Deploying artificial intelligence to detect and respond to the use of digital technology by perpetrators of human trafficking

Barry Cartwright
Richard Frank
Tiana Sharifi
Mandeep Pannu

April 2022

We gratefully acknowledge the financial support of the Crime Reduction Research Program, administered in partnership between the Ministry of Public Safety and Solicitor General, the Combined Forces Special Enforcement Unit of British Columbia, and RCMP 'E' Division.

Contents

1 Introduction

The International CyberCrime Research Centre (ICCRC), located in the School of Criminology at Simon Fraser University (SFU), set out to develop an artificial intelligence solution that would assist law enforcement agencies in detecting and responding to the use of digital technology by those who engage in human trafficking for the purposes of sexual exploitation. In particular, the ICCRC focused on adult classified websites and how they are being used to market the services of victims of human trafficking for the purposes of sexual exploitation. We have made significant progress in this regard, and hope to continue our work with various law enforcement agencies and community stakeholders to improve upon and finalize this artificial intelligence solution. We also attempted as part of this research project to investigate Internet-based communication tools such as social media and online games, and how they might be used to lure, groom and recruit minors and young adults for the purposes of sexual exploitation/trafficking; however, we found this to be unworkable, for reasons outlined later in this report.

Human trafficking for the purposes of sexual exploitation has been described as a “heinous crime” and a form of “modern-day slavery” (Public Safety Canada, 2018; Standing Committee on Justice and Human Rights, 2018). Ninety-five percent of victims are female, with 72% of those being under 25 years of age (Ibrahim, 2018). Teenagers or older children are at greater risk for online recruitment, as they often have unfettered Internet access with limited monitoring by parents or adult guardians (Johnson, 2019). Youth also tend to be more susceptible to flattery, gifts, promises of money and false promises of friendship or romance (French, 2019). While Indigenous peoples make up only 4% of the Canadian population, they are thought to account for 52-60% of sexually exploited youth in Vancouver, and 50% of all victims of sex trafficking in Canada (Mapp, 2020; Native Women’s Association, 2014). In particular, young indigenous females—especially those moving to urban centres—may be recruited online through promises of food, shelter, employment and protection.

Given that this research project was funded by the Government of British Columbia, and that it involved provincial stakeholders such as the Combined Forces Special Enforcement Unit-BC, the ICCRC concentrated primarily on activities taking place within British Columbia (and across Canada), bearing in mind that the Internet knows no boundaries, and that minors and young adults who are lured, groomed and recruited via the Internet in British Columbia may well be trafficked in other parts of Canada or in other countries around the world via the Internet, most notably across the border in the United States. Similarly, minors and young adults who are lured, groomed, and recruited from other parts of the world such as The Philippines, Hong Kong, South Korea, Taiwan, parts of Eastern Europe or even the United States may be trafficked in British Columbia or across Canada. To illustrate, indigenous females may initially be recruited from First Nations territories in the north of British Columbia, but subsequently find themselves being moved around between Prince Rupert, Vancouver, Calgary and Seattle (Oxman-Martinez, Lacroix & Hanley, 2005).

For this project, the ICCRC refined its custom-designed web crawler and its more recently developed AI (machine-learning) technology to scan classified advertisement websites that are thought to be used for the marketing of victims/services of victims of sexual exploitation or trafficking, retrieving and analyzing that evidence. The ultimate objective of the web-crawling and machine-learning technology that we are developing is to provide law enforcement agencies in British Columbia with the tools required to scan the Internet continuously, along with the ability to identify—and take prompt action against—the misuse of digital technology and Internet-based communication technology by the perpetrators of human trafficking for the purposes of sexual exploitation.

This research would not have been possible without the support of and funding from the British Columbia Ministry of Public Safety and Solicitor General's Crime Reduction Research Program, which is governed by the Joint Management Team of the Office of Crime Reduction and Gang Outreach (ORC-GO). We are deeply indebted to Marlies Dick and Julius Yuen of the RCMP's "E" Division Major Crime Section, Provincial Counter Exploitation Unit, who met with us on many occasions via Zoom and also corresponded with us extensively via email and telephone to provide us with the guidance required to program our web-crawling infrastructure and our machine-learning algorithms. Tiana Sharifi of Sexual Exploitation Education provided us with many invaluable insights into sexual exploitation and human trafficking, eventually joining our research team to help with our research design and qualitative coding, not to mention with arranging key meetings with various community stakeholders. Mandeep Pannu of PanVan CyberSecurity also joined the research team, to assist with the development of the web-crawling technology and to help us to scrape data from adult classified websites that were thought to be hosting content pertaining to the marketing of the services of victims of sexual exploitation and human trafficking. More is said about the research team below.

2 The Research Team

The principal investigator, Dr. Richard Frank, has a PhD in Criminology and a PhD in Computing Science and is the Director of the ICCRC. In the past, Dr. Frank has conducted research on hacking forums, child exploitation websites and extremist websites. He is the inventor of The Dark Crawler (TDC), a tool for collecting and analyzing data from the Internet, and the author of more than 50 publications on cybercrime. Dr. Frank oversaw this entire project and performed a number of the technical tasks himself.

Dr. Barry Cartwright, Associate Director of the ICCRC, has a BA (Hons.) and MA in Sociology, and a PhD in Criminology. Until recently, he was a Senior Lecturer in the School of Criminology at SFU, where he taught courses such as Current Perspectives in Cybercrime, Morality, Ethics and Law in Cyberspace, and the Sociology of Law. Dr. Cartwright has published in the areas of cyber-research, cyber-law, cyber-bullying, cyber-warfare and cyber-terrorism. He was in charge of the qualitative research team and assisted Dr. Frank in overseeing this project.

Drs. Frank and Cartwright were joined in this project by PanVan CyberSecurity, which is owned and operated by Dr. Mandeep Pannu. Dr. Pannu holds an MSc and PhD in Computing Science from Coventry University in the UK. Apart from operating PanVan, she is also an Assistant Professor in Computer Information Systems at the University of the Fraser Valley and an Instructor in Computer Science & Information Technology at Kwantlen Polytechnic University. She and her team at PanVan CyberSecurity helped to design a customized web-crawler as well as to scrape, organize format and save data from adult classified websites.

Tiana Sharifi, a recognized national expert in the area of human trafficking/sexual exploitation, and the founder of Sexual Exploitation Education, provided extensive information to the research team during the funding application phase, and was then invited to join the team once the funding was obtained and the project was underway. Tiana has a B.A. in Psychology with a minor in Counselling. After working frontline and as a Program Director for a provincial non profit in the field, Tiana Sharifi created the agency, Sexual Exploitation Education (SEE), which strives to enhance the safety of children and youth through education, training and consulting services. Through SEE, Tiana has been involved with law enforcement training, non profit training, student and school district presentations, and serving as the Chair of the Board of Directors for the Breakfree Collective. Tiana helped with the research design and qualitative coding, as well as with setting up numerous consultation meetings with various community stakeholders.

Karmvir Padda, a member of the Indo-Canadian community, has an MA in Criminology, a diploma in Police Foundations and a BSS (Hons.) in Criminal Justice. Karmvir, who is the holder of a SSHRC PhD grant, is enrolled in the PhD program at the University of Waterloo, where she is studying the propagation of online disinformation by alt-right and White Supremacist groups. She has extensive experience in qualitative coding, having worked on several different projects with the ICCRC. Karmvir assisted with the formatting of the datasheets and drop-down menus, and with the NVivo and SPSS analysis of the qualitative data.

Sarah-May Strange, a member of the LGBTQ+ community, has an MA in Criminology and has conducted graduate research with the ICCRC on the topic of dis/misinformation on social media. She has extensive experience in qualitative research, in particular dealing with dis/misinformation attacks against LGBTQ+ communities in North America. She is planning to pursue PhD studies in areas relating to radicalization and de-radicalization, i.e., what makes people who have been indoctrinated into homophobia, Islamophobia and extreme right-wing conservatism alter or modify their views. Sarah-May was primarily responsible for conducting the literature review and helping to write up the literature review.

Soobin Rim, a Korean-Canadian female, has a BSc in Computing Science, and is presently enrolled in an MSc program in Computing Sciences at SFU with a focus on cyber-security. Soobin has worked on several different projects with the ICCRC and has extensive experience in dealing with machine-learning algorithms. She was primarily responsible (with the assistance of Jason Wang, below) for training the machine-learning algorithms to identify adult classified ads that exhibited clear signs of sexual exploitation/human trafficking and to teach the algorithms to distinguish them from ads posted by consensual sex workers.

Jason Wang, a Taiwanese-Canadian male, has a BSc in Computing Science, and has previously worked with the ICCRC on the development of its web-crawling infrastructure and on several disinformation warfare projects. Jason was primarily responsible (with the assistance of Soobin Rim, above) for training the machine-learning algorithms to identify adult classified ads that exhibited clear signs of sexual exploitation/human trafficking and to teach the algorithms to distinguish them from ads posted by consensual sex workers.

Dr. George Weir, an international member of the ICCRC and the designer of the Posit Toolkit, is a Lecturer in Computing Science at the University of Strathclyde in Glasgow. Posit facilitates quantitative analysis of large text corpora, applying a Part-of-Speech tagger and outputting statistical details of the text in terms of individual words (tokens) and word types. Dr Weir has worked previously with the ICCRC on various research projects. More will be said about his contributions to this project later in the report.

3 Literature Review

In 2000, the United Nations created the United Nations Protocol to Prevent Suppress and Punish Trafficking in Persons, Especially Women and Children, Supplementing the UN Convention Against Transnational Organized Crime, often referred to as the Palermo Protocol (Siller & van Doore, 2019). Globally, the most commonly used definitions of human trafficking in both law and academic research derive from those set out in the Palermo Protocol (Mapp, 2020). The Palermo Protocol states that:

> *"Trafficking in persons" shall mean the recruitment, transportation, transfer, harbouring or receipt of persons, by means of the threat or use of force or other forms of coercion, of abduction, of fraud, of deception, of the abuse of power or of a position of vulnerability or of the giving or receiving of payments or benefits to achieve the consent of a person having control over another person, for the purpose of exploitation. Exploitation shall include, at a minimum, the exploitation or the prostitution of others or other forms of sexual exploitation, forced labour or services, slavery, or practices similar to slavery, servitude or the removal of organs (Protocol to Prevent, Suppress and Punish Trafficking in Persons Especially Women and Children, Supplementing the United Nations Convention against Transnational Organized Crime, 2000)*

The *Palermo Protocol* identifies three elements as essential to qualify as human trafficking of adults: the act, the means, and the purpose (Millar & O'Doherty, 2015; Siller & van Doore, 2019). This definition "require[es] that a stipulated 'act' be committed through application of specific 'means' for a number of stipulated exploitative 'purposes'" (United Nations Office on Drugs and Crime, 2018, p. 22). When it comes to the trafficking of children, however, only the act and purpose are required, and not the means (Millar & O'Doherty, 2015). An act might consist of recruiting, harbouring, or moving a person. The term "means" refers to illicit methods to complete these actions or to obtain "consent" from the trafficked person: those means might consist of threats or coercion, deceiving a person, or taking advantage of a position of power over the person. Lastly, the purpose—the reason for the trafficking—must be exploitative in nature (Siller & van Doore, 2019; United Nations Office on Drugs and Crime, 2018b).

This last element or requirement, the presence of an exploitative purpose, has at times created issues and "wiggle room" in legislation, academic research, and global views of what constitutes trafficking (Millar & O'Doherty, 2015; Siller & van Doore, 2019; United Nations Office on Drugs and Crime, 2018b). The Palermo Protocol notably does not fully define the term "exploitation", choosing instead to offer potential examples of minimum actions that might count as exploitation, rather than setting out a solid definition (Siller & van Doore, 2019; Protocol to

Prevent, Suppress and Punish Trafficking in Persons Especially Women and Children, Supplementing the United Nations Convention against Transnational Organized Crime, 2000). Some perspectives, for instance, may frame all pimps as sex traffickers, seeing them as inherently and always exploitative, a view which many consensual sex workers see as reductive and potentially dangerous, or to express it differently, another way in which sex trafficking is sometimes conflated inaccurately and thoughtlessly with sex work (Canadian Alliance for Sex Work Law Reform, 2019; DeLateur, 2016; Roots, 2019). However, some experts in the field of anti-human trafficking feel that the very nature of a pimp demonstrates the legal elements necessary to be identified as a human trafficker.

(If we can not incorporate this, then I would delete this last paragraph as it is not objective).

The creation of the Palermo Protocol undoubtedly moved global anti-trafficking efforts forward, for example, mandating the creation of anti-trafficking legislation and action plans within signatory nations such as Canada (Siller & van Doore, 2019; Protocol to Prevent, Suppress and Punish Trafficking in Persons Especially Women and Children, Supplementing the United Nations Convention against Transnational Organized Crime, 2000). Canadian law addresses trafficking in persons via both the Criminal Code, and the Immigration and Refugee Protection Act, the latter applying primarily to international trafficking (Government of Canada, 2021). Trafficking may be domestic or international, with the majority of cases in Canada involving Canadian victims. The Canadian Centre to End Human Trafficking released a 2019-2020 report that found 86% of survivors or victims to be Canadian (Canadian Centre to End Human Trafficking, 2021). Domestic trafficking consists of crimes of exploitation which take place entirely inside a country, without the crossing of international borders. (That said, domestic trafficking may—and often does—involve the crossing of provincial or state borders within a given country). International trafficking by definition involves the crossing of international borders, and in particular, taking a person from one country to another in order to commit a crime of exploitation (Ministry of Public Safety and Emergency Preparedness, 2019). Human trafficking is sometimes divided into labour trafficking and sex trafficking, although the two are frequently viewed as part of the same continuum (Arifin et al., 2021; Keskin et al., 2021; Volodko et al., 2020).

3.1 Prevalence

Accurate statistics regarding sex trafficking can be difficult to come by for a variety of reasons. The intensely criminal nature of sex trafficking means that it is inherently secretive and covert (Ibanez & Gazan, 2016b; Keskin et al., 2021). Variations in definitions of sex trafficking—between jurisdictions and between government and nongovernmental organizations (NGOs) engaged in its measurement and prevention—create the potential for flawed figures, which can be exacerbated further by the conflation of sex trafficking with consensual sex work (Mackenzie & Clancey, 2020; Naqvi, 2020; Siller & van Doore, 2019). Additionally, estimates of the prevalence and nature of sex trafficking are affected by the source and data used to create the estimates. If one looks at police-reported sex trafficking cases, for instance, the numbers are much lower than those supplied by assorted NGOs (Farrell & de Vries, 2019; Ibrahim, 2021; Khodarkovsky et al., 2021; United Nations Office on Drugs and Crime, 2021b). This is understandable, given that victimization surveys such as the Canadian General Social Survey routinely reveal that seven out of ten crimes are not reported to the police, and that crimes with a sexual component are among the least likely to be reported (Cotter, 2019). Moreover, victims of sex trafficking are by definition controlled and coerced to one extent or another, often living in fear of their controller, and thus may lack the ability or wherewithal to report their victimization to law enforcement agencies. Lastly, due to this coercion, many victims may not self-identify as victims, and therefore may not seek help.

In 2021, the United Nations Office on Drugs and Crime (2021) released its Global Report on Trafficking in Persons for the year 2020. Researchers who created the report examined official (primarily police-recorded) figures about human trafficking from 148 countries. They found that the victims of (officially-recorded) human trafficking were primarily female (see Figure 1). In 2018, approximately 46% of recorded victims of trafficking were adult women, and 19% of all victims of trafficking were underage girls. Adult men made up 20% of trafficking victims, and underage boys made up a further 15% (United Nations Office on Drugs and Crime, 2021a). Additionally, approximately half of the victims of human trafficking detected in 2018 experienced sexual exploitation, with most women having been trafficked for that purpose, while most men were trafficked for the purpose of forced labour (United Nations Office on Drugs and Crime, 2021a).

According to the Global Report on Trafficking in Persons, in 2018, approximately 50,000 victims of human trafficking were detected globally (United Nations Office on Drugs and Crime, 2021a). However, these figures must be treated with caution, as they only reflect individuals who have come into contact with authorities, which may in turn impact the recorded number of victims and their demographics (Cockbain & Bowers, 2019; United Nations Office on Drugs and Crime, 2021b). Many victims of trafficking may not come into contact with authorities, especially as laws, preventative measures, commitment to intervention and funding can vary significantly from nation-to-nation and from jurisdiction-to-jurisdiction (Barrett & Shaw, 2013; United Nations Office on Drugs and Crime, 2021b).

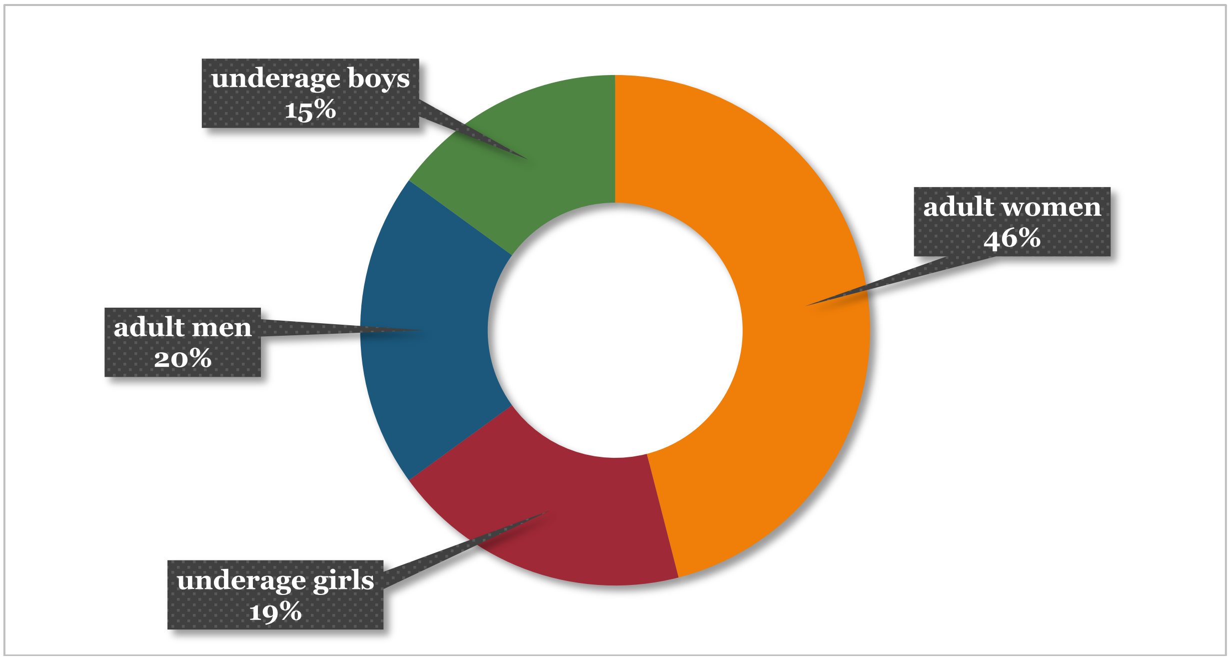


Figure 1 - Gender and age group of global victims of human trafficking who came into contact with authorities in 2018, from statistics compiled by UNODC (2021)

Additionally, the overrepresentation of women and children in these statistics may reflect the tendency of anti-trafficking efforts—and concomitant legislation—to focus on women and children (Bouché et al., 2018; Cockbain & Bowers, 2019). This focus may result in adult male victims receiving less help, and women and children being subjected to greater scrutiny, and in some cases, unwelcome intervention from and involvement with authorities (Cockbain & Bowers, 2019; De Shalit et al., 2021). It is important to note that in many instances, suspected victims of human trafficking are subjected to harsh and punitive treatment by authorities, including measures such as incarceration or deportation or both (De Shalit et al., 2021; Millar & O'Doherty, 2020; Musto et al., 2021). To complicate matters further, legislative initiatives, detection and intervention efforts by governmental and nongovernmental agencies, the availability of resources and support mechanisms and anti-trafficking initiatives in general are often focused on sex trafficking, quite possibly to the detriment of victims of labour trafficking (Clancey & Wiseman, 2020; Cockbain & Bowers, 2019). This uneven focus may be reflected in the data. Moreover, if detection is focused on sex trafficking, then it seems likely that the victims of sex trafficking victims will be detected more often and in greater numbers (Clancey & Wiseman, 2020; Rodríguez-López, 2018).

With respect to Canada, data regarding the prevalence of human trafficking is similarly limited, with the 2019 report by Statistics Canada on human trafficking mostly relying on police-reported incidents (Ibrahim, 2021). This means that only those cases of human trafficking that were detected by the police and/or brought to the attention of the police <u>and</u> that were made the subject of an official police report were counted. For 2019, Canadian law enforcement reported 511 such incidents, representing a slight increase from prior years; however, this increase may be attributable to the greater focus on trafficking in recent times (Ibrahim, 2021; United Nations Office on Drugs and Crime, 2018a). According to the cases of human trafficking that were reported to (and recorded by) the police, 95% of victims of human trafficking in Canada in 2019 were female. However, these numbers may also reflect the predominant focus on sex trafficking, as well as on women as victims (Cockbain & Bowers, 2019; Ibrahim, 2021). Notably, human trafficking cases faced a significant winnowing-down or attrition process in the Canadian criminal justice system: of the cases of human trafficking reported by police, only one-third resulted in charges. Of those charges, 89% were stayed, withdrawn, discharged, or dismissed (Ibrahim, 2021).

3.2 Identifying Trafficking

In 2020, the United Nations Office on Drugs and Crime (UNODC) published a list of Human Trafficking Indicators, with a cautionary note that the absence or presence of one or more of these indicators does not prove conclusively that human trafficking is or is not occurring—instead, they are intended to indicate that further investigation is warranted in a given situation (United Nations Office on Drugs and Crime, 2020a). These indicators are split into different sections: general indicators, those relating specifically to children, and those relating to domestic servitude, along with sections relating to sexual exploitation, labour exploitation and exploiting victims for the purposes of begging and engaging in petty crimes (Giommoni & Ikwu, 2021; United Nations Office on Drugs and Crime, 2020a). A selection of these indicators is presented in Figure 2a and 2b. Recent research into human trafficking has attempted to operationalize these indicators for the online environment in general, and for online advertisements of sexual services in particular (Giommoni & Ikwu, 2021).

That said, these indicators can be difficult to operationalize effectively, as researchers must convert general ideas into specific, actionable signs, which can prove difficult when attempting to detect and make sense of these indicators in online advertisements (de Vries & Radford, 2021; Giommoni & Ikwu, 2021). Additional issues with these indicators were identified by De Vries and Radford (2021), who pointed out that many of the indicators might be applicable to any number of employment situations. To illustrate, Volodko, Cockbain, and Kleinberg (2020) applied some of the UNODC's indicators to a dataset of posts taken from a website that advertised job positions. Of the 430 advertisements examined, 98.4% contained at least one of the UNODC's indicators of labour trafficking, which according to the authors of the study suggests that certain "indicators" of labour trafficking may simply be commonplace characteristics associated with a variety of occupations (Volodko et al., 2020).

While Volodko et al. (2020) examined the UNODC indicators of labour trafficking, specifically within online labour advertisements, their observations can be extended to the general indicators of human trafficking as well as to indicators that are thought to be specific to sex trafficking. A number of the UNDOC indicators or human trafficking in general—e.g., showing fear or anxiety, being distrustful of authorities or living in poor or substandard accommodations—may be common to many marginalized individuals, not just to victims of human trafficking. Some of the UNODC's sex trafficking indicators—e.g., working in various locations, working long hours, having the type of clothing mostly associated with sex work, or having unprotected sex—may also be characteristics of sex trade work in general, and of the type of language found in online advertisements posted by sex trade workers (Mackenzie & Clancey, 2020; Marcus et al., 2013; Millar & O'Doherty, 2020).

Therefore, caution must be exercised when attempting to operationalize these UNODC indicators as concrete variables, as well as when using the presence of these indicators as determinants of the presence of trafficking (de Vries & Radford, 2021; Mackenzie & Clancey, 2020; Volodko et al., 2020). Giommoni & Ikwu (2021) suggest that researchers could improve the likelihood of correctly identifying advertisements linked to sex trafficking by searching for posts that contain multiple indicators. Some researchers have created systems in which advertisements might be flagged as potentially showing signs of being related to sex trafficking in cases where the ad features key words related to a possible underage victim as well as words or symbols suggesting that the victim frequently travels from location to location (Ibanez & Gazan, 2016b). For this present study, we used search terms consisting of key words and key phrases identified by law enforcement agencies and community outreach organizations who specifically seek to educate about, prevent and interdict sex trafficking, we winnowed those down into key words and key phrases where we could confirm through careful manual inspection that they applied exclusively (or almost exclusively) to the marketing of the services of sex trafficking victims, and for the sake of precision, decided for the qualitative inspection that at least three of those key words or key phrases had to be present in order to flag an online advertisement as involving sex trafficking.

General Indicators

- Believe that they must work against their will
- Be unable to leave their work environment
- Show signs that their movements are being controlled
- Feel that they cannot leave
  Show fear or anxiety
- Be subjected to violence or threats of violence against themselves or against their family members and loved ones
- Suffer injuries that appear to be the result of an assault
- Suffer injuries or impairments typical of certain jobs or control measures
- Suffer injuries that appear to be the result of the application of control measures
- Be distrustful of the authorities
- Be threatened with being handed over to the authorities
- Be afraid of revealing their immigration status
- Not be in possession of their passports or other travel or identity documents, as those documents are being held by someone else
- Have false identity or travel documents
- Be found in or connected to a type of location likely to be used for exploiting people
- Be unfamiliar with the local language
- Not know their home or work address
- Allow others to speak for them when addressed directly
- Be forced to work under certain conditions
- Be disciplined through punishment
- Be unable to negotiate working conditions
- Receive little or no payment
- Have no access to their earnings
- Act as if they were instructed by someone else
- Work excessively long hours over long periods
- Not have any days off
- Live in poor or substandard accommodations
- Have no access to medical care
- Have limited or no social interaction
- Have limited contact with their families or with people outside of their immediate environment
- Be unable to communicate freely with others
- Be under the perception that they are bonded by debt
- Be in a situation of dependence
- Come from a place known to be a source of human trafficking
- Have had the fees for their transport to the country of destination paid for by facilitators, whom they must pay back by working or providing services in the destination
- Have acted on the basis of false promises

Sexual Exploitation

May Also Indicate Sexual Trafficking of Children

- Be of any age, although the age may vary according to the location and the market.
- Move from one brothel to the next or work in various locations
- Be escorted whenever they go to and return from work and other outside activities
- Have tattoos or other marks indicating "ownership" by their exploiters
- Work long hours or have few if any days off
- Sleep where they work
- Live or travel in a group, sometimes with other women who do not speak the same language
- Have very few items of clothing
- Have clothes that are mostly the kind typically worn for doing sex work
- Only know how to say sex-related words in the local language or in the language of the client group
- Have no cash of their own
- Be unable to show an identity document
- There is evidence that suspected victims have had unprotected and/or violent sex.
- There is evidence that suspected victims cannot refuse unprotected and/or violent sex.
- There is evidence that a person has been bought and sold.
- There is evidence that groups of women are under the control of others.
- Advertisements are placed for brothels or similar places offering the services of women of a particular ethnicity or nationality.
- It is reported that sex workers provide services to a clientele of a particular ethnicity or nationality.
- It is reported by clients that sex workers do not smile.

Figure 2a - UNODC indicators related to sexual trafficking of humans (UNODC, 2020)

Sexual Exploitation Education, a Canadian Anti-Human Trafficking Agency, notes the following additional factors:

1 Feel indebted to their trafficker for supplying their accommodation, food and living expenses
2 Have someone managing their finances
3 Experience control over their behaviours and actions that leads them to engage in sexual services, often in exchange for a need
4 Experience a different work environment than what was promised to them
5 Believe that their trafficker is their boyfriend or manager
6 Possess items that they cannot typically afford, including nail acrylics or eyelash extensions
7 Be under the age of 18

Figure 2b - Sexual Exploitation Education indicators of sexually exploited youth (Sexual Exploitation Education, n.d.)

7.1 Online Sex Trafficking

FOSTA-SESTA, an American law passed in 2018, has had a major impact upon both the landscape of sex trafficking and upon the lives and safety of sex workers (Reynolds, 2021). Referred to fully as the Allow States and Victims to Fight Online Sex Trafficking Act and the Stop Enabling Sex Traffickers Act, the law was passed with support from Republicans and Democrats alike. FOSTA-SESTA made Internet service providers and websites liable for sex trafficking that was facilitated by use of their services. The stated intent of the law was to curtail sex trafficking and to make the Internet safer, although it has also had the (perhaps unintended) effect of shutting down the most popular online venues—*Backpage* and *Craigslist*'s ads for sexual services—used by legitimate, consensual sex workers to advertise their work (Blunt & Wolf, 2020; Musto et al., 2021; Reynolds, 2021). After *Craigslist* and *Backpage*'s sex work advertisements were shut down, advertising and marketing of sexual services dispersed, and now cannot be found easily in one central online location, thereby creating further challenges for current research and investigation (Blunt & Wolf, 2020; Khodarkovsky et al., 2021). The fall of *Backpage* in particular resulted in the dispersion of online advertisements for sexual services with numerous alternatives appearing (Khodarkovsky et al., 2021; United States Government Accountability Office, 2021). Before one can begin examining advertisements for signs of human trafficking one must first locate the online advertisements.

According to Keskin et al. (2021), sex traffickers embrace and implement technology in order to expand, streamline and conceal their illicit operations. At the same time, governmental bodies, law enforcement, and NGOs look to technology to provide more effective means of identifying, finding, and fighting against trafficking (Keskin et al., 2021; O'Brien & Li, 2020; Raets & Janssens, 2021; United Nations Office on Drugs and Crime, 2020b). "[I]t is difficult … for [law enforcement] to continuously monitor the available data, identify patterns and translate them into meaningful information" (Keskin et al., 2021, p. 1111). Therefore, many law enforcement and legislative bodies have encouraged research into detection using AI, machine-learning, and other technological methods (Hundman et al., 2018; Landman, 2020; Simonson, 2021), along the lines of what we are pursuing with this present research project. Despite the intensity of the present-day focus on sex trafficking, research into the use of the Internet and other digital technology in the commission of sex trafficking is relatively limited, with more recent studies beginning to examine the Internet's role and impact in greater depth (Raets & Janssens, 2021). Many assume that the Internet is deeply intertwined with human trafficking, including sex trafficking, but to date, empirical research has been somewhat lacking (de Vries et al., 2020; Raets & Janssens, 2021).

Raets and Janssens (2021) carried out an extensive literature review of extant studies on the topic of digital technology and human trafficking, supplementing their investigation with a number of qualitative interviews of anti-trafficking workers and interviews with offenders actually convicted of human trafficking. Their literature review included 82 studies, and along with the qualitative interviews, provided insight into the ways that the Internet is used both to commit and combat human trafficking. Raets and Janssens (2021) explored four areas in which the internet, and digital technology in general, are used by human traffickers, reflecting domains identified by Aronowitz (2009, as cited in Raets and Janssens, 2021): recruitment, transportation, managing finances, and exploitation (see Figure 3).

Recruitment involves using the Internet as a place to search for both victims and customers (Simonson, 2021; Volodko et al., 2020). Traffickers use a variety of online locations, from social media platforms to online classifieds/marketplaces to dating sites; on these sites, the traffickers scope out possible victims, often using the personal information posted by the site users in order to identify those who appear easy to access. (Middleton, 2020; Raets & Janssens, 2021; United Nations Office on Drugs and Crime, 2021a). Once victims are identified, the trafficker moves on to connecting with the target, creating trust and using strategies to gain control, including coercion, threats, persuasion, blackmail, and/or violence (Public Safety Canada, 2021; Raets & Janssens, 2021).

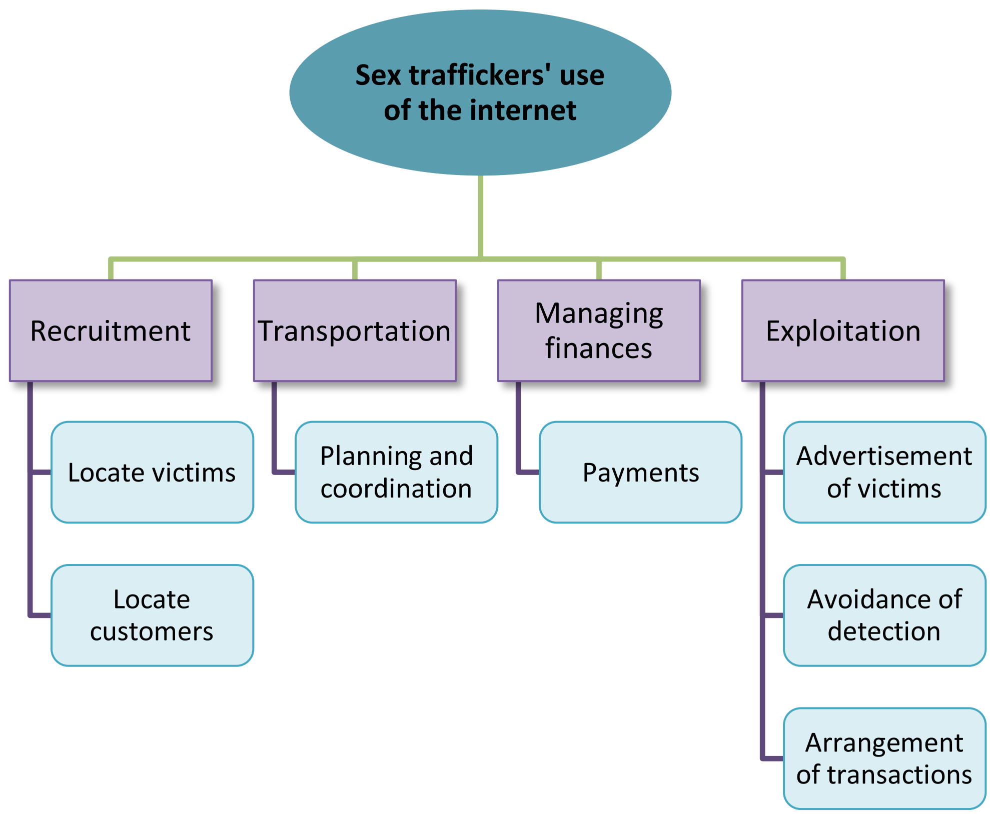


Figure 3 - The use of the internet by sex traffickers, adpated from Raets and Janssens (2021)

Transportation, especially in international trafficking, may require moving a victim from their original location into other destinations, which use of the Internet helps to coordinate by allowing traffickers to direct operations, even from far away (Oxman-Martinez et al., 2005; Raets & Janssens, 2021; United Nations Office on Drugs and Crime, 2021a). The Internet is also valuable to traffickers for the purpose of managing their finances and indeed for collecting their ill-gotten gains. Some traffickers accept digital payments while others have begun to accept cryptocurrencies such as Bitcoin (Di Nicola et al., 2017; United Nations Office on Drugs and Crime, 2021a).

7.2 The Online Marketplace

Sex traffickers also use the Internet as a tool for exploiting their victims, including marketing the services of those victims through online advertisements (Inter-Agency Coordination Group against Trafficking of Persons, 2019; United Nations: UN News, 2021). In 2020, 80% of sex trafficking cases prosecuted by the United States Justice Department involved the use of online advertisements (United Nations: UN News, 2021). This number must be viewed with caution, however, as the use of online advertisements can create useful evidence in the case, a technological trail of linkages and data, which may result in prosecutors being more likely to take the case to court (Bach & Litam, 2017; Siouxsie Q, 2018; United Nations: UN News, 2021). Nonetheless, using the Internet allows traffickers to reach a large audience through their advertisements, which can create a high impact for relatively low effort (Inter-Agency Coordination Group against Trafficking of Persons, 2019; Middleton, 2020; United Nations: UN News, 2021).

By using online advertisements, traffickers can capture the attention and business of potential buyers, and also use the online platform to arrange the transaction (Inter-Agency Coordination Group against Trafficking of Persons, 2019; United Nations: UN News, 2021). Advertising online permits greater secrecy and anonymity for both traffickers and customers, allowing them to discuss terms, prices, and other details in a relatively removed manner, with a perceived reduction in possible exposure (Raets & Janssens, 2021; United Nations Office on Drugs and Crime, 2021a). Traffickers can create online profiles and post advertisements without having to reveal their true identity, be it their name, face, or other personal information (Middleton, 2020; Parise, 2020).

Sex traffickers are flexible and quick to adapt and may use a variety of advertising methods and locations (Raets & Janssens, 2021; United Nations Office on Drugs and Crime, 2021a). Advertisements may appear on well-known websites or on smaller niche sites with a limited audience; they may appear on social media platforms or on websites specifically designed to host escort advertisements or reviews and on dating sites, and online classified/marketplace sites (Di Nicola et al., 2017; Ibanez & Suthers, 2014; Raets & Janssens, 2021; United States Government Accountability Office, 2021). Especially since *Backpage* was shut down, each country or general geographic area may have its own websites which are primary locations of sex trafficking ads, because advertisements relating to sex work in general have been compelled to decentralize (Blunt & Wolf, 2020; Keskin et al., 2021; Khodarkovsky et al., 2021; Raets & Janssens, 2021). One of the first tasks for researchers, therefore, is to identify potential online sites where such advertising might be taking place, based upon geographical location. Once those are identified, investigators can begin analyzing posts for indicators of sex trafficking (Raets & Janssens, 2021).

As *Backpage* and Craigslist's advertisements for sexual services were closed relatively recently, and as they provided a fairly easy collection point for data, much current research is based on data samples taken from these sites (Alvari et al., 2016; Asencion, 2017; Nagpal et al., 2017). In the years since *Backpage* and Craigslist's sexual service advertising platforms closed, and advertisers have been forced to relocate, some researchers have scraped and analyzed data from alternative sites (Giommoni & Ikwu, 2021; Volodko et al., 2020). Some of these sites go unnamed: for instance, Giommoni and Ikwu (2021) refer to their data source obliquely as being the biggest platform in the UK for sex workers to advertise their services. However, other research has mentioned collecting advertisements from specific locations, such as Keskin et al. (2021) who analyzed advertisements posted on the recently-popular platform skipthegames.com, collecting 434,720 posts and 749,503 images.

Traffickers may use multiple websites, and indeed multiple types of online platforms (United Nations Office on Drugs and Crime, 2021a; United States Government Accountability Office, 2021). In addition to classifieds-style sites intended to host advertisements, and marketing-focused escort pages, traffickers make use of many social media platforms as well as free-standing websites that they create themselves (Mensikova & Mattmann, 2018; Tong et al., 2017; United Nations Office on Drugs and Crime, 2021a). Child victims of trafficking are purportedly more commonly advertised on social media (United Nations Office on Drugs and Crime, 2021a). Social media platforms used by traffickers both to recruit and to advertise may include sites and applications such as Facebook, WhatsApp, and VKontakte (Siouxsie Q, 2018; United Nations Office on Drugs and Crime, 2021a).

When setting out to analyze the data they have collected, there are a number of potential signs that researchers suggest might indicate the involvement of human trafficking, one of the most common places being indicators found within the actual text of the advertisement (Esfahani et al., 2019; Raets & Janssens, 2021; Tong et al., 2017). Analysis of the text may include searching for key words and key phrases believed to relate to exploitation and trafficking, such as euphemistic code terms for victims that are underage, like "fresh" (Hultgren et al., 2016; Keskin et al., 2021; Whitney et al., 2018). Researchers also specifically search for facets of grammar in the phrasing of the posts, both in terms of third-person pronouns and in terms of repeated use of similar language and phrasing (Alvari et al., 2016; Kennedy, 2012). However, it is important to heed the caution of Raets and Janssen (2021): "One critical issue…is the lack of ground truth in identifying human trafficking cases" (p. 227).

To reduce the danger of being found and prosecuted—and particularly to avoid detection by text-focused research or programs—online advertisements by human traffickers often feature evasive language (Giommoni & Ikwu, 2021). The conclusions of the research and literature review conducted by Raets and Janssens (2021) suggest that when traffickers use the Internet to advertise, they usually "hide in plain sight through the use of cryptic terms or coded language" (p. 223). Sex traffickers, like other businesses, must advertise to reach customers; however, they must do so while avoiding detection by law enforcement (Keskin et al., 2021). Keskin et al. (2021) identify some of key challenges that law enforcement officials face when looking for sex trafficking online, including obfuscation, changes of location, and the high volume of posts.

These methods of obfuscation often appear to be designed to avoid detection by artificial intelligence programs and other automated text-reading programs (Kejriwal & Kapoor, 2019; Parise, 2020). Advertisements may appear "noisy", and intentionally filled with elements that would render the ad difficult for computer programs to comprehend, while remaining intelligible for human readers (Nagpal et al., 2017; Raets & Janssens, 2021). Methods thought to be common to online sex trafficking ads include use of coded key words/phrases to convey information without using words that are recognisable as being associated with illicit activity (Bounds et al., 2020; Hultgren et al., 2016; Keskin et al., 2021). As these coded key words have become more well-known, traffickers have begun to use emojis to represent illicit meanings, as will be discussed (Nagpal et al., 2017; Whitney et al., 2018). Other methods of obfuscation may include insertion of letters or words into phone numbers, unusual spellings, special characters, and the use of icons/images in place of letters or words (Kejriwal & Szekely, 2017; Keskin et al., 2021; Raets & Janssens, 2021). In our current research, for example, we found the work "young" spelled as "YOunG," and code words that do not exist, such as "PetiteCute" and "cutegirl" to signify youth, as well as many variations of "BBBJ," "BBBj," "bxxJ," and "b88j", all intended to signify the availability of unprotected oral sex.

The work of Whitney, Jennex, Elkins and Frost (2018) has been a basis for many discussions about the use, significance, and meaning of emojis in sex trafficking advertisements. Although their research was primarily presented as part of conference proceedings in 2018, it has been cited by a number of current reports and studies, including government reports, as the primary (and sometimes the only) source for claims that emojis are often used in specific ways for specific illicit meanings (Kejriwal & Kapoor, 2019; Parise, 2020; United States Government Accountability Office, 2021). While Whitney et al.'s (2018) work is valuable, its use as an often-solitary source illustrates the need for greater academic research into online sex trafficking advertisements, as discussed by Raets and Janssen (2021). We observed extensive use of emojis in our present study, including for example the airplane emoji indicating that the person had "just arrived" or was "newly arrived" or was in town for a "short stay only," and the lollipop and growing heart emojis to signify youth inexperience or childlike innocence.

Whitney et al. (2018) used knowledge management principles as well as natural language processing to analyze the coded language in sex trafficking advertisements that they gathered from *Backpage*. By analysing the number of times where known key words/terms appeared in the same advertisements as specific emojis, they identified emojis that appeared to be statistically significant, suggesting that these emojis could serve as a coded reference for those elements or features usually indicated by key words. For instance, they found statistically significant connections between the appearance of the "growing heart", "cherry", and "cherry blossom" emojis alongside key terms believed to indicate a victim who was a minor as compared to the relative absence of these emojis in ads that did not feature any of those key words. According to Whitney et al. (2018), some of the key terms believed to be associated with minors included words such as "fresh", "young", "new", "tiny", "little", "new in town", "girl", and "college" (Whitney et al., 2018, p. 4277). New key words and key phrases that we have come across in our own research and that are also intended to signify minors include "Pettit", "PetiteCute", "babygirl", "100% YOUNG" and "rope bunny", to name but a few.

Some of these key words, euphemisms, and codes are mutually understood by both sellers and buyers, and quite possibly originated and became well-known to those involved in trafficking (including customers) when classified websites such as *Backpage* and *Craigslist* were still in operation (Ibanez & Gazan, 2016a; Parise, 2020). To discourage traffickers, these webpages began to filter out ads that featured particular forbidden terms associated with or assumed to be associated with sex trafficking. The steps taken by sex traffickers to make advertisements more creative and more difficult to detect for law enforcement and site operators (Nagpal et al., 2017; Parise, 2020) demonstrates the ability of the traffickers to adapt quickly to efforts to curtail their activities, and at the same time, the need for online hosting platforms, law enforcement agencies and social researchers to be equally flexible and quick to adapt (Keskin et al., 2021).

7.3 Fighting Technology with Technology

Giommoni and Ikwu (2021) examined indicators of human trafficking in online advertisements for sexual services, primarily analyzing the text used by the advertisers. Focusing specifically on services of female sex workers in the United Kingdom, Giommoni and Ikwa (2021) created a data-scraping tool which gathered over 17,000 advertisements from a very popular online platform for "adult entertainment"'. This unnamed online platform offered a multitude of services, such as cam shows and chatting, but a large part of it focused on "hosting" advertisements. Posts on the site were required to include fairly detailed textual information, thereby providing the potential for data scraping by researchers and the assembly and assessment of considerable amount of textual content. Their analysis in particular focused on advertisements by individuals who offered "in-person" sexual services (Giommoni and Ikwu, 2021). Notably, the sample purposively excluded advertisements featuring male sex workers, or any ads from female sex workers who accepted non-male clients, which reflects a common gap in sex trafficking research (Cockbain & Bowers, 2019; Giommoni & Ikwu, 2021).

After harvesting the dataset, Giommoni and Ikwu (2021) set out to identify advertisements that featured indicators of sex trafficking. To analyze their data, they based their variables on the UNODC (2020) list of indicators of human trafficking that are set out at length above. They found the indicators to be complex and difficult to operationalize in the realm of online advertisements. For instance, the UNODC (2020) indicator— "evidence that groups of women are under the control of others" —was operationalized on the basis of searching for potential indicators of third-party involvement. Therefore, if multiple advertisements contained the same phone number, or if advertisements used plural pronouns such as "we" or "they", and there was noteworthy similarity between the wording of multiple advertisements, then they considered that to be evidence of sex trafficking (Giommoni and Ikwu, 2021). Our qualitative research team also observed the presence or plural pronouns such as "we" and "they" in our manual inspection of 1,840 online ads drawn mostly from Canadian sites, and also, similarity between the wording of multiple advertisements; however, we treated the interpretation of those ads with caution, because in many cases, there appeared to be an equally likely chance that the individuals being described in this manner were consensual sex workers, and not necessarily sex trafficking victims.

Keskin et al. (2021) sought to identify online human trafficking based upon high-level patterns, in contrast to the primary method that is currently in use, that being the comparison of new data to databases of historically posted advertisements. To achieve this goal, Keskin et al. (2021) used a dataset of 10 million posts harvested from a popular venue for advertising sexual services, skipthegames.com, which is thought to have been associated with multiple cases of sex trafficking. The researchers aimed to detect patterns and connections within the data, grouping posts based upon post text, listed phone numbers and photos (Keskin et al., 2021).

Analysis of photos represents an interesting and promising step in the detection of advertisements related to sex trafficking, as most computer-based efforts have tended to focus on textual content, which is easier to access and analyze (Alvari et al., 2017; de Vries & Radford, 2021). By employing a technique known as image hashing, Keskin et al. (2021) were able to identify when the same photo appeared in different advertisements within their dataset, as well as it appearing at different times and/or on different platforms. Image hashing uses an algorithm to analyze an image and produce a specific string of characters based upon the qualities of that images, including pixel values. An identical string of characters will result from the picture, even if the picture is harvested from different posts and/or from a different website. Thus, even if the ads featured different text and different phone numbers, or had no text or phone numbers at all, the presence of the same photo would group those posts together, indicating they may originate from the same source (Keskin et al., 2021; Portnoff et al., 2017). That said, it should be borne in mind that generally speaking, there is nothing to prevent a complete stranger from copying (stealing) a photo that is posted on the Internet by somebody else and then using the copied photo for their own purposes.

The discovery of groups who use online classified ads to market the sexual services of their victims is considered valuable, as it offers researchers and law enforcement agencies the additional insights required to engage in activity tracking as well as the identification of traffickers, their networks, and their modus operandi (Andrews et al., 2016; Keskin et al., 2021). Raets & Janssens (2021) describe an entity resolution process which enabled the location of similar potential groups. The entity resolution process focused on creating ad clusters based upon traditional textual information, such as phone numbers and e-mail addresses, but also incorporated relatively new developments for the prospective grouping and tracing on illicit activities online, namely cryptocurrencies such as Bitcoin (Portnoff et al., 2017; Raets & Janssens, 2021).

In summary, sex traffickers use online platforms and digital technology to exploit victims, increasing the field of potential customers and thereby enhancing their profit potential (Inter-Agency Coordination Group against Trafficking of Persons, 2019; Tundis et al., 2019). At the same time, however, data harvested from these online platforms can be highly valuable in efforts to combat sex trafficking (Raets & Janssens, 2021; Simonson, 2021). The use of online platforms and marketplaces by sex traffickers allows researchers insight into this highly covert crime, facilitating the study of methods, strategies and networks that are usually not visible, while also creating potential evidence with which to build and successfully prosecute criminal cases (Dubrawski et al., 2015; Raets & Janssens, 2021; Siouxsie Q, 2018).

The presence of sex trafficking indicators within online advertisements creates an opportunity for identification and intervention, but for this opportunity to be truly useful, it must be approached with precision, care, rigorous methodology, willingness to question commonly accepted conclusions, and ethical consideration at the forefront (Clancey & Wiseman, 2020; Cockbain & Bowers, 2019; Musto et al., 2021). Academic research and government initiatives that target sex trafficking require increased rigour, more stringent measures to ensure that the results are truly empirical, and a concerted effort to ensure that anti-trafficking programs are properly aimed at harm prevention, not at marginalized populations and consensual sex workers (Cockbain & Bowers, 2019; DoCarmo, 2019; Raets & Janssens, 2021).

8 Methodological considerations

In essence, there were three different research teams, all working together but also working individually on quite different tasks. The qualitative research team, led by Dr. Barry Cartwright, was tasked with identifying websites that contained sex trafficking ads, defining the search terms (key words, key phrases, key symbols) for the web-crawling and machine-learning teams, manually inspecting and classifying large samples of the data downloaded by the web-crawling team, refining the search terms as the data were being manually inspected and classified, and conducting in-depth analysis of the data in order to identify the major themes, patterns and narratives associated with human trafficking/sexual exploitation. The web-crawling team, led by Dr. Mandeep Pannu, was tasked with building a stand-alone crawler designed specifically to harvest material from Canadian ad sites that were identified as being likely hosts of advertisements for the sexual services of victims of human trafficking. The machine-learning team, led by Dr. Richard Frank, was tasked with the building, testing and validation of machine-learning algorithms that would have the capacity to automatically read, analyze and correctly classify datasets gathered from these adult classified websites.

8.1 The qualitative research team

In the initial stages of the project, the qualitative research team exchanged numerous emails and met several times via Zoom with Marlies Dick, Julius Yuen of the RCMP's "E" Division Major Crime Section, Provincial Counter Exploitation Unit, and Tiana Sharifi of Sexual Exploitation Education. The RCMP directed the team to a number sites that were thought to host advertising for the sexual services of victims of human trafficking. Websites identified by the RCMP included *adultlook*, *adultsearch*, *girl-directory*, *megapersonals*, *leolist*, *onebackpage*, *skipthegames* and *yesbackpage*. This information was then provided to Dr. Mandeep Pannu and her team at PanVan CyberSecurity, who were tasked with building a customized web-crawler in order to retrieve relevant data from these sites for further analysis by the qualitative research team and the machine-learning team.

As noted in our literature review, these websites tend to be a moving target, with new ones always opening up to fill the gap created by those that are forced to close or that close voluntarily due to the attention they attract from law enforcement agencies and other governmental regulatory agencies (*cf.* Blunt & Wolf, 2020; Giommoni & Ikwu, 2021; Keskin et al., 2021; Khodarkovsky et al., 2021; Raets & Janssens, 2021; Volodko et al., 2020). The web-crawling team advised the qualtiative research team that some sites that were identifed as potential targets, such as *adultsearch*, *vancouverbedpage*, *vancouver2bedpage*, no longer existed. Thus, in consultation with the RCMP and various community stakeholders, Tiana Sharifi and Dr. Cartwright generated a new, updated list of sites that included *megapersonals* and *locanto*.

There was considerable consultation between Dr. Pannu's team at PanVan, Dr. Frank's machine-learning team and Dr. Cartwright's qualitative research team with respect to the layout of the columns and contents of these spreadsheets, as they had to be formatted in a fashion that was amenable to manual inspection and coding, downloading into NVivo for computer-assisted qualitative analysis, and inputting to the machine-learning algorithms. It was discovered during the early stages of the research process that these datasheets had to be modified on a site-by-site basis, as none of the sites used exactly the same layout, and different sites contained quite different content.

The RCMP, alongside Sexual Exploitation Education, also provided the qualitative team with a number of the key words, key phrases and key symbols, such as "young", "24/7", "no restrictions" and no "black gents". They provided us with the type of altered photographs used by sex traffickers in their online advertising, as well pictures of some of the tattoos found on the victims, e.g., currency symbols (indicating that their services are for sale), a crown or a crown with initials (signifying that they are owned or controlled), and even barcodes. Examples of the symbols and emojis provided to us by the RCMP included the cherry or cluster of cherries (signifying virginity or being underage), the airplane symbol (new in town or visiting), the growing heart (minor, childlike or inexperienced), the strawberry (juicy or young), and again the crown (indicating that they have a pimp or a controller). This information was then relayed to Dr. Mandeep Pannu and her team at PanVan CyberSecurity (who were tasked with building the customized web-crawler) to assist them in searching for and retrieving relevant data for further analysis by the qualitative team and the machine-learning team.

The web-crawling team started to download, organize and save (thousands) of online classified ads, forwarding the data to the qualitative team in Excel spreadsheets that had been pre-formatted for this purpose. Two members of the qualitative team, Dr. Barry Cartwright and Tiana Sharifi, started reading and classifying these advertisements together through a series of Zoom meetings, to develop a rudimentary coding schematic and to ensure consistency in how they were coding the ads (inter-rater-reliability). Next, they started coding the ads on their own in batches of 100 messages, and then meeting regularly via Zoom to review their coding decisions, modifying the coding schematic as well as the case-by-case coding where warranted.

They consulted with the RCMP and other community stakeholders in cases where there was uncertainty, to ensure that their final interpretations were correct. To illustrate, it was observed that some of the key words, key phrases and key symbols appeared with regularity in the online ads for consensual sex trade workers as well as in the ads for individuals deemed to be potential victims of human trafficking/sexual exploitation. The RCMP was asked how much weight should be accorded to the plane emoji and expressions such as "catch me while you can" or "short stay only", in the absence of any other evidence of trafficking. The "24/7" phrase also appeared frequently in these ads, but in many cases appeared to be associated with body rub parlours, where it was not clear whether the women were being trafficked. It was ultimately decided that in order to ensure the most accuracy for law enforcement agencies in British Columbia, one or two of these words, phrases or symbols on their own or in combination with each other would not be sufficient indicators of sex trafficking.

A prime example of this would be the plane symbol and/or the expressions "short stay only," "new to town" or "just arrived", because these also appeared in ads where the individuals described themselves as "mature" or "professional" and as a "woman", and talked about “companionship” or “dine and date” scenarios; such ads also set out all sorts of restrictions, such as “safe services only”, “no car dates”, “limited availability” “deposit in advance” or “no back door”. Upon closer inspection, and in consultation with the RCMP and other community stakeholders, along with extensive discussion between members of the qualitative team, it was concluded that ads where both sex trafficking narratives and sex worker narratives were present, would not be coded as trafficking for the purpose of this study.

Some sex workers may travel from town-to-town and/or from province-to-province, possibly in search of new trade, or following major events such as the Stanley Cup Playoffs, the Grey Cup, or the Calgary Stampede, and/or simply using language (quite possibly copied from other online ads) that they thought would be appealing to potential customers.

In fact, there were multiple observable instances of ads that evidently been copied and pasted, or that had been assembled by copying and pasting parts of different ads.

~~In some cases, the individuals who composed such ads had not given a great deal of thought to what they were doing, because they might start off the ad by saying that they were a “mature” or “professional” lady, and then saying that they were “young” and “just arrived”.~~

In some cases, the individuals who composed such ads had not given a great deal of thought to what they were doing, and how one statement that they were making might contradict another statement made in the same ad. One ad, for example, included the words “young” and “petite,” code words used by traffickers to signify youth and tininess, but then in the age, height and weight sections of the same ad went on the describe the person as 27 yrs old , 5' 6” and 136 lbs. Another also advertised the word “petite”, but then went on to describe the person as “6ft 177 Ilbs.” It is plausible that they were using the word “petite” because they saw it mentioned many times in other ads, without realizing the meaning of the word or what it was usually intended to signify. It is also possible that they used words such as “young” and “petite,” knowing what the words meant and that they did not necessarily apply to them, hoping that use of this language would attract customers.

One person who had seemingly assembled her ad from other ads said that she was “just landed” and mentioned it being her “First time here” and her “First work.” The word “first” and the phrase “first time” are used in trafficking ads to signify youthfulness and inexperience. However, it may have been true that it was this person’s “First time here” and her “First work” in Canada, as she then proceeded to demonstrate her lack of fluency in written English by describing herself as a “sweat Chinese girl” with “Big boots”. She also used the phrase “open minded,” but did not seem to grasp what this was usually intended to convey, as she went on to provide a fairly typical, detailed menu that she likely did not compose herself and which contradicted the notion of her being open-minded by adding a number of restrictions: “100% G.friend E. Shower together, ; *69*; ATY, multi position, Sensual and good massage skill.......and more., , F.Service 140~30mins 180~45mins 220~60mins EXTRA C1M+40 R1M+40 SORRY no back door and SAFETY PLAY ONLY...Add 60 for second shot...”

This was also evident in ads that were suspected of involving sex trafficking, where they started out with descriptors such as “100% YOUNG 100% Pretty”, “code words” thought to be used by the traffickers to signify a very young/underage victim, but these ads then went on the describe them as a “horny mature girl” and as “honest sincere lady,” without recognizing the apparent contradiction in terms.

While some researchers have regarded copied-and-pasted ads as evidence of sex traffickers at work—using the same language repeatedly to advertise the services of the same victims (cf. (Alvari et al., 2016; Kennedy, 2012)—we decided that this was not necessarily the case. Rather, it appeared upon closer manual inspection that the advertisers simply looking at other online ads, copying and pasting the “templates”, and then altering the templates in a manner that more or less conveyed what they were trying to get across, much like lawyers use the same template when drawing up a will (they change the names addresses, heirs, assets and division of assets in accordance with the unique circumstances but otherwise use exactly the same language over and over again), or when college students copy and paste the same language that they find online or in the scholarly journals that they consult (changing a few words here or there to make it look like their own academic writing). While copying and pasting may be indicative of sex traffickers at work, it should not on its own be regarded as conclusive evidence of sex trafficking.

Another research decision made by the qualitative team was to avoid counting the same indicator of sex trafficking more than once. For example, a given advertisement might feature the phrase 24/7” and “available all day and all night”, but they essentially signified the same thing. Similarly, the online ads might use phrases like “no restrictions” or “no limits” or “do as you wish” to tell customers that anything was on the table. If you counted instances of this, and found that it was mentioned twice, once in the tagline and then again in the detailed description of services offered, then you would be measuring the same indicator twice.

In all, the qualitative research team manually coded—and cross-validated the scoring of—1,840 online sex ads, breaking them down in accordance with their likelihood of representing sex trafficking or sex trade work. The team derived extensive lists of (ever-evolving) key words and key phrases and made note of the presence (or absence) of specific emojis or symbols. All this information was made available to the web-crawling and machine-learning teams for the purposes of refining their searches and/or the development of the machine-reading/machine-learning algorithms.

Also, this information was used to search the data once it was downloaded into NVivo. Essentially, NVivo is an advanced code and retrieve program that facilitates the formulation and testing of hypotheses and the building of theory (Bryman, Teevan & Bell, 2009; Popping, 2000; Richards, 1999). It is particularly helpful for narrative and discourse analysis because apart from coding, categorization, and coding storage, it facilitates advanced word searches, proximity searches, organization of data into sets, and assaying of nodes and attributes, to mention several of its many features. Equally important, NVivo makes it possible to link documents to each other, and to write memos that can be linked to a specific document or to as many documents to which they might be relevant (Bryman et al., 2009: Weitzman, 2003). In summary, NVivo facilitates codification and visualization of data and allows for data queries and automatic provisional coding of an entire dataset. It makes analysis of textual content more systematic, thorough and transparent. It also allows for team members working from different locations to access and code the shared dataset and compare and merge their findings (Elg & Ghauri, 2019; Wiltshier, 2011).

One of the original research objectives of the qualitative team was to gain greater understanding of, and insights into, the processes involved in the luring, grooming and recruitment of youth for the purposes of sexual exploitation/human trafficking. With this in mind, the qualitative researchers on the team planned to search for and harvest chat-room content from popular gaming sites such as Fortnite, Minecraft, League of Nations, Roblox, Call of Duty, and Discord, as well as text messaging from social media sites such as Twitter, Whisper, Kik Messenger, Live.Me, Tik Tok, and Google Hangouts. In particular, the team planned to look for early invitations to go on webcam or to meet in person, requests for information about age, gender and location, the presence of flattery, requests to undress, and the sharing of sexually explicit comments or sexually explicit material, as well as party invitations, promises of employment (like modelling gigs with photographers) or offers of free video gaming cards, free trips, and free accommodations.

However, we soon found this approach to be highly impractical, not to mention non-productive and fraught with ethical concerns. Our search for evidence of luring, grooming and recruitment of youth in chat rooms and on gaming sites or social media sites could be likened to a social researcher sitting in a cul-de-sac in a quiet suburban residential area, watching and waiting for a crime to occur—you could wait for hours, days, or even weeks, and never see anything of interest. Tiana Sharifi possessed by far and away the most experience in identifying this sort of behaviour by virtue or her work with Sexual Exploitation Education. Thus, we assigned her to this research task. She spent countless days visiting any number of these venues in search of the type of luring, grooming and recruitment discourse that would permit us to extract key words and key phases to build search terms that would in turn assist us in customizing our web-crawling infrastructure to automatically search for such illicit online content, but found nothing. Apart from that, we determined that continuing to monitor the online activities of youth was clearly unwarranted if there was no evidence of predatory criminal behaviour taking place.

We turned to Cybertip.ca, but they were either unable or unwilling to provide us with this type of data. We also consulted with members of the RCMP's "E" Division Major Crime Section, Provincial Counter Exploitation Unit, but they either did not have data of this nature or felt that it was too sensitive or case-specific for them to release. We made a similar request to members of the Victoria Police Department's Special Victims Unit (introduced to us by Tiana Sharifi) who were similarly tasked with investigating cases of sex trafficking, but they too were unable to offer us anything. We then attempted to glean this information from court cases involving the prosecution of high-profile sex traffickers but found that the type of discourse employed in the luring, grooming and recruitment process was either not included in the court transcripts or had been expunged entirely. Thus, we made a research decision to abandon this fruitless search, and instead focus our efforts on areas where we could have some actual impact.

While most forms of human trafficking are covert, sex trafficking includes one public point, that being where these services are advertised (Silva et al. 2014). These sites serve as "virtual red-light districts," where traffickers can advertise their victims as escorts, masseuses or companions (Middleton, 2019). This afforded us an opportunity to gain greater understanding of and insights into how classified websites are used by traffickers to market the victims/services of victims of sexual exploitation/trafficking. We had originally planned to harvest data from *vancouver2backpage* and *vancouver.bedpage.com*, but both sites disappeared by the time that our research got underway, suffering the same fate as their infamous predecessor, *Backpage* (Blunt & Wolf, 2020; Khodarkovsky et al., 2021). Nevertherless, we identified new and equally suitable sites with the assistance of Tiana Shariai and the RCMP members of the Provincial Counter Exploitation Unit and were able to harvest the material required for in-depth qualitative analysis and for the training and testing of our machine-reading/machine-learning algorithms.

8.2 The web-crawling infrastructure

The ICCRC possess its own web-crawler, known as The Dark Crawler (TDC), designed by Dr. Richard Frank, who was the principal investigator for this project. We had originally considered using TDC for this project, as it is readily extensible and has massive scraping and storage capacity (Mei & Frank, 2015). However, we elected instead to involve PanVan CyberSecurity, which is owned and operated by Dr. Mandeep Pannu. As noted previously, Dr. Pannu holds an MSc and PhD in Computing Science from Coventry University in the UK and is also an Assistant Professor in Computer Information Systems at the University of the Fraser Valley and an Instructor in Computer Science & Information Technology at Kwantlen Polytechnic University. She and her team at PanVan CyberSecurity already had their own crawler infrastructure and were tasked with building a stand-alone crawler that was specifically designed to harvest material from Canadian online adult classified ad sites identified as being likely hosts of advertisements for the sexual services of victims of human trafficking. It was, and remains, our goal to either integrate this new crawling infrastructure into TDC or append it to TDC.

The team that was building the crawling infrastructure started by crawling the type of sites mentioned above, searching for online content containing known advertising “code words” provided to them by the qualitative research team, such as “young, “youthful,” “new,” “fresh,” “sweet,” “barely legal”, and “daddy’s little girl” (International Center for Missing and Exploited Children, 2018). The qualitative research team also provided them with a number of emojis that were thought to be associated with the trafficking of humans for the purposes of sexual exploitation. Many of these key words, key phrases and key symbols were originally provided to us by Tiana Sharifi and RCMP’s “E” Division Major Crime Section, Provincial Counter Exploitation Unit.

That said, in the longer term, we found it more effective (and more beneficial from a knowledge-generation perspective) to instruct the team that was building the crawling infrastructure to download all of the content found on a given site, leaving the extraction of relevant key words and key phrases to the qualitative research team. In this regard, we took into consideration the ever-shifting nature of the key words and key phrases, and the ever-increasing use of emojis or deliberately misspelled words (or previously non-existent words)—or to express it differently, the lengths to which the purveyors of such services will go in order to obscure their illicit activities (Bounds et al., 2020; Kejriwal & Kapoor, Keskin et al., 2021; 2019; Parise, 2020; Raets & Janssens, 2021).

The design of the customized web-crawling infrastructure commenced with an examination of the page sources for each site under investigation, to consider how the online content could best be extracted and filtered down into Excel spreadsheets. As noted above, there was extensive consultation between the three research teams with respect to the lay-out of the columns and contents of these spreadsheets, to ensure that they were in a format that was open to manual inspection and coding, downloading into NVivo for computer-assisted qualitative analysis, and inputting to the machine-learning algorithms.

With respect to the online classified ad sites, Dr. Pannu's team at PanVan started by exploring whether the site in question was "static" or "dynamic". Static sites have their HTML rendered to the backend server, after which the HTML is then sent to the client. For static sites, the full HTML content can be retrieved using Linux commands such as *wget*/*curl*, which facilitates the recursive downloading of content as a file. Dynamic sites send minimal HTML to the client (as required), instead using asynchronous scripts to retrieve the requested data and fill in the HTML content on the client's browser, at the same time allowing users to create their own content. The content of static sites proved to be reasonably easy to download and parse in order to retrieve the desired content. In the case of dynamic sites, it sometimes proved difficult to retrieve the entire HTML content without the involvement of an automated web browser, in this case, Chrome, automated by Selenium, another open-source software program. The dynamic sites required the team to program a browser to open, orchestrate, render and download the HTML links.

Due to differences in the way that the sites were constructed (dynamic or static), it was deemed necessary to create a new crawler for each site, requiring considerable "trial-and-error" manual work to search for the right tags upon which to build the different filters. For the dynamic sites, Dr. Pannu's team at PanVan turned to Scrapy—an open-source Python framework—to build out their crawlers. Scrapy has a user-friendly Application Programming Interface (API) that permits you to select and filter the content that you actually want and then use that data to create a CSV file (which in turn can be converted easily into an Excel file for manual qualitative inspection or for inputting to NVivo).

The PanVan team also had to specify search terms for each site and for each person being advertised, such as location, age, height, body type, eye colour, ethnicity, and the presence of tattoos, to list a few of the search terms. To illustrate, the code built by the PanVan team to scrape content from a dynamic website would look something like the code presented in Appendix A.

When it came to static websites such as *yesbackpage*, the PanVan team found that Scrapy was not as effective at collecting the requisite data, because it was necessary to make changes to the code every time that the content of the website was changed. Thus, it proved easier for the team to render the page by adding the plugins for a headless Chrome browser and developing their own Python code for static websites. The team also employed Beautiful Soup, which is a Python library that helped to parse the HTML returned by the server, and also to ascertain whether they were logged in or not. To illustrate, one of the codes built by the PanVan team to scrape content from a static website would look something like the code presented in Appendix B.

It was also observed that the crawlers for the individual sites had to be modified on a day-by-day basis, due to the fact that the sites were being updated on a daily basis and that new content was being added regularly. In some instances, the websites had restricted or blocked content, which necessitated additional spiders, a spider being a program that visits websites and reads their pages and other information to create entries for a search engine index. In some instances, the PanVan team found it necessary to change their login credentials on a regular basis, possibly because the site operators were on the lookout for Internet hosting platforms and law enforcement agencies and/or academic researchers that might be monitoring their activities and scraping data; thus, it is likely that the site operators were changing their parameters to thwart such activities.

The crawls could take anywhere from a few hours to several days, sometimes involving multiple crawls of the same sites, depending upon whether the site was static or dynamic, and the rate limitations set by the websites themselves.

8.3 The Posit Toolkit

One of the main objectives of the qualitative research team was to manually inspect and classify as much data as possible, to provide the machine-reading/machine-learning team with as much pre-classified data as possible, for purposes of training the machine-learning algorithms, and for cross-validation of the output of the machine-learning algorithms. In a sense, the classifications assigned by the qualitative team may be regarded as a “gold standard” against which the output of the machine-learning algorithms can be judged. As noted previously, the qualitative team manually inspected 1,840 online sex ads, extracting key words, key phrases, and key symbols, and breaking the ads down in accordance with their likelihood of representing sex trafficking or sex trade work. However, it is impossible for two, three (or even four of five) qualitative researchers to manually inspect and classify thousands or hundreds of thousands of data items as they are being scraped.

Thus, we decided to turn to Dr. George Weir, an international member of the ICCRC, and to his Posit Toolkit. Dr. Weir has worked with the ICCRC on a number of occasions, including a research project that investigated extremist content on the Internet (cf., Weir, dos Satos, Cartwright & Frank, 2016), another that examined COVID-19 mis/disinformation on social media, and a more recent (ongoing) investigation into the influence operations of hostile foreign actors on social media. It was anticipated that Dr. Weir and his Posit Toolkit (Weir, 2007, 2009) might be able to automate the pre-classification of large datasets of online sex ads with a high degree of precision, thereby eliminating the need for manual inspection and classification of thousands or hundreds of thousands of data items, while at the same time providing us with a second “gold standard” against which the output of the machine-learning algorithms could be judged.

Total words	Adjective types	Nouns
Total unique words	Preposition types	Prepositions
Type-token ratio	Possessive types	Possessive pronouns
Number of sentences	Personal types	Personal pronouns
Average sentence length	Determiner types	Particles
Number of characters	Adverb types	Interjections
Average word length	Particle types	Determiners
Noun types	Interjection types	Adverbs
Verb types	Verbs	Adjectives

Figure 4 - Posit summary values

Posit is a set of software tools designed for quantitative textual analysis of single texts or sets of texts. The result of Posit analysis is a detailed account of the words and part-of-speech types present in the sample data. Applying such analysis to sets of related texts affords a means of contrasting various textual dimensions across the texts. Such characteristics have proven to be highly effective as a means of generating feature sets from texts for use in classification, especially in the context of machine learning.

Among its actions on any text sample, Posit applies a Part-of-Speech tagger and notes statistical details of the text, from total word and character counts through to individual words (tokens) and word types (based upon Parts-of-Speech). A convenient summary is generated that indicates values for 27 aspects of any analysed text. Figure 4 lists the standard set of 27 attributes reported by Posit.

Based upon this analytical approach, when applied to sets of textual data, the generated features and their values serve to characterize the corresponding text items and afford 27 dimensions of comparison across the texts.

In the context of textual data classification, where a pre-classified representative sample set is available, Posit can be used as an analytical tool to generate the aforementioned features for the pre-classified data. Thereafter, using supervised machine learning, the resultant Posit analyses are used as features in conjunction with classification algorithms to match the known manual classification. In principle, a high degree of match suggests that the generated model (mapping the identified features to the manual classification) could be used to automate the classification of future unseen text samples.

The Posit analysis processed and classified three datasets, those being *leolist_initial*, *leolist_recoded*, and *yesbackpage*. These three datasets had been scraped by the web-crawling team, and then manually classified by the qualitative team with each item assigned the value of "yes" or "no" for suspected sex trafficking, one of them (*leolist_recoded*) having been re-coded by the qualitative team after Dr. Weir pointed out that a few of the cases had been left out in the initial manual coding process, and that a number of them had been marked as "undecided". With each of these datasets, a subset of manually classified items was provided. This manual classification identified each item with the value "yes" or "no".

The task was to apply Posit analysis to inform subsequent supervised machine-learning across each of the three sample datasets. The resultant feature values would then be used as a basis for classification, with a view to calculating the effective degree of match that could be achieved between classification based upon the Posit analysis and machine-learning on the one hand and the pre-determined (manual) classification on the other. More is said about the results of the Posit analysis in Section 0 of this report, where we set out our research results in detail.

8.4 The machine-learning algorithms

Below we describe the model-building or machine-learning algorithms that were used to analyze the datasets gathered from adult classified websites. Each algorithm was implemented and evaluated on each of the datasets. Each algorithm takes as input a dataset similar to an Excel spreadsheet, with each row being a different webpage, and each column a different feature extracted from that webpage. One of the columns in this dataset must be the label upon which the models are constructed, which can then be used to predict whether or not a given ad is indicative of human trafficking. The machine-learning team would extract the textual content of each ad as it had been scraped by the web-crawling team, and then use the column where the qualitative team had assigned the value of "yes" or "no" for suspected sex trafficking, thus creating separate datasets that were labeled as either "Human Trafficking" (HT) or "Not Human Trafficking" (NHT). A sample from one of the datasets is shown in Figure 5.

age	availability	Services	Title/tagline	description	Is it HT?	Sex Trafficked Narratives	Sex Work Narrati	Images	id	name	Comments (e.g.,	image_urls	images	link	stats	weight	Trafficking (yes/r	eye	hair	height
25	Incall & Outcall & Online			,Hello , My name is Priya ., ,I'M DOING INCALIS & OUTCALLS., ,DON'T TEXT IF YOU'RE NOT ON YOUR BUDGET, ,LUXURY SERVICE / HIGHLY SKILLED / PSE , ,SEXY GIRL. NEW IN TOWN., ,SEXY SEDUCTIVE CLASSY, ,: Surrey , , GOOD REVIEWS search for me or Look at my ad , ,MULTIPLE HOUR SPECIALS. Only for the right Gentleman ,	No	None	3 of the above		ad6577711	Priya	While "New in town"	https://i.leolist.cc/8{	[{'url': 'https://i.leolist	https://www.leolist.c			no			
20	Incall & Outcall & Online		NEW GIRL IN TOWN FETISH NO RUSH + FREE SHOWER SHORT STAY!	,I am KATRINA, 20, Real Asian college student in town, ,REAL pics . Real sweet heart just right here., ,BEST BBB! BEST BLOW IN TOWN GUARANTEED, ,Open minded and shower Together (No Rush), ,Amazing Mouth** Deep & Sloppy! ((*))((*)) the best month ever,WET&JUICY.., ,I am here to relieve any stress or desires you may have!, ,NON-Rushed service, ,Clean and Discreet Private place, ,Tight and clean pu$$y, ,Safe Services,	Yes	2 of the above	None		ad6548921	KATRINA 604*265*2	"new and short stay'	https://i.leolist.cc/11	[{'url': 'https://i.leolist	https://www.leolist.c	36D / 24 /	110 lbs	yes	Brown	Black	5'4"
20	Incall & Outcall & Online		Upscale , Best service, Best Relaxation Girl, 100% Real Pics	,NEW SEXY UPSCALE SINGAPORE GIRL (Best Service & No Rush), ,KATRINA,, ,5'4, 36D, 110lbs (20).ASIAN Model. Now is the time to, ,pamper yourself!If you are looking for an enjoyable massage (deep tissue, ,or relaxation), ,Just off plane , BEST HOT OIL MASSAGE & HAPPY MENU., ,Are you looking for a high quality deep tissue massage or on the softer more sensual side?, ,Let me help you release your stress and fulfill your deepest desires., ,Come and join me at my clean, spacious,	Yes	2 of the above	None		ad6557955	KATRINA 604*670*5	"first time and new"	https://i.leolist.cc/11	[{'url': 'https://i.leolist	https://www.leolist.c	36D / 24 /	110 lbs	yes	Brown	Black	5'4"
				,New Asian KOREAN LISA, Ready to bring all Her Spiciness to your Engine. Caring, Sweet and Playful Personality.,																

Figure 5 – A sample from one of the datasets gathered

1.1.1 Random Forest

A decision tree is one of the basic, and probably most understandable, classification algorithms (Myles et al., 2004). Given a dataset, with a structure of Figure 5, it evaluates the values within each feature (column), and attempts to find "split points" which would create two groups, each containing a purer grouping of class labels. A "split point" is a value (say *n*) of one of the features, which splits the set of records into two (one larger than *n*, the other smaller than *n*). The "split point" is chosen in such a fashion that the resulting two groups (on each side of the split) is purer than the group before, with purity measured on the basis of the count of each different class label present in the group. A group with 100 rows, 50 labelled as HT and 50 NHT—if split into two groups, both containing 25 HT and 25 NHT—does not isolate the class labels, because the HT/NHT ratio is the same. Thus, the algorithm would not split on this value. Another value might create two groups: one with 30 HT and 10 NHT and the second with 20 HT and 40 NHT, in which case both groups are "purer" than the group before, thereby making this a valid split. The goal is to evaluate every feature and every value of the feature to find the split that yields the optimal purity of groups. The output of this, when graphed, looks like a tree, with each node of the tree representing a split point. A sample decision tree, as output in *WEKA*, is shown in Figure 6, while another visualization type of another decision tree is shown in Figure 7.

While decision trees have a habit of overfitting to their training data (Hastie et al., 2008), the randomization introduced by *Random Forest* solves this problem whereby decisions at each step are made only on a random subset of all the features present. In this fashion, multiple decision trees (i.e., a forest) are built on a dataset simultaneously, and one picked at each step with the use of randomization (hence *Random Forest*). We built models using the *Random Forest* classification algorithm implementation by *TensorFlow*'s SciKit-Learn toolkit (Rokach & Maimon, 2008).

The Random Forest (RF) algorithm was used by Dr. George Weir in his Posit analysis, and was also used by the machine-learning team. In the Random Forest method, classification trees are independently constructed by employing a bootstrap sample of the entire dataset, and then relying on a simple majority vote for predictive purposes, rather than relying on earlier trees to boost the weight of successive trees (Breiman, 2001; Liaw & Wiener, 2002). As a result, rules are made using a series of "*If….Then*" statements. So, "If *this is true,* then, *this*. Otherwise, *this."* For instance, "*If* all previous statements until this point are true, And *this* is also true, Then *this*. Otherwise, *this*." The predicted label of RF's input data is a vote by the trees in the forest, weighted by their probability estimates. Thus, the prediction probabilities of RF can be computed as the mean predicted class probabilities of the trees in the forest, and the class probability of a single tree is the fraction of samples of the same class in a leaf (Pedregosa, Varoquaux, Gramfort, Michel, Thirion, Grisel & Duchesnay, 2011).

To determine accuracy of a forest, the standard approach samples all the records of the input and extracts 90% of the dataset upon which a classification model is built. This classification model is then used to predict the label on the remaining (hidden) 10% of the input data. Once predictions are complete, they are compared against the (hidden) original labels, and the accuracy determined. To ensure representative accuracy, this process is normally repeated 10 times, each time with a different hidden 10% subset.

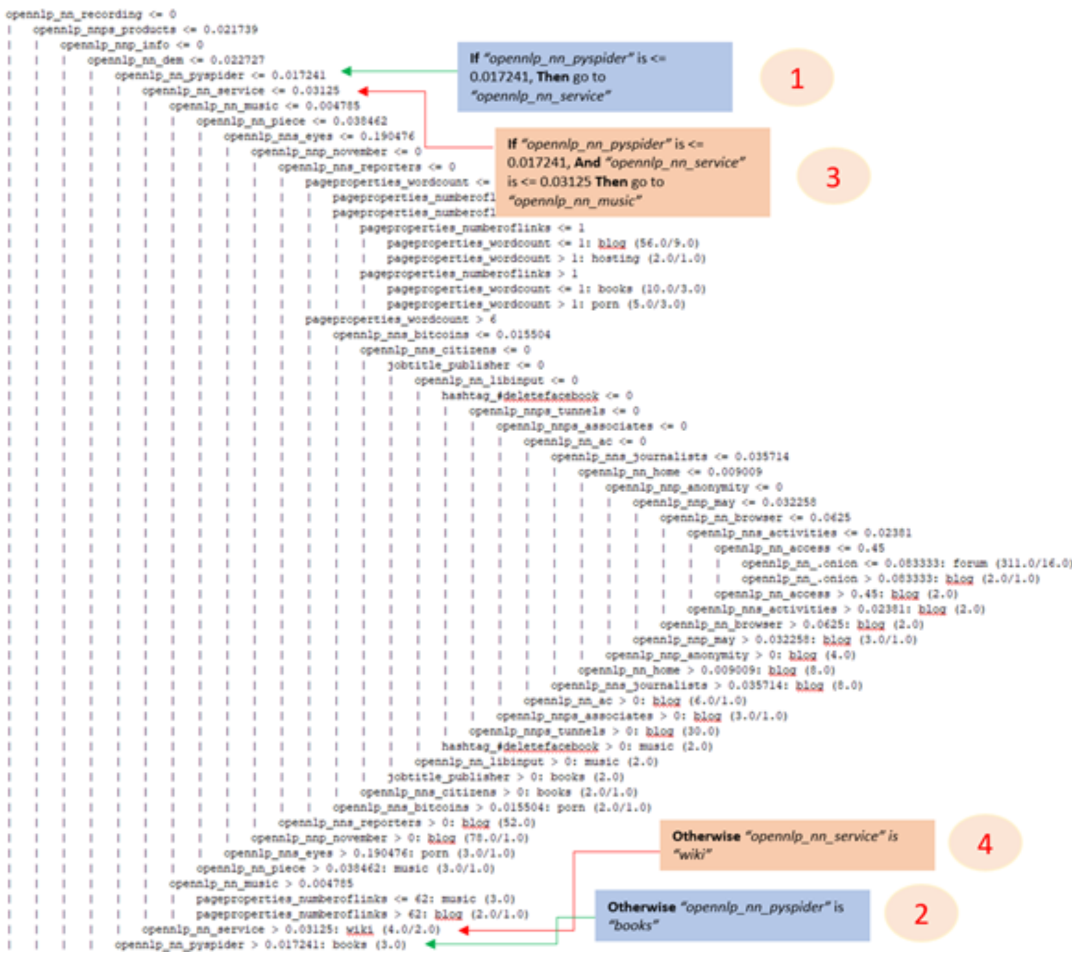


Figure 6 – The decision tree for the model is read from top to bottom and left to right.

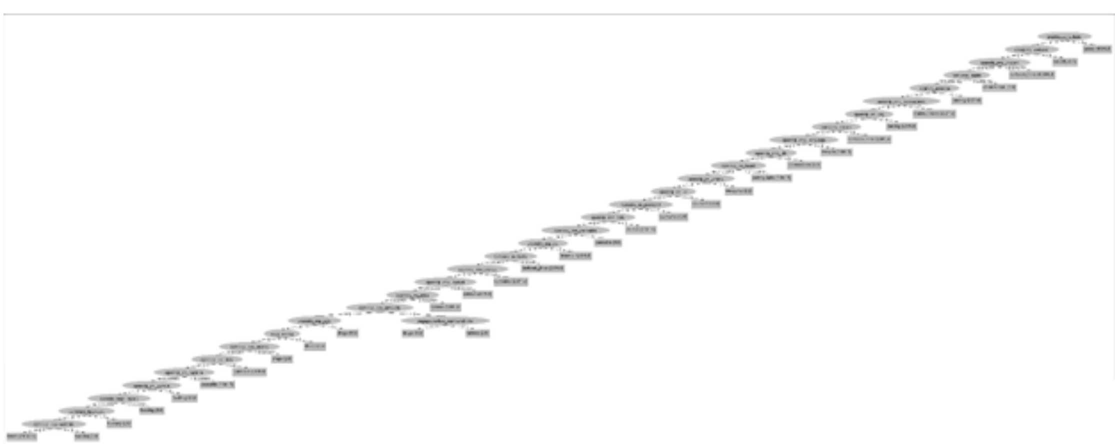


Figure 7 – Another visualization of a decision tree. This tree is very unbalanced which is not at all ideal

1.1.2 Deep Neural Networks

Deep Neural Networks (DNN), are a set of machine learning systems inspired by real-life neural systems. The learning algorithms are designed to excel in pattern recognition and knowledge-based prediction by training sensory data through an artificial network structure of neurons (nodes) and neuronal connections (weights). The network structure is usually constructed with an input layer, one or more hidden layers, and an output layer. Each layer contains multiple nodes, with connections between the nodes in the different layers. As data is fed into this neural system, weights are calculated and repeatedly changed for each connection (Abadi, Barham, Chen, Chen, Davis & Dean, 2016; Kietzmann, McClure & Kriegeskorte, 2019). We used DNN as implemented within *TensorFlow*, originally developed by the Google Brain Team.

To elaborate, *Deep Neural Networks* (DNN) constitute a network of neurons, or nodes, which are organized into rows, each of which represents a layer. Layers are identified as *input*, *hidden*, and *output*, respectively. The *Input Layer* takes information directly from the data, as an input value, and passes it through to the DNN. *Hidden layers* are layers that exist between the *Input Layer* and the *Output Layer*. DNNs can use any number of hidden layers for the network. The greater the number of hidden layers, the deeper the DNN becomes. Multiple hidden layers allow DNNs to solve more complex problems, by preventing the Network from relying on linear separability, as would be the case with decision trees. Where decision trees follow a linear rule pattern, establishing which class values exhibit specific characteristics, DNNs can generate patterns that are not limited to a single dimension. The *Output Layer* displays the various outputs required for the problem. In this case, the class values (HT, NHT) would be presented in the *Output Layer*.

Changing the number of neurons in each layer affects the accuracy. While there is no basic configuration for all datasets, as different layers are ideal for DNNs solving different problems, a higher number of hidden layers is usually found to improve accuracy. Figure 8 demonstrates what a DNN looks like. In this example the DNN is attempting to predict the probability of a specific type of Iris (plant genus). In the *Input Layer*, information about specific characteristics of Irises are exposed to the DNN, after which the *Hidden Layers* attempt to group the characteristics into categories. The *Output Layer* then produces the probability that an Iris will be a certain species. For this project, we used the default DNN setting by having two *Hidden Layers*, each with 500 nodes.

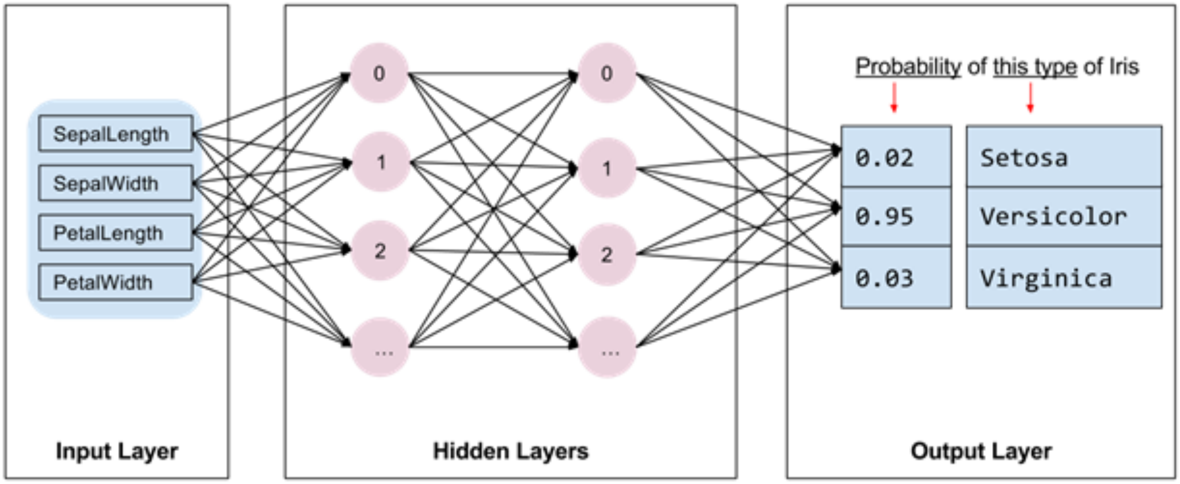


Figure 8 – Neural Nodes of a Deep Neural Network (DNN) used to predict probability of an Iris type

A learning rate determines the rate at which the model converges to the local minima. Usually, a smaller learning rate means that it takes longer for the model to converge at the local minima. Each iteration determines how much closer the model gets to this convergence point. With a larger learning rate, naturally, the model gets closer to this convergence point much more quickly. Unfortunately, speed compromises accuracy. While a larger learning rate gets you closer to the convergence point faster than a smaller learning rate, it also runs the risk of overshooting the convergence point. However, with too small a learning rate, the learning algorithm could take billions of iterations to converge on the local minima, resulting in a model building process that proves too costly in terms of time. With the ideal learning rate, the learning algorithm will be able to get much closer to the convergence point, within a reasonable period of time, thus improving the overall accuracy of the model. Figure 9 shows the difference between using various learning rates.

Figure 9 – A depiction of various learning rates
(Retrieved from – https://www.jeremyjordan.me/nn-learning-rate/)

There are two methods of regularization used in the model training process – L1 & L2. Both methods help improve accuracy but use different methods. The best way to understand how the regularization process works is to envision data plots on a graph, with some plotted in a linear fashion, and others scattered far from this linear progression (Figure 10). To achieve a greater degree of accuracy, we have to prevent overfitting and underfitting. Thus, rather than having our data generate a model of all the data points, and yielding a model that overfits, we must generate a model that adapts to as much of the data as possible and yields the greatest accuracy. Outliers tend to be less informative of data and thus result in decreased accuracy. L1 regularization proves to be robust with respect to these outliers and generates a model that is more consistent with the majority of the data (represented by the red line), while L2 regularization accounts for the outliers, and produces a best fit line with respect to those outliers (represented by the blue line). An L1 regularization technique essentially helps eliminate features that are not important for categorization. Model training parameters can use both methods for regularization and result in a slight improvement to accuracy than just using one regularization technique on its own. Regularization values, for L1 and L2 have been deemed generally optimal at 1.4E-6 and 7.0E-6, respectively.

TensorFlow has a robust community of developers and supports GPU processing. In this project, *TensorFlow* was adopted for processing the data with a DNN. A large data set was initially fed into *TensorFlow*, in order to conduct DNN learning. The DNN results either updated an existing model or created a new model. *TensorFlow* then compared the same data against the constructed DNN model and utilized that model to predict the category for each data entry. We input into *TensorFlow* the pre-processed text along with 128 features extracted from the OpenNLP tool.

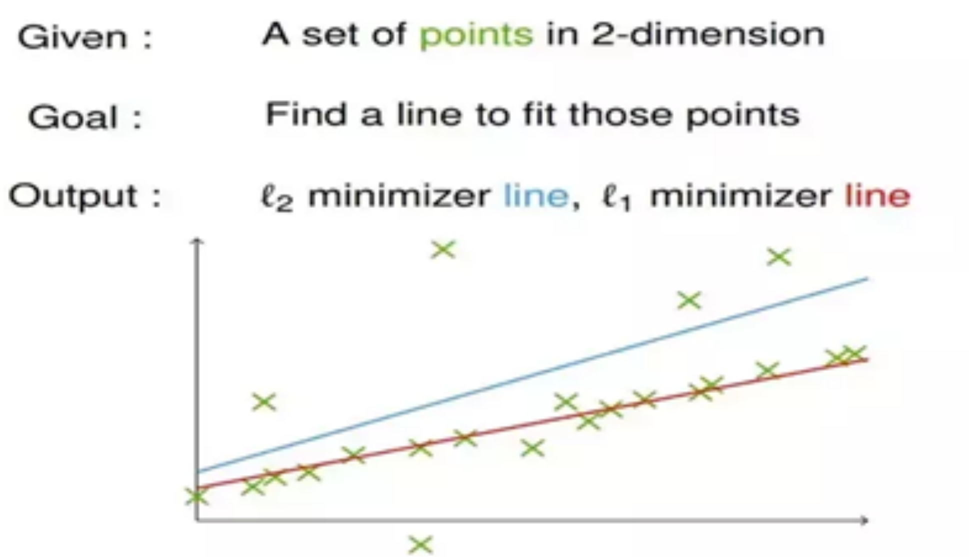


Figure 10 – L1 and L2 regularization
(Retrieved from: https://www.quora.com/Why-is-L1-regularization-better-than-L2-regularization-provided-that-all-Norms-are-equivalent)

For DNN analysis to operate effectively, a larger dataset is desirable (it thrives on big data). The more data that is available for training a model, the better the accuracy should be. However, we found that the individual data files were inconsistent, since they were collected from various online resources, and were formatted in very different ways. The model was evaluated for its accuracy in predicting class values for the “real”, “fake” or “troll”news category. The DNN model used was *TensorFlow*’s *Estimator.DNNClassifier*.

In the early stages of experimentation, we employed *TensorFlow*’s default settings for the parameters pertaining to the number of partitions, epochs, layers, learning rate, and regularization. With respect to regularization, data was partitioned into groups according to the order in which it appeared in the dataset. Thus, if the majority of “fake news” appeared in the beginning of the dataset, it would be difficult to maintain consistent accuracy when conducting X-fold cross validation. To overcome this issue, the data was randomized as it became partitioned. Furthermore, each partition maintained the same data across all X-fold cross validation tests, so that the accuracy of the results could be compared properly.

With *TensorFlow*, epochs refer to the number of times the dataset is processed during training. The greater the number of epochs, the higher the accuracy tends to be. The learning rate determines the rate at which the model converges to the local minima. Usually, a smaller learning rate means it that it would take longer for the model to converge at the local minima. With a larger learning rate, the model would get closer to this convergence point more quickly. The values for these parameters—i.e., the number of partitions, epochs, layers, learning rate, and regularization (L1 & L2)—were then tested to identify an optimal set of parameter values.

TensorFlow next compared the same data against the constructed DNN model and utilized that model to predict the category for each data entry. Many parameter values (for each parameter: number of partitions, epochs, layers, learning rate, and regularization) were then tested to identify an optimal set of parameter values.

1.1.3 LibShortText/LibLinear

LibShortText is an open-source software package, developed by the Machine Learning Group at National Taiwan University. The use of *LibShortText* was recommended in a 2018 paper by William Yang Wang of the University of California at Santa Barbara, wherein he also described (and provided access to) his benchmark LIAR dataset. This LIAR dataset, which included 12,836 statements labeled for their subject matter, situational context, and truthfulness, was broken down into training, validation and test sets, and was accompanied by instructions for automatic fake news detection.

LibShortText is said to be more efficient and more extensible than other generalized text-mining tools allowing for the conversion of short texts into sparse feature vectors, and also for micro- and macro-level error analysis (Yu, Ho, Juan & Lin, 2013). On a typical computer, for example, processing and training with 10 million short texts requires only half an hour or so, whereas *Posit* might require a day or more. *LibShortText* includes an interactive tool for error analysis, and the program’s default options usually work well, without tedious fine-tuning.

For our research project, we built a model using the default settings that came with the *LibShortText* software. We started by running "$ python text-train.py trainfile," which generated a "trainfile.model" for our given "trainfile." Working with this previously built model, we set out to predict the classification labels of the test set, or "trainfile," using the instructions: "$ python text-predict.py -f testfile trainfile.model predict_result," followed by "Option –f" to overwrite the existing model file and predict_result

The initial *LibShortText* results were considerably more encouraging than the *SentiStrength* results, and even a bit better than the *TensorFlow* results, with a credible classification accuracy of 90.2% for the IRA tweets, comparable to the 90.12% classification accuracy yielded by *Posit* for the same dataset. Thus, we decided not to use *SentiStrength* or *Posit*, and instead incorporated *LibShortText* into our processes.

LibLinear is a companion open-source software package to *LibShortText*, developed by the Machine Learning Group at National Taiwan University that developed *LibShortText* (Fan et al, 2008). *LibShortText* is a text analysis program, while *LibLinear* is a classification program. *LibLinear* predicts the accuracy of the classification performed by *LibShortText*, much like *WEKA* predicts the accuracy of the classification performed by *Posit*. Another advantage to *LibLinear* is that is supports incremental and decremental learning, or to express it differently, the addition and removal of data in order to improve optimization and decrease run time. *LibShortText*, on the other hand, does not readily support updating of the model.

1.1.4 SVM

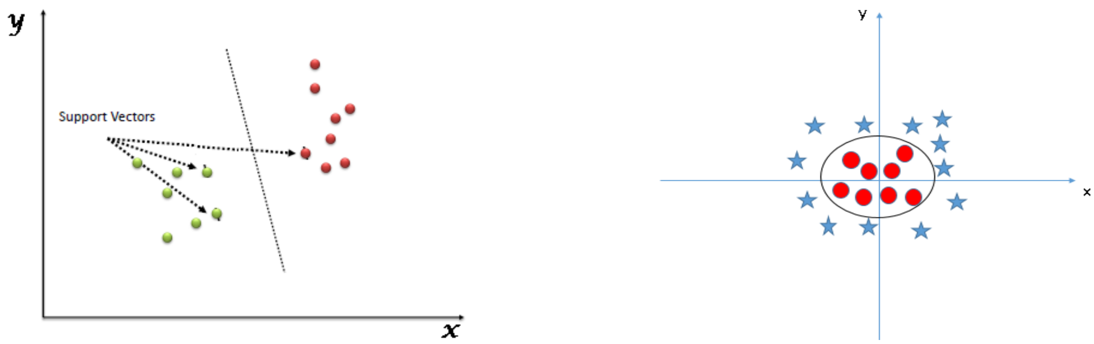
SVM stands for Support Vector Machine. It is an algorithm used for supervised learning in both classification and regression. *SVM* performs better with a limited number of samples and takes less time as compared to other algorithms like *neural networks*. In this algorithm, each data item is plotted as a point in an n-dimensional space (where *n* is the number of features you have), with the value of each feature being the value of a certain coordinate in the *SVM* algorithm (*TMLHB, 2021).* Then classification is accomplished by locating the hyper-plane that distinguishes the two classes as best as possible and has the maximum margin (Figure 11a). The margin is the distance between the nearest data point (of either class) and the hyperplane. The higher the margin, the higher is the robustness of the hyperplane. So, the hyperplane that best separates the two classes and is the one that maximizes the margins from both classes is chosen.

If one of the coordinates is an outlier in the region of the other class, a straight line cannot be used to separate the two classes. In these situations, the *SVM* algorithm offers a feature that allows it to disregard outliers and select the linear hyperplane with the greatest margin. As a result, it can be concluded that *SVM* classification is resistant to outliers.

Sometimes, if the classes are such that no linear hyperplane can distinguish them (Figure 11b), then the *SVM* method employs a technique known as the "kernel" trick. The *SVM* kernel is a function that changes a not separable problem into a separable problem by taking a low-dimensional input space and transforming it to a higher-dimensional space. Hence, non-linear separation problems are solved with *SVM* kernels. *SVM* kernels perform some complex data transformations and use labels and outputs to separate the data.

The model performance can be improved by fine-tuning the parameters for machine learning algorithms. Some of the most significant hyperparameters in *SVM* are *kernel*, *Gamma*, and *C*.

The kernel parameter can be set to "*Linear", "Poly"*, "*rbf"*, etc. The "Gamma" parameter can be used to adjust the gamma value. The "Cost" argument in R controls the "C" value in Python. However, if there is a large data collection, or if the dataset contains more noise, *SVM* does not perform very well.



a) *Linear hyperplane* **b)** *Circular hyperplane*

Figure 11 – Hyperplanes distinguishing two classes
(Retrieved from- https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/)

1.1.5 Naïve Bayes

Naïve Bayes is a supervised machine learning model based on *Bayes' theorem*. Unlike *SVM*, *Naïve Bayes* is used for huge amounts of data. It even works efficiently with data that has millions of records. When it comes to Natural Language Processing tasks like sentimental analysis, it performs very well. It's a simple and quick classification algorithm which classifies binary and multi-class data. Compared to other algorithms, it performs well in multi-class predictions.

Likelihood Class Prior Probability

$$P(c \mid x) = \frac{P(x \mid c)P(c)}{P(x)}$$

Posterior Probability Predictor Prior Probability

$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Figure 12 - Equation of Bayes Theorem
(Retrieved from- https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/)

Bayes' theorem is a probability theorem that works with conditional probabilities (Figure 12). The probability that something will happen if something else has already happened is known as conditional probability. Using prior knowledge, conditional probability can calculate the probability of an event. The *Naïve Bayesian* equation is used to determine the posterior probability for each class. The outcome of prediction is the class with the highest posterior probability. *Naïve Bayes* simplifies the calculation of probabilities for each hypothesis to make it tractable.

In this algorithm, a variable's presence or absence has no bearing on the presence or absence of any other variable. Hence, it is called *Naïve* because it makes some assumptions about certain features being independent of other features and *Bayes* because it depends on the principle of *Bayes'* Theorem. But sometimes, the assumption of independent predictors can act as a flaw in *Naïve Bayes*. In actual life, getting a collection of predictors that are totally independent is nearly impossible. Nonetheless, the technique works very well on data that contradicts this premise.

Naïve Bayes classifiers have a limited number of parameter tuning options, such as *alpha* = 1 for smoothing, *fit prior* = *[True | False]* to learn or not learn class prior probabilities, and a few others. There are three types of *Naïve Bayes* Model - *Gaussian* (assumes normal distribution), *Multinomial* (used for discrete counts) and *Bernoulli* (used if feature vectors are binary). The Gaussian (or Normal) distribution is the easiest to deal with as it just needs to estimate the mean and standard deviation from the training data.

9 Research Results

The ultimate objective of the three research teams was to collaborate in the building of an artificial intelligence tool that was capable of scanning adult classified advertisements and automatically flagging those ads that likely involved human trafficking for the purposes of sexual exploitation. That said, the three teams performed quite different tasks and along the way, producing three different (yet complementary) sets of research results.

As noted in the preceding section on methodological considerations, the research results of the team that was tasked with the building of the customized web-crawling infrastructure, led by Dr. Mandeep Pannu, made significant strides toward the development of this customized crawler. In fact, they found it necessary to develop two unique crawler infrastructures, one for static websites and the other for dynamic websites, and also, found it necessary to recalibrate or "fine-tune" the crawlers (sometimes on a daily basis) due in large part to the steps taken by the site operators to restrict access to the illicit content on their sites and/or to restrict the monitoring activities of Internet hosting platforms, and law enforcement agencies and academic researchers. Nevertheless, the team built out the capacity to scrape large samples from these adult classified ads, and the ability to modify their login credentials and search terms in near-real-time, to keep up with these sites, which, as noted earlier, are constantly shifting targets. We plan to integrate this new crawling infrastructure into TDC, or in the alternative, append it to TDC.

The remainder of this section on our research results will report first on the findings of the qualitative research team, and finally, on the finding of the machine-learning team. In between, we will report on the research findings of Dr. George Weir and his Posit Toolkit, as these were intended to serve as a "bridge" between the work of the qualitative team and the work of the machine-learning team.

9.1 Qualitative research results

Two members of the qualitative research team, Tiana Sharifi and Dr. Barry Cartwright, manually inspected and classified a total of 1,840 advertisements from four different Canadian sites, *leolist* (1,000 ads), *yesbackpage* (682 ads), *locanto* (97 ads) and *megapersonals* (61 ads). This coding was used as a baseline for computer-assisted qualitative analysis in NVivo, for Posit analysis, and also, to cross validate the results of the machine-learning algorithms.

The team focussed primarily on adult sex ads that ostensibly targeted prospective clientele in the Province of British Columbia, although for comparison purposes, the sample of 682 ads from *yesbackpage* included 79 classified sex ads from Toronto and 530 ads from Washington State (mostly Bellingham, Seattle and Tacoma). In this regard, the team was looking for similarity between the wording of multiple advertisements across provincial and international boundaries, as some researchers consider this to be evidence of sex trafficking (Giommoni and Ikwu, 2021). The team also focussed almost exclusively on ads for "female escorts", as the extant literature indicates that most of the individuals who are trafficked for the purposes of sexual exploitation are in fact females (United Nations Office on Drugs and Crime, 2021a).

Table 1 - Key words, key phrases and key symbols-sex trafficking

	Key Words	Key Phrases	Key Symbols
Sweet, compliant, obedient	Sweet; sweet; sweetest; polite; personality; cheerful	Sweet and polite; sweet and polite; super sweet; Great Personality; cheerful mood; nice and friendly	
Availability	24/7; 24hrs; anytime	Always available; always available; available right now; all day; any time; Day or night; all day and night; available day & night; Ready Now; always ready; READY ANY TIME; always ready and available; available right now; Available 24hrs; 24 hours; 7 days a week; 7days a week	
Flexibility	CARDATE; CAR; CARPLAY	INCALL AND OUTCALL AND CARDATE; CAR DATE; Car-Date; Car Dates; CAR DATES; CARPLAY; incall and outcall; Incall and Outcall; INCALL AND OUTCALL: any where;	
No restrictions	openminded	Open Minded; open-minded; open minded; as much as you want; Low restriction; No Restriction; NO RESTRICTION; NO restrictions; no limit; No limit; NO LIMIT; no limits; all service; unlimited fun; MINIMUM RESTRICTIONS; ALL STYLE; ALL STYLE ARE WELCOME; What you like; your own style; all you can play; do as you wish; anything you want; whatever you desire; TRY EVERYTHING; ALL YOUR NEED; everything you want; all your needs	💗 💋 💟
Unprotected sex	Bareback; raw; BAREBACK; Bbj; BBBJ; BBBj; BBBJ+; B-B-B-J;	bBlowj,ob; BARE BACK; With Out Condom; without condom; without condoms; no condom; NO CONDOM; No Condom; raw; No Cover; no cover;	

	b88j; bxxJ; B*BJ; B-B-B-; BB Be J; BEBEFS		
New in town, short visit	town; new; NEW; plane; arrived; arrival; ARRIVE; ARRIVED; ARRIVAL	Catch me while you can; a few days; new in town; COMING WEEKLY; NEW IN TOWN; NEW here in town; new to town; NEW ARRIVED; NEW ARRIVAL; New Arrival; New arrival; new arrived; Newly Arrived; just arrived; short visit; Short-Term; Short Term Only; just off plane; limited time	✈️
Young girl; new girl, fresh and juicy	Baby; baby; BABY; fresh; Girl; GIRL; GURL; Newly; New; new; NEW; YOUNG; yOunG; Youth; tight; Juicy; daddy; Daddy; Pettit; Petite; PETITE; Cherry; student; Student; students; college;	Rope bunny; babygirl; YOUNG GIRL; super tight; Tight Juicy; Sweet and Juicy; Sweet & Juicy; New girl; New girls; Brand new; SUPER TIGHT; SUPER TIGHT FIT; 100% PRETTY; PRETTY GIRL; Pretty Girl; 100% YOUNG; YOUNG girl; young girls; SEXY YOUNG GIRL; REAL YOUNG; First Time; PetiteCute; SEXY YOUNG GIRL; daddy; NeW YouNG Girl; FIRST TIME; 1st time; my first; First Day; first day; first customer; PetiteCute; cutegirl; student girl; school girl; shaved and sweet; sweet, young; SHOWERED & SHAVEN; 19yes; 19yo; 19 Year old; 19 yrs; 19year old; 19yrs old; 20yo; 20/yo; 20 yrs old; 20yrs; 20 years old; 20 year old; 20yrs old.; 19-22; 21 yrs; 21 old yrs; 21 years; 22yo: 22 yrs old; 22 yrs.; 22 year old; 22 years old; 22yrs; 22yrs old	💖 🐱 🍭 💐 🍓 ✨ 🌷 🌸
Low prices	CHEAP	NO DEPOSIT; price is low; PAYMENT ON ARRIVAL; 100 dollars an hour; $100 1 hour; FOR CHEAP; cheap rate; CHEAP~RATE; very cheap rate; low rate	
Multiple girls	girls;	COME AND PICK; Our selling-points; different girls; Different new girls;	
Male Operator		male operator	

During the coding process, Tiana Sharifi and Dr. Cartwright developed a highly elaborated list of key words, key phrases and key symbols, which were shared with the web-crawling team and with the machine-learning team for the purposes of machine-training and calibration (see Table 1). This table (or list) sets out the many and varied expressions that the qualitative team identified as being employed by sex traffickers to describe their victims and the services provided by their victims, wherein they emphasize obedience and compliance (e.g., "sweet and polite" or "cheerful mood"), their unlimited time ranges of availability (e.g., "24/7" or "always available"), flexibility in the many types of venues covered (e.g., "incall and outcall and car date"), the lack of restrictions (e.g., "no limits" or "no restrictions"), the availability of unprotected sex (e.g., "bareback" or "without condom"), their "newness" (e.g., "new to town" or "just arrived"), their youth (e.g., "student" or "young girl"), the low prices (e.g., "price is low" or "very cheap rate"), the availability of multiple "girls" (e.g., "come and pick" or "different new girls"), and finally, the presence of a male operator.

It was also observed during the coding process that there were key words and key phrases that were associated with consensual sex workers. Indeed, of the 1,840 advertisements that were manually inspected and classified by the qualitative research team, only 554 (30%) were deemed to clearly involve sex trafficking (see Table 2). Recall that for a final decision to be made in this regard, the team required at least three key words, key phrases or key symbols associated with sex trafficking to be present in the ad (including in the tagline and/or the description or both).

Table 2 - Presence of sex trafficking indicators by website

	Yes	No	TOTAL
Leolist	380	620	1000
Yesbackpage	162	520	682
Locanto	4	93	97
Megapersonals	8	53	61
TOTAL	554	1284	1840

It can be said that *leolist* (at 38%, or 380 of the 1,000 *leolist* ads manually classified) had the highest percentage of ads deemed to involve sex trafficking, while *yesbackpage* (at 23%, or 162 of the 682 *yesbackpage* ads manually classified) had the second highest. The percentage of sex trafficking ads on *locanto* (at 4%) and *megapersonals* (at 13%) were negligible in comparison.

Table 3 - Key works and key phrases-sex trade workers

	Key Words	Key Phrases
Availability		Available 2-8pm; Available till Midnight; pre booking; Pre-booking; available daytime or evenings; limited availability; make an appointment; 2 hour notice; 9am to 5pm
Flexibility	Uber; etransfer; deposit	no car dates; No Car Dates; etransfer deposit; email money transfer
Restrictions	subscribers; Subscribers; SUBSCRIPTION;	GENTLEMEN'S ONLY; mature respectful gentlemen; NO DEGREDATION; OVER 30 ONLY; respectful clients only; NO AN..AL; no back door; NO BEBEFS; NO BBBF; RESTRICTIONS: NO C1M; NO BB8S; NO BeBeF$; No Greek; NO Greek; NO gr33k; NO GREEEK; NO C0.F; NO $wall*w; SUBSCRIPTION REQUIRED; No dr_gs
Protected sex only	safe; SAFE	no bbfks; safe play; *safe play*; Safe play always; SAFE PLAY ONLY; SAFE SERVICES ONLY; SAFETY PLAY ONLY; EverySafe; No B4re services; NO B*RE SERVICES; SAFE, CLEAN, & COVERED services; COVERED services; with a condom; All services covered; NO FLUID EXCHANGE
Mature, older	mature; woman; MOMMY; MILF	REAL WOMAN;
Upscale	LUXURY; CLASSY; Upscale; UPSCALE; class;	LUXURY SERVICE; UPSCALE LOCATION; Classy mature
Realistic prices		deposit in advance; Deposits are required; no low ballers; NO LOWBALLERS; NO LOWBALLERS OR DISCOUNTS; No low balling; 1 hr - $400; 1.5hr - $550; 2hr - $750: 400/hr outcall; $450 - 1hr; $450 per hour for outcalls; 500 hr; 2/600; $600 - 1.5 hr; $800 - 2hr; $1100 - 3hr; 1800/6hrs; Overnights 2500; NO NEGOTIATIONS;
Experienced, professional	Experience; professional;	

More than just sex	Cuddle; cuddles; companion; companions; romance	dine and date; Dinner Date; dinner dates; events/dinner dates; true companion; Counselling sessions; make a date

On the other hand, there were 1,284 of the 1,840 online ads that were inspected and coded manually that did not appear to involve sex trafficking. While some of these ads did contain features that were similar to sex trafficking ads, they had an insufficient number of sex trafficking features, and these were usually offset by entirely different language. Thus, during the coding process, Tiana Sharifi and Dr. Cartwright developed a highly elaborated list of key words and key phrases that appeared with regularity in the type of online ads posted by sex workers (see Table 3).

In the case of online ads posted by sex trade workers, we see a greater emphasis on restricted availability (e.g., "pre booking" or "make an appointment"), less flexibility in terms of venues or how they expect to be paid (e.g., "etransfer deposit" or "no car dates"), greater restrictions on services provided (e.g., "no Greek" or "no degradation"), more stringent requirements for safe sex practices (e.g., "safe play" or "all services covered"), an advisory not to expect an inexperienced "young girl" (e.g., "mature" or "real woman"), the posting of "realistic" hourly or overnight prices (e.g., "$1,000 – 3hr" or "Overnights 2500"), the offer of more than just sex (e.g., "cuddles" or "counselling sessions"), and an overall emphasis on being experienced., upscale, classy and professional.

With the assistance of team-member Karmvir Padda, the qualitative team in input the 1840 manually classified ads into NVivo, along with 4,260 ads from *leolist* that had not been manually classified by the qualitative team, giving us a larger dataset of 6,100 ads. We then used NVivo to search for and retrieve a number of the key phrases that we had identified during manual inspection and coding, to determine how often they appeared in the larger (6,100 item) dataset. As illustrated in Table 4, the key phrases related to sex trafficking that appeared with the greatest frequency in this larger dataset were "male operator" (n = 1011), "20yo" (n = 984), "student girl" (n = 830), "7 days a week" (n = 648) and "different girls" (n = 635). In all, there were 7,526 mentions of phrases thought to be related to sex trafficking, demonstrating clearly that these websites—especially *leolist* and *yesbackpage*—are indeed marketing venues for this type of criminal activity.

Table 4 - NVivo analysis of key phrases associated with sex trafficking

Key Phrases	Frequency
male operator	1011
20yo	984
student girl	830
7 days a week	648
24 hours	637
different girls	635
TRY EVERYTHING	306
incall and outcall	300
19yo	191
open minded	160
all day	145
Ready Now	137
all day and night	121
just arrived	104
Always available	100
INCALL AND OUTCALL	84
any time	79
my first	78
Open Minded	78
open-minded	76
always available	74
NO RESTRICTION	71
ALL STYLE	67
NO DEPOSIT	65
all service	64
22 year old	59
first customer	57
all your needs	56
19 yrs	53
ALL STYLE ARE WELCOME	53
cheerful mood	52
without condom	51
NO LIMIT	50
INCALL AND OUTCALL AND CARDATE	50
TOTAL	7,526

We used NVivo to break this down into more specific key phrases as they applied to such advertised categories as availability, youth and low prices. As illustrated in Table 5, the key phrases related to availability that appeared with the greatest frequency in this larger dataset were "7 days a week" (n = 648), "24 hours" (n = 637), "all day" (n = 145), "ready now"(n = 137), "all day and all night" (n = 121) and "always available (n = 100). Apparently 24/7 availability was considered to be an important selling point.

Table 5 - NVivo analysis of key phrases regarding availability

Key Phrases	Frequency
7 days a week	648
24 hours	637
all day	145
Ready Now	137
all day and night	121
Always available	100
any time	79
always available	74
Day or night	24
always ready	22
available right now	16
available right now	15
Available 24hrs	9
7days a week	9
available day & night	7
READY ANY TIME	2
always ready and available	1
TOTAL	2046

Another important selling point identified by the qualitative team when it came to possible sexual trafficking was "youthfulness" (Table 6). There were myriad ways in which this was advertised, the most prominent being "20yo" (n = 984), "student girl" (n = 637) and "19yo" (n = 191). Although one might expect to see younger victims being advertised, e.g., 18 years old, 17 years old, or even younger, the team inferred that sex traffickers would avoid language suggesting that their victims might be under the legal age of consent, for fear of attracting the unwanted attention of law enforcement agencies.

Table 6 - NVivo analysis of key phrases regarding "youthfulness"

Key Phrases	Frequency
20yo	984
student girl	830
19yo	191
my first	78
22 year old	59
first customer	57
19 yrs	53
Daddy	35
New girl	31

22yo	26
22 years old	24
21 years	21
Sweet and young	19
Brand new	17
22 yrs old	17
YOUNG GIRL	14
New girls	14
school girl	14
20 years old	14
100% PRETTY	12
Pretty Girl	12
100% YOUNG	12
PetiteCute	12
PetiteCute	12
SUPER TIGHT	11
SEXY YOUNG GIRL	11
SEXY YOUNG GIRL	11
NeW YouNG Girl	11
REAL YOUNG	10
first day	8
20 yrs old	8
SUPER TIGHT FIT	7
FIRST TIME	6
20yrs	6
super tight	5
19 Year old	5
20yrs old	5
young girls	4
First Time	4
Babygirl	3
Sweet and Juicy	3
PRETTY GIRL	3
First Day	3
Cutegirl	3
SHOWERED & SHAVEN	3
22 yrs.	3
Tight Juicy	2
1st time	2
shaved and sweet	2
19yes	2

19yrs old	2
20 year old	2
22yrs old	2
Rope bunny	1
Sweet & Juicy	1
19year old	1
19-22	1
21 yrs	1
21 old yrs	1
TOTAL	2711

Yet another important selling point identified by the qualitative team as being associated with sexual trafficking was typically the lack of restrictions when it came to the range of services provided by their victims (Table 7). As noted above, sex trade workers may stipulate any number of restrictions, such as "with a condom" or "always safe play". Again, there were a myriad ways in which the lack of restrictions was advertised, the most prominent being "TRY EVERYTHING" (n = 306), and various different ways of spelling the term "open minded" (n = 225).

Table 7 - NVivo analysis-lack of restrictions

Key Phrases	Frequency
TRY EVERYTHING	306
open minded	160
Open Minded	78
open-minded	76
NO RESTRICTION	71
ALL STYLE	67
all service	64
all your needs	56
ALL STYLE ARE WELCOME	53
NO LIMIT	50
your own style	30
no limit	27
no limit	27
no limits	21
do as you wish	18
unlimited fun	16
anything you want	14
What you like	11
MINIMUM RESTRICTIONS	9
Low restriction	7
everything you want	5

as much as you want	3
NO restrictions	2
all you can play	2
No Restriction	1
ALL YOUR NEED	1
TOTAL	1175

From this information, two members of the qualitative team, Karmvir Padda and Dr. Barry Cartwright, set out to develop a series of search terms that each contained a sufficient number of indicators to designate a given online ad as a likely case of sex trafficking, as set out below:

- "new arrived" OR "new arrival" AND "$ALL Inclusive" AND "24/7" AND "IN & OUTCALL"
- "non-rushed" or "No Rush" AND "Open-Minded menu" AND "Party Friendly" OR "All you can play"
- "New girl" AND "All Inclusive" OR "Low Restriction" AND "No RUSH GIRL"
- "Never Rushed" OR "Non-rushed" AND "Juicy & Tight" AND "Open minded"
- "new arrived" OR "new arrival" AND "sweet and polite" OR "super sweet" AND "24/7" OR "24hrs" OR "available 24 hours"
- "Student" AND "Party Friendly" AND "Multi-hours available"
- "Student" AND "Young" AND "New"
- "New opening" AND "Open-minded menu" AND "No Rush"
- "24/7 Incall" AND "Free Shower" AND "2 times play"
- "New Girl" AND "Open-Minded" AND "Never Rush service" OR "non-rushed"
- "Open minded" AND "Clubbing Style" AND "Real Young" OR "Student"
- "in or out call" AND "Never Rush" AND "Duo" OR "Threesome" AND "any position"
- "deepthroat" OR "BSMD" AND "No Cover" AND " Mouth explosion"
- "try everything" OR "everything you need" OR "all you can play" AND "young girl" OR "PetiteCute" AND "24/7" OR "24hrs" OR "available 24 hours"
- "new arrived" OR "new arrival" AND "young girl" OR "PetiteCute" AND "24/7" OR "24hrs"
- "great personality" OR "cheerful mood" AND "24 hours" OR "always available" AND "try everything" OR "everything you need" OR "all you can play"
- "young girl" OR "young girls" AND "no restrictions" OR "no limit" AND "sweet and polite" OR "super sweet"
- "open-minded" OR "open minded" OR "no restrictions" AND "24/7" OR "24hrs" OR "available 24hrs" AND "new arrived" OR "new arrival"
- "100% young" OR "100% pretty" AND "no limit" OR "no restrictions" OR "do as you wish" AND "without condom" or "no cover"
- "rope bunny" OR "sweet and juicy" OR "new girl" AND "no limit" OR "no restrictions" AND "new arrived" OR "new arrival"
- "first time" OR "first day" OR "first customer" AND "come and pick" AND "no condom" OR "no cover"

To counterbalance this, we developed a number of search terms that contained combinations of exclusion words or terms that did not indicate the presence of sex trafficking, and that were more likely indicative of language used by consensual sex workers:

- NOT “make an appointment” OR “limited availability”
- NOT “deposit” OR “etransfer” OR “safe play”
- NOT “subscribers” OR “respectful gentlemen”
- NOT “subscribers” OR “dine and date”
- NOT “protected sex only” OR “safe services only”
- NOT “covered services” OR “protected sex only”
- NOT “no car dates” OR “safe play only”

These search terms were provided to the Dr. Frank and the machine-learning team, to assist with the refinement or calibration of the machine-reading technology.

When the larger dataset of 6,100 ads was searched in NVivo, NVivo found, retrieved and coded 3,566 of them as potential cases of sex trafficking (i.e., 59.4% of the dataset). However, we identified one particular ad that appeared 995 times, sometimes with slightly altered wording, that we thought could have skewed the results. If that ad had only been counted once, then the NVivo results (at 42% for sex trafficking) would have been closer to the 30% of ads that were scored as possible trafficking through the manual classification process, and even closer after accounting for the fact that ads from *leolist* comprised 4,260 ads (i.e., 69% of the larger dataset of 6,100 ads) under consideration, and that *leolist* also had more suspected sex trafficking ads than the other websites selected for study.

9.2 The Posit results

For analysis with the Posit Tookit, Dr Weir considered three datasets, those being what were labelled as *leolist_initial*, *leolist_recoded*, and *yesbackpage*. The *leolist_initial* dataset included 1,000 online ads that had been manually classified by the qualitative research team along with 4,260 ads from *leolist* that had not been manually classified by the qualitative team. The hope was that if the Posit classification matched the manual classification for the first 1,000 ads, then the remainder of the *leolist* ads (as well as other ads harvested from similar sites) could be pre-classified in Posit as a means of breaking them into two datasets (human trafficking and not human trafficking) and then using that information to cross-validate the output of Dr. Frank’s machine-learning algorithms. The *leolist_recoded* dataset contained the same ads as the *leolist_initial* dataset but had been re-coded slightly by the qualitative research team with an eye to improving the accuracy of the Posit analysis. The *yesbackpage* dataset was the original (605-ad) dataset, that was coded by the qualitative team prior to the inclusion of 79 more (late-arriving) ads from the Toronto *yesbackpage* site.

Each of the datasets was provided as an individual Excel file with the individual data items listed in rows along with an extensive set of attributes recorded in associated columns. Only specific attributes (containing textual content) were included in the Posit analysis. Therefore, the first step in pre-processing was the removal of extraneous columns for each source file. In step two of the pre-processing, for each case, two columns containing distinctive textual content were retained. These were “title/tagline” and “description”. To enrich the scope for textual analysis, the content in each of these columns was merged to form a composite text, “title/description”. The third column retained from each source data file was the field indicating the manual classification for each data item, as this was required as a basis for testing the effectiveness of the Posit classification.

After reducing the content for each data file and merging the two desired text columns, the third step in data pre-processing was cleaning of the text content. The requirement here was to eliminate aspects in the texts that could obstruct the Posit analysis. Since each text data item is eventually discriminated on the basis of tab separators, extraneous tabs or newlines contained within the text data fields needed to be eliminated. In this cleansing step, they were replaced with a single space character and the three data fields were saved in tab-delimited form.

After pre-processing, the output files were suitable for Posit analysis. Once Posit had completed the analysis stage, the output—containing features and values for each data item—was reformatted into the (arff) layout required for input to WEKA, a popular machine-learning (knowledge acquisition) toolkit. The final steps set the target feature (i.e., the manual classification field), selected a classification algorithm, and attempted to match the pre-defined manual classification assigned by the qualitative research team. The complete sequence of steps from pre-processing through to the Posit results is listed below:

1. Removal of extraneous columns
2. Columns merged to form a composite text
3. Cleansing of the text content
4. Posit analysis
5. Conversion of Posit output to *arff* format
6. Input to WEKA operation
7. Review of classification performance

The algorithms selected for comparison were J48 and Random Forest. Results for each of these algorithms against each of the datasets are set below.

1.1.6 Leolist_initial

When applying J48 to the *leolist_initial* dataset, Posit achieved an overall match of 87.46% compared to the manual classification. This equates to 865 correctly classified instances and 124 incorrectly classified instances from the total of 989 available items (or ads) (see Table 8). Detailed accuracy of the classification match for each class is given in Table 9 and the comparison of match performance is shown as a confusion matrix in Table 10.

Table 8 - Summary of Posit results for leolist_initial/Random Forest

Correctly Classified Instances	87.4621%
Incorrectly Classified Instances	12.5379%
Kappa statistic	0.7461

Mean absolute error	0.1432
Root mean squared error	0.3424
Relative absolute error	28.9631%
Root relative squared error	68.8502%
Total Number of Instances	989

Table 9 - Detailed accuracy of Posit by class (leolist_initial/J48)

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.851	0.106	0.867	0.851	0.859	yes
0.894	0.149	0.881	0.894	0.887	no
Weighted Avg.	0.875	0.130	0.875	0.875	

Table 10 - Confusion matrix for Posit analysis (leolist_initial/J48

a	b	<--classified as
377	66	a = yes
58	488	b = no

When applying *Random Forest* to the *leolist_initial* dataset, Dr Weir's Posit toolkit was able to achieve a greater overall match against the manual classification, in this case 91.6%. This equates to 906 correctly classified instances and 83 incorrectly classified instances from the total of 989 items (Table 11). Detailed accuracy of the classification match for each class is given in Table 12 and the comparison of match performance is shown as a confusion matrix in Table 13.

Table 11 - Summary of Posit results (leolist_initial/Random Forest)

Correctly Classified Instances	91.6077%
Incorrectly Classified Instances	8.3923%
Kappa statistic	0.8297
Mean absolute error	0.1514
Root mean squared error	0.2564
Relative absolute error	30.6069%
Root relative squared error	51.5646%
Total Number of Instances	989

Table 12 - Detailed accuracy of Posit analysis by class (leolist_initial/Random Forest)

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.887	0.060	0.923	0.887	0.904	yes
0.940	0.113	0.911	0.940	0.925	no
Weighted Avg.	0.916	0.089	0.916	0.916	

Table 13 - Confusion matrix for Posit analysis (leolist_initial/Random Forest

a	b	<--classified as
393	50	a = yes
33	513	b = no

1.1.7 Leolist_recoded

The initial version of the *leolist* dataset contained 11 items labelled as "undecided". Following further manual inspection by the qualitative team, these items were classified either as "yes" or "no". This allowed us to consider the effect of this additional insight on the Posit-based classification performance. In this case, the *leolist_recoded* dataset was identical to the *leolist_initial* dataset, save for this change in additional manual classification. Clearly, the expectation was that extra insight offered by further examples of manual classification could enrich the Posit-based analysis and improve the overall classification performance using the supervised-learning approach.

Contrary to expectations, the effect of the manual classification on the previously excluded 11 data items was to impair the overall performance in Posit classification. Using J48, the overall correctly classified instances reached 86.5%, with 865 instances correctly classified of the 1000 data items included. This compares with 87.46% overall correctly classified instances for the *leolist_initial* dataset, with 989 items. A similar change in performance was observed when applying *Random Forest*. In this case, the overall correctly classified instances reached 89.2%, compared with 91.6% for the *leolist_initial* dataset. The *Random Forest* performance details are illustrated in Tables 14-16, below.

Table 14 - Summary of Posit results (leolist_recoded/Random Forest)

Correctly Classified Instances	89.2%
Incorrectly Classified Instances	10.8%
Kappa statistic	0.7692
Mean absolute error	0.1759
Root mean squared error	0.2888
Relative absolute error	37.3164%
Root relative squared error	59.4999%
Total Number of Instances	1000

Table 15 - Detailed accuracy of Posit results (leolist-recoded/Random Forest)

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.839	0.076	0.872	0.839	0.855	yes
0.924	0.161	0.904	0.924	0.914	no
Weighted Avg.	0.892	0.128	0.892	0.892	

Table 16 - Confusion matrix for Posit results (leolist_recoded/Random Forest)

a	b	<--classified as
319	61	a = yes
47	573	b = no

The drop in overall performance after the inclusion of these 11 additional manual classifications may come as a surprise, but there are several possible factors that may explain this result. Firstly, 11 samples are a proportionally small addition (1.1%) to the overall set of items. While this may account for the absence of improvement in classification, how could it explain a drop in performance? Ultimately, the criteria employed by the individuals who undertake the manual classification is certainly different from the criteria applied by the classification algorithms, especially given that the latter depends solely upon quantitative textual features within the sample set. The newly introduced 11 samples may have textual characteristics that are not wholly sympathetic with those identified through Posit analysis of the 989 items in *leolist_initial*. As a result, there may be a weakening of association between specific feature combinations and the corresponding classification.

1.1.8 Yesbackpage

The final dataset addressed via Posit analysis was *yesbackpage*. This is a set of 605 items for which all instances were pre-classified manually. Following the same procedures as outlined above, this dataset was pre-processed before being analyzed in Posit and the resultant feature details being input to WEKA.

In this case, using J48, the overall instances correctly classified by Posit reached 79.5%. Full details for this set of results are provided in Tables 17-19, below.

Table 17 - Summary of Posit results for yesbackpage/J48

Correctly Classified Instances	79.50%
Incorrectly Classified Instances	20.49%
Kappa statistic	0.3752
Mean absolute error	0.2387
Root mean squared error	0.4164
Relative absolute error	70.2355%
Root relative squared error	91.0865%
Total Number of Instances	605

Table 18 - Detailed accuracy of Posit results by class (yesbackpage/J48)

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.481	0.118	0.529	0.481	0.504	yes
0.882	0.519	0.860	0.882	0.871	no

Weighted Avg.	0.795	0.432	0.788	0.795	

Table 19 - Confusion matrix for Posit results (yesbackpage/J48)

a	b	<--classified as
63	68	a = yes
56	418	b = no

When *Random Forest* was applied as the classification algorithm to the *yesbackpage* dataset, the overall correctly classified proportion was slightly better than J48, at 80.82%. The full set of performance details for *yesbackpage* are shown below in Tables 20-22.

Table 20 - Summary of Posit results (yesbackpage/Random Forest)

Correctly Classified Instances	80.82%
Incorrectly Classified Instances	19.17%
Kappa statistic	0.3443
Mean absolute error	0.2424
Root mean squared error	0.3621
Relative absolute error	71.3318%
Root relative squared error	87.9022%
Total Number of Instances	605

Table 21 - Detailed accuracy of Posit results by class (yesbackpage/Random Forest)

TP Rate	FP Rate	Precision	Recall	F-Measure	Class
0.366	0.070	0.593	0.366	0.453	yes
0.930	0.634	0.842	0.930	0.884	no
Weighted Avg.	0.808	0.511	0.788	0.808	

Table 22 - Confusion matrix for Posit analysis (yesbackpage/Random Forest)

a	b	<--classified as
48	83	a = yes
33	441	b = no

While the performance of the Posit-based analysis and classification accuracy were lower for *yesbackpage* than the results for the *leolist* datasets, this may in part be explained in terms of the size of the *yesbackpage* dataset. At 605 samples, *yesbackpage* has around half of the sample size of the *leolist* data. In other words, the analysis has less data to work from when generating the machine classifications. Also, while the Posit results were (upon careful consideration) deemed a bit too inaccurate for the purposes of machine-driven pre-classification of larger datasets, or for cross-validation of the output of Dr. Frank's machine-learning algorithms, the application of Posit to the above datasets demonstrates the promise of automated text-reading and automated data classification. In fact, if the classification accuracy of Posit could be improved to the ~95% range, which might be achievable by programming it to read the UTF-8 coding for emojis, then it might become possible to use Posit to pre-classify very large datasets of online "adult" ads.

9.3 Machine-Learning Results

Advertisements from four platforms, LeoList, YesBackPage, MegaPersonals and Locanto was captured, and analyzed using five machine learning algorithms: LibLinear, Naïve Bayes, Random Forest, SVM, and TensorFlow. Each of these models were built, and the results are presented below.

1.1.9 LeoList

There were 1000 classified advertisements downloaded and manually coded for *LeoList*, with 620 advertisements labelled as "Not Human Trafficking", and 380 identified as "Human Trafficking", resulting in a somewhat unbalanced dataset used for training. Each classification algorithm was run on this data, with 90% (i.e., 900 randomly selected records) of the data used to train the model, and the remaining 10% (i.e., 100 remaining records) used to evaluate the model accuracy. Figure 13 shows the results of each algorithm.

Algorithm	Accuracy	Macro Avg	Weighted Avg
LibLinear	0.84	0.84	0.84
Naïve Bayes	0.83	0.83	0.83
Random Forest	0.87	0.87	0.87
SVM	0.78	0.77	0.78
TensorFlow	0.43	0.34	0.29

Figure 13: Classification accuracies on LeoList

Random Forest achieved the overall best results on this dataset, with an accuracy of 87%, classifying 87 of the 100 testing records accurately, and mis-identifying 13 (7 real Human Trafficking advertisements were misidentified as not related to human trafficking, while 6 advertisements manually identified as Not Human Trafficking were identified as human trafficking by the algorithm). With the exception of TensorFlow, a Deep Neural Network-based algorithm, the other algorithms were able to achieve an accuracy similar to that of Random Forest.

1.1.10 YesBackPage

There were a total of 605 advertisements scraped and manually coded from *YesBackPage*. Of this, 474 were classified as "Not Human Trafficking" after manual review, and the remaining 131 were classified as "Human Trafficking". This was another unbalanced dataset, from a classification point of view. Each classification algorithm was run on this data, with 90% (i.e., 545 randomly selected records) of the data used to train the model, and the remaining 10% (i.e., the remaining 60 records) used to evaluate the model accuracy. Figure 14 shows the results of each algorithm.

Algorithm	Accuracy	Macro Avg	Weighted Avg
LibLinear	1.00	1.00	1.00
Naïve Bayes	1.00	1.00	1.00
Random Forest	0.98	0.97	0.98
SVM	0.85	0.46	0.78
TensorFlow	0.15	0.13	0.14

Figure 14: Classification accuracies on YesBackPage

On this dataset TensorFlow produced very similar and disappointing result as it did on the *LeoList* platform, as it classified all 52 non-human-trafficking advertisements as human trafficking. For some reason, TensorFlow defaulted to a model that said "everything is human trafficking", resulting in the very low classification accuracy of 15%. SVM, although producing an accuracy of 85%, did the opposite, and defaulted to "everything is not human trafficking", and mis-classified all 9 human trafficking cases. Neither of these models were useful (as seen by the more meaningful macro average).

Random Forest, on the other hand, created a model with a 98% accuracy, and mis-classified only a single instance of human trafficking as non-, a false negative. Even more impressive though, both Naïve Bayes and LibLinear were able to produce a model that was fully accurate, able to identify all 52 non-human-trafficking advertisements as such, as well as correctly identify the 9 human trafficking advertisements.

Although these results were impressive, it is unexpected that a model produces 100% accuracy, but in this case three models produced almost-100% accuracy, supporting the idea that these advertisements can be found automatically through machine learning.

1.1.11 MegaPersonals

MegaPersonals was a relatively small website, with only 61 advertisements captured and evaluated for traces of sex trafficking, which, after manual coding, resulted in 49 advertisements categorized as *not* sex trafficking, 8 as sex trafficking, and 4 as undecided. Although on the surface *MegaPersonals* has a relevant subset of its advertisements categorized as sex trafficking, for machine learning, the number of advertisements was simply not sufficient. All models were trained on 90% (54) of the advertisements, leaving only 7 advertisements to test the model on, which can result in a wide variance in accuracy, which is not representative of the details of the model but simply random chance. This type of variance is easily seen with some models producing an accuracy of 0%, others 100% (Table 15).

Algorithm	Accuracy	Macro Avg	Weighted Avg
LibLinear	1.00	1.00	1.00
Naïve Bayes	1.00	1.00	1.00
Random Forest	0.86	0.46	0.79
SVM	1.00	1.00	1.00
TensorFlow	0.00	0.00	0.00

Figure 15: Classification accuracies on MegaPersonals

1.1.12 MegaPersonals – Trained with LeoList Ads

Given that training and testing a model on *MegaPersonals*' 61 advertisements resulted in a model that was not practically useful, a different approach was taken, whereby the models were instead trained on *LeoList*'s 1000 advertisements (620 *non* sex trafficking, 380 sex trafficking), and then the same model used to predict whether *MegaPersonals'* 61 advertisements were sex trafficking or legitimate sex advertisements.

Algorithm	Accuracy	Macro Avg	Weighted Avg
LibLinear	0.81	0.30	0.74
Naïve Bayes	0.81	0.30	0.76
Random Forest	0.83	0.30	0.76
SVM	0.83	0.30	0.76
TensorFlow	0.83	0.30	0.76

Figure 16: Classification accuracies on MegaPersonals with the models trained on LeoList

As seen in Figure 16, model accuracies were quite stable across all models, but even with this approach the macro average was not too impressive. However, these results do demonstrate the viability of training these advertisements on one platform, and then deploying them onto another platform to analyze advertisements.

1.1.13 Locanto

Locanto was another small platform, with only 94 advertisements available for data capture and model building. After manual evaluations, 90 advertisements were labelled as *not* sex trafficking, and only the remaining 4 advertisements were labelled as sex trafficking. The small number of samples for sex trafficking would make model building very difficult, with a single mis-classification resulting in a large change in model accuracy. Especially if 90% of the data is used for training, leaving only 10% (9 records) available for prediction. However, the models were built, and, as seen in Figure 17, four of the models were not usable (predicting all as non sex trafficking – as Naïve Bayes, Random Forest and SVM did – is not practically useful, while the reverse, TensorFlow's approach of predicting everything as sex trafficking is not practically useful either).

Algorithm	Accuracy	Macro Avg	Weighted Avg
LibLinear	0.90	0.47	0.95
Naïve Bayes	1.00	1.00	1.00
Random Forest	1.00	1.00	1.00
SVM	1.00	1.00	1.00
TensorFlow	0.00	0.00	0.00

Figure 17: Classification accuracies on Locanto

1.1.14 Locanto – Trained with LeoList Ads

As with *MegaPersonals*, *Locanto* was too small a dataset to build useful models on, and thus another approach was taken: build the models on the *LeoList* dataset then apply that model to classify the advertisements of *Locanto*. This approach yielded the same set of model accuracies as those built on *Locanto* itself, probably as a result of the small number of actual sex trafficking advertisements found on the platform.

Algorithm	Accuracy	Macro Avg	Weighted Avg
LibLinear	0.90	0.47	0.95
Naïve Bayes	1.00	1.00	1.00
Random Forest	1.00	1.00	1.00
SVM	1.00	1.00	1.00
TensorFlow	0.00	0.00	0.00

Figure 18: Classification accuracies on Locanto with the models trained on LeoList

10 Conclusions

In this study, we have demonstrated the viability of using web-crawling to scrape adult classified ads, as well as the viability of using machine-learning algorithms to read and automatically classify those ads in accordance with their likelihood of involving sex trafficking. We have also demonstrated the promise of both Posit and NVivo when it comes to pre-classifying the ads, and in particular, the possibility of using NVivo to extract the ever-changing key words and key phrases that are associated with the online marketing of victims, so that the machine-learning algorithms can be re-calibrated as required, and their output and classifications can be cross-validated. For machine-learning models, we were successful in creating quite accurate models on *LeoList* and *YesBackPage*, both of which have 620 and 1000 advertisements respectively, however the models built on advertisements captured from the smaller platforms MegaPersonals and Locanto (61 and 94 advertisements respectively) were not successful. In the future, we plan to explore the possibility of programming Posit and NVivo to read UTF-8 coding, as both are text-reading programs, and thus cannot read emojis. However, the UTF-8 coding for emojis can be downloaded readily, and we have reason to believe that either NVivo or Posit or both can be programmed to read UTF-8.

The completion of this project was slowed considerably by the COVID-19 pandemic. While Zoom meetings and Microsoft Team meetings have become commonplace, they are no substitute for being able to meet in person, brainstorm, and work things out. Throughout the entire one-year period, for example, the team that was responsible for developing the customized web-crawler never got to meet in person with the machine-learning team. The same can be said for the qualitative research team, which became dispersed across the country, including the Okanagan, Alberta and Ontario.

Nevertheless, we are poised at this point to integrate our custom-designed web-crawler and our recently re-calibrated machine-learning algorithms to complete the building of an AI tool that will detect and alert to the presence of sex-trafficking content in Canadian online classified ads, and ideally, alert to sex trafficking content on the Internet in general. Over the next two weeks, we are scheduled to meet with several entities, including non-governmental organizations, academic researchers and law enforcement agencies, both Canadian and American, with an eye to mutual cooperation and hopefully obtaining some of the sex trafficking data that they have assembled. We hope to use that additional (and more varied) data for the further testing and cross-validating of our web-scraping and machine-learning algorithms.

We expect—subject to us being able to source some additional (modest) funding—to bring this project to fruition within the next six months or so. From here on in, it is more a question of integrating the web-crawling and machine-reading/machine-learning technology into a cohesive software package, doing a proof-of-concept with one or more Canadian law enforcement agencies, making whatever refinements might be required, and designing a graphic user interface that will permit law enforcement agencies to put this AI tool to good use.

11 References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., & Kudlur, M. (2016, November). *TensorFlow*: A System for Large-Scale Machine Learning. In OSDI (Vol. 16, pp. 265-283).

Alvari, H., Shakarian, P., & Snyder, J. E. K. (2016). A Non-Parametric Learning Approach to Identify Online Human Trafficking. *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*.

Alvari, H., Shakarian, P., & Snyder, J. E. K. (2017). Semi-supervised learning for detecting human trafficking. *Security Informatics*, *6*(1), 1–14. https://doi.org/10.1186/s13388-017-0029-8

Andrews, S., Brewster, B., & Day, T. (2016). Organised Crime and Social Media: Detecting and Corroborating Weak Signals of Human Trafficking Online. In O. Haemmerle, G. Stapleton, & C. Faron-Zucker (Eds.), *Graph-based representation and reasoning. Lecture Notes in Computer Science* (pp. 137–150). Springer. https://doi.org/10.1007/978-3-319-40985-6_11

Arifin, R., Lemuel, Y., & Nte, N. D. (2021). International Legal Instruments in Responding to Human Trafficking. *Lentera Hukum*, *8*(3), 417. https://doi.org/10.19184/ejlh.v8i3.22137

Asencion, M. L. (2017). Classified Websites, Sex Trafficking, and the Law: Problem and Proposal. *Intercultural Human Rights Law Review*, *12*, 227–266.

Bach, J. E., & Litam, S. D. A. (2017). "Kind regards": an examination of one buyer's attempt to purchase a trafficked child for sex. *Journal of Sexual Aggression*, *23*(2), 222–233. https://doi.org/10.1080/13552600.2017.1323124

Barrett, N. A., & Shaw, M. J. (2013). Laws to Combat Sex Trafficking: An Overview of International, National, Provincial and Municipal Laws and their Enforcement.

Blunt, D., & Wolf, A. (2020). Erased: The impact of FOSTA-SESTA and the removal of Backpage on sex workers. *Anti-Trafficking Review*, *14*, 117–121. https://doi.org/10.14197/atr.201220148

Bouché, V., Farrell, A., & Wittmer-Wolfe, D. E. (2018). Challenging the Dominant Frame: The Moderating Impact of Exposure and Knowledge on Perceptions of Sex Trafficking Victimization*. *Social Science Quarterly*, *99*(4), 1283–1302. https://doi.org/10.1111/ssqu.12492

Bounds, D., Delaney, K. R., Julion, W., & Breitenstein, S. (2020). Uncovering Indicators of Commercial Sexual Exploitation. *Journal of Interpersonal Violence*, *35*(23–24), 5607–5623. https://doi.org/10.1177/0886260517723141

Breiman, L. (2001). Random Forests. *Machine Learning*, 45:5-32. http://doi.org/10.1023/A:1010933404324

Bryman, A., Teevan, J. J., & Bell, E. (2009). *Social research methods* (Second Canadian Edition ed.). Toronto: Oxford University Press.

Canadian Alliance for Sex Work Law Reform. (2019). Moving Backwards in the Fight Against Human Trafficking in Canada: An analysis of and response to the report of the Standing Committee on Justice and Human Rights.

Canadian Centre to End Human Trafficking. (2021). Retrieved from https://www.canadiancentretoendhumantrafficking.ca/human-trafficking-trends-in-canada-2019-2020/#:~:text=The%20vast%20majority%20of%20victims,victims%2Fsurvivors%20were%20foreign%20nationals.

Canadian HIV/AIDS Legal Network. (2018). Brief to the House of Commons Standing Committee on Justice and Human Rights on Human Trafficking in Canada.

Choo, K., Jang, J. O., & Choi, K. (2010). Methodological and Ethical Challenges to Conducting Human Trafficking Studies: A Case Study of Korean Trafficking and Smuggling for Sexual Exploitation to the United States. *Women & Criminal Justice*, *20*(1–2), 167–185. https://doi.org/10.1080/08974451003641461

Clancey, A., & Wiseman, A. (2020). Transforming Anti-Trafficking Sentiment into Effective Action.

Cockbain, E., & Bowers, K. (2019). Human trafficking for sex, labour and domestic servitude: how do key trafficking types compare and what are their predictors? *Crime, Law and Social Change*, *72*(1), 9–34. https://doi.org/10.1007/s10611-019-09836-7

Cosentino, G. (2020). From Pizzagate to the Great Replacement: The Globalization of Conspiracy Theories. *Social Media and the Post-Truth World Order*, 59–86. https://doi.org/10.1007/978-3-030-43005-4_3

Cotter, A. (2019). Criminal victimization in Canada, 2019. *Juristat*, 85-002-X, 1-37. Downloaded from: https://www150.statcan.gc.ca/n1/en/pub/85-002-x/2021001/article/00014-eng.pdf?st=gvkvIB0W

De Shalit, A., van der Meulen, E., & Guta, A. (2021). Social service responses to human trafficking: the making of a public health problem. *Culture, Health & Sexuality*, *23*(12), 1717–1732. https://doi.org/10.1080/13691058.2020.1802670

de Vries, I., Kafafian, M., Goggin, K., Bouchard, E., Goldfarb, S., & Farrell, A. (2020). Enhancing the Identification of Commercial Sexual Exploitation Among a Population of High-Risk Youths Using Predictive Regularization Models. *Child Maltreatment*, *25*(3), 318–327. https://doi.org/10.1177/1077559519889178

de Vries, I., & Radford, J. (2021). Identifying online risk markers of hard-to-observe crimes through semi-inductive triangulation: The case of human trafficking in the United States. *The British Journal of Criminology*, 1–20. https://doi.org/10.1093/bjc/azab077

DeLateur, M. (2016). From Craigslist to Backpage.com Conspiracy as Strategy to Prosecute Third-Party Websites for Sex Trafficking. *Santa Clara Law Review*, *56*(3), 531–592.

Di Nicola, A., Baratto, G., Martini, E., Antonopoulos, G. A., Cicaloni, M., Damian, A., Diba, P., Dimitrov, D., Faion, M., Ferrari, V., Ivanova, S., Mancheva, M., Rusev, A., Papanicolaou, G., Toderita, A., & Toma, B. (2017). *Surf and Sound: The Role of the Internet in People Smuggling and Human Trafficking*.

DoCarmo, T. E. (2019). Ethical considerations for studying human trafficking. In J. Winterdyk & J. Jones (Eds.), *The Palgrave International Handbook of Human Trafficking* (pp. 177–194). Springer International Publishing. https://doi.org/10.1007/978-3-319-63058-8_8

Dubrawski, A., Miller, K., Barnes, M., Boecking, B., & Kennedy, E. (2015). Leveraging Publicly Available Data to Discern Patterns of Human-Trafficking Activity. *Journal of Human Trafficking*, *1*(1), 65–85. https://doi.org/10.1080/23322705.2015.1015342

Elg, U. and Ghauri, P. N. (2019). Systematic Data Anaylsis for Multinational Case Study Research. *Sage Research Methods Cases*. Sage Publications, London, 2019.

Esfahani, S. S., Cafarella, M. J., Pouyan, M. B., DeAngelo, G. J., Eneva, E., & Fano, A. E. (2019). Context-specific Language Modeling for Human Trafficking Detection from Online Advertisements. *ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, 1180–1184. https://doi.org/10.18653/V1/P19-1114

Fan, R. E., Chang, K. W., Hsieh, C. J., Wang, X. R., & Lin, C. J. (2008). LIBLINEAR: A library for large linear classification. *the Journal of machine Learning research*, *9*, 1871-1874.

Farrell, A., & de Vries, I. (2019). Measuring the nature and prevalence of human trafficking. In J. Winterdyk & J. Jones (Eds.), *The Palgrave International Handbook of Human Trafficking* (pp. 147–162). Springer International Publishing. https://doi.org/10.1007/978-3-319-63058-8_6

French, K. (2019). *Sexual Exploitation and Trafficking of Youth in Canada: A Prevention and Early Intervention Toolkit for Parents*. Children of the Street Society. Downloaded from: http://www.kristenfrenchcacn.org/wp-content/uploads/2019/08/Parents-Toolkit-on-Sexual-Exploitation-and-Trafficking.pdf

Giommoni, L., & Ikwu, R. (2021). Identifying human trafficking indicators in the UK online sex market. *Trends in Organized Crime*. https://doi.org/10.1007/s12117-021-09431-0

Global Network of Sex Work Projects. (2019). Policy Brief: The Impact of Anti-trafficking Legislation and Initiatives on Sex Workers.

Government of Canada. (2021). *Human Trafficking: Canadian Legislation*. https://doi.org/10.2307/1945872

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, *11*(1), 10-18.

Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome (2008). *The Elements of Statistical Learning (2nd ed.)*. Springer. ISBN 0-387-95284-5.

Hu, R. (2021). Problematizing the Educational Messaging on Sex Trafficking in the US "End-demand" Movement: The (Mis)Representation of Victims and Anti-Sex Work Rhetoric. *Affilia: Feminist Inquiry in Social Work*, 1–18. https://doi.org/10.1177/08861099211058827

Hultgren, M., Persano, J., Jennex, M. E., & Ornatowski, C. (2016). Using Knowledge Management to Assist in Identifying Human Sex Trafficking. *2016 49th Hawaii International Conference on System Sciences (HICSS)*. https://doi.org/10.1109/HICSS.2016.539

Hundman, K., Gowda, T., Kejriwal, M., & Boecking, B. (2018). Always Lurking. *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, 137–143. https://doi.org/10.1145/3278721.3278782

Ibanez, M., & Gazan, R. (2016a). Virtual indicators of sex trafficking to identify potential victims in online advertisements; In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*. https://doi.org/10.1109/ASONAM.2016.7752332

Ibanez, M., & Gazan, R. (2016b). Detecting sex trafficking circuits in the U.S. through analysis of online escort advertisements. *Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2016*, 892–895. https://doi.org/10.1109/ASONAM.2016.7752344

Ibanez, M., & Suthers, D. D. (2014). Detection of domestic human trafficking indicators and movement trends using content available on open internet sources. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 1556–1565. https://doi.org/10.1109/HICSS.2014.200

Ibrahim, D. (2018). Trafficking in persons in Canada, 2016. Canadian Centre for Justice Statistics. *Juristat* 85-005-X. Downloaded from: https://www150.statcan.gc.ca/n1/en/pub/85-005-x/2018001/article/54979-eng.pdf?st=ez8Pg2BI

Ibrahim, D. (2021). Trafficking in persons in Canada, 2019. In *Juristat Bulletin — Quick Fact*.

Inter-Agency Coordination Group against Trafficking of Persons. (2019). *Human Trafficking and Technology: Trends, Challenges and Opportunities*.

Johnson, C. (2019). Human traffickers using online video games to lure children. St Augustine Record, Jan 2019. Downloaded from: https://www.staugustine.com/news/20190129/human-traffickers-using-online-video-games-to-lure-children

Johnson, J. (2018). The Self-Radicalization of White Men: “Fake News” and the Affective Networking of Paranoia. *Communication, Culture and Critique*, *11*(1), 100–115. https://doi.org/10.1093/CCC/TCX014

Kejriwal, M., & Kapoor, R. (2019). Network-theoretic information extraction quality assessment in the human trafficking domain. *Applied Network Science*, *4*(1). https://doi.org/10.1007/s41109-019-0154-z

Kejriwal, M., & Szekely, P. (2017). An Investigative Search Engine for the Human Trafficking Domain. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10588 LNCS, 247–262. https://doi.org/10.1007/978-3-319-68204-4_25

Kennedy, E. (2012). Predictive patterns of sex trafficking online. In *Dietrich College Honors Theses* (Issue April). Carnegie Mellon University.

Keskin, B. B., Bott, G. J., & Freeman, N. K. (2021). Cracking Sex Trafficking: Data Analysis, Pattern Recognition, and Path Prediction. *Production and Operations Management*, *30*(4), 1110–1135. https://doi.org/10.1111/poms.13294

Khodarkovsky, J., Russo, A. N., & Britsch, L. E. (2021). Prosecuting Sex Trafficking Cases in the Wake of the Backpage Takedown and the World of Cryptocurrency. *Department of Justice Journal of Federal Law and Practice*, *69*(3), 101–126.

Kietzmann, T.C., McClure, P., & Kriegeskorte, N. (2019). Deep neural networks in computational neuroscience. In *Oxford Research Encyclopedia of Neuroscience. Oxford University Press.*

Landman, T. (2020). Measuring Modern Slavery: Law, Human Rights, and New Forms of Data. *Human Rights Quarterly*, *42*(2), 303–331. https://doi.org/10.1353/hrq.2020.0019

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, *2*(3), 18-22. https://doi.org/10.4236/oalib.1103108

Mackenzie, K., & Clancey, A. (2020). Im/migrant sex workers, myths and misconceptions SWAN Vancouver Society.

Mapp, S. C. (2020). Domestic Sex Trafficking of Children. In J. Winterdyk & J. Jones (Eds.), *The Palgrave International Handbook of Human Trafficking* (pp. 353–370). Springer International Publishing.

Marcus, A., Snajdr, • Edward, Marcus, A., Snajdr, Á. E., & Jay, J. (2013). Anti-anti-trafficking? Toward critical ethnographies of human trafficking. *Dialect Anthropol*, *37*, 191–194. https://doi.org/10.1007/s10624-013-9307-0

Mei, J., & Frank, R. (2015). Sentiment crawling: Extremist content collection through a sentiment analysis guided webcrawler. *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, August 2015. https://doi.org/10.1145/2808797.2809373

Mensikova, A., & Mattmann, C. A. (2018). Ensemble Sentiment Analysis to Identify Human Trafficking in Web Data. *GTA3*.

Middleton, J. (2020). From the Street Corner to the Digital World: How the Digital Age Impacts Sex Trafficking Detection and Data Collection. In J. Winterdyk & J. Jones (Eds.), *The Palgrave International Handbook of Human Trafficking* (pp. 467–480). Springer International Publishing. https://doi.org/10.1007/978-3-319-63058-8

Millar, H., & O'Doherty, T. (2015). The Palermo Protocol & Canada: The Evolution and Human Rights Impacts of Anti-Trafficking Laws in Canada (2002-2015).

Millar, H., & O'Doherty, T. (2020). Canadian Human Trafficking Prosecutions and Principles of Fundamental Justice: A Contradiction in Terms ?

Millar, H., O'Doherty, T., & Roots, K. (2017). A Formidable Task: Reflections on obtaining legal empirical evidence on human trafficking in Canada. *Anti-Trafficking Review*, *8*, 34–49. https://doi.org/10.14197/atr.20121783

Ministry of Public Safety and Emergency Preparedness. (2019). *National strategy to combat human trafficking 2019-2024*.

Musto, J., Fehrenbacher, A. E., Hoefinger, H., Mai, N., Macioti, P. G., Bennachie, C., Giametta, C., & D'Adamo, K. (2021). Anti-Trafficking in the Time of FOSTA/SESTA: Networked Moral Gentrification and Sexual Humanitarian Creep. *Social Sciences*, *10*(58). https://doi.org/10.3390/socsci10020058

Nagpal, C., Miller, K., Boecking, B., & Dubrawski, A. (2017). An Entity Resolution Approach to Isolate Instances of Human Trafficking Online. *Proceedings of the 3rd Workshop on Noisy User-Generated Text*, 77–84. https://doi.org/10.18653/v1/W17-4411

Naqvi, Z. B. (2020). Critical Perspectives on Trafficked Persons in Canada and the US: Survivors or Perpetrators? *Feminist Legal Studies*, *28*(1), 107–112. https://doi.org/10.1007/s10691-019-09405-2

Native Women's Association of Canada. (2014). *Sexual Exploitation and Trafficking of Aboriginal Women and Girls: Literature Review and Key Informant Interviews*. Downloaded from: https://www.nwac.ca/wp-content/uploads/2015/05/2014_NWAC_Human_Trafficking_and_Sexual_Exploitation_Report.pdf

O'Brien, J. E., & Li, W. (2020). The role of the internet in the grooming, exploitation, and exit of United States domestic minor sex trafficking victims. *Journal of Children and Media*, *14*(2), 187–203. https://doi.org/10.1080/17482798.2019.1688668

Oxman-Martinez, J., Lacroix, M., & Hanley, J. (2005). Victims of trafficking in persons: Perspectives from the Canadian Community Sector.

Parise, O. (2020). Little Pink Flower with a Darker Story to Tell: The Role of Emojis in Online Human Trafficking and Potential FOSTA-SESTA Liability. *University of Miami Race and Social Justice Law Review For*, *11*(1), 52–75.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. T*he Journal of Mchine Learning Research*, *12*, 2825-2830. https://doi.org/10.48550/arXiv.1201.0490

Peterson, M., Robinson, B., & Shih, E. (2019). The New Virtual Crackdown on Sex Workers' Rights: Perspectives from the United States. *Anti-Trafficking Review*, *12*, 189–193. https://doi.org/10.14197/atr.2012191212

Popping, R. (2000). *Computer-assisted text analysis*. London, Thousand Oaks, CA: Sage.

Portnoff, R. S., Huang, D. Y., Doerfler, P., Afroz, S., & McCoy, D. (2017). Backpage and Bitcoin: Uncovering Human Traffickers. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1595–1604. https://doi.org/10.1145/3097983.3098082

Public Safety Canada. (2018). *The Way Forward to End Human Trafficking: National Consultations Discussion Paper*. Downloaded from: https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/wy-frwrd-nd-hmn-trffckng-ppr/wy-frwrd-nd-hmn-trffckng-ppr-en.pdf

Public Safety Canada. (2021). *Sex Trafficking*. https://doi.org/10.4324/9781003124672-5

Raets, S., & Janssens, J. (2021). Trafficking and Technology: Exploring the Role of Digital Communication Technologies in the Belgian Human Trafficking Business. *European Journal on Criminal Policy and Research*, *27*(2), 215–238. https://doi.org/10.1007/s10610-019-09429-z

Reynolds, C. (2021). “Craigslist is Nothing More than an Internet Brothel”: Sex Work and Sex Trafficking in U.S. Newspaper Coverage of Craigslist Sex Forums. *The Journal of Sex Research*, *58*(6), 681–693. https://doi.org/10.1080/00224499.2020.1786662

Richards, L. (1999). *Using NVivo in qualitative research*. London: Sage Publications.

Rokach, L., & Maimon, O. (2008). Data mining with decision trees: theory and applications. *World Scientific Pub Co Inc*. ISBN 978-9812771711.

Rodríguez-López, S. (2018). (De)Constructing Stereotypes: Media Representations, Social Perceptions, and Legal Responses to Human Trafficking. *Journal of Human Trafficking*, *4*(1), 61–72. https://doi.org/10.1080/23322705.2018.1423447

Roots, K. (2019). Human trafficking in Canada as a historical continuation of the 1980s and 1990s panics over youth in sex trade. In J. Winterdyk & J. Jones (Eds.), *The Palgrave International Handbook of Human Trafficking* (pp. 97–113). Springer International Publishing. https://doi.org/10.1007/978-3-319-63058-8_114

Russo, R. A. (2020). Online sex trafficking hysteria: Flawed policies, ignored human rights, and censorship. *Cleveland State Law Review*, *68*(2), 314–345.

Sexual Exploitation Education. (n.d.). Retrieved from https://sexualexploitationeducation.com/resources.

Siegel, D. (2016). *Ethical Concerns in Research on Human Trafficking* (D. Siegel & R. de Wildt (eds.)). Springer International Publishing.

Siller, N., & van Doore, K. E. (2019). Establishing the constituent elements of trafficking in persons: Conceptualizing “transnationality” and “involvement by an organized criminal group.” In J. Winterdyk & J. Jones (Eds.), *The Palgrave International Handbook of Human Trafficking* (pp. 271–283). Springer International Publishing. https://doi.org/10.1007/978-3-319-63058-8_12

Simonson, E. (2021). Semi-Supervised Classification of Social Media Posts: Identifying Sex-Industry Posts to Enable Better Support for Those Experiencing Sex-Trafficking. Massachusetts Institute of Technology.

Siouxsie Q. (2018). Anti-sex-trafficking advocates say new law cripples efforts to save victims. *Rolling Stone*.

Standing Committee of Justice and Human Rights. (2018). *Moving Forward in the Fight Against Human Trafficking in Canada*. Ottawa: House of Commons. Downloaded from: https://www.ourcommons.ca/Content/Committee/421/JUST/Reports/RP10243099/justrp24/justrp24-e.pdf

Tong, E., Zadeh, A., Jones, C., & Morency, L.-P. (2017). Combating Human Trafficking with Deep Multimodal Models. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*.

Tundis, A., Böck, L., Stanilescu, V., & Mühlhäuser, M. (2019). Limits in the data for detecting criminals on social media. *Proceedings of the 14th International Conference on Availability, Reliability and Security*, 1–8. https://doi.org/10.1145/3339252.3341483

United Nations: UN News. (2021). *Traffickers abusing online technology, UN crime prevention agency warns*. https://news.un.org/en/story/2021/10/1104392

Protocol to Prevent, Suppress and Punish Trafficking in Persons Especially Women and Children, supplementing the United Nations Convention against Transnational Organized Crime, (2000).

United Nations Office on Drugs and Crime. (2018a). *Global Report on Trafficking in Persons 2018*.

United Nations Office on Drugs and Crime. (2018b). The International Legal Definition of Trafficking in Persons: Consolidation of research findings and reflection on issues raised.

United Nations Office on Drugs and Crime. (2020a). *Human Trafficking Indicators*.

United Nations Office on Drugs and Crime. (2020b). Impact of the Covid-19 Pandemic on Traffickign in Persons: Preliminary findings and messaging based on rapid stocktaking. https://doi.org/10.21608/geget.2020.100470

United Nations Office on Drugs and Crime. (2021a). Global Report on Trafficking in Persons 2020. In *The Palgrave Encyclopedia of Global Security Studies*. https://doi.org/10.1007/978-3-319-74336-3_221-1

United Nations Office on Drugs and Crime. (2021b). *Human Trafficking FAQs*.

United States Government Accountability Office. (2021). Sex Trafficking: Online Platforms and Federal Prosecutions.

Volodko, A., Cockbain, E., & Kleinberg, B. (2020). "Spotting the signs" of trafficking recruitment online: exploring the characteristics of advertisements targeted at migrant job-seekers. *Trends in Organized Crime*, *23*(1), 7–35. https://doi.org/10.1007/s12117-019-09376-5

Weir, G. R. S. (2007). The posit text profiling toolset. *12th Conference of Pan-Pacific Association of Applied Linguistics*, pp. 106-109.

Weir, G. R. S. (2009). Corpus profiling with the Posit tools. *Proceedings of the 5th Corpus Linguistics Conference,* July 2009. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.159.9606&rep=rep1&type=pdf

Weir, G. R. S., Dos Santos, E., Cartwright, B. & Frank, R. (2016). Positing the problem: Enhancing classification of extremist web content through textual analysis. In *IEEE International Conference on Cybercrime and Computer Forensic*, ICCCF 2016, 67–69, 2016.

Weitzer, R. (2020). The Campaign Against Sex Work in the United States: A Successful Moral Crusade. *Sexuality Research and Social Policy*, *17*(3), 399–414. https://doi.org/10.1007/s13178-019-00404-1

Whitney, J., Jennex, M., Elkins, A., & Frost, E. (2018). Don't Want to Get Caught? Don't Say It: The Use of EMOJIS in Online Human Sex Trafficking Ads. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 4273–4282. https://doi.org/10.24251/HICSS.2018.537

Wijkman, M., & Kleemans, E. (2019). Female offenders of human trafficking and sexual exploitation. *Crime, Law and Social Change*, *72*(1), 53–72. https://doi.org/10.1007/s10611-019-09840-x

Wiltshier, F. (2011). Researching with NVivo. *Qualitative Social Research*, 12(1). https://doi.org/10.17169/fqs-12.1.1628

Yu, H.-F., Ho, C.-H. , Juan, Y.-C., and Lin, C.-J. *LibShortText*: A Library for Short-text Classification and Analysis (2013). Retrieved from https://www.csie.ntu.edu.tw/~cjlin/*LibShortText*/

12 Appendix A: Code to retrieve Dynamic HTML

```
import time
from scrapy_selenium import SeleniumRequest
from scrapy import Selector
import scrapy
from ..items import SecretBenefitsItem
import re

URL_RE = re.compile(r'.+url\("(.+)\"\);')
# username narit6496
# email narit64962@d3ff.com
# password B1gChungus
class SecretbenefitsSpider(scrapy.Spider):
 name = 'secretbenefits'
 allowed_domains = ['www.secretbenfits.com']
 root_url = 'https://www.secretbenefits.com'
 login_url = 'https://www.secretbenefits.com/login'
 search_url = 'https://www.secretbenefits.com/next/search' \
 '?page={page}' \
 '&per_page=45' \
 '&manual_search=false' \
 '&location=' \
 '&geo_country=Canada' \
 '&geo_country_code=CA' \
 '&geo_city=Vancouver' \
 '&geo_state=BC' \
```

```
'&geo_latitude=' \
'&geo_longitude=' \
'&sort_order=Newest' \
'&radius=100' \
'&min_height=48' \
'&max_height=84' \
'&min_age=18' \
'&max_age=49' \
'&has_photos=false' \
'&has_videos=false' \
'&has_verifications=false' \
'&location_range=city'
page = 21

def start_requests(self):
yield SeleniumRequest(
url=self.login_url,
wait_time=3,
callback=self.login,
dont_filter=True,
)

def login(self, response):
time.sleep(3)
driver = response.meta['driver']
email_input = driver.find_element_by_id('emailInput')
email_input.send_keys('narit64962@d3ff.com')
time.sleep(1)
pass_input = driver.find_element_by_id('passwordInput')
pass_input.send_keys('B1gChungus')
time.sleep(1)
login_button = driver.find_element_by_id('loginButton')
login_button.click()
# Wait for page to load
time.sleep(3)

yield SeleniumRequest(url=self.search_url.format(page=self.page),
                       callback=self.parse, dont_filter=True)

def parse(self, response):
time.sleep(3)
driver = response.meta['driver']
profiles = driver.find_elements_by_css_selector('sb-profile-entry')

profiles_requests = []

try:
no_result = driver.find_elements_by_css_selector('.issue.cloud-border.no-results')
if no_result:
return
except Exception as e:
pass

next_page_exists = False
next_button = driver.find_element_by_css_selector('.pagination-action.next')
if next_button:
next_page_exists = True

for profile in profiles:
result = Selector(text=profile.get_attribute('innerHTML'))
link = result.css('a.profile-username::attr(href)').get()
profiles_requests.append(link)
```

```
for link in profiles_requests:
item = self.parse_item(driver, link)
yield item

if next_page_exists:
self.page += 1
yield SeleniumRequest(url=self.search_url.format(page=self.page),
                       callback=self.parse, dont_filter=True)

def parse_item(self, driver, link):
driver.get(self.root_url + link)
time.sleep(5)
profile = driver.find_element_by_css_selector('.profile.ng-star-inserted')
response = Selector(text=profile.get_attribute('innerHTML'))

item = SecretBenefits()
item['link'] = self.root_url + link
item['id'] = link.split('/')[-1]
item['name'] = response.css('div.header__details > h2 > span::text').get()
print(item['name'], ", page=", self.page)
item['tagline'] = response.css('div.profile-tagline::text').get()

city = response.css('span#display-city::text').get()
country = response.css('span#display-country::text').get()

item['location'] = city + ", " + country

sections = response.css('section.mb-4-r.mt-4-r.vapor-bg.round-big.pv-5.ph-5.ng-star-
                       inserted')
for section in sections:
key = ''.join(section.css('h3::text').getall())
if 'About' in key:
item['about'] = section.css('p *::text').get()
elif 'Wants' in key:
item['wants'] = section.css('p *::text').get()

image_urls = []
avatar = response.css('div.avatar')
avatar_photo = avatar.css('div.img.clickable::attr(style)')
if avatar_photo:
avatar_link = avatar_photo.get()
image_urls.append(URL_RE.match(avatar_link).group(1))
photos = response.css('div.slideable-grid__tile')
for photo in photos:
link = photo.css('div.img::attr(style)').get()
image_urls.append(URL_RE.match(link).group(1))

item['image_urls'] = image_urls

attributes = response.css('div.details__fields > div.details__row')
for attribute in attributes:
key = attribute.css('h5::text').get().strip()
if 'Age' in key:
item['age'] = attribute.css('p::text').get().strip()
elif 'Height' in key:
item['height'] = attribute.css('p::text').get().strip()
elif 'Body Type' in key:
item['body_type'] = attribute.css('p::text').get().strip()
elif 'Ethnicity' in key:
item['ethnicity'] = attribute.css('p::text').get().strip()
elif 'Hair Color' in key:
```

```
 item['hair_color'] = attribute.css('p::text').get().strip()
 elif 'Eye Color' in key:
 item['eye_color'] = attribute.css('p::text').get().strip()
 elif 'Piercings' in key:
 item['piercings'] = attribute.css('p::text').get().strip()
 elif 'Tattoos' in key:
 item['tattoos'] = attribute.css('p::text').get().strip()
 elif 'Smokes' in key:
 item['smokes'] = attribute.css('p::text').get().strip()
 elif 'Drinks' in key:
 item['drinks'] = attribute.css('p::text').get().strip()
 elif 'Education' in key:
 item['education'] = attribute.css('p::text').get().strip()
 else:
 print("New key found: ", key)

 return item

class SecretBenefitsItem(scrapy.Item):
 id = scrapy.Field()
                         name = scrapy.Field()
                         tagline = scrapy.Field()
                         about = scrapy.Field()
                         wants = scrapy.Field()
                         image_urls = scrapy.Field()
                         images = scrapy.Field()
                         age = scrapy.Field()
                         height = scrapy.Field()
                         body_type = scrapy.Field()
                         ethnicity = scrapy.Field()
                         hair_color = scrapy.Field()
                         eye_color = scrapy.Field()
                         piercings = scrapy.Field()
                         tattoos = scrapy.Field()
                         smokes = scrapy.Field()
                         drinks = scrapy.Field()
                         education = scrapy.Field()
                         link = scrapy.Field()
                         location = scrapy.Field()
```

13 Appendix B: Code to retrieve Static HTML

```
import requests
from bs4 import BeautifulSoup
import csv

# These proxies below may not work.

proxies = {
 "https" : "http://74.208.187.198:3128",
 "http" : "http://74.208.187.198:3128"
}

f = open("yesbackpage.csv", "w", encoding='UTF8')
writer = csv.writer(f)
```

```
header = ["name", "sex", "age", "orientation", "services", "phone_number",
                        "services_provided_for", "email_address", "body", "location",
                        "post_name", "post_link", "date"]

writer.writerow(header)
url = "https://www.yesbackpage.com/620/posts/8-Adult/122-Female-Escorts/"

data = requests.get(url, proxies=proxies)

soup = BeautifulSoup(data.text,'html.parser')

links = soup.find_all("a", {"class":"posttitle"})

def cfDecodeEmail(encodedString):
 r = int(encodedString[:2],16)
 email = ''.join([chr(int(encodedString[i:i+2], 16) ^ r) for i in range(2,
                        len(encodedString), 2)])
 return email

for link in links:
 # DATA
 post_name = link.text
 post_link = "https://www.yesbackpage.com/" + link["href"]
 services = ""
 name = ""
 sex = ""
 age = ""
 orientation = ""
 services = ""
 phone_number = ""
 services_provided_for = ""
 email_address = ""
 location = ""
 date = ""
 body = ""
 # END OF DATA

 link_data = requests.get(post_link, proxies=proxies)

 soup_link = BeautifulSoup(link_data.text,'html.parser')

 rows = soup_link.find_all("tr")
 for row in rows:
 _name = row.find("td", text="Name")
 if _name:
 name = (_name.find_next_sibling("td").text)[1:].strip()

 _sex = row.find("td", text="Sex")
 if _sex:
 sex = (_sex.find_next_sibling("td").text)[1:].strip()

 _age = row.find("td", text="Age")
 if _age:
 age = (_age.find_next_sibling("td").text)[1:].strip()

 _orientation = row.find("td", text="Sexual Orientation")
 if _orientation:
 orientation = (_orientation.find_next_sibling("td").text)[1:].strip()

 _services = row.find("td", text="Services")
 if _services:
 services = (_services.find_next_sibling("td").text)[1:].strip()
```

```
_phone_number = row.find("td", text="Phone Number")
if _phone_number:
phone_number = (_phone_number.find_next_sibling("td").text)[1:].strip()
_services_provided_for = row.find("td", text="Services Provided For")
if _services_provided_for:
services_provided_for =
                       (_services_provided_for.find_next_sibling("td").text)[1:].stri
                       p()

_email_address = row.find("td", text="Email Address")
if _email_address:
__email_address = _email_address.find_next_sibling("td").find("a")
if "data-cfemail" in str(__email_address):
email_address = cfDecodeEmail(__email_address.get("data-cfemail"))

_location = row.find("td", text="Specific Location")
if _location:
location = (_location.find_next_sibling("td").text)[1:].strip()

for _body in row.find_all("div", {"class":"wrap"}):
body = _body.find_all("p")[1].text

_dates = soup_link.find_all("div")
for _date in _dates:
if "Posted on" in _date.text:
date = _date.text[:(_date.text.find("Expire"))]

data = [name, sex, age, orientation, services, phone_number, services_provided_for,
                       email_address, body, location, post_name, post_link, date]
writer.writerow(data)

print(date)
#print(name)
#print(sex)
#print(age)
#print(orientation)
#print(services)
#print(phone_number)
#print(services_provided_for)
print(email_address)
#print(body)
#print(location)
#print(post_name)
#print(post_link)
```