ZADÁNÍ DIPLOMOVÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Hanzl** Jméno: **Petr** Osobní číslo: **420866**

Fakulta/ústav: **Fakulta informačních technologií**

Zadávající katedra/ústav:

Studijní program: **Informatika**

Studijní obor: **Znalostní inženýrství**

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

Detekce temných vzorů v českých internetových obchodech

Název diplomové práce anglicky:

Detection of Dark Patterns on Czech Webshops

Pokyny pro vypracování:

The goal of the thesis is to analyze content on selected Czech webshops in order to detect so called dark patterns.
1. Analyze and describe existing methods for dark patterns detection in the Czech Web environment as well as in the world.
2. Design a crawler to retrieve Czech Webshops content and identify relevant product pages.
3. Implement the crawler and a method for dark patterns detection on selected Webshops.
4. Evaluate and describe results of your method.

Seznam doporučené literatury:

Jméno a pracoviště vedoucí(ho) diplomové práce:

doc. Ing. Tomáš Vitvar, Ph.D., katedra softwarového inženýrství FIT

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) diplomové práce:

Datum zadání diplomové práce: **16.02.2021** Termín odevzdání diplomové práce: _____

Platnost zadání diplomové práce: _____

doc. Ing. Tomáš Vitvar, Ph.D. podpis vedoucí(ho) práce	podpis vedoucí(ho) ústavu/katedry	doc. RNDr. Ing. Marcel Jiřina, Ph.D. podpis děkana(ky)

III. PŘEVZETÍ ZADÁNÍ

Diplomant bere na vědomí, že je povinen vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v diplomové práci.

. _____ _____

Datum převzetí zadání Podpis studenta

Czech Technical University in Prague

Faculty of Information Technology

Department of Software Engineering

Master's thesis

Detection of Dark Patterns on Czech Webshops

Bc. Petr Hanzl

Supervisor: doc. Ing. Tomáš Vitvar, Ph.D.

16th of May, 2019

Acknowledgements

I wish to express my sicere thanks to my supervisor doc. Ing. Tomáš Vitvar, Ph.D. for the continuos encouragement.

I also thank my whole family, especially my parents for the support and attention.

Declaration

I hereby declare that the presented thesis is my own work and that I have cited all sources of information in accordance with the Guideline for adhering to ethical principles when elaborating an academic final thesis.

I acknowledge that my thesis is subject to the rights and obligations stipulated by the Act No. 121/2000 Coll., the Copyright Act, as amended. In accordance with Article 46(6) of the Act, I hereby grant a nonexclusive authorization (license) to utilize this thesis, including any and all computer programs incorporated therein or attached thereto and all corresponding documentation (hereinafter collectively referred to as the "Work"), to any and all persons that wish to utilize the Work. Such persons are entitled to use the Work in any way (including for-profit purposes) that does not detract from its value. This authorization is not limited in terms of time, location and quantity.

In Prague on 16th of May, 2019 .....

Czech Technical University in Prague

Faculty of Information Technology

Citation of this thesis

HANZL, Petr. *Detection of Dark Patterns on Czech Webshops.* Master's thesis. Czech Technical University in Prague, Faculty of Information Technology, 2021. Available also from WWW: ⟨`https://github.com/Lznah/DarkPatterns`⟩.

Abstrakt

Klíčová slova Temné vzory, Výpočty zaměřené na člověka, Strojové učení, Hierarchické hnízdění, Webové obchody

Abstract

Keywords Dark patterns, Human Centered Computing, Machine Learning, Hierarchical clustering, Webshops

Contents

List of Tables

List of Figures

Introduction

Dark patterns[7, 11, 14, 12] are ways of designing a user interface of websites, apps or any other computer system in a specific way to trick, confuse or coerce a user in doing unwanted actions like confirming to share more information than is needed to use the service, signing up for things that the user did not mean to, buying unwanted products and more.

Typically, when the user reads a website or uses an app, he does not read all the words and makes quick assumptions[7]. Dark patterns then trick the user by hiding information of unpleasant truth. The user also trusts in his experience that he has gained from using other websites or apps and expects specific actions to happen or not to happen by using a similar pattern in the user interface. The user is tricked here by excepting this user interface behaviour, but in reality, it does something more or less than what the user expects[14]. Dark patterns are not only able to take advantage of the user not paying enough attention. Another dark pattern uses psychological methods to make users feel bad and guilty for not doing what the dark pattern wants them to do[14].

Research into tricky user interface designs and deceptive practices has surprisingly much history, but it was neglected for many years. In 1999, Hanson and Kysar were the first who examined how companies abuse customers' cognitive limitations and profit from them. The rapid growth of the Internet and e-commerce increased more serious discussions and analyses of this topic. The term Dark Pattern itself was introduced by user interface expert Harry Brignull

in 2010 to create a library of different types of dark patterns and to shame websites using them[8].

In March 2021, the state of California added new regulation that now bans dark patterns that prevent users from opting out of the sale of their personal data[3]. Therefore, the topic of dark patterns becomes more and more relevant.

In 2019, a group of scientist from Princeton University introduced an automated approach that enables experts to identify dark patterns used on websites at scale[12].

This thesis's primary goal is to build on top of their research to analyse the prevalence of dark patterns on Czech webshops, also described in the Princeton study[12]. This thesis focuses on product pages and product purchase flow only because these are the most promising pages, where all the buying happens. Several subgoals need to be done to fulfil the primary goal:

- Build an automated mechanism of gathering data from numerous Czech webshops at scale.
- These extracted data needs to be analysed in order to train a model that can detect dark patterns.
- Evaluate and describe results.

The thesis does not aim to study the prevalence of dynamic dark patterns that display transients values over time.

Chapter 1

State of the art

Most studies[7, 11, 5] in the field of dark patterns have only described known existing types of dark patterns. Also, literature often proposes different dark pattern taxonomies. To find these patterns, scholars did manual research, analysing page by page.

In contrast to this approach, which requires much manual work, there is a study from Princeton University[12], and it proposes an entirely new taxonomy. Not only the researchers recategorised and made more accurate the currently known types from the literature, but they were able to find new types of dark patterns; thus, they extended the literature about these new types.

Princeton researchers also note that only textual information on webshops was analysed. Continues that the set of the found dark patterns is restricted in this manner[12].

In an attempt to find these new types, researchers focused on product pages of webshops, because as they say, these pages are the most promising to contain dark patterns at any level of purchase flow[12]. Princeton Researchers did much work to find these dark patterns. They separated it into three steps, as can be seen in figure 1.1.

Corpus Creation is the first step; there are several scripts to get domain names of webshops. They gathered websites with the highest Alexa Rank via Alexa Rank API. Then, they used paid service Webshrinker to filter out

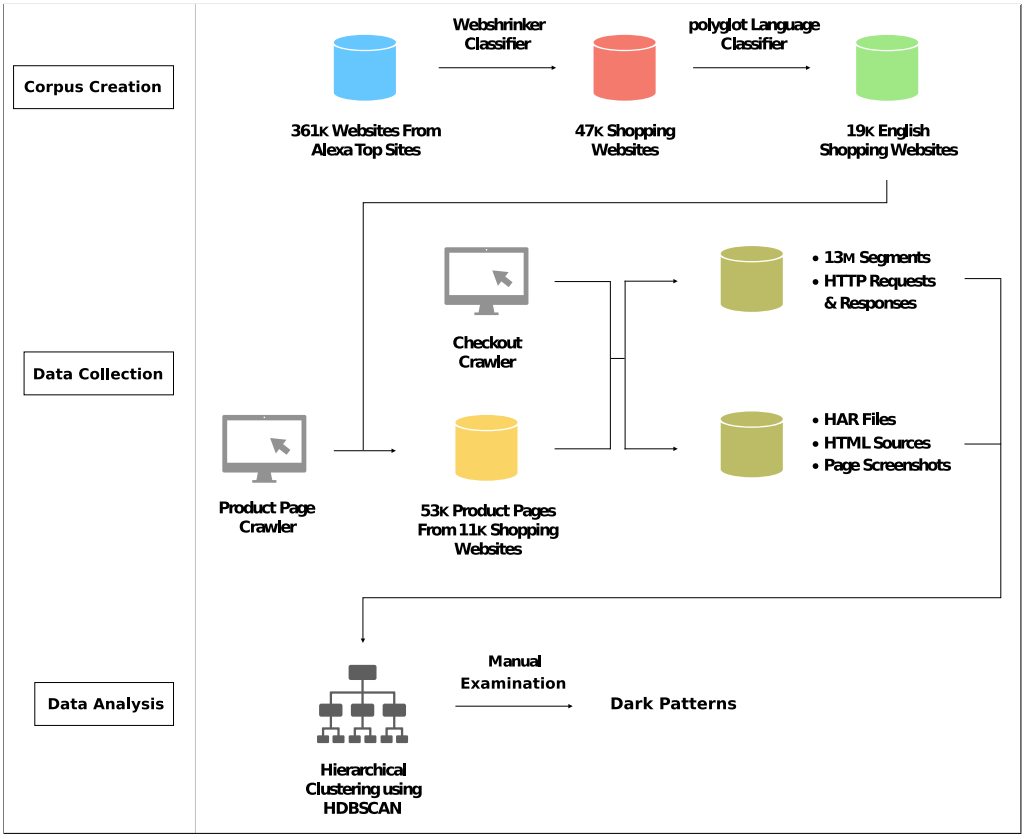

Figure 1.1: Overview of the shopping website corpus creation, data collection using crawling, and data analysis, as proposed by Princeton University researchers.[12].

only those websites that are webshops. The list of domain still contained non-English websites. They used a language classifier library Polyglot to filter them out of the list. Overall, researchers gathered a list of 19K English shopping websites[12].

Data Collection is the second step. It consists of two crawlers created by Princeton researchers. The first crawler is meant to find product links on a single website. To speed up the process of finding these product pages, they trained a classifier of Logistic Regression on a dataset of 1000 URL links manually labelled by the researchers. The first crawler found 53K pages in 11K domain names.

The second crawler, also referred to as checkout crawler, is meant to simulate users' shopping flow. This ability to simulate users' flow means that the crawler

follows the buying process steps, including selecting product options (e.g., size or colour), adding the product to the card, viewing the cart, and checking out. To evaluate whether or not this crawler can simulate users' shopping flow, the researchers randomly sampled 100 product pages and examined whether the crawler successfully reached the checkout page.

This crawler is built on OpenWPN, which is a web privacy measurement framework for privacy studies on a large set of websites. Princeton researchers implemented additional features to this framework. For example, they created a feature to store HAR files, which contain all the HTTP communication and Javascript calls. All these collected data are further utilised in an analysis phase by researchers. These data help researched recognise whether or not a found pattern is one of the types of dark patterns.

The checkout crawler also divides visited pages into meaningful textual segments. Researchers define this textual segment and an algorithm to split the page's HTML code into these segments[12]. Also, the checkout crawler extracts data about text and background colours, positions and dimensions of the segments and others. With this algorithm, there were able to capture approximately 13 million segments across the previously noted 53K product URL pages.

Data Analysis is the last step of the research. It consists of data preprocessing, hierarchical clustering, examining and analysing the found clusters. The data cleansing phase reduced 90% of all segments to 1.3 million segments.

Data were transformed into a representation of Bag of Words (BoW)[15]. Then, Principal Component Analysis was performed on the BoW matrix. The outcome was three components, which together represented 95% of the variance in the data.

Researchers chose an algorithm called Hierarchical Density-Based Spatial Clustering of Application with Noise (HDB-SCAN)[1] to find clusters in data. They tried different hyperparameters of this clustering algorithm and picked the most promising results.

Then, they did two passes examining the clusters. In the first pass, they manually tagged clusters that can manifest as dark patterns. This pass reduced

the number of the clusters from 10,277 to 1,768. During the second examination, researchers manually examined which of these 1,768 clusters contain dark patterns[12].

Lastly, the researchers discussed the results, and they iteratively grouped the discovered dark patterns into categories. They revealed 15 types of dark patterns in 7 categories on 1,254 websites, representing 11,1% out of 10,277[12].

Chapter 2

Dark Patterns

The 'Dark Pattern' is a relatively new term. This neologism was firstly used by Harry Brignull in 2010[10] when he registered a domain darkpatterns.org. In this domain, Brignull created an online library to share user interface patterns with deceptive characteristics that intentionally confuse and enrol users in unwanted situations. Another purpose of this online library is to shame websites that use dark patterns.

2.1 Definition

Brignull described dark patterns like so: 'Dark Patterns are tricks used in websites and apps that make you do things that you did not mean to, like buying or signing up for something.'[7] Brignull's definition is simplified to understand what dark patterns with ease. However, it does not include all the dark patterns that Brignull describes. For example, there is a dark pattern that purposely focuses users attention on doing one action and distracts their attention from alternatives. Brignull's definition does not imply this example.

A more accurate definition is the one used in the study made by Princeton researchers. They suggest this definition: 'Dark patterns are user interface design choices that benefit an online service by coercing, steering, or deceiving users into making decisions that, if fully informed and capable of selecting alternatives, they might not make.' [12]

2.2 Taxonomy

Brignull also defined the first types of dark patterns. This list of types is continuously updated when a new type of dark pattern is found. In April 2021, there were twelve different types of dark patterns defined[9].

The researchers from Princeton University have redefined this list considering the results of their study. This list consists of fifteen types of dark patterns and seven broad categories. Their work also differs from the prior work[7, 2, 5] by the new proposed taxonomy. This new taxonomy focuses on the characteristics of dark patterns and cognitive biases that they exploit in users. They used their taxonomy to classify and describe discovered dark patterns.

This thesis uses the same taxonomy defined by Princeton researchers. This taxonomy consists of five dimensions:

Asymmetric
: The user interface presents more alternatives to a user. It is an asymmetric characteristic of a dark pattern if the user interface requires less effort to continue with the alternative that might be disadvantageous for users. A typical example is buttons for accepting and rejecting cookies on websites. Usually, the rejecting button is less noticeable. Also, if users want to reject saving cookies, the user interface forces them to read much more text and click many buttons for every single cookie.

Covert
: The user interface shows evidence of covert characteristics if users may fail to recognise the intended outcome of a specific action. Users have experience with other user interfaces, and they may predict a similar outcome from the interface that shows similar traits as a decoy to influence their decision-making process. For instance, most of the websites offer a subscription to a newsletter in the process of registration. Usually, this subscription to the newsletter is done by ticking a checkbox in the registration form. When users start to read a sentence mentioning the subscription, they automatically expect that not ticking the checkbox means not subscribing to the newsletter.

Deceptive
: The user interface induces false beliefs in users by presenting them

misleading information. For instance, a website may offer a discount for a limited period of time, but in reality, the discount is permanent. Another example is a website that shows how many users are watching the given product and how many products are in stock. This information can take advantage of the deal by steering users into making quick decisions or inducing false beliefs of the product's exclusivity.

Hides Information
: The user interface intentionally delay presenting necessary information in places or in time, where or when users do not expect them to be presented. For instance, a website may present extra fees for a bought product at the very last step of the checkout.

Restrictive
: The user interface restricts the set of choices available to users and takes advantage of it. For example, a website may require to sign up only with Facebook to collect additional personal information.

In addition to these dimensions, Princeton researchers define six different effects on users through exploiting different cognitive biases by specific dark patterns:

- **Anchoring Effect**: The tendency of users to over-rely on the first piece of information in the future decision-making process.
- **Bandwagon Effect**: The tendency of users to value more or believe in something simply because others do.
- **Default Effect**: The tendency of users to stick with default options.
- **Framing Effect**: The tendency of users to choose different options with knowledge of the same information, but with a different way of presenting the options.
- **Scarcity Bias**: The tendency of users to value more things that are more sparse.
- **Sunk Cost Fallacy**: The tendency of users to continue an action because they already invested time or other resources in it. Users tend to continue even if that action is capable of putting them in an even worse situation.

2.3 Categories and types of Dark Patterns

The types introduced in this section are the same defined in the paper from Princeton university[12]. They are based on the types firstly published by Harry Brignull[7]. Princeton researchers discovered 15 types of dark patterns in total, and they divided them into seven broader categories. The summarisation of these types is in table 2.1 at the end of this section.

2.3.1 Sneaking

It is an attempt to hide, disguise, or delay information that is relevant to users. Users would likely change their action future action if they knew about this information. There are three types of dark patterns in this category: Sneak into Basket, Hidden Costs, and Hidden Subscription. Examples of these dark patterns can be seen in figure ??

2.3.1.1 Sneak into Basket

This type of dark pattern adds additional products into the user's basket without their consent. Usually, he is not aware of this fact. The added products are bonuses or additional services — for example, an additional year of warranty or a gift card. The essential for this dark patterns is that it raises the total price, and users might not be aware of this fact.

This dark pattern exploits the *default effect* of cognitive bias in users that was described earlier in this thesis. The literature says that this dart pattern is not *covert* because users can see the added products in their baskets.

2.3.1.2 Hidden Cost

This pattern is an attempt to add additional charges, typically at the end of the purchase process. Typical examples of this type of dark pattern are additional service fees or handling costs.

This type of dark pattern is also not *covert*, but it may be considered partially *deceptive* because the information is delayed from users. Also, this dark pattern can be classified into *hides information* dimension, as it attempts to hide information from users.

2.3.1.3 Hidden Subscription

This pattern signs up users into a subscription with a recurring fee. Users may not be aware of this subscription because the subscription is presented as a one-time payment or a free trial. This type of dark pattern usually appears together with another dark pattern named 'Hard to Cancel'.

This dark pattern is classified to be partially *deceptive* because it may confuse and mislead users. Also, it can be said that this dark pattern *hides information* from users.

2.3.2 Urgency

Dark patterns from this category speed-up users decision-making process by exploiting scarcity bias in users. For example, this can be done by showing more beneficial or time-limited discounts to users. As a result, users value products more than they would normally do. These dark patterns usually keep signalling that the special offer may be lost to users if they do not react promptly. This dark pattern is usually combined together with 'Social Proof' and 'Scarcity' types of dark patterns defined below. Examples of 'Urgency' dark patterns are shown in figure ??.

2.3.2.1 Countdown Timers

This dark pattern is usually in the form of an indicator of a deadline, counting down to the end of the deadline.

This dark pattern is classified as partially *covert* because it evokes untrue feelings of immediacy in users and is sometimes classified as *deceptive* because the indicator sometimes shows false information. For example, the timer can reset every time it reaches the deadline.

2.3.2.2 Limited-time Messages

The 'Limited-time message' dark pattern differs from 'Countdown Timer' by static urgency message and not showing the exact time of the deadline.

With the taxonomy defined before, this dark pattern is classified as *covert* because of the same reason as 'Countdown Timer' dark pattern and *information hiding* because it does not show the deadline in its offers.

2.3.3 Misdirection

This category of dark patterns uses visuals and language to distract users' attention on other possible presented choices. Also, some types from this category use users' emotions to invoke bad feelings of being guilty or ashamed for not making a specific choice. Users trust or feel that the other choices are unavailable or less beneficial for them. Essential for this dark pattern is that other choices are not hidden. Users are aware of the other choices, but this category of dark patterns steers users away from the other choices. Princeton researchers discovered four types of dark patterns from this category: 'Confirmshaming', 'Visual Interference', 'Trick Questions', and 'Pressured Selling'.

2.3.3.1 Confirmshaming

The 'Confirmshaming' dark pattern uses language and emotions to focus the attention of users on one choice in order to distract attention on other choices. Researchers point out that this dark pattern usually appeared in popup dialogues that asked for an email address in exchange for a discount. Users who did not want this discount had to choose a button (and the website did not want to choose this button) with the text 'No, I want to pay full price' or 'No thanks, I hate saving money'. This dark pattern exploits the framing effect of cognitive bias in users by presenting choices differently to users.

Thus, this dark pattern is classified as *asymmetric*. However, it is not *covert* since all the possible choices are presented to users.

2.3.3.2 Visual Interference

The 'Visual Interference' dark pattern uses different styles and visuals to draw users' attention to certain choices - the choices that the website wants users to choose. A typical example of this dark pattern is two buttons in different styles for opting-in and opting-out for the website's newsletter subscription. One of the buttons - the one that the website wants users to click on - looks more promising, more attractive to users' eyes than the other one. The different styles steer users attraction to the opting-in choice.

By provided taxonomy, this type of dark pattern is partially classified as *asymmetric* because it sometimes unequally present choices to users. Users may not realise that the effect of the dark pattern influenced them. Because of this fact,

this dark pattern is also classified as *covert*. Some instances can be also classified as *deceptive*, and Princeton researchers give an example of an option "lucky draw" among others that are deterministic and not random.

2.3.3.3 Trick Questions

The 'Trick Question' dark pattern uses confusing language to confuse users and their ability to make decisions. A typical trick in the English language for this dark pattern is double negatives. For example, websites, using this type of dark pattern, invert the meaning of a check subscription checkbox, usually seen in registration forms, followed with confusing language 'Uncheck this box if you prefer not to receive email updates'. Users need to pay more attention to properly understand which state of the checkbox means the subscription for the newsletter and which not. This type of dark pattern exploits the default effect in users, who erroneously believe that to them presented user interface follows traditional patterns. Also, this dark pattern exploits the framing effect by presenting the same information in a different, more confusing way to influence users in choosing different choices.

Therefore, Princeton researchers classify this type of dark pattern as *asymmetric* because opting out takes more effort than opting in. Also, researchers classify this dark pattern as *covert* because users may falsely understand the effect of their choice.

2.3.3.4 Pressured Selling

Princeton researchers define the 'Pressured Selling' dark pattern as pre-selecting more expensive variations of the same product as default. Additionally, pressuring users into choosing the more expensive variations or buying related products is also considered as a tactic of this dark pattern. More cognitive biases are triggered and exploited by this dark pattern, such as the default effect, the anchoring effect (users may tend to overlook the other choices) and the scarcity bias (more expensive variations may seem to be more exclusive).

This dark pattern is for some instances classified as *asymmetric* (i.e., steering users and their acceptance towards more expensive options), and partially *covert* (users may fail to realise that the firstly shown price of the less expensive variation of the product is not the same price, as the more expensive default variation).

2.3.4 Social Proof

The 'Social Proof' category of dark patterns is based on a social proof principle. Those hesitating individuals, who do not know what to do in a given situation, tend to observe others and mimic their moves, actions, and behaviour [4, 13]. This category of dark patterns misuses this behaviour of individuals, and it exploits the bandwagon effect of cognitive bias to its advantage. Princeton researchers define two types from this category: Activity Notifications and Testimonials of Uncertain Origin.

2.3.4.1 Activity Notifications

The 'Activity Notifications' dark pattern is information on product pages that indicate other users' activity. The message can have different forms. It can be a number of other users watching the same product or a number of sold products to other users. Messages displaying recent purchases of other users (e.g., 'User X just bought a product Y') also counts as 'Activity Notifications' dark pattern. Princeton researchers point out that some websites claim activity that is deceptive and not true. These websites use a misleading random number instead of factual information. This number also changes after some time, making it even more challenging to recognise as deceptive.

Some instances of this dark pattern can be classified as *covert* because users fail to understand that this dark pattern influences their decision-making process in a way that they tend to buy a product, which is sold more often or is viewed more by other users. Also, some instances are classified as *deceptive* because they present made up untruthful information and users are not aware of this fact.

2.3.4.2 Testimonials of Uncertain Origin

This type of dark pattern refers to the use of customer testimonials whose origin is unclear and not sourced enough. The result of such testimonials is that users' decision-making process is influenced by untrue information, and they erroneously believe in the quality of products.

Using taxonomy defined by Princeton researchers, this dark pattern is classified as sometimes *deceptive*, and it depends on the truthfulness of the testimonials, which can be determined by scanning the website and looking for a submission form for sending testimonials.

2.3.5 Scarcity

The 'Scarcity' category contains such types of dark pattern that implement messages indicating limited availability or high demand for a product. Thus, the value of the product increases because of its exclusivity. This dark pattern forces users to make quicker decisions. Users may feel intimidated by losing the chance to buy this very desirable product because it could be sold out soon. Princeton researchers define two types of dark patterns: 'Low-stock Message' and 'High-demand Message'.

2.3.5.1 Low-stock Message

The 'Low-stock Message' dark pattern informs users about the limited availability of a product; thus, users want to prevent losing the chance to buy the product by making quicker decisions than they normally do. Some instances of this dark pattern show exact quantities left on the stock. Others only show a message that stock is almost empty. This dark pattern exploits scarcity bias in users - making products more valuable only because it is low on stocks. Some websites use untruthful data to keep arousing the feelings of need in users all the time.

Princeton researchers classify the 'Low-stock Message' dark pattern as partially *covert* because users fail to realise that these messages influenced their decision-making process. Some instances of this dark pattern are classified as *deceptive* for displaying false information to users about being low on stock, but it is not. Some other instances are classified as *information hiding* for hiding the exact quantities of the product on stock.

2.3.5.2 High-demand Message

The 'High-demand Message' dark pattern informs users that a product is in high demand and can be sold out soon.

Similarly to 'Low-stock Message' dark pattern, 'High-demand Message' is also classified as partially *covert*.

2.3.6 Obstruction

This 'Obstruction' category contains only one type of dark pattern, which is 'Hard to Cancel'. This type of dark pattern refers to make specific actions harder to complete than other actions. For instance, signing up for a subscription to an annually paid service is often much more straightforward than cancelling the

subscription [6]. Also, Princeton researchers mention examples when cancellation of a subscription is available only by calling customer service[12].

This dark pattern is sometimes classified as *restrictive* with the defined taxonomy because it restricts the available choices to cancel the previous subscriptions. The 'Hard to Cancel' dark pattern becomes *information hiding* when the website does not inform users how to cancel the subscription or about the fact that cancellation is not as easy as signing up.

2.3.7 Forced Action

The 'Forced Action' dark pattern category forces users to take additional action, even though they might not normally take it to finish their task. The 'Forced Enrollment' is the only type of dark pattern discovered and defined by Princeton researchers in this category. This type of dark pattern forces users (that want to use the service) into enrolling for a marketing newsletter or into creating accounts, which gives the website more information than is needed to use the service. Princeton researchers describe an example when users have to simultaneously sign up for a marketing newsletter alongside their consent to terms of service.

Princeton researchers define this type of dark pattern as *assymetric*, because of the requirement of the additional actions to complete users' tasks, which creates asymmetrically balanced choices, and *restrictive*, because it forces users into creating accounts and signing up for marketing newsletters.

Table 2.1: Summarisation of categories and types of dark patterns with their description, definition and cognitive biases they exploit [12].

Legend: ●= Always, ◐= Sometimes, ○= Never

Category	Type	Description	Asymmetric?	Covert?	Deceptive?	Hides Info?	Restrictive?	Cognitive Biases
Sneaking	Sneak into Basket	Adding additional products to users' shopping carts without their consent	○	○	◐	●	○	Default Effect
	Hidden Costs	Revealing previously undisclosed charges to users right before they make a purchase	○	○	◐	●	○	Sunk Cost Fallacy
	Hidden Subscription	Charging users a recurring fee under the pretense of a one-time fee or a free trial	○	○	◐	●	○	None
Urgency	Countdown Timer	Indicating to users that a deal or discount will expire using a counting-down timer	○	◐	◐	○	○	Scarcity Bias
	Limited-time Message	Indicating to users that a deal or sale will expire will expire soon without specifying a deadline	○	◐	○	●	○	Scarcity Bias
Misdirection	Confirmshaming	Using language and emotion (shame) to steer users away from making a certain choice	●	○	○	○	○	Framing Effect
	Visual Interference	Using style and visual presentation to steer users to or away from certain choices	◐	●	◐	○	○	Anchoring & Framing Effect
	Trick Questions	Using confusing language to steer users into making certain choices	●	●	○	○	○	Default & Framing Effect
	Pressured Selling	Pre-selecting more expensive variations of a product, or pressuring the user to accept the more expensive variations of a product and related products	◐	◐	○	○	○	Anchoring & Default Effect, Scarcity Bias
Social Proof	Activity Message	Informing the user about the activity on the website (e.g., purchases, views, visits)	○	◐	◐	○	○	Bandwagon Effect
	Testimonials	Testimonials on a product page whose origin is unclear	○	○	◐	○	○	Bandwagon Effect
Scarcity	Low-stock Message	Indicating to users that limited quantities of a product are available, increasing its desirability	○	◐	◐	◐	○	Scarcity Bias
	High-demand Message	Indicating to users that a product is in high-demand and likely to sell out soon, increasing its desirability	○	◐	○	○	○	Scarcity Bias
Obstruction	Hard to Cancel	Making it easy for the user to sign up for a service but hard to cancel it	○	○	○	◐	●	None
Forced Action	Forced Enrollment	Coercing users to create accounts or share their information to complete their tasks	●	○	○	○	●	None

Chapter 3

Methodology

3.1 Logistic Regression

3.2 Bag of Words

3.3 Principal Component Analysis

3.4 HDB-SCAN

Conclusion

Bibliography

1. BERBA, Pepe. *Understanding HDBSCAN and Density-Based Clustering* [online]. 2020 [visited on 2021-04-04]. Available from: https://towardsdatascience.com/understanding-hdbscan-and-density-based-clustering-121dbee1320e.
2. BÖSCH, Christoph; ERB, Benjamin; KARGL, Frank; KOPP, Henning; PFATTHEICHER, Stefan. Tales from the dark side: Privacy dark strategies and privacy dark patterns. *Proceedings on Privacy Enhancing Technologies*. 2016, vol. 2016, no. 4, pp. 237–254. Available from DOI: 10.1515/popets-2016-0038.
3. *California bans 'dark patterns' that trick users into giving away their personal data* [online]. 2021 [visited on 2021-03-30]. Available from: https://www.theverge.com/2021/3/16/22333506/california-bans-dark-patterns-opt-out-selling-data.
4. CIALDINI, Robert B. *Influence: Science and practice*. Pearson education Boston, MA, 2009.
5. CONTI, Gregory; SOBIESK, Edward. Malicious Interface Design: Exploiting the User. In: *Proceedings of the 19th International Conference on World Wide Web*. Raleigh, North Carolina, USA: Association for Computing Machinery, 2010, pp. 271–280. WWW '10. ISBN 9781605587998. Available from DOI: 10.1145/1772690.1772719.

6. CRESTODINA, Andy. *The Long Goodbye: 7 Sites That Make It Hard to Unsubscribe* [online]. 2016 [visited on 2021-04-10]. Available from: https://unbounce.com/conversion-rate-optimization/when-friction-is-good/.

7. *Dark Patterns* [online]. 2010 [visited on 2021-03-28]. Available from: https://www.darkpatterns.org/.

8. *Dark Patterns - About us* [online]. 2010 [visited on 2021-04-01]. Available from: https://www.darkpatterns.org/about-us.

9. *Dark Patterns - Types of Dark Pattern* [online]. 2010 [visited on 2021-04-05]. Available from: https://www.darkpatterns.org/types-of-dark-pattern.

10. GRAUER, Yael. *Understanding HDBSCAN and Density-Based Clustering* [online]. 2016 [visited on 2021-04-05]. Available from: https://arstechnica.com/information-technology/2016/07/dark-patterns-are-designed-to-trick-you-and-theyre-all-over-the-web/.

11. GRAY, Colin M.; KOU, Yubo; BATTLES, Bryan; HOGGATT, Joseph; TOOMBS, Austin L. The Dark (Patterns) Side of UX Design. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Montreal QC, Canada: Association for Computing Machinery, 2018, pp. 1–14. CHI '18. ISBN 9781450356206. Available from DOI: 10.1145/3173574.3174108.

12. MATHUR, Arunesh; ACAR, Gunes; FRIEDMAN, Michael; LUCHERINI, Elena; MAYER, Jonathan; CHETTY, Marshini; NARAYANAN, Arvind. Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. *Proc. ACM Hum.-Comput. Interact.* 2019, vol. 1, no. CSCW.

13. NODDER, C. *Evil by Design: Interaction Design to Lead Us into Temptation*. Wiley, 2013. ISBN 9781118422144. Available also from: https://books.google.cz/books?id=ytwgZ%5C_QElT4C.

14. *The Year Dark Patterns Won* [online]. 2010 [visited on 2021-03-28]. Available from: https://www.fastcompany.com/3066586/the-year-dark-patterns-won.

15. ZHANG, Yin; JIN, Rong; ZHOU, Zhi-Hua. Understanding bag-of-words model: A statistical framework. *International Journal of Machine Learning and Cybernetics*. 2010, vol. 1, pp. 43–52. Available from DOI: 10.1007/s13042-010-0001-0.

Appendix A

List of Acronyms

DP	Dark Pattern
JS	Javascript
HTTP	Hyper Text Transfer Protocol
HAR	HTTP Archive
HTML	Hyper Text Markup Language
CSS	Cascading Style Sheets
API	Application Programming Interface
HDB-SCAN	Hierarchical Density-Based Spatial Clustering of Application with Noise
BoW	Bag of Words
PCA	Principal Component Analysis

Appendix B

Supplemental Material

The source code of the thesis and the implementation can be found on the attached medium or online at GitHub.

Thesis `https://github.com/Lznah/master-thesis`

GraphEvolution `https://github.com/Lznah/DarkPatterns`

```
├── README.md ..... the file with a brief contents description
├── MT_Petr_Hanzl_2019.pdf ..... the thesis text in PDF format
├── DarkPatterns/ ..... repository for the prototype
│   └── src/ ..... source code of the prototype
└── thesis/ ..... the directory of LaTeX source codes of the thesis
```

Directory structure B.1: Contents of the attached medium