using information in a way that could be sustainable to both businesses as well as consumers and society as a whole.

Bibliography

- Advertising Becomes Amazon's Newest Low-Price Weapon*. Advertising Age. den 08 10 2012.
- Alexanderson, Björn, interviewed by John Tengberg. *Director Sales & Marketing, Extenda* (den 18 04 2013).
- Altair; Loyalty360. "Data Management: The DNA to Loyalty Program Success." 2012.
- American Express. <https://www209.americanexpress.com/merchant/marketing-data/pages/home>.
- Ayliffe, Peter. "More than just a way to pay: A new perspective for retail." Industry Report, VISA Europe, 2011.
- Barret, Stephanie, and Benn Konsynski. "Inter-organization Information Sharing Systems." *MIS Quarterly/Special Issue*, 1982.
- Bashir, Yusuf. "Next Generation Business Intelligence Software, Areas for Growth & Opportunities for Innovation." Cambridge, MA: Massachusetts Institute of Technology, 2011.
- Batiz-Lazo, Bernardo. "Emergence and Evolution of Proprietary ATM Networks in the UK, 1967-2000." University of Leicester, 2007.
- Berman, Barry. *How Retailers Can Unlock the Challenge of Amazon's New Delivery Lockers*. den 25 08 2012. <http://upstreamcommerce.com/blog/2012/08/25/retailers-unlock-challenge-amazons-delivery-lockers> (använd den 17 04 2013).
- Brandenburger, Adam, and Barry Nalebuff. *Co-opetition*. New York: Currency Doubleday.
- Brijs, Tom, Bart Goethals, Gilbert Swinnen, Koen Vanhoof, and Geert Wets. "A Data Mining Framework for Optimal Product Selection in Retail Supermarket Data: The Generalized PROFSET Model." Limburg University Centre.
- Bulik, Beth. *AmEx Opens Up Data Treasure Trove*. den 30 11 2009. <http://adage.com/article/news/amex-opens-data-treasure-trove/140752/>.
- Chen, Injazz, and Karen Popvich. "Understanding customer relationship management (CRM): people, process and technology." *Business Process Management Journal*, 2003.
- Chiu, Susan, and Domingo Tavella. *Data Mining and Market Intelligence for Optimal Marketing Returns*. Elsevier, 2008.
- Close, Angeline, Monika Kukar-Kinney, and Timothy Benusa. "Towards a Theory of Consumer Electronic Shopping Cart Behavior." i *Online consumer behavior: theory and research in social media, advertising and e-tail*, av Angeline Close. New York: Routledge, 2012.
- Davenport, Thomas. "Competing on Analytics." *Harvard Business Review*, 2005.
- Davenport, Thomas, Paul Barth, and Randy Bean. "How 'Big Data' is different." *MIT Sloan Management Review*, den 12 06 2012.
- Dawes, Sharon. "Interagency Information Sharing: Expected Benefits, Manageable Risks." *Journal of Policy Analysis and Management*, 1996.
- Delo, Cotton. "Facebook to partner with Acxiom." *Advertising Age*, den 22 02 2013.
- Duhigg, Charles. *New York Times*. den 16 02 2012.
http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=all&_r=1&.
- Dutta, Suomitra, and Irene Mia. *The Global Information Technology Report 2008-2009*. World Economic Forum, INSEAD, World Economic Forum, 2009.
- Dyché, Jill. *The CRM Handbook: A Business Guide to Customer Relationship Management*. USA: Addison-Wesley, 2002.
- Dyché, Jill, and Evan Levy. *Customer Data Integration: Reaching a single version of the truth*. Hoboken, New Jersey: John Wiley & Sons, 2006.
- Eckerson, Wayne. "In Search of A Single Version of the Truth: Strategies for Consolidating Analytic Silos." The Data Warehousing Institute, 2004.
- Eriksson, Ingrid, and Lisette Wallman. "Belongin to a Network - collarboation and its effects for airlines in Star Alliance(Att tillhöra ett nätverk)." Södertörns Högskola, 2011.
- FBI. *FBI Information Sharing Report 2011*. Federal Bureau of Investigation, 2011.

Fjermestad, Romano (editors). *Electronic customer relationship management*. Redigerad av Advancement in management information systems. Armonk, New York: ME Sharpe, 2006.

Forum, World Economic. *Personal Data: The Emergence of a New Asset Class*. 2011: World Economic Forum.

Fournier, Breazeale, and Fetscherin. *Consumer-brand relationships: Theory and practise*. New York, New York: Routledge, 2012.

Friedman, George, interviewed by John Tengberg. *CEO, Qapital AB* (den 10 04 2013).

Habegger, Jay. *Why Amazon Is About to Become a Force in Online Advertising. Expect Other Retailers to Follow, Leveraging Their Data for Sales and Profit*. Advertsing Age. den 11 08 2011.

Hagel, John. *The strategic value of Web Services*. McKinsey Quarterly, McKinsey & Company, 2002.

Han, Jiawei, Micheline Kamber, and Jian Pei. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, 2012.

Hill, Kasmir. *forbes.com*. den 6 2 2012. <http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/>.

Hirshorn, Susan. "Sharing Secrets: Peer groups can help businesses get ahead." *The Costco Connection*, March 2013.

Hughes, Jeff. *Amazon Customer Data To Fuel Companies New Ad Network*. den 28 06 2011. <http://www.digitaltrends.com/computing/amazon-customer-data-to-fuel-companies-new-ad-network/>.

Institute, Strategic Planning. *PIMS Online*. <http://pimsonline.com>.

James, Anthony. *Amazon begins ad network using customer buying data*. Puget Sound Business Journal. den 28 06 2011.

Kohavi, Ron, Llew Mason, Rajesh Parekh, and Zijian Zheng. "Lessons and Challenges from Mining Retail E-Commerce Data." *Machine Learning Journal*.

Kubat, Rony Daniel. "Will they buy?" Massachusetts Institute of Technology, 2012.

Kulp, Susan, Hay Lee, and Elie Ofek. "Manufacturer Benefits from Information Integration with Retail Customers." *Management Science*, April 2004.

Learmonth, Michael. *Advertising Becomes Amazon's Newest Low-Price Weapon*. den 8 10 2012. <http://adage.com/article/digital/advertising-amazon-s-newest-low-price-weapon/237630/>.

Leber, Jessica. "Amazon Woos Advertisers with What It Knows about Consumers." *MIT Technology Review*, den 21 01 2013.

Linden, Greg, Brent Smith, and Jeremy York. *Amazon.com Recommendations: Item-to-Item Collaborative Filtering*. Amazon.com, IEEE Internet Computing, 2003.

Lindgren, Klas. *Director of Sales, Ikano Bank* (den 03 04 2013).

Madden, Sean. *How Companies Like Amazon Use Big Data To Make You Love Them*. www.fastcodesign.com. den 02 05 2012.

Madnick, Stuart, and Michael Siegel. "Seizing the opportunity: Exploiting web aggregation." MIT Sloan School of Management, 2001.

Martin, Steven, interviewed by John Tengberg. *Nectar Insights Director* (den 16 04 2013).

Matthews, Cristopher. *Will Amazon Take Over the World?* den 12 06 2012. <http://business.time.com/2012/07/16/will-amazon-take-over-the-world/>.

Mayer-Schönberger, Viktor. *Delete*. Woodstock, Oxfordshire: Princeton University Press, 2009.

Mercier, Pierre, Rune Jacobsen, and Andy Veitch. *Retail 2020: Competing in a changing industry*. The Boston Consulting Group, 2012.

Miller, Amalia, and Catherine Tucker. "Health Information Exchange, System Size and Information Silos." MIT Sloan School of Management, 2011.

Nygren, Malin, interviewed by John Tengberg. *Director Strategic Partners, Coop (KF)* (den 04 04 2013).

Olausson, Martin. *Director CRM and Financial services, SIBA* (den 03 04 2013).

Oracle. *6 Data Mining Models in Oracle Retail Data Model*. http://docs.oracle.com/cd/B19306_01/bi.102/e10084/data_mining.htm.

Out-Law. *Proposed amendments to data protection reforms would be unfair to companies, say business groups*. den 10 01 2013. <http://www.out-law.com/en/articles/2013/january/proposed-amendments-to-data-protection-reforms-would-be-unfair-to-companies-say-business-groups/>.

Pentland, Alexander. "Society's Nervous System: Building Effective Government, Energy and Public Health Systems." *IEEE Computer*, 2012.

Peppers, Don, and Martha Rogers. *The one to one future*. New York: Doubleday Dell Publishing Group, 1993.

Piller, Frank, and Christian Schaller. "Individualization Based Collaborative Customer Relationship Management: Motives, Structures and Modes of Collaboration for Mass Customization and CRM." München: Technische Universität München, 2002.

Postrel, Richard, and Kelly Hlavinka. *An Open Economy: The Evolution of Loyalty in the United States*. PartnerTalk, Swift Exchange, Colloquy, LoyaltyOne, 2012.

Pridmore, Jason. "Loyalty ambivalence in the US and Canada: the GDP survey, the focus groups and the context of those wonderfully intrusive loyalty cards." i *Surveillance, Privacy and the Globalization of Personal Information*. McGill-Queen's University Press, 2010.

Ross, Dan. "Best practises for Mining Retail Transactions." *BeyeNETWORK*, 11 2005.

Shapiro, Varian. *The Information Economy*.

Sheridan, Barrett. "A trillion points of data." *Newsweek Magazine*, den 09 03 2009.

Shy, Oz. *The Economics of Network Industries*. Cambridge: Cambridge University Press, 2001.

Sidel, Robin. "Payment Network Takes on Google." *Wall Street Journal*, den 15 08 2012.

Stratten, Scott. *Un-marketing: Stop marketing. Start engaging.* . Hoboken, New Jersey: John Wiley & Sons, 2012.

Strehl, Alexander, and Joydeep Ghosh. *Value-based customer grouping from large retail data-sets*. Vol. 4057, i *Data Mining and Knowledge Discovery: Theory, Tools and Technolgoiy II*, redigerad av Belur V. Dasarathy. 2000.

Svadkooho, Farnaz Barary. "Personalized Online Promotions: Long-term Impacts on Customer Behavior." MIT Sloan School of Management, 2012.

the Guardian. *The card up their sleeve*. den 12 09 2012.

Thelen, Shawn, Sandra Mottner, and Barry Berman. "Data Mining: On the trail to marketing gold." *Business Horizons*, November-December 2004.

Thomas, Lewison, Hauser, Foley. *Direct Marketing in Action*. Westport, Connecticut: Praeger Publishers, 2007.

Thomson, Rebecca. "How Amazon changed retailing." *RetailWeek*, den 11 06 2012.

Tsiptsis, Konstantinos, and Antonius Chorianopoulos. *Data Mining Techniques in CRM: Inside Customer Segmentation*. Chippenham, Wiltshire: John Wiley & Sons, 2009.

US-CERT. *Protecting Aggregated Data*. US-CERT, 2005.

Walker, Jan, Eric Pan, Douglas Johnston, Julia Adler-Milstein, David Bates, and Blackford Middleton. "The value of Health Care Information Exchange and Interoperability." *HealthAffairs*, 2005.

Wall Street Journal. 2010. <http://online.wsj.com/public/page/what-they-know-2010.html>.

Wasserman, Todd. *Mashable*. den 12 09 2012. <http://mashable.com/2012/09/24/facebook-tracking-retail-purchases/>.

Wessman, Malin, interviewed by John Tengberg. *CRM Manager, RNB Retail* (den 10 04 2013).

Winkler, Rolfe. *Facebook Adds to Google Battle*. den 04 03 2013.

<http://online.wsj.com/article/SB10001424127887324678604578340403379945798.html>

—. *Facebook Adds to Google Battle*. den 04 03 2013.

<http://online.wsj.com/article/SB10001424127887324678604578340403379945798.html>

Yan, Yinghui. *The Online Customer: New Data Mining and Marketing Approaches*. Youngstown, New York: Cambria Press, 2006.