

Seminar Paper

# **Determining Prices in the Information Technology Age: How Can "Just Prices" Be Achieved for Good?**

by

**Mariusz Nitecki, h11711554**

Supervisor

**Univ. Prof. Mag. Dr. Rony G. Flatscher**

LV

Seminar aus BIS, 4167

Erklärung:

Ich versichere:

dass ich die Seminararbeit selbstständig verfasst, andere als die angegebenen Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfe bedient habe.

dass ich dieses Seminararbeitsthema bisher weder im In- noch Ausland (einer Beurteilerin/ einem Beurteiler) in irgendeiner Form als Prüfungsarbeit vorgelegt habe.

dass diese Arbeit mit der vom Begutachter beurteilten Arbeit übereinstimmt.

Datum: 03.06.2020

Unterschrift: Mariusz Nitecki

A handwritten signature in black ink, reading 'Nitecki Mariusz', written in a cursive style.

## Table of Contents

List of Tables.....	4
Table of Figures .....	4
1. Introduction.....	5
2. Just Prices .....	6
2.1 What Are “Just Prices” .....	6
2.1.1 The Idea of Justice in Pricing.....	6
2.1.2 “Just Prices” in Recent History .....	7
2.2 Economic Viability of Just Prices.....	8
3. Currently Practiced Pricing Schemes .....	11
3.1 Mobile Data Providers.....	11
3.1.1 Incidental Costs .....	12
3.1.2 Overage Charges .....	13
3.2 Streaming Service Provider .....	16
4. Use of Modern Technologies for Price Setting .....	19
4.1 Processing of Personal Data to Determine Fair Prices.....	19
4.1.1 Customer Profiling.....	20
4.1.2 Behavioral Data .....	21
4.1.2.1 Log Files.....	21
4.1.2.2 CGI-based Files .....	23
4.1.2.3 Cookies.....	24
4.1.3 Non-Behavioral Data .....	25
4.2 Prediction Models .....	27
4.3 Data Protection Concerns and Ethics.....	32
5. Conclusion .....	34
Literature .....	35

## List of Tables

Table 1: Pricing of providers with different bandwidth, (Tristan, 2013).....	13
Table 2: Pricing of providers for one Gigabyte of additional wireless data, (Tristan, 2013)...	14
Table 3: Example of an ECLF log file, (Wiedmann, 2002) .....	22

## Table of Figures

Figure 1: The most important data collection procedures used in customer profiling, (Wiedmann, 2002).....	26
Figure 2: K Nearest Neighbor Algorithm classification visualized (Maklin, 2019) .....	29
Figure 3: Overview over Machine Learning subcategories (Machine Learning in MATLAB, 2020) .....	31

## 1. Introduction

For centuries economists and society wonders about what prices should be demanded a product or service. Whether it is ancient Greek philosophers hundreds of years BC, political theorists like Karl Marx, or religious figures like Saint Thomas Aquinas; all of them tried to define what prices are morally and economically appropriate to be demand from a potential buyer. Many of those theories don't just advise on the costs of goods, but also instruct on how to adapt them accordingly to the demand and supply. Asking the same price for goods in different economical stages or periods is inadequate if considering dynamic price levels, inflations, and other major aspects.

With the rise of digital technology and the internet, many of those principles and advice don't apply anymore. Buying a digital product is not connected to receiving material goods or services in traditional terms. The product often is a simple copy of the original. Be it movies, music, software, or even usage of the internet itself; none of those have a material medium for which a price is charged. It's the right to (sometimes only single) use of information and data.

The question is; if no material costs arise and service costs only emerge once with the development of the original copy, what should the buyer pay for it?

This work will firstly introduce just prices. In that, a short historical overview will be given about the evolution of prices for goods and services and then the economic viability of "Just Prices" will be explained. Next, the currently widely used price-setting methods in product areas like streaming services and wireless internet providers will be analyzed.

Further, it will explore the current uses of technology for determining product prices. Those in focus will be customer profiling and the use of new technologies to predict user behavior. Additionally, ethical aspects behind pricing based on private customer information and data protection concerns will be assessed to highlight the dangers behind storing, processing, and the use of very confidential personal data.

## 2. Just Prices

### 2.1 What Are "Just Prices"

#### 2.1.1 The Idea of Justice in Pricing

The idea of a fair relationship between the amount an individual pays and the value it receives is nothing new. Yet, it is difficult to define just prices scientifically and it is certainly not regarded as a technical term in the international economic environment. Even though the subject of fair costs for a product and service is debated for over two thousand years, no general meaning for the term "Just Price" was agreed on. It was overall understood as an idea of linking justice with the exchange of goods (Schachter, 1975).

Despite that, certain general correlations have been observed if determining the setting of prices. One of them is, that industrial goods are more often subject to overpricing due to surcharges of manufacturers, monopolistic market behavior, or even cartel price-fixing agreements. On the other hand, prices of primary products, raw materials without a manufacturing process, and other resources like crops, wood, coal, iron, etc. are tied to the strict law of supply and demand. Thus, primary products, unless a rupture in the supply and demand market occurs, are rarely subject of over- or under-pricing.

A similar difference reflects on the private sector and the government, with the private sector tending to overprice goods (Schachter, 1975).

However, the approach to pinpoint a meaning changed over the decades and the debate about the need for "justice" in pricing shifted, even in the last few years. Often, it is subject to abuse to the political reality and a powerful tool to gain following in times of crisis; to gain power, political leaders pledge positive change in the pricing of goods or even change them according to their ideological belief.

### 2.1.2 “Just Prices” in Recent History

After the second world war the debate about the legitimacy of the relation, what the customer pays, and what value he receives in return, was not questioned in industrialized countries. The debate about unfair prices was in poor countries much further spread and a big issue amongst the population. That changed during the oil crises in the '70s and '80s of the 20th century, after which the price for oil and minerals soared, sparking the debate for fairness in pricing and bringing back attention to this topic in developed countries.

Today, the conversation shifted the focus on goods and services as a “human right”, often in a clash with big corporations on the other side of the spectrum. A prominent example of this is the debate, whether water, and air should be free to the general public. Many companies like “Nestle” make efforts to privatize big natural water supplies to capitalize on selling naturally occurring water (McCauley, 2016). Similar scenarios can be observed in Asian countries, where even fresh air is canned and sold to the population of heavily polluted regions (The dystopian business of bottled air , 2018).

Accordingly, many different consumer protection groups are calling those corporations out on privatizing and capitalizing on something, that should be free to the public.

A similar debate happens in the political environment; Is healthcare a human right? Should education in the form of universities have a price tag? Is it fair that people not using certain services still must pay for them in the form of taxes (Hockett, 2020)?

Questions like these are not directly associated with the topic “Just Prices”, but they are too asking what the fair price for healthcare and education should be. They focus on the justified ratio of costs to received value and contribute to the debate about justice in price setting.

Especially interesting is the debate of fair and just prices in information technology, even though the public interest often ignores this important area. With many services now provided through the internet the question for adequate costs is more important than ever, be it education, entertainment, data transfer, or the service exponentially growing in importance, communication.

Due to the nature of digital products and services the average consumer has no sense of comparison, what he should pay for something. Lending a digital book for one Euro might not seem too expansive. But when analyzing it, the person paying does not receive any materialistic value. The producer or seller made no effort or bared costs in creating the good; he merely lends out a digital copy of the original good.

Similar unfairness seems to be when signing a contract with a mobile internet provider. The customer pays for the internet in a specific country. As soon as the person however uses the service abroad, the new costs in some cases are ten times higher than normal. Another fee that surpasses the value received for it is additional data units for internet usage after the limit of the contract is exceeded. Single further units consumed by the customer sometimes exceed costs of entire monthly fees.

Knowing that the coming years will be strongly affected and shaped by information technologies, business sectors profiting from digital markets should be under special observation during the “Just Price” debate.

## 2.2 Economic Viability of Just Prices

Christian scholastics in the 17th century constructed a system of ethics rested on the virtue of justice to provided conditions for a good and harmonious life among all citizens. With just prices being one of the major components of that system, the scholars tried to figure out a value measurement of goods that is also compatible with the economic market and transaction laws at that time. Despite the ethical reflections being hundreds of years old, most of those theories are in principle applicable to this day and represent a value system that can be learned from a lot. However, even the system of the scholastics was not without flaws. Many controversies and disagreements upon some established theories arose and were not unanimously agreed on. The cost-of-production versus the subjective-utility theory for example was one of the major discords within the just price system which is to this day a discussed subject when talking about justice in pricing (Monsalve, 2012).



The scholastic tradition had two theories of value. The first is the cost-of-production theory. This value approach links the price of a good with the production and labor cost to create the product. This pricing system thus ensured the coverage of expenses for the sold goods and the proper reward for the seller, so he can support his family accordingly to the level of the profession. The reward exceeding the cost for production also is an important incentive to keep the amount of supply and suppliers in the economic market. With too small additional profits the number of sellers willing to provide goods would be too low and therefore the open market could not cover the demand, consequently, threatening the well-being of the republic (Monsalve, 2012).

The issue of too high profit margins is today especially in big technological corporations prevailing. Companies like “Apple” price their products much higher than the costs of components, composition, and labor. Due to their monopolistic or oligopolistic market position, the prices of products are sometimes twice as high as the production cost and thus don’t have an “adequate reward for the seller” as it is required in the cost-of-production theory.

The second approach for achieving just prices is the subjective utility theory of value. This theory determines the value of the goods that are exchanged by comparing the utility it will provide the buyer to fulfill human needs. Here the true source of value of a product is the human need for them and not intrinsic values (Monsalve, 2012).

Today's equivalent of this approach would be the reservation price of an individual. The reservation price on the demand side of the market presents the highest price that a buyer is willing to pay; on the supply side, it is the lower limit of the price scale at which a seller is willing to sell a good (Reservation Price, 2020). Companies try to cover as much of the reservation price of the demand side as possible with one single price. With modern technology, however, targeted price discrimination is possible. Thus, an approximation of reservation prices of every individual can be set as the price. On the other side, such targeted pricing schemes the reservation price on the supply side of the market is not always reached. Thus, there would have to be a compromise between individual customer and product provider reservation prices.

Just price is the price that should guarantee equality in a transaction, yet every transaction is different and needs certain elements to recognize unfairness. Similar to the subjective utility theory, two important aspects were also a major aspect of the just price philosophy.

“Just price is two-fold: the legal and the natural. The rightful or legal price is that set by law or by decree of the prince or magistrate after taking into consideration the quality of the object and all other buying and selling circumstances. The vulgar or natural price, on the other hand, is based upon the common estimation and judgment; not upon the particular appreciation of someone, but upon the community” (Lugo 1848 [1642] 26-38).

Many later scholars thus believed that the just price is nothing more than the competitive market price because the natural price is based on the “common estimation and judgment” of the value. This philosophy is also the core statement of Adam Smith’s concept of the invisible hand and the world economy claims to live that natural price system. But despite the principle of the competitive market implemented in stock exchanges and auctions, the current neo-capitalistic economical system seems to fail to create an environment of fairness. With stock values and commodity prices often being subject to investment gambles, prices rapidly grow despite no change in real value.

An attempt to explain why neither of this century-old ethical structures was effectively implemented in society and didn’t prevail for the next hundreds of years was made by Fabio Monsalve; “The problem arose when this ideal vision was overshadowed by commercial and financial practices” (Monsalve, 2012, p. 1). Even four centuries later these commercial practices still cannot be restricted to provide justice on the market.

### 3. Currently Practiced Pricing Schemes

#### 3.1 Mobile Data Providers

Mobile data providers in the days of the information age are one of the most important services. Not only for private clients is an internet service of utmost importance. Especially businesses rely on a fast and fully functional internet connection with the world. Particularly times of a global pandemic, like the recent COVID-19 virus, show that internet providers are an essential business that allows staff to carry out business activities from their homes, lowering the chances of a further spread of the virus.

Recent statistics even show, how the usage of internet data increased drastically. According to statistics presented in the "The New York Times", the usage of internet entertainment services like "Netflix" or "YouTube" saw an increase in traffic as big as 15 percent. Even video chat websites and programs like "Zoom", "Houseparty" or "Google Duo" rose by about 70 percent on average visitors (Koeze & Popper, 2020).

For that reason, it is very important to analyze and question the pricing techniques these companies apply. Although it might be very difficult to compare certain services with other, longer existing industries and define a fair price for providing access to a worldwide network, it is very interesting to examine the different prices some big players are currently setting.

At the beginning of the information age, many internet providers relied on a usage-based pricing model, meaning the user paid for data he actively used during a month. Fixed prices were set for every Megabyte or Gigabyte the customer consumed. But with a rapidly growing demand amongst the private clientele and the business world for bigger bandwidth and a sharp increase in services provided via the internet, the technologies also changed and got more sophisticated with time. Providing wireless data was getting cheaper and so the revenue of companies providing internet data was simultaneously shrinking. Companies were forced to change their pricing models to keep a steady turnover amongst strong competition.

The new pricing model established in the following years, that prevails today, is the packaging of mobile data. The user is charged a monthly fixed service charge and is provided with the possibility to access the internet. However, the amount of data the customer can use is capped by a data usage limit based on a fixed monthly limit of Megabyte usage across the broadband connection (including uploads and downloads). Usually, the packages are structured so that more expensive ones have a higher set usage limit (Choose an internet service provider for your business, 2020).

This pricing scheme is very appalling to customers; thus, it is the most popular price setting in the industry. The user pays a monthly fee and can use data until he reaches the cap, without having to pay for every single Megabyte he consumes. But two aspects are especially concerning when talking about fair pricing and reasonable exchange of value. The first being the incidental costs and the second being additional fees for exceeding the data limit.

### 3.1.1 Incidental Costs

The monthly fee for an internet data package is not the only cost that has to be paid when buying a new mobile internet service product. The consumer is obligated to pay many hidden fees that are often not clarified upfront. It is usual for mobile providers to have up to 33 different additional fees, some of them being activation cost for the new SIM card or yearly service charges. For these services, in particular, it can be argued that those fees are necessary for administrative reasons.

What cannot be argued for are administrative tasks that require no special costs. For, some internet providers charge customers when changing their residential address or even changing bank account information. These obvious modifications of information should not cost a user up to 15 Euro. Even a simple process of locking a sim card can cost a customer a significant amount of money (Arbeiterkammer, 2020).

In these examples, the paid price is not equal to the value the person receives. For the above-mentioned services emerge no additional costs for the company. Despite that, the customer is charged for them significant amounts of money. Often such unfair prices are not clarified before the signing of a contract and mostly overlooked by lawmakers.

### 3.1.2 Overage Charges

Depending on what data package the customer chooses to buy, a different data limit will be set, with the more expensive ones having a higher set limitation. When this limit, or also called a cap, is reached, no more data is available to the user and the internet provider stops the data transfer. In order to keep using the internet on the mobile device, a fee is charged. The person must pay for every additional Gigabyte (with some providers every Megabyte), which is called “Overage Charges”. However, the prices for the added possibility to access the internet are beyond any relation to the actual value of the service.

Bandwidth (Gb)	AT&T	Verizon	Sprint	T-mobile
.5				\$20
1				
2				
2.5				\$30
3			\$34.99	
4	\$30	\$30		
4.5				\$40
6	\$40	\$40	\$49.99	
6.5				\$50
8		\$50		
8.5				\$60
10	\$60	\$60		
10.5				\$70
12		\$70	\$79.99	Not Available
14		\$80		
15	\$90			
16		\$90		
18		\$100		
20	\$110	\$110		
30	\$185	\$185		
40	\$260	\$260		
50	\$335	\$335		

Table 1: Pricing of providers with different bandwidth, (Tristan, 2013)

To visualize how high some companies set overage fees, we are going to compare the four biggest mobile internet providers in the USA and their prices. In the table above (Table 1) the different prices for the packaged mobile data can be seen. These prices regard the monthly fees for the fixed amount of data the user can use, before being limited through the cap. For a bandwidth of 4.5 Gigabyte, the company “T-Mobile” charges 40 US-Dollars. The company “Sprint” on the other hand charges 49,99 US-Dollars for 6 Gigabyte data. (Tristan, 2013)

The next table below (Table 2) depicts how high the costs of one additional Gigabyte of bandwidth in the respective packages are. So, for a package of 6 Gigabyte bought at the company “Sprint”, the user must pay 8.33 US-Dollars for an additional Gigabyte. But what is interesting is the fact, that for a package of 2.5 Gigabyte bandwidth, the user is required to pay 14 US-Dollars if he wants to use one Gigabyte more data. Two completely different prices for the same value the customer receives. The other companies’ prices per additional gigabyte decrease with bigger basis packages, even though the amount of internet data provided stays the same. (Tristan, 2013)

Bandwidth (Gb)	AT&T	Verizon	Sprint
.5	60	\$60	\$69.98
1	\$30	\$30	\$34.99
2	\$17.50	\$17.50	\$16.67
2.5	\$12	\$12	<b>\$14</b>
3	\$10	\$10	\$11.66
4	<b>\$7.50</b>	<b>\$7.50</b>	\$21.55
4.5	\$10	\$10	\$30.53
6	<b>\$6.67</b>	<b>\$6.67</b>	<b>\$8.33</b>
6.5	\$8.46	\$8.46	\$15.57
8	\$8.75	<b>\$6.25</b>	\$19.05
8.5	\$10	\$7.65	\$22.62
10	<b>\$6</b>	<b>\$6</b>	\$25.48
10.5	\$7.14	\$7.14	<b>\$27.82</b>
12	\$7.50	<b>\$5.83</b>	\$6.67
14	\$8.57	<b>\$5.71</b>	\$13.03
15	<b>\$6</b>	\$6.33	\$15.57
16	\$6.88	<b>\$5.63</b>	\$17.80
18	\$6.11	<b>\$5.56</b>	\$21.51
20	\$5.50	<b>\$5.50</b>	\$24.48
30	<b>\$6.17</b>	<b>\$6.17</b>	\$33.39
40	<b>\$6.50</b>	<b>\$6.50</b>	\$37.84
50	<b>\$6.70</b>	<b>\$6.70</b>	\$40.51

Table 2: Pricing of providers for one Gigabyte of additional wireless data, (Tristan, 2013)

In this example, it can be seen that the same service does not have the same price. If subscribed to a package with more available data, the customer gets additional Gigabytes cheaper. However, they are none the less overpriced. The price of a 4.5 Gigabyte package is about 40 US-Dollars, so with that logic, one Gigabyte costs the user on average 8.90 US-Dollars. Yet an additional unit over the limit costs the customer approximately 30 US-Dollars. When speaking about fair prices the question arises, why one Gigabyte suddenly is 3 times as valuable than before in the eyes of the provider.

Some companies have withdrawn from overage charges amongst heavy criticism from consumer advocacy groups and some lawmakers. One example is the internet provider “T-Mobile”, who has dropped overage charges, observable in the table above (Table 2) with the lack of the firm in the comparison. Nonetheless, the majority of the telecom industry keeps the controversial pricing model, which especially in times of a global pandemic is very costly for people who are forced to work from home, while internet bills pose an existential threat to some who are out of work (Ron, 2020).

This is a very unfair, yet common pricing strategy. Packages with limits and overage charges exist for many years now. The company provides the same service because after the initial installation of powering equipment the cost of delivering one Gigabyte of wireless bandwidth should be stable. Thus, it should set the price for the first unit of internet data equal to the last one. Despite that, mobile internet providers charge a multiple of the initial price without having an equally valued effort.

### 3.2 Streaming Service Provider

When talking about fair prices and equal value exchange, one must analyze the services provided on the market that seem to give a bigger value than the price paid as well. A relatively young business model with a very appalling pricing model and a seemingly good value provided are streaming services.

Streaming became popular with “Netflix” launching a video streaming platform in the early 2010s. Providing movies online was not a new idea, in fact, “Netflix” was doing it for many years before. However, users had to rent or buy every movie separately, equivalent to a normal DVD store pricing scheme (William, 2020).

What changed in 2010 was the for that time very unusual and innovative subscription model. The model itself already existed for very long, however not in combination with online entertainment. Internet users can enter a subscription contract, in which a constant flat rate would be charged every month. In return, the paying customer had access to all movies provided on the portal. What made “Netflix’s” approach so special is that it granted access to a big database of content with monthly payment. The streaming site itself had at this time a very large offer of movies, series, and shows of different genres. A few years later “Netflix” itself became a respected film producer, exclusively providing its users with their own produced and critically acclaimed movies and shows.

With the success of “Netflix”, many other companies made use of the subscription-based payment model. Just in the last year dozens of film studios and big online conglomerates have incepted their movie streaming service competing with “Netflix”. Some of them are “Disney+”, “Prime Video”, “Apple TV+”, “HBO”, “ESPN+”, “YouTube Red” and many more (Handley, 2019).



Movie streaming platforms are however not the only online entertainment industry that adapted a subscription-based pricing scheme to their business model. With rising dissatisfaction with the high prices people had to pay online to buy songs individually in the mid-2000s, the company “Spotify” introduced a then-novel way of pricing music. 2008 they introduced, like “Netflix”, the possibility to enter a contract, that can be canceled any time, to pay monthly fees to gain unlimited access to their database of songs.

A few years later, the user base grew to over 160 million users with 35 million songs available on their platform, meaning, virtually any popular song can be accessed (How Spotify came to be worth billions, 2018).

Soon after “Spotify’s” massive success, many other streaming services emerged.

Platforms like “Apple Music”, “Amazon Music” or “Tidal” entered the music streaming industry as well to get a share of customers this new pricing scheme was attracting.

When looking at the criteria for “fair prices”, namely the exchange of equal valued goods, it should also be noted that too low costs for a product are not fair either. Looking at the prices of these subscription-based streaming services, the opposite extreme of the spectrum can be recognized; very low prices for the received value.

Subscription fees of 13,99 Euros (varying between different resolution options) every month for the ability to access thousands of movies without additional cost is considerably low, considering that a cinema ticket for only one movie is priced equally high. The same is true with the price of one month of “Spotify”, which is with 9,99 Euro even lower than the average cost of a traditional music cd.

If compared to the costs of renting or buying movies and shows individually in digital form, renting it costs around 3 to 4 Euros. That means that the customer gets access to thousands of movies for a month for the equal price of renting 4 movies (Prime Video, 2020). Similar ratios can be calculated with music streaming portals (Spotify, 2020).

Yet, given the statistics, the average “Netflix” user watches 1 hour and 11 minutes of content every day in 2018, which is almost an entire movie daily (Pesce, 2018).

Comparing the price and actual usage of the service gives an impression of a strongly under-priced product. On closer inspection of the cost structure and strategy these companies follow to continue to exist despite the low fees, the bigger picture gets clearer.

For streaming portals to make revenue, the amount they will take in subscription fees has to exceed the sums needed to buy licenses for movies and to make new content.

Leaving out the exact math of how many subscription fees have to be collected to cover the expenses for acquiring new licenses and creating new content, it is evident that the number of active customers has to be high. How high is kept very secret by the companies.

Streaming services make many years no profit at all until their business model is profitable.

The revenue might be very high, “Netflix” for example made 20,156 million US-Dollars in 2019 b (Netflix Revenue 2006-2020 | NFLX, 2020); but the expenditures often exceed those numbers, leaving the company no profit. The firms are solely liquid in cash because of the monetary investments of stakeholders. It is estimated that the newly founded movie streaming platforms like “Disney+” or “Apple TV +” will not make any profits for at least another 5 years. This period gets bigger with growing competition, as it is more difficult to convince customers of their product (Handley, 2019).

Similar problems are observable in music streaming platforms like Spotify. In their case, instead of licenses, the companies are obligated to pay royalties to the original creators of the songs.

Both sides of a trading agreement have to receive a value equivalent. For that reason, the question in “fair pricing” about the company receiving an adequate value in return for the product it provides as well, is equally important. Not every subscription-based business model turns positive in revenue. One prominent company that failed at sustaining a growing number of new customers is “MoviePass”, which allowed clients to watch movies in cinemas without limitations for a sole monthly payment of 10 US-Dollars (Guerrasio, 2019).

## 4. Use of Modern Technologies for Price Setting

### 4.1 Processing of Personal Data to Determine Fair Prices

It is well known that companies collect big amounts of data from their users. Considering some estimates indicate that around 7.5 septillion Gigabytes of data are generated every day, and about half are useful for information retrieval, it is self-evident that corporations make use of this valuable information (Companies Collect a Lot of Data, But How Much Do They Actually Use?, 2020).

Most of the data collected are used to improve the service or product, gather statistics about customers' behavior, and for marketing purposes. Back in the case of "Netflix", a large percentage of the user's viewing behavior is recorded and analyzed. From the resulting information, the platform expects to recommend movies and shows that statistically fit the users' past viewing experience and thus have a higher chance to be watched by them. For that, advanced algorithms are processing and analyzing millions of Gigabytes of data daily. New technologies like machine learning and artificial intelligence are used in order to bind the client to the service and raise his average time spent on the streaming platform (Gomez-Uribe & Hunt, 2020).

Most of the data collected are behavioral and non-behavioral information about the user itself. Often a profile based is created on the data the user gives the service provider with his consuming behavior. Besides improving the service and marketing strategies however, the collected information could be used to set prices. With the users' input, efficient, attractive, and fair prices could be achieved for the offered products.

In the following, first, the type of data that is relevant for such efforts is presented. Next, the possibilities and the basic concept of such an approach are explored.

#### 4.1.1 Customer Profiling

The general goal in pricing a product is to set it under or equal to the limit the majority of the target group is ready to pay for it. That limit is very individual and depends on multiple aspects of the person itself. Coefficients that contribute to the reservation price are culture, demographic profile, social class, income, preferences, geographical aspects, and many more. Therefore, it is of utmost importance to find how these aspects correlate with the reservation price for a product, predict it with the given coefficients, and thus set a fair price.

As mentioned earlier, fair prices are very subjective and are strongly linked to the personal values and beliefs of the spectator. One contested view on fair prices is that no fixed amount should be paid. Every customer should rather pay different prices, according to their reservation price. Of course, such price discrimination was very difficult in the past, because the only possibility to find the limit was by imprecise estimations based on sociodemographic profiles or negotiations which often resulted in suboptimal price settings for either the seller or the buyer.

That changes with the age of information, in which every relevant aspect of a user's life is recorded and analyzed. Every person leaves a data fingerprint everywhere on the internet. Cookies and other behavioral tracking technologies are quick to record and forward data to so-called data warehouses for analysis. Information like what websites were visited, how long a user stayed on a certain site, are very valuable. The analysis of the data and the creation of a customer profile happens automatically and creates no additional costs for the company.

From the gathered information about a user, the reservation price can be estimated and thus a unique price can be set (Wiedmann, 2002).

#### 4.1.2 Behavioral Data

Person related data can be divided into two groups; the information collected about the user's behavior and activities, called behavioral data, and information that relate to characteristics or features of a person that cannot be changed. The former focuses more on the recording of data resulting from the usage behavior of a user on a specific website. No personal information is used, the sole focus is what links are clicked, how long specific sites are looked at, scrolling activities, frequency of visits, and sometimes even the history of visited internet pages in the past. Additionally, the data collection can be divided further, namely in non-reactive and reactive data collection.

Reactive data collection implies that the user is aware that his activities are recorded, for example, if the person registers with an account and uses it to regularly browse the page, which could alter his actual "surf"-behavior. The second is non-reactive data collection, in which the user does not create a unique identifiable account and thus is not aware of the collection and storage of his activities (Wiedmann, 2002).

Common methods and technologies for that procedure include the collection of log files, CGI-based files, and especially for that purpose designed applications. In the following these technologies will be explored, the type of data collected explained, and attempted to use this data for creating a fair price set for the evaluated user.

##### 4.1.2.1 Log Files

If a software accesses a service via the internet, the entity is called a "client". Meanwhile, the software providing it is referred to as "provider". So, the relationship between those two applications is called the client-server principle, on which the entire World Wide Web is based on. Log files are constantly recording the connection and data exchange between client and provider. There are various log file formats, however, most of them save the same type of information (with the current format being ECLF).

These files contain valuable information about the connection and the client, which can be used by the provider. Examples of what data is saved in the log files are firstly the IP address of the user's computer, from which the client gets data from the server. Secondly, the user ID is recorded, which is used by the website to uniquely identify the client. Additionally, the time and date, when the content from the provider was accessed, is noted in the ECLF. Requests of what type of data was retrieved from the server and the number of bytes that were sent to the user's PC is manifested in the log as well. The agent tells the provider what browser the client is using in order to retrieve data from the server (Wiedmann, 2002).

One additional very valuable information for the companies analyzing log files that are recorded is the Referrer. It is the previous URL visited by the user. The reason for the high importance of this data is, that with the help of the Referrer the provider keeps better track of which sites redirect higher traffic to the own page and contain valuable information for marketing purposes. Additionally, it supports conclusions about significant relationships about the interests of relevant clients and to other websites (Grace, Maheswari, & Nagamalai, 2001).

In the table below (Table 3) it is shown how and what type of data is recorded and saved in a database by the latest log format ECLF. Important to note is that the IP address information is not always reliable. So-called "Proxy Servers" are intermediate servers, which forward the request from the user's PC to the website. If the client is connected to a "Proxy", the log files only record the exchange between the Proxy itself and the provider. Considering that multiple users can use one proxy server, the log files show only one unique IP address despite the fact, that multiple users were using the service (Grace, Maheswari, & Nagamalai, 2001).

IP Address	User id	Time	Request (Method/URL/Protocol)	Status	Size	Referrer	Agent
123.456.78.8	--	[09/May/2001:03:04:41 - 0500]	"Get Buxel.html HTTP/1.0"	200	3290	--	Mozilla/3.04 (Win95,I)
123.456.78.8	--	[09/May/2001:03:04:51 - 0500]	"Get Wiedmann.html HTTP/1.0"	200	5450	Buxel.html	Mozilla/3.04 (Win95,I)
123.456.78.8	--	[09/May/2001:03:05:32 - 0500]	"POST/cgi-bin/p1HTTP/1.0"	200	5096	Wiedmann.html	Mozilla/3.04 (Win95,I)
123.456.78.8	--	[09/May/2001:03:05:41 - 0500]	"Get Buxel.html HTTP/1.0"	200	3290	--	Mozilla (IE4.2,WinNT)
123.456.78.8	--	[09/May/2001:03:05:59 - 0500]	"Get Wiedmann.html HTTP/1.0"	200	5450	Buxel.html	Mozilla (IE4.2,WinNT)
123.456.78.8	--	[09/May/2001:03:06:30 - 0500]	"Get Frenzel.html HTTP/1.0"	200	1000	Wiedmann.html	Mozilla (IE4.2,WinNT)
123.456.78.8	--	[09/May/2001:03:07:11 - 0500]	"Get Buckler.html HTTP/1.0"	200	2020	F.html	Mozilla/3.04 (Win95,I)
123.456.78.8	--	[09/May/2001:03:07:45 - 0500]	"Get Halstrup.html HTTP/1.0"	200	3030	Frenzel.html	Mozilla (IE4.2,WinNT)
123.456.78.8	--	[09/May/2001:03:12:23 - 0500]	"Get Meissner.html HTTP/1.0"	200	4040	Wiedmann.html	Mozilla/3.04 (Win95,I)
123.456.78.2	--	[09/May/2001:05:05:11 - 0500]	"Get Buxel.html HTTP/1.0"	200	3290	--	Mozilla/3.04 (Win95,I)
123.456.78.3	--	[09/May/2001:05:06:03 - 0500]	"Get Walsh.html HTTP/1.0"	200	4040	Buxel.html	Mozilla/3.04 (Win95,I)
123.456.78.5	--	[09/May/2001:05:06:05 - 0500]	"Get robots.txt"	200	1020	--	Mozilla/3.04 (Win95,I)
233.999.79.4	--	[09/May/2001:05:06:07 - 0500]	"Get Buxel.html HTTP/1.0"	200	3290	--	Ultraseek

Table 3: Example of an ECLF log file, (Wiedmann, 2002)

The collected data from the log now can be used to recognize patterns within the data. The surf- and buying behavior of the user can be linked to information collected by the ECLF and predictions can be made (Grace, Maheswari, & Nagamalai, 2001). One example that is used very often is, that if a group of people that are accessing the server with a certain browser, say "Safari", are more likely to buy more or spent more money, that means that future clients using the same browser have a higher probability to have a higher reservation price, and thus a price can be calculated referencing data from the log.

Another possibility to read information from the log files is to look at the recorded times a client accessed the site. The analysis of usage periods and time spaces shows at which time of day the user accessed the service, how much time he spent on the site, and how frequently he visits it (Wiedmann, 2002). This information reveals interests in the relevant offers. Potential customers with higher interests have higher reservation prices and are willing to pay more. On the other hand, infrequent users, who spent on average very little time on the site have no high interest in the offers and thus can be offered lower prices to attract them for the product.

#### *4.1.2.2 CGI-based Files*

Alternatively, to request and retrieve static information and documents, the client can run a CGI script on the provider's server. The CGI protocol, or "Common Gateway Interface" protocol, is used between web forms and programs for communication and recording of user activities. Similar to the log files, the CGI-based files contain information events caused by the client. This includes spent time on the website, users' access patterns, but also further information about the client certificate and its issuer, location of the client certificate, session information's and data about the services requested by the web client. The CGI protocol supplements the log very well, which is why it is very common among web-based services (Oracle, 2020). Hence very similar to log files analysis can be carried out with the information obtained with the help of CGI files.

#### 4.1.2.3 Cookies

Besides CGI and log files, website providers have many additional options to collect further data that are not covered by the first two protocols. Special software applications like cookies, packet-sniffer, and web bugs are designed for enhanced recording of the client's behavior and information. Often these technologies are combined to supplement the other protocols as well as possible. (Wiedmann, 2002)

One of the most useful and prominent special applications are cookies. According to Peters and Sikorski (1997), the definition of cookies is: "Cookies are small data structures sent from a Web server to your browser and saved on your hard drive in a text file. They are nothing more than a string of characters (letters and numbers) that store certain pieces of information about you" (p. 1486). The cookies itself collect browser usage-related information and first saved at the hard disc of the client. Later, with sufficient data collected, the cookies are sent back to the server provider. After that, every time the user accesses the same website, the service will recognize the client with the matching cookies and can further trace his behavior. (Peng & Cisna, 2000)

The type of information cookies collect includes user settings, like the preferred language and special preferences of the client. Further, the time spent on the website and the individual sub-pages and data entered via web forms like search terms, address, and phone number is also recorded in a cookie.

Another very useful information for the provider, that is gathered by the cookie technology, are the sub-page types the user was visiting (What are Cookies?, 2018). Especially the knowledge which sub-product pages in online stores were visited is of high importance. The understanding of visited sub-pages and the time spent on them indicate the interests of the customer. Analyzing the visiting behavior not only can help the website provider to improve marketing efforts but also to adjust the prices according to the interest in the product. Price discrimination in regards to interest and reservation price could improve costs for offered goods and services.



#### 4.1.3 Non-Behavioral Data

The technologies in the previous chapter (4.1.2 Behavioral Data) can be used for the analysis of non-behavioral data, except that information regarding characteristics of a person often have to be derived from the recorded behavior. If talking about non-behavioral information, it is often closely related to reactive data collection. That means that the user is actively creating an account and providing it with personal information and thus is aware of the service recording his identity. Good examples of technologies collecting personal information are cookies (Peng & Cisna, 2000). Besides behavioral data, cookies also gather data about a person's non-behavioral information. Even information that at first doesn't seem to reveal any characteristics or properties of the user can still contain valuable clues about sociodemographic status, psychological profile, and many more traits of the person using the service.

With cookies recording data entered in web forms, it contains identification data of the user like address, age, or e-mail. This data can be used in a number of ways to calculate and set a fair price for the customer based on their characteristics. Taking the age as an example, analysis can be made on the average age of active clientele. Accordingly, if the statistics show a younger target group the prices of goods or services can be set lower due to usually limited budgets of that demographic.

Additionally, demographic clues about the user can be derived from the location data of the client's certificate which is recorded in CGI-based files or the address entered in a web form and recorded by cookies. The geographic location of the person currently using the web service can be compared with socioeconomic statistic information of the given region or country. Districts with lower average income for example are very likely to have on average a lower reservation price than areas with statistically higher income. With this information, the provider can set prices lower for poorer regions or adjust the costs for a service or product to the average income in a certain country.

Switzerland for example has a much higher average income with simultaneously more expensive prices in their country. Austria on the other hand has a much lower level of income and average prices. Thus, charging customers from both countries the same would be unfair. An adjustment of costs for goods based on the client's location would be a improvement towards fair prices.

Some services gather sociographic information about a person. The sociographic characteristics of a person include education, occupation, and income. Information of that kind can be used as well for setting fair prices. For, the average income of the customers can be statistically evaluated and incorporated into the price for goods and services to reach the estimated reservation price of the client.

It is of course once more a question of ideology and different value concepts if a wealthier person should pay more for the same product than a person with lower income. Yet the same data can be used for many other price differentiation models. At this point, the technology for elaborate pricing schemes is already availed but further contributions are a matter of companys' business policies.

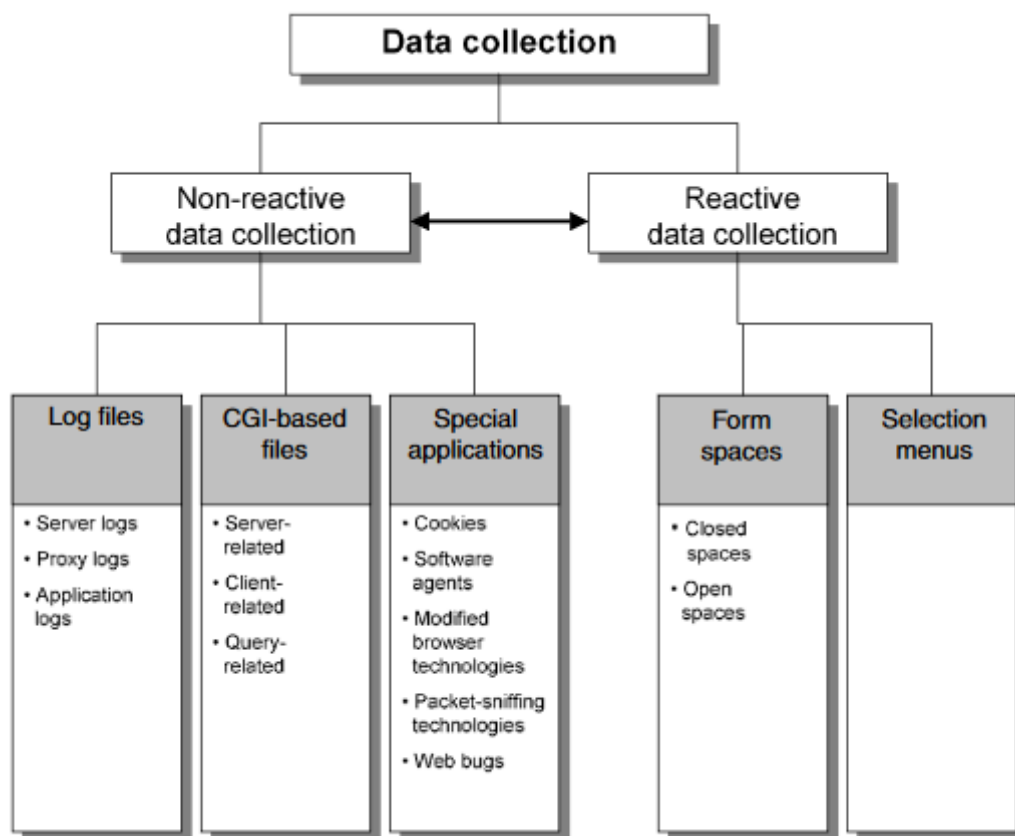


Figure 1: The most important data collection procedures used in customer profiling, (Wiedmann, 2002)

## 4.2 Prediction Models

The big amounts of data collected with different technologies from users can be used in a variety of ways to improve the service and marketing efforts. Especially new sophisticated algorithms like machine learning, which recognize patterns in data structures that are often not possible to be found by a human, open many possibilities for providers and companies. These patterns can be then used to forecast the behavior or preferences of new customers. Major areas that are optimized with prediction technologies are recommendation systems, like the “Netflix” service proposing movies and shows to watch next (Gomez-Uribe & Hunt, 2020). These algorithms however also can be used for price setting aspects to create more adequate and just pricing of products.

To analyze entire data lakes of collected information, machine learning algorithms are used. These algorithms build a mathematical model based on the sample data that was previously collected and cleaned and based on that model predictions are made about new input that is unknown to the program. The interesting part is that these machine learning algorithms get more precise and automatically improve with experience and bigger amounts of data. The most important part is thus the learning process of the algorithm. Those can be divided into three categories; supervised, unsupervised, and reinforced machine learning (Alpaydin, 2020).

Supervised learning algorithms need an initial classification of data before learning them. This means that before giving the algorithm the vast amounts of data for analysis it has to be first made clear what the individual inputs mean and how we want the model to classify similar data points. An unsupervised learner on the other side looks for patterns in the dataset without the need for pre-existing labels on data. Thus, the program does no prediction based on input parameters with a certain class as output but rather allows for modeling of probability density. Neither does reinforced learning algorithms need labeled data for learning, but rather is a mixture of unknown data with some preexisting knowledge of statistical patterns.

One of the most common and efficient supervised machine learning classifiers is the “K-Nearest Neighbors Algorithm”. It is a simple algorithm that stores all the available cases and classifies the new data based on similarity measures and features of the previous data points. To classify a data point it compares how its neighbors are classified. The amount of neighbors that are checked to classify the new point is a variable named “k”; hence the name “K-Nearest Neighbors Algorithm” (Cunningham & Delany, 2020).

In the figure below (Figure 2), it is visualized how this machine learning algorithm works. First, the data points with the according labels are loaded into the algorithm. In this case, the two classes are green and red crosses. Then, a new data point, in the figure the gray cross, is inputted into the algorithm to classify it. Assuming the k value is five, the five closest entries are referred to, to predict the class of the unknown point. To determine the closest objects, a distance measure is calculated between the unknown and all other points, and the data points with the smallest distance are taken into consideration. For the distance measurement, all relevant features from the dataset can be used. In this case, there are two features; x1 and x2. However, there is no limit for the number of attributes, because the distance measure can be calculated on multidimensional data as well. The new entry then is classified with the same class the majority of its neighbors have (Cunningham & Delany, 2020).

For distance measurement, many different approaches can be followed. The Euclidian distance is the most common and simplest mathematical measurement to use for the algorithm. It is simply the Pythagorean formula used on the difference of, in this case, the feature x1 and x2 between the neighbor and the unknown point. The result is the hypotenuse between those two data points, which simultaneously is the distance. The same methodical formula is applicable to calculate the distance between objects with more than two features as well. There are also many different distance measures methods like the Minkowski distance, Cosine similarity, and Chi-Square. Which one should be used depends on the data structures and what the purpose of the classifier is (The distance function effect on k-nearest neighbor classification for medical datasets, 2016).



Figure 2: K Nearest Neighbor Algorithm classification visualized (Maklin, 2019)

“Netflix”, “Spotify” and many other online platforms use the K-Nearest Neighbors algorithm amongst other prediction techniques. The company uses individual customers as data points. The features of these entries are information about the user. Average watching time, watching behavior, history of watched movies, location data, preferences in languages, movie ratings, most-watched genres, and much other information are defined as the attributes of the data points. With the help of advanced distance measures, the nearest neighbors amongst the multidimensional set are searched. The closest neighbors are also individuals with similar preferences and behavior to the person we want to propose new movies. Based on what content people with similar experiences watched, movies are then recommended to the analyzed user. How the user reacts to the recommended content is also saved and analyzed by the machine learning algorithm. Whether the analyzed customer watches the proposed movie and liked it or not is remembered and important to optimize the recommendation system. Based on this indirect feedback from the user the algorithm learns what methods are more effective and improves (partly) itself (Gomez-Urbe & Hunt, 2020).

The kind of features that are used for this sort of recommendation system is always different. Some suggestions for movies are based on geographic location. One example is the section on the online service with top viewed movies in the corresponding country. For this, the K-Nearest Neighbors algorithm uses the geographic location data to find similarities amongst users and recommends their often-viewed movies. A similar analysis is made with movie ratings, where customers with similar rating behavior are categorized into a class, and predictions are made about the probability of users liking certain movies, they didn't watch yet (Hoogendoorn & Funk, 2018).

K-Nearest Neighbors Regression is another tool for predicting user behavior. Rather than classify customers as the classifier does, however, regression is implemented to predict a certain missing number or value based on neighbors that contain that same value. For example, if the analyzed user didn't watch a certain movie and thus did not leave a review about it yet, the algorithm takes all k nearest neighbors that have reviewed that movie. Again, the data points with the minimal distance to the individual of interest can be based on any feature that seems appropriate for the task. Based on those reviews, a certain review is then predicted for you. If the results are a positive review, it is more likely to be recommended to the analyzed user (Hoogendoorn & Funk, 2018).

The unsupervised machine learning algorithm's major method is clustering. For a data set that is not labeled in advance but rather is raw, cleaned data, the classes are not yet determined. This step however is important to perform any classification or regression analysis. What clustering algorithms like "K-Medoids" do in principle is to find the classes, also called clusters, completely by itself, and determines which objects belong to that cluster. However, it has to be first declared how many clusters should be created before running the algorithm. This method is especially important to draw boundaries between the data points in order to later accurately predict behaviors (Hoogendoorn & Funk, 2018).

Besides the k-Nearest Neighbors and k-Medoids, there is a wide variety of algorithms that serve similar or even more advanced problems, as seen in the overview graphic below (Figure 3). Machine learning models like Hierarchical Clustering, Support Vector Machines, Linear Regression, and Naïve Bayes are also very powerful tools when analyzing big amounts of data. Each method has some strengths and some weaknesses, which is why the choice of a prediction algorithm strongly depends on the purpose and previously set goals. Companies have special data science departments in order to work with collected data to extract as much important knowledge about the customers as possible.

With the same approach to predict user behavior and suggest content or goods, these data processing tools can be used to set fair and just prices for products and services. As in the previous chapter elaborated (Processing of Personal Data to Determine Fair Prices, p. 19-25), certain information can be used for individual pricing to create an equal exchange of goods between provider and customer. The technological tools at the disposal of companies are more powerful than before and thus completely new opportunities arise to challenge the traditional form of pricing goods and services.

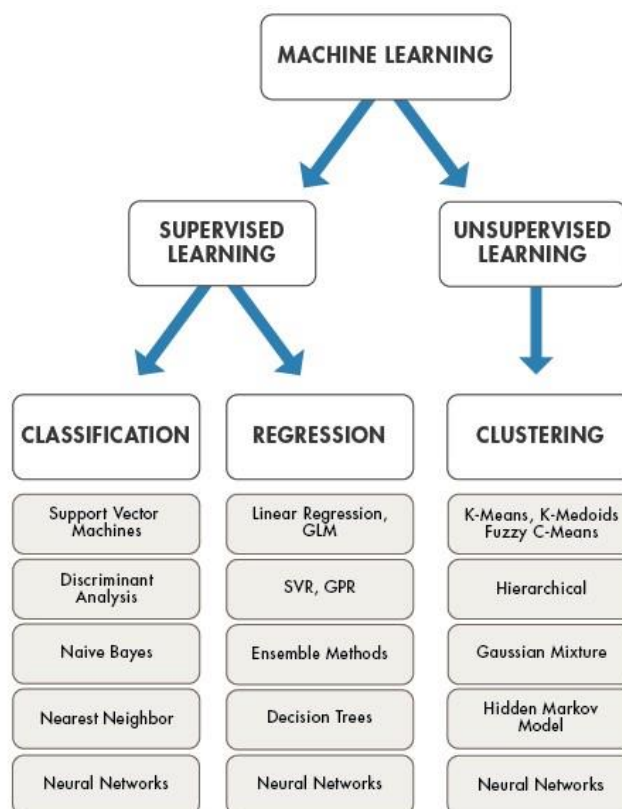


Figure 3: Overview over Machine Learning subcategories (Machine Learning in MATLAB, 2020)

#### 4.3 Data Protection Concerns and Ethics

The collection of large quantities of sensitive data raises the question of the legitimacy of such procedures. Most of the information like the behavior on a website or the previously visited subsites indicates many characteristics and interests of users. Thus, many data protection organizations claim that this information should not be held by website providers which often are a big corporation with the sole goal of profit maximization. The collection of private information about a person is according to many privacy activists an invasion of privacy and a violation of fundamental human rights like the right to privacy (The EU General Data Protection Regulation, 2018).

Especially when considering that data is often collected without the knowledge of the user. Despite attempts to force companies and website providers to request the page visitor for permission before recording behavior, like the General Data Protection Regulation (EU) 2016/679 (GDPR) tried, often social media platforms like “Facebook” are very non-transparent about their data policies and even break protection laws.

Further, some argue that collection of identifiable data is an infringement on the human freedom (of action), basing it on the reasoning that in countries with troubled socio-political situations authoritarian governments can use the collected information to identify individuals and infringe their freedom based on religion, politics and other reasons. China is a good example of a regime silencing political enemies and cracking down on countermovements with the help of data collection via the internet. The countries government goes even so far as to monitor the behavior of nearly all citizens to encourage pro-communist behavior and punish individuals that cry out against authoritarian practices. Currently, an artificial-intelligence-powered policing platform analyses the data of millions of Chinese citizens. Thousands of Muslim Uighurs and other minorities were sent into detention camps solely based on the computer-generated findings resulting from the monitoring of behavior online, indicating that religious persecutions take place in the communist country (Allen-Ebrahimian, 2019).



The storage and protection of collected data is a major concern as well. Despite the importance and sensitivity of personal information collected about the user, many companies and website providers don't securely store them. Especially with smaller companies with lower cybersecurity budgets, there is a high risk of theft, alteration, or deletion of personal data. Since reactive data collection often includes the storage of very important information like home addresses, credit card numbers, and social security numbers, the loss of such data is very dangerous. In the wrong hands, it can cause individuals loss of money or various other attacks including blackmailing and extortion. Even with appropriate security measures, data leaks and attacks on databases are not completely preventable, which fuels the argument that companies should not own information about users in the first place. A prominent example of massive data leaks is the case of the consumer credit reporting agency "Equifax". Millions of social security numbers were stolen; one of the most valuable information to a person in the United States (Gervais, 2020). With the number, an individual gains access to a wide range of critical information and services, such as court records, medical services, requests for new credit, and government benefits.

Since user's behavior is used to create tailored advertisements and products for specific users, many raise concerns about the danger of influencing people by these forms of technology. Information about personal interests and values can be used to convey a certain message or influence a person. The "Cambridge Analytica" scandal shows, how it already is common practice for political parties to harvest over 50 million "Facebook" profile's information in order to influence people. A powerful software used personal information to build a system that could profile individual US voters and target them with personalized political advertisements. The data used for the analysis was recorded without authorization. Supposedly the very same system helped to win voters in the Brexit campaign and was a major factor for the leave of Great Britain from the European Union (Cadwalladr & Graham-Harrison, 2018).

## 5. Conclusion

Despite many attempts over the history to define rules on how to set fair prices, no definition of just prices was unanimously agreed on. Some theories defined by ancient Christian scholars like the “Natural Price” are in some form applied in today's economic system, but seem despite that not to grant equal exchange of value.

The question for fair prices gets even more important in the age of information technology. Major parts of people's lives are now dependent on digital goods and services, like entertainment, productivity, or communication. With the rise of such technologies the question, what these products should be charged for, is more important than ever. Many companies either set too high or too low prices. Mobile providers for example seem to approach the need for internet connection opportunistic and set prices beyond any equal value exchange. On the other hand, many entertainment platforms have too ambitious pricing models often generating big losses and thus preventing new companies from entering the market.

With the possibility of collecting big amounts of data from customers, companies can use modern technologies to set fair prices for customers. Millions of Terabytes of collected data stored on big data warehouses contain valuable information about the customer base. Advanced machine learning algorithms can predict certain behavior or characteristics of a user. The resulting estimations are often used for service optimization, product recommendations, and marketing efforts, but also can be used to set fair and just prices.

The technology for elaborate pricing schemes is already availed but despite that, the future of fair prices for goods and services is dependent on the company's business policies. It is a matter of agreeing on rules of an equal exchange of values and determining what implications such pricing of goods and services involve.

## Literature

- Allen-Ebrahimian, B. (2019, November 24). *Exposed: China's Operating Manuals for Mass Internment and Arrest by Algorithm*. Retrieved from International Consortium of Investigative Journalists: <https://www.icij.org/investigations/china-cables/exposed-chinas-operating-manuals-for-mass-internment-and-arrest-by-algorithm/>
- Alpaydin, E. (2020). Introduction to machine learning. *MIT press*.
- Arbeiterkammer. (2020, May 5). Retrieved from AK Test: „Bunte Welt“ der Nebenspesen bei Handyanbietern! : [https://www.arbeiterkammer.at/beratung/konsument/HandyundInternet/Handy/Nebenkosten\\_Mobilfunk.html](https://www.arbeiterkammer.at/beratung/konsument/HandyundInternet/Handy/Nebenkosten_Mobilfunk.html)
- Bhattacharjee, J. (2017, November 10). *Popular Machine Learning Algorithms*. Retrieved from Medium: <https://medium.com/technology-nineleaps/popular-machine-learning-algorithms-a574e3835ebb>
- Cadwalladr, C., & Graham-Harrison, E. (2018, March 17). *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach* . Retrieved from The Guardian: <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
- Choose an internet service provider for your business. (2020, April 25). Retrieved from NIBusiness Info: <https://www.nibusinessinfo.co.uk/content/how-isp-pricing-models-work>
- Companies Collect a Lot of Data, But How Much Do They Actually Use? (2020, May 8). Retrieved from Priceonomics: <https://priceonomics.com/companies-collect-a-lot-of-data-but-how-much-do/>
- Cunningham, P., & Delany, S. J. (2020). k-Nearest Neighbour Classifiers: 2nd Edition (with Python examples). *arXiv preprint arXiv:2004.04523*.
- Gervais, J. (2020, May 15). *Equifax Data Breach: Why You Should Care If a Hacker Has Your Social Security Number*. Retrieved from LifeLock: <https://www.lifelock.com/learn-data-breaches-equifax-data-breach-why-you-should-care-if-a-hacker-has-your-social-security-number.html>
- Gomez-Uribe, C. A., & Hunt, N. (2020). The Netflix Recommender System. *ACM Transactions on Management Information Systems, Volume 6, Issue 4*, 1-19.

- Grace, L. K., Maheswari, V., & Nagamalai, D. (2001). Analysis of web logs and web user in web mining. *arXiv preprint arXiv:1101.5668*.
- Guerrasio, J. (2019, August 6). *How MoviePass went from a Hollywood disrupter to flat broke in 18 months*. Retrieved from Business Insider:  
<https://www.businessinsider.com/investigation-into-the-rise-and-fall-of-moviepass-2019-8?r=DE&IR=T>
- Handley, L. (2019, December 9). *Streaming services like Disney+ aren't likely to make money 'anytime soon,' analyst says*. Retrieved from CNBC:  
<https://www.cnbc.com/2019/12/09/streaming-services-arent-likely-to-make-money-anytime-soon.html>
- Hockett, R. (2020, Mai 26). *Whatever Happend To "Just Prices"?* Retrieved from Law And Political Economy: <https://lpeblog.org/2018/01/10/whatever-happened-to-just-prices/>
- Hoogendoorn, M., & Funk, B. (2018). In *Machine Learning for the Quantified Self*. Springer.
- How Spotify came to be worth billions*. (2018, March 1). Retrieved from BBC:  
<https://www.bbc.com/news/newsbeat-43240886>
- Koeze, E., & Popper, N. (2020, April 7). *The Virus Changed the Way We Internet*. Retrieved from The New York Times:  
<https://www.nytimes.com/interactive/2020/04/07/technology/coronavirus-internet-use.html>
- Lugo, J. (1848 [1642]). *De Iustitia et Iure*. Lyon: Vive's, L. (Reprinted from: 1868), pp. 26-38.
- Machine Learning in MATLAB*. (2020, May 20). Retrieved from Matlab:  
<https://in.mathworks.com/help/stats/machine-learning-in-matlab.html?w.mathworks.com>
- Maklin, C. (2019, July 22). *K Nearest Neighbor Algorithm In Python*. Retrieved from Toward Data Science: <https://towardsdatascience.com/k-nearest-neighbor-python-2fccc47d2a55>
- McCauley, L. (2016, November 2). *Nestlé Plans Dramatic Expansion of Water Privatization in Michigan, Just 120 Miles From Flint*. Retrieved from EcoWatch:  
<https://www.ecowatch.com/nestle-bottled-water-flint-2075968508.html>

Monsalve, F. (2012, January 21). Scholastic just price versus current market price: Is it merely a matter of labelling? *European Journal of The History of Economic Thought - EUR J HIST ECON THOUGHT* Vol. 21, pp. 1-17.

*Netflix Revenue 2006-2020 | NFLX*. (2020, May 10). Retrieved from Macrotrends: <https://www.macrotrends.net/stocks/charts/NFLX/netflix/revenue>

Oracle. (2020, May 13). *Using CGI*. Retrieved from Oracle: <https://docs.oracle.com/cd/E19857-01/817-6250/pgcgi.html>

Peng, W., & Cisna, J. (2000). HTTP cookies—a promising technology. *Online Information Review*.

Pesce, N. L. (2018, September 13). *We now spend more time on Netflix than we do bonding with our kids*. Retrieved from MarketWatch: <https://www.marketwatch.com/story/we-now-spend-more-time-on-netflix-than-we-do-bonding-with-our-kids-2018-09-13-12882032>

Peters, R., & Sikorski, R. (1997). Cookie monster? *Science*, Vol. 278, pp. 1486-1487.

*Prime Video*. (2020, May 10). Retrieved from Amazon: <https://www.amazon.de/Amazon-Video/b?ie=UTF8&node=3010075031>

*Reservation Price*. (2020, May 20). Retrieved from Negotiations: <https://www.negotiations.com/definition/reservation-price/>

Ron, G. (2020, April 27). *Northerners 'unfairly dinged' with data overuse fees – Angus*. Retrieved from The Daily Press: <https://www.timminspress.com/news/local-news/northerners-unfairly-dinged-with-data-overuse-fees-angus>

Schachter, O. (1975). Just Prices in World Markets: Proposals De Lege Ferenda. *American Journal of International Law*, pp. 101-109.

*Spotify*. (2020, May 10). Retrieved from Spotify: <https://www.spotify.com/at/premium/>

*The distance function effect on k-nearest neighbor classification for medical datasets*. (2016, August 9). Retrieved from National Center for Biotechnology Information: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4978658/>

*The dystopian business of bottled air*. (2018, October 27). Retrieved from The Hustle: <https://thehustle.co/the-dystopian-business-of-bottled-air/>

*The EU General Data Protection Regulation*. (2018, June 6). Retrieved from Human Rights Watch: <https://www.hrw.org/news/2018/06/06/eu-general-data-protection-regulation>

Tristan, L. (2013, September 22). *The Real Price of Wireless Data*. Retrieved from Forbes:  
<https://www.forbes.com/sites/tristanlouis/2013/09/22/the-real-price-of-wireless-data/#1b8485c7749f>

*What are Cookies?* (2018, September 20). Retrieved from IONOS:  
<https://www.ionos.com/digitalguide/hosting/technical-matters/what-are-cookies/>

Wiedmann, K. B. (2002). Customer profiling in e-commerce: Methodological aspects and challenges. *J Database Mark Cust Strategy Manag* 9, 170-184.

William, L. H. (2020, March 18). *Netflix*. Retrieved from Britannica:  
<https://www.britannica.com/topic/Netflix-Inc>