LINCOLN LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

The Human Trafficking Technology Roadmap: A Targeted Development Strategy for the Department of Homeland Security

July 2019

M.P. Daggett
F.R. Waugh
H.J. Davison Reynolds
C.D. Engholm
J.A. Sloboda
C.E. Heine

EXECUTIVE SUMMARY

Human trafficking is a form of modern-day slavery that involves the use of force, fraud, or coercion for the purposes of involuntary labor and sexual exploitation. It affects tens of million of victims worldwide and generates tens of billions of dollars in illicit profits annually. While agencies across the U.S. Government employ a diverse range of resources to combat human trafficking in the U.S. and abroad, trafficking operations remain challenging to measure, investigate, and interdict. Within the Department of Homeland Security, the Science and Technology Directorate is addressing these challenges by incorporating computational social science research into their counter-human trafficking approach. As part of this approach, the Directorate tasked an interdisciplinary team of national security researchers at the Massachusetts Institute of Technology's Lincoln Laboratory, a federally funded research and development center, to undertake a detailed examination of the human trafficking response across the Homeland Security Enterprise. The first phase of this effort was a government-wide systems analysis of major counter-trafficking thrust areas, including law enforcement and prosecution; public health and emergency medicine; victim services; and policy and legislation. The second phase built on this systems analysis to develop a human trafficking technology roadmap and implementation strategy for the Science and Technology Directorate, which is presented in this document.

The roadmap consists of twenty nine targeted findings and recommendations across five major focus areas: law enforcement interdiction operations; trafficking investigations and prosecutions; information sharing and collaboration; measurement of trafficking; and training. Each recommendation includes a set of programmatic actions and an assessment of operational impact, scope, technological maturity, and development timescale. The recommendations are sequenced into a prioritized and phased implementation strategy that provides a flexible plan of action depending on resource availability. Nine high-priority recommendations focus on short-term, small-scale technology development and implementation, including: developing capabilities that automate time-consuming analytical tasks and reduce workloads for investigators and prosecutors; enhancing financial, telecommunication, and forensic device evidence analysis; establishing repositories of tools, evidence templates, and trafficking signatures to be shared among law enforcement organizations; forming a data standards working group to improve the completeness and consistency of trafficking data collection; and using novel engagement methods to train members of the counter-human trafficking community. Ten recommendations focus on long-term, large-scale, potentially nationally disruptive capabilities, including: employing statistical techniques to quantify the impact of counter-human trafficking efforts; developing sophisticated technologies for determining buyer identities; leveraging advances in secure computing to facilitate information sharing among diverse organizations; and using remote sensing techniques to detect signatures of human trafficking activity globally.

To construct the roadmap, a comprehensive literature review was undertaken to understand the approach, lessons learned, and limitations of previous work in the counter-human trafficking domain. Next, over thirty five stakeholder organizations including federal, state, and local law enforcement and prosecutors; federal and academic researchers; private sector companies; and nongovernmental organizations were interviewed to gain first-hand knowledge of their roles and functions in countering human trafficking. Operations research methodologies were employed to

identify how stakeholders currently use technology and to elicit the attributes of their work that are significantly time- and human-intensive. A subset of stakeholders were assessed in greater detail with multiple interviews and in-person work-domain observations, in order to further deconstruct their investigative processes and analytical workflows. Throughout this process, a comprehensive capability inventory was developed across government, commercial, open-source, and academic sectors. This capability inventory was assessed to determine what technology gaps exist between the desired capabilities of stakeholders and existing capabilities available today. From this technology gap analysis and the stakeholder interviews, findings and recommendations were developed and were adjudicated by an interdisciplinary inter-agency human trafficking advisory group. The resulting technology roadmap represents an informed and actionable research strategy for the Department of Homeland Security to transform capabilities and operations across the Homeland Security Enterprise to more effectively protect victims and bring human traffickers to justice.

TABLE OF CONTENTS

Page

1. PROCESS AND METHODOLOGY

1.1 PURPOSE AND SCOPE

Human trafficking, also known as trafficking in persons, is a vast global problem that respects no boundaries, and whose scope requires a large scale response from thousands of organizations worldwide, including cooperation among and between governments and law enforcement agencies in different countries [1, 2]. Within the United States (U.S.) the U.S. government combats human trafficking through a whole-of-government approach that employs a variety of interdisciplinary organizations and efforts [3, 4]. One such organization is the Department of Homeland Security (DHS), whose components perform many functions to combat trafficking, such as performing federal investigations through the U.S. Immigration and Customs Enforcement's Homeland Security Investigations component; providing awareness and outreach through the DHS Blue Campaign; and enabling intelligence sharing through the Human Smuggling and Trafficking Center and collaborative partnerships with state-level information fusion centers. DHS orchestrates these counter-human trafficking activities through its Homeland Security Enterprise, which incorporates mission partners and stakeholders at global, national, and local levels [5].

Within the DHS Science and Technology Directorate (S&T), an effort was undertaken to advance DHS capabilities to address human trafficking, beginning with two cornerstone applied social science projects: the *Counter-human Trafficking and Modern Slavery Foundational Effort* and the *Human Trafficking System Analysis and Technology Roadmap* [6]. DHS S&T tasked an interdisciplinary team of national security researchers at the Massachusetts Institute of Technology's Lincoln Laboratory (MIT LL) to address the latter project. This document serves as the technology roadmap component of the project, and a companion document addresses the system analysis component [7]. The purpose of the roadmap is to provide recommendations both for *"short-term, small-scale technology for quick implementation, as well as long-term, large-scale nationally disruptive technology,"* in areas including law enforcement; information sensing and collaboration; and development of software and hardware [6]. DHS S&T will use these recommendations to develop a strategy that will directly enable, inform, and assist the development and deployment of technology to operational partners, decision makers, and policy makers across the Homeland Security Enterprise.

Given the scale of the problem of human trafficking, the breadth and depth of the roles of DHS and its partners in the Homeland Security Enterprise, and the level of effort and timeline allocated to this effort, several critical design decisions were made to reduce the scope of study, in order for the roadmap to be thorough and actionable. Some of the key determinations, assumptions, and limitations that underpin this scope reduction are as follows:

Sex trafficking focus: While there are many identified types of human trafficking in the U.S. [8], this roadmap is largely focused on addressing three major types of sex trafficking (outlined in Section 2.2). While labor trafficking continues to be a critical and often underserved issue, we found significantly fewer practitioners working in this area and less data and resources, limiting deeper study. However, it was determined during subsequent review of the

findings of this roadmap with labor trafficking-focused organizations that the majority of the recommendations also address labor trafficking processes and needs.

Domestic operation focus: Given the focus on the aforementioned areas of sex trafficking, we also primarily focused on domestic actors, operations, and coordination. While we included the study of cross-border issues such as immigration and human smuggling as part of our broad stakeholder interview process, there are not findings and recommendations that specifically target these issues. Subsequent review of the findings with federal agencies focused on cross-border issues indicated that the many of the domestically focused recommendations also address limitations and gaps in multinational investigations, including several U.S. bilateral trafficking initiatives with other nations.

Inclusion of the Homeland Security Enterprise: We aimed to look across the entire Homeland Security Enterprise and included a wide range of federal, state, and local partners as part of the study. Additionally, we focused on the interplay between federal and state organizations, operations, and resources, including hybrid organizations such as the 70+ state-owned information fusion centers. The intent of this organizational focus was to ensure that technological capabilities developed and advanced at the federal level would have broad impact and multiple echelons across the Homeland Security Enterprise, yielding higher return on investment.

Technology focused: The findings and recommendations in the roadmap address areas of technological need in order to improve the response to human trafficking. Issues identified during counter-trafficking community interviews that could be addressed with cultural, organizational, and non-technological interventions are not covered or included in this document.

While these decisions served to narrow the scope of this study, many of the findings and recommendations are also more broadly applicable to other non-human trafficking areas of complex investigation and prosecution that involve information discovery, analysis, and fusion of multiple heterogeneous data sources. Similarly, some of the information sharing recommendations are relevant to any group of organizations seeking to gain access to each other's data while also addressing legal, proprietary, privacy, and other concerns.

1.2 STUDY PROCESS

The approach for this roadmap was modeled after similar previous studies performed by MIT LL for other areas of the U.S. Government, which culminated in successful research and implementation strategies. A summary of the process appears in Figure 1.

To begin, a comprehensive literature review was performed to understand the approach, lessons learned, and limitations of previous studies and work in the counter-human trafficking domain. These included studies focused on counter-human trafficking technologies or investigative methods; on marketplace supply and demand; and on measuring illicit behavior and crime using incomplete or inconsistent data. Next, a broad range of over 35 stakeholder counter-human trafficking organizations including federal, state and local law enforcement and prosecutors; federal
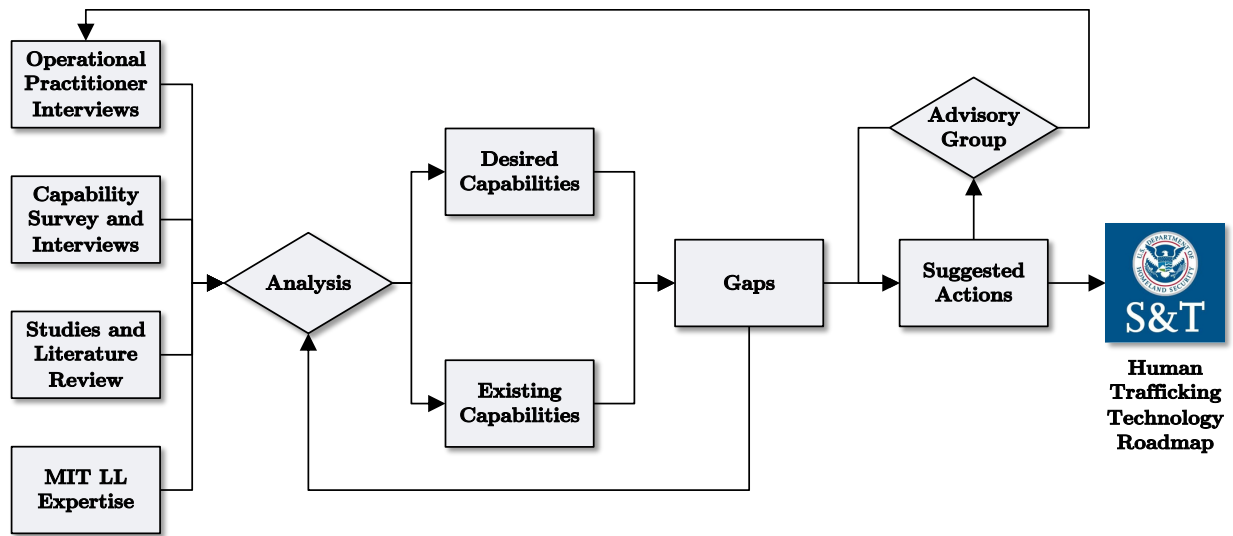
Figure 1. Roadmap development process.

and academic researchers; private sector companies; and nongovernmental organizations (NGOs) were interviewed to gain first-hand knowledge of their roles and functions in countering human trafficking. NGOs included a wide range of organizations such as victim services providers; commercial sex marketplace search platform developers; demand deterrence organizations; and others. A full list of these organizations can be found in Appendix A. Operations research methodologies were employed to identify how some of these stakeholders currently use technology and to elicit the attributes of their work that are significantly time- and human-intensive. A subset of the stakeholders were assessed in greater detail with multiple interviews and in-person work-domain observations, in order to further deconstruct their investigative processes and analytical workflows. Throughout this interview process, a comprehensive capability inventory was developed from across government, commercial, open-source, and academic sectors, which can be found in Appendix B. After the interviews and capability inventories were completed, an analysis was performed to determine the technology gaps that exist between desired capabilities of stakeholders and existing capabilities available today. From this analysis a series of targeted findings and recommendations were developed according to the process detailed in Section 1.3.

As part of this study, an interdisciplinary inter-agency human trafficking advisory group was formed to help identify a broad scope of experienced stakeholders; to provide feedback on the study approach and methodology; and to adjudicate the resulting findings and recommendations. This panel, which consisted of federal government personnel from multiple agencies, representatives from victims service organizations, and academic experts, contributed expertise in law enforcement and prosecution; public health and emergency medicine; victims services; and policy and legislation. The panel met at the midpoint and end of the study, and their feedback and suggested prioritization of recommendations have been incorporated into the implementation strategy detailed in Section 4.2.

3

1.3 FOCUS AREAS AND RECOMMENDATION STRUCTURE

After the interviews, literature reviews, capability surveys, and subsequent analysis described in Section 1.2, the focus of the roadmap coalesced around technology recommendations in five major areas, as follows:

Marketplace interdiction operations: Technologies for augmenting law enforcement in disrupting supply and demand in commercial sex marketplaces, in order to reduce human trafficking.

Investigations and prosecution: Technologies that can assist law enforcement and prosecutorial organizations in investigating complex cases involving the analysis of large quantities of complex heterogeneous data.

Information sharing: Technologies that enable sharing among the counter-human trafficking community while addressing such concerns as victim retraumatization, disclosure of personally identifiable information, and sharing protected data.

Measurement of trafficking: Methods and studies needed to understand supply and demand in commercial sex marketplaces given incomplete and inconsistent data, in order to measure impact of counter-human trafficking efforts.

Training: Technologies that can be used to improve counter-human trafficking awareness training, and training about technology aspects of counter-human trafficking efforts.

For each of these five focus areas, Section 3 presents findings and recommendations using a predefined format containing the following elements:

Recommendation: A short description of the proposed action.

Finding: A summary of the need or gap observed.

Discussion: A detailed description of the issue, with first-hand details obtained through the stakeholder interviews described in Section 1.2.

Suggested actions: A description of actions that DHS S&T can take to implement technologies that address the finding, including specific technologies that should be considered.

Assessment: A categorical estimate of the impact, scope, maturity, and timescale associated with implementing the recommendation as follows:

- **Impact:** *incremental* or *significant*. An estimate of what the perceived operational impact would be if the recommendation were adopted. An incremental improvement advances the state of practice in a positive direction, where-as a significant improvement would likely be "game-changing" with major advances in operational capability.

– **Scope:** *narrow*, *medium*, or *wide*. An estimate of the width of the implementation of the suggested action. For example, making a change within a single organization would be narrow, where-as coordinating many disparate organizations to collaborate on a data sharing protocol would be wide in scope.

– **Maturity:** *low*, *medium*, or *high*. An estimate of the technological readiness of the proposed solutions. For example, for a software solution, high maturity might indicate that robust commercial or open-source implementations are readily available, while low maturity might indicate that further research and development are needed.

– **Timescale:** *short*, *medium*, or *long*. An estimate of the time horizon or how quickly a recommendation could be implemented. A short timescale suggests that the path to an implementation is known or straight-forward, where as a long timescale is something that will need enduring support and advocacy in order for measurable progress to take place or for a capability to come to fruition.

Each of the findings, recommendations, and their assessments were combined with feedback from the advisory group and other stakeholders to form a prioritized and phased development and implementation strategy, which appears in Section 4.

1.4 COMPARISON TO RELATED WORK

Increased public awareness of human trafficking over the last two decades has led researchers at universities, government laboratories, and NGOs to propose technology solutions to the problem. As part of this study, we reviewed the existing literature on technologies for counter-human trafficking and used it as a basis on which to build our recommendations. We found that this literature falls into two broad categories: descriptions of individual technologies that address specific facets of human trafficking, and studies (like this one) that offer a range of technology solutions to many facets.

In the first category, a number of researchers have proposed software systems and algorithms for analyzing data associated with human trafficking. For example, Szekely *et al.* [9] propose a system that crawls online commercial sex advertisements, extracts and correlates that data, applies entity resolution and similarity analytics, and builds a knowledge graph that can be queried by an end user. Alvari *et al.* [10], Li *et al.* [11] and Tong *et al.* [12] describe applications of machine learning to advertisements, respectively to detect individual advertisements likely associated with trafficking and to detect correlations among advertisements to uncover large trafficking networks. This research informs recommendations in Sections 3.1 and 3.2.

Works in the second category offer a broad array of technology solutions, often concentrating on specific technology domains or approaches. For example, Konrad *et al.* [13] propose applying a variety of operations research techniques to identify victims and traffickers, interdict networks, analyze heterogeneous prosecution data, and evaluate policy effectiveness. McKenzie [14] focuses on data mining analytics that financial institutions could implement to uncover transactions associated with trafficking. Among research efforts in this second category, two 2011 reports by researchers at Pacific Northwest National Laboratory (PNNL) stand out for their similarities to our work

[15, 16]. In the first of these reports, the authors used comprehensive interviews of representatives from counter-human trafficking organizations in the greater Seattle Washington area to identify 13 key issues. They then offer largely technology-based recommendations for addressing these issues [15]. In the second report, the authors identify general law enforcement technology needs in areas including data sharing, mobile device exploitation, and audio and video analysis [16]. This roadmap builds on many of the themes in these reports by describing concrete, phased steps to achieving the technology goals.

In addition to research proposing technological solutions to trafficking, a number of studies address the increasing use of technology by traffickers themselves. For example, Latonero *et al.* [17], Boyd *et al.* [18], and Morris [19] all note that the increasing reliance of traffickers on social media, smartphones, and the internet presents opportunities for deploying technologies to discover and disrupt their activities. Although they recognize the role of technology, these studies typically recommend legislative, policy, educational, and other non-technology approaches.

Finally, there is a set of literature that identifies many of the challenges hindering human trafficking research and data analysis. Examples include work by Aghazarm and Laczko [20], Weitzer [21], Farrell [22] and others. We incorporate many of their observations into our discussion of metrics in Section 3.4.

2. HUMAN TRAFFICKING OVERVIEW

2.1 DEFINITIONS AND TERMINOLOGY

Human trafficking, also referred to as trafficking in persons, is a form of modern-day slavery and a crime under international law. In 2016, the International Labour Organization estimated that 40.3 million people worldwide were victims of trafficking [23] and that approximately $150 billion was illegally generated annually due to trafficking, with $99 billion from commercial sexual exploitation [24]. While the term trafficking often implies the movement of persons, the United Nations Office on Drugs and Crime has found that worldwide victims are often trafficked within their region rather than across regions, and within North America 76% of victims were sourced domestically [1]. In the U.S., human trafficking is a crime under federal law, as established by the Trafficking Victim Protection Act (2000). While there are many definitions of trafficking, federal law (22 USC 7102) defines *"severe forms of trafficking in persons"* to include sex trafficking and labor trafficking, as follows:

> **Sex trafficking** is defined as, *"the recruitment, harboring, transportation, provision, obtaining, patronizing, or soliciting of a person for the purpose of a commercial sex act"* in which *"a commercial sex act is induced by force, fraud, or coercion, or in which the person induced to perform such act has not attained 18 years of age"*.

> **Labor trafficking** is defined as, *"the recruitment, harboring, transportation, provision, or obtaining of a person for labor or services, through the use of force, fraud, or coercion for the purpose of subjection to involuntary servitude, peonage, debt bondage, or slavery"*.

Human trafficking is a crime in every state in the U.S., and each state has a range of definitions, statutes, and approaches to countering human trafficking. For the purposes of this document, the federal definition is largely used for simplicity. In addition to the definition of human trafficking, labor trafficking, and sex trafficking, the following sex trafficking-related terms are used throughout the document and are defined as follows:

> **Victim:** Someone involuntarily compelled by force, fraud, or coercion to perform a labor or commercial sex act.

> **Provider:** A person, often a victim, providing a sex service.

> **Buyer:** A person purchasing a sex service, sometimes self-identifying as a "hobbyist" and referred to in popular culture as a "john."

> **Trafficker:** Someone who uses force, fraud, or coercion to obtain labor or commercial sex acts.

> **Facilitator:** An intermediary who promotes and knowingly benefits from connecting buyers and providers, also referred to as a "pimp" or "panderer," and often a trafficker.

Advertising site: An online marketplace with a classified advertising model which facilitates the supply in the commercial sex industry.

Hobby board: An online forum where buyers can rate and review providers.

Supply: The availability of providers.

Demand: The desire of buyers to purchase services from providers.

Commercial sex marketplace: The supply, facilitation, and demand markets for illegal prostitution in the U.S.

Law enforcement: Government agencies that investigate and prosecute criminal behavior, including local, county, and state police; federal agencies such as the FBI; and prosecutors.

Prosecutor: A lawyer responsible for representing the government in criminal trials or for presenting evidence to grand juries; includes county/city prosecutors, district attorneys, state attorneys general and federal prosecutors in the U.S. Attorney General Office.

Decoy: A commercial sex advertisement designed and placed by undercover law enforcement, as opposed to by a provider or facilitator, in order to lure buyers to sting operations.

Sting: A law enforcement operation designed to disrupt the supply, demand, or both.

John school: A diversion program and form of educational intervention aimed at buyers arrested for soliciting commercial sex.

Counter-human trafficking community: Individuals and organizations including law enforcement, other government agencies, and non-governmental organizations that seek to reduce or eliminate trafficking or mitigate its effects.

2.2 TYPES, ROLES, AND ATTRIBUTES

In 2017, the Polaris Project analyzed over 32,000 cases between 2007-2016 from their National Human Trafficking Hotline and text service, and developed a classification system that identified 25 types of human trafficking in the U.S. [8]. Each of these types has a business model; facilitator (trafficker) and provider (victim) profiles; and recruitment and control strategies that enable the facilitation of human trafficking. For the purposes of this document, we largely focus on the following trafficking typologies:

Escort Services: Commercial sex services that take place at a temporary location. Locations can be *out-calls*, where facilitators deliver providers to a buyer's hotel room or residence, or *in-calls* where buyers cycle in and out of a hotel room or other location where the facilitator has housed the provider for an extended stay. These types of operations often move from city to city depending on demand, and are largely popular with online classified advertising websites. These service networks may be run by a single facilitator exploiting one or more providers or by coordinated networks of facilitators affiliated with gangs and organized crime.

Illicit Massage, Health, & Beauty: These are illicit operations, often called massage parlors, which present a facade of a legitimate spa business, but often house providers who are being held against their will inside, unable to leave and sleeping on the massage tables at night. They appear as individual storefronts with unique business names, but are often part of a network of facilities run by multiple owners with centralized command and control of operations. Victims are often smuggled into the U.S. from China and Korea [25, 26] and research estimates there are at least 7,000 such parlors in the U.S. [8].

Outdoor Solicitation: Unlike other facets of commercial sex that have largely moved their soliciting operations to the internet, there are still many traditional operations where facilitators force providers to find buyers in a public setting, such as on a particular block [26] (often referred to as a *track*, *stroll*, or *blade*) or at places like truck stops along highways in more rural areas [27]. Due to the nature of the settings, these tend to be individual facilitators, with no larger network of operations.

Within each of these three typologies, there are a range of similar attributes for the roles in these operations, as follows:

Facilitators / Traffickers: These may be male or female, with an age range of 25-40 and a median age of 31 [28]. They often share similar national or ethnic backgrounds as their victims and use that cultural understanding as leverage for force, fraud, and coercion [8]. They are often involved in other types of criminal operations, such as drug distribution, money laundering, or gang activity.

Providers / Victims: There is a wide demographic of victims across gender, age, ethnicity, and sexual orientation. It is estimated that globally 75% of victims are women and girls, and 25% are children [23]. In the U.S. the majority of sex trafficking involves adult victims with an average age of 19, with exploitation typically beginning in an age range of 15-17 [29]. There is also high correlation with specific familial attributes: studies show that 50%-90% of child victims of child sex trafficking were involved with child welfare services, and a significant portion have a history of running away [30].

Buyers: Sex buyers are primarily men with a diverse distribution of race, sexual orientation, and income. A recent study found that 20.6% of men will buy sex at some point in their lifetime, and 6.2% bought sex in the last year [31]. Additionally, it was found that *high-frequency* buyers account for 75% of the market transactions, even though they are only 25% of the buyer population. Also, buyers spent an average of $100 per transaction and used a diverse range of venues and information sources to connect with providers [31]. Buyers typically use a set of normalized beliefs about the sex industry to justify purchasing sex; for example, that providers are making a life choice or enjoy their work, and that prostitution is a "victimless crime."

2.3 STAGES OF TRAFFICKING

In addition to the prevalence, types, and attributes of sex trafficking, its also important to understand the *cycle* of trafficking and how people become victims and ultimately *survivors*. There are many ways to define the cycle, and individuals in different roles in the trafficking process may describe it with different terminology. For example, the provider's view of the process is often called *"the life"*, where as the facilitator will often refer to it as *"the game"*, and frequent buyers colloquially refer to their role in the process as *"the hobby"*. Combining these points of view, the cycle can be understood in seven different stages, as follows:

i. **Vulnerability:** Individuals who become trafficking victims often have life experiences that make them more vulnerable to exploitation. Common experiences include maltreatment at home, especially sexual abuse, and involvement with the child welfare or juvenile justice systems, especially those with a history of running away. They can also have financial, substance, or mental health problems that make them more at risk, and these can be amplified by issues with lack of identity and self-esteem.

ii. **Identification:** Traffickers look for individuals who have some of the previously mentioned vulnerability attributes. This identification increasingly takes place online through social media and smartphone applications [17–19, 32].

iii. **Recruitment:** The trafficker seeks to exploit the vulnerabilities of the identified person by appealing to a spectrum of unmet intangible needs, such as affection, love, belonging, or protection. The trafficker uses psychological manipulation and financial incentives. Techniques include posing as an intimate partner or benefactor and employing false promises to coerce the recruitment of the individual [30].

iv. **Breaking:** The trafficker gains control over the individual using a range of tactics such as isolation from family, friends, and other support systems; physical confinement; emotional abuse and coercive behaviors; and forms of leverage such as economic indebtedness [29].

v. **Sustainment:** Once a victim is under control, traffickers use a spectrum of threats to maintain the power dynamic, including physical, often non-sexual, abuse. In addition, they create dependencies for food and shelter by taking all of the proceeds earned from sex services rendered. Creating alcohol or drug dependence is also a common method for maintaining control [33].

vi. **Extraction:** The victim is able to break out of the cycle of trafficking through reintroduction to various support systems, such as interactions with friends and family; access to telecommunications services and the internet; or contact with professional organizations such as law enforcement or medical services.

vii. **Rehabilitation:** The victim maintains contact with service organizations to prevent relapse into being trafficked again by the same or another trafficker. The rehabilitation process is long and can be constrained by the availability of resources in a given geographics area, and hampered by restrictions to certain services due to a criminal record.

3. FINDINGS AND RECOMMENDATIONS

This section details the findings, recommendations, and assessments using the methodology and format described in Section 1.3. In all, 29 recommendations are made across the five overarching focus areas of the roadmap. These recommendations and their assessments are summarized in Figure 2 and described in more detail in the remainder of this section.

Focus Area		Recommendation	Impact	Scope	Maturity	Timescale
Marketplace Interdiction Operations	O1	Improve online persona management	Incremental	Medium	Medium	Short
	O2	Determine buyer identities	Significant	Medium	Medium	Medium
	O3	Establish buyer repository	Incremental	Narrow	High	Short
	O4	Improve sting operations	Incremental	Narrow	Medium	Short
Investigations and Prosecution	I1	Create trafficking signature repository	Significant	Narrow	High	Short
	I2	Monitor online vulnerability indicators	Incremental	Medium	Medium	Medium
	I3	Develop advertisement platform federated search	Incremental	Wide	High	Medium
	I4	Develop case knowledge management system	Significant	Wide	High	Long
	I5	Establish data template repository	Significant	Wide	High	Short
	I6	Develop data enrichment capabilities	Significant	Wide	Medium	Medium
	I7	Use geospatial information effectively	Incremental	Narrow	High	Medium
	I8	Improve call record analysis	Significant	Narrow	High	Short
	I9	Enhance financial record analysis	Incremental	Wide	Medium	Short
	I10	Utilize surveillance video analytics	Incremental	Wide	Medium	Long
	I11	Employ speech transcription tools	Significant	Wide	Medium	Medium
Information Sharing	S1	Leverage existing sharing mechanisms	Incremental	Wide	High	Medium
	S2	Encourage novel data sharing	Incremental	Narrow	High	Medium
	S3	Encourage local data sharing	Incremental	Narrow	High	Medium
	S4	Leverage secure computing	Significant	Wide	Low	Long
	S5	Implement P2P marketplaces for victim services	Incremental	Narrow	High	Short
Measurement of Trafficking	M1	Quantify impact of actions	Significant	Medium	High	Long
	M2	Develop geospatially hierarchical methods	Incremental	Medium	High	Medium
	M3	Establish data standards working group	Significant	Wide	High	Short
	M4	Build victim prevalence models	Significant	Medium	Medium	Long
	M5	Build buyer demand models	Incremental	Narrow	Medium	Medium
	M6	Investigate remote sensing	Significant	Wide	Low	Long
Training	T1	Train using serious games	Significant	Wide	High	Long
	T2	Implement training for online investigations	Incremental	Narrow	High	Short
	T3	Implement training for technical evidence gathering	Incremental	Narrow	High	Short

Figure 2. Summary of all recommendations.

3.1 MARKETPLACE INTERDICTION OPERATIONS

One of the major reasons the counter-human trafficking community has had more success in combating sex trafficking as opposed to labor trafficking, is that commercial sex services rely on an open marketplace where demand can find supply. While nowhere near fully transparent, the commercial sex marketplace is observable, measurable, and like any market, able to be disrupted. For example, for an escort services typology there are four major phases of the marketplace, as follows:

   i. **Advertisement:** A facilitator or provider posts a sex service advertisement to a commercial sex advertising site. This posting usually includes a description of the service offered; links to associated websites, locations, and contact information; and other information, such as unique identification numbers on hobby boards. Example advertising sites include: *Adult Look, Adult Search, Bedpage, City X Guide, Eros, Listcrawler, Megapersonals, One BackPage, Plenty of Fish, Skip The Games, Switter*, and others.

  ii. **Negotiation:** A buyer views a posted advertisement and typically sends a text message to the point of contact number listed in the advertisement. The buyer and the provider or facilitator negotiate an in-call or out-call location, a price for the service, and other details.

 iii. **Service:** For in-calls, the buyer will be asked to use the two-call system where the objective is to have the next buyer available on-site to ensure minimum down time, but at the same time ensure the previous buyer has left. The first call will provide general location information (such as the name of a hotel) while the second call (which occurs after arrival at the location) will provide more detailed information such as a room number. Services take place and payment changes hands. A slightly different process is used for out-calls.

 iv. **Review:** Some buyers will rate and review their experience with the provider in graphic detail on forums or hobby boards. Example review sites include: *ECCIE, Erotic Monkey, The Erotic Review, Rubmaps, TNA Board, USA Sex Guide*, and others.

In this construct, the first three phases (advertisement, negotiation, and service) represent roles associated with market supply, and the last three phases (negotiation, service, and review) represents the roles associated with market demand.

3.1.1 Supply-Side Interdiction

Historically, law enforcement has focused primarily on the supply-side of the commercial sex marketplace, using stings to arrest and charge providers under prostitution laws. In recent years, the distinction between prostitution and sex trafficking has become so blurred that many in law enforcement are taking a different approach to countering the supply-side of the commercial sex marketplace. The effect of arresting and prosecuting providers is essentially re-victimization: an arrest record makes it difficult to obtain employment, to secure federal benefits such as subsidized housing, and ultimately to escape *the life*. Stakeholder interviews (see Appendix A) indicated that the common supply-side interdiction operations are as follows:

Facilitator stings: The objective of these operations is to gather evidence to arrest a facilitator or trafficker. Law enforcement agents pose as buyers and respond to advertisements to set up an out-call meeting. At the location, law enforcement surveillance teams are watching to determine if the facilitator has arrived with the provider. Often real-time intelligence-gathering takes place to identify the owner of the vehicle and other information pertinent to the operation. If the sting moves forward, a team meets the provider to discuss their situation and, if possible, to gather evidence by gaining consent to search their phone. At the same time, another team detains the driver and gathers additional information before any arrests are made. The intent of these operations is to gather evidence and intelligence on the commercial sex operation, to determine whether trafficking is present, and to open a non-punitive dialog with the provider, as first connections can be critical to establishing a collaborative relationship as a case against a trafficker is built and moves toward trial.

Counter-recruitment operations: These operations target traffickers or facilitators during the identification and recruitment stages of the trafficking cycle. Law enforcement undercover officers create social media personas that express many of the vulnerability indicators outlined in Section 2.3 combined with common vernacular associated with *the life* in order to entice facilitators to contact them. After initial contact, the facilitator tries to groom and recruit the undercover officer; during this time, the agent collects transcripts as evidence. Once the undercover officer has sufficient evidence to arrest, a meeting is arranged that requires the trafficker or facilitator to travel to the local jurisdiction, where they are arrested upon arrival. An advantage of this approach is that it targets individuals seeking to recruit vulnerable populations, especially minors, without involving actual victims. Such operations are heavily dependent on the statutes in place in a given geographic area. For example, in some states an individual can be arrested for *pandering* an undercover officer; the practice of procuring a person to be used for commercial sex; but not for attempting to *pimp* an undercover officer; such as asking for or receiving money in exchange for soliciting for a person to be used for commercial sex.

Platform takedowns: This approach targets websites that knowingly facilitate the commercial sex marketplace, in order to disrupt the marketplace. The most notable example is the seizure of *Backpage.com* by the United States Department of Justice in April 2018. Backpage was the second largest classified advertisement listing service on the internet in the U.S. after Craigslist, and the largest provider of adult service classified advertisements. The National Center for Missing and Exploited Children reported in 2014 that the majority of child sex trafficking cases it handled involved advertisements on Backpage [34]. While the seizure of Backpage led initially to a dramatic reduction in advertisement activity, total advertisement volume returned to pre-seizure levels within a year as new sites filled the void [35]. An additional downside to this disruption is that Backpage was a U.S.-based corporation that responded to judicial orders for records, while many of the sites that replaced Backpage are largely hosted in foreign countries outside the jurisdiction of the U.S. justice system.

3.1.2 Demand-Side Interdiction

Historically, the demand-side of the commercial sex marketplace has rarely been the focus of law enforcement operations, but in the recent years there has been a dramatic change in approaches toward targeting buyers. Stakeholder interviews (see Appendix A) identified the following common demand-side interdiction operations:

Buyer stings: These operations target buyers knowingly prepared to engage in commercial sex services. A law enforcement agent posts decoy sex advertisements and poses as a provider when transacting with the targeted buyer via text message or voice phone call. Specialized telecommunication platforms allow law enforcement to call and text buyers from their work phones while protecting their real identities and the origin of the calls and messages. The undercover officer and the buyer negotiate an in-call service at a location such as a hotel room; upon arrival, the buyer is immediately detained and brought to another room for processing. Often, these decoy advertisements draw so many buyers that the operations are scaled down to handle the volume, and use a two-call scheduling system. These operations can yield 20-30 arrested buyers a day with a modest complement of law enforcement agents.

Deterrence platforms: Several NGOs have developed buyer deterrence platforms that automate elements of buyer sting operations. One such platform combines a text message back-end with an autonomous chat bot trained on conversations between sex buyers and undercover agents. When a buyer sends a text message to a number from a decoy sex advertisement, the chat bot engages with the most relevant response from the transcripts it has been trained on. Subsequent interaction may lead to negotiating a service, price, location, and possibly establishing that the buyer is aware the provider is likely a minor. Once this phase is reached, the chat bot sends a deterrence message informing the buyer that purchasing sex is a crime and then provides internet links to educational resources. Several days later local police send a message to the buyer informing them their number is now known to law enforcement and reminding them of possible arrests and fines. This system uses the threat of law enforcement action to deter buyers from seeking commercial sex in the future [36]. A similar system uses a distributed network of human operators who respond to the buyers responding to decoy advertisements via a cloud-based text message and call center back-end. Its intent is to have one-on-one conversations with buyers to raise their awareness of facets of human trafficking and to change their attitude towards purchasing sex [37].

Education and awareness: Over the past decade, several programs have been developed in the U.S. to educate sex buyers arrested in sting operations about sex trafficking. The intent is to reduce recidivism by changing perceptions and beliefs about engaging in commercial sex. In some jurisdictions, training completion and fine payment may be offered as an alternative to prosecution. In some areas, attending such training is often a condition of probation.

All of these operations attempt to affect the demand for commercial sex; however, little research has been done to date to assess their effectiveness over time.

3.1.3 Recommendations

Recommendation O-1: Improve online persona management

Finding: The use of social media personas is widespread in supply and demand interdiction operations. These personas are time-intensive to create and maintain, and many agencies use poor operational security practices that could compromise undercover law enforcement assets.

Discussion: Social media is commonly used by facilitators to look for vulnerable individuals to connect with in hopes of getting them to work for them. Additionally, with increased pressure on classified advertising sites, social media is also leveraged as a solicitation mechanism to connect buyers and providers. For these reasons, law enforcement has an increasing presence on these platforms. To infiltrate the social networks of some of these actors, officers need profiles, or personas, that appear similar to those of other members of the social network. To do this, undercover officers create social media profiles with pictures, posts, and other information to build out a history of a fake persona. Once that persona is accepted into the social network, they often try to friend as many other profiles as they can, in order to collect as much as intelligence as possible [38].

Creating and maintaining multiple personas is time-intensive and requires constant bookkeeping. In addition, officers often manage these personas from department computers, which have identifying information such as IP addresses and can also be associated with other social media accounts used from the same computer, such as personal accounts. Personas, and sometimes entire investigations, can be compromised if this information becomes visible to the targeted individuals or network. For example, Facebook's *People You May Know* feature [39], has been responsible for outing several undercover officers to their suspects, without the officers realizing that Facebook recommended their personal profile to their targets.

Suggested Actions: Develop an integrated system to address persona management in two ways. First, design a capability to semi-automatically generate social media personas based on user-defined attributes such as race, age, gender, affiliation, location, and other information, and a generate full persona with all required account information including pictures and a non-attributable email address. This email address is used to forward to an official email account, and can be used to alert law enforcement more quickly to activity of interest, across a range of personas, while also masking identifying information from the social media service. Second, a virtual computing environment should be leveraged, along with a "managed attribution" or virtual private network service, which would allow an officer to use a virtual computer hosted on a cloud service provider, that does not have IP addresses and other identifying information associated with a law enforcement organization. Lastly, training would be required to ensure law enforcement personnel do not use the virtual persona environment with any other accounts or internet activity that could uniquely identify them or their organization. Many of these component technologies and services exist, but would need to be integrated into a turn-key system for law enforcement, including features like a dashboard interface to display the status of multiple personas at a single glance.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Medium	Medium	Short

Recommendation O-2: Determine buyer identities

Finding: Law enforcement would prefer to prioritize the targeting of high-frequency buyers during counter-demand operations, but positively identifying buyers is difficult.

Discussion: Law enforcement engaging in counter-demand operations seek to have the biggest impact possible by interdicting and deterring high-frequency buyers, as opposed to first-time or infrequent buyers. Recent studies suggest that high-frequency buyers account for 75% of the market transactions, even though they are only 25% of the buyer population [31]. The challenge for law enforcement is to identify the frequency of a buyer before a sting or counter-demand operation takes place. High-frequency buyers also are known to frequent review forums and hobby boards, where they engage in a community of like-minded buyers, and are often candid about their activities due to the perceived belief of anonymity offered by the review or board sites, many of which are hosted on foreign servers outside of U.S. legal jurisdiction.

Suggested Actions: Research technologies to monitor hobby boards and profile high-frequency buyers, with the intent of developing signatures that could be matched with activity on the open internet, leading to possible identification of the buyer. Authorship analysis and language-based modeling has been used previously to characterize other illicit online activities [40], and could be adapted for sex buyer profiling, in order to create language-based fingerprints for individual buyers. The same techniques could be used to analyze other internet forums, websites, and social media to look for potential similarities to known buyer language fingerprints. This type of language-based persona linking across multiple forums or networks is an active area of research, with promising potential [11, 41].

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Medium	Medium	Medium

Recommendation O-3: Establish buyer repository

Finding: No nation-wide repository exists for data collected during demand deterrence operations, limiting the impact of such operations to target high-frequency sex buyers

Discussion: Across the U.S., law enforcement agencies run counter-demand operations targeting buyers, without prior knowledge regarding the history of a buyer at the time they respond to a decoy advertisement – for example, is the person who is responding to the advertisement an infrequent buyer or a high-frequency buyer. Local law enforcement agencies often keep track of the phone numbers that have previously responded to decoy advertisements, but those numbers are not shared with other agencies. In addition to law enforcement operations, multiple buyer demand deterrence platforms collect potential buyer phone numbers, but have no mechanism to make these numbers more widely available. The inability of law enforcement and NGOs to store and share phone numbers and other attributes of buyers known to respond to advertisements limits the long-term impact of counter-demand operations. For example, high-frequency buyers could be purchasing locally where they reside as well as when on business travel. If the buyer is in another area, sharing that buyer's history with local law enforcement, can provide operational value.

Suggested Actions: Develop a federated national-level sex buyer repository that allows storing, searching, and sharing of data collected during buyer stings and from demand deterrence platforms, such as phone numbers that replied to decoy advertisements. Authorized users have the ability to both query and submit phone and other information to the repository, and the application should be mobile friendly in order to support time-critical tactical operations. The technology to implement such a database is readily available and should be developed by an NGO, to minimize or avoid federal policy and governance issues.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Narrow	High	Short

Recommendation O-4: Improve sting operations

Finding: Buyer sting operations are time and personnel intensive, specifically the development and posting decoy advertisements and negotiation with buyers.

Discussion: Sting operations are an important tool in countering demand, but can significantly burden law enforcement agencies when employed effectively. Operations in busy marketplaces can yield up to 20-30 buyers per operation, which requires several undercover officers to be in constant negotiation with buyers via text and voice, sometimes juggling up to

ten different unattributed cell phones [38]. Law enforcement focused call and text platforms exist that generate phone numbers in an area code of choice and route texts and calls to an officer's own phone. This eases some of the technical burden, but the human-intensive negotiation process is still required.

Suggested Actions: Develop an integrated capability for workflow automation to reduce the time- and human-intensive nature of buyer sting operations. Explore partnering with demand deterrence platform NGOs to cross-leverage technology of mutual benefit in order to scale operations. For example, chat bots could be employed to field some of the initial steps of buyer negotiation, and alert undercover officers at the point at which a human-operator must take over the final steps of negotiation. The system should have an integrated scheduling system to provide a sequence of buyers, and a dashboard showing negotiation and arrival status to make sting operations more efficient and effective.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Narrow	Medium	Short

3.2 INVESTIGATIONS AND PROSECUTION

A key element of combating human trafficking is bringing to justice those who knowingly facilitate and benefit from it. More recently this includes those who support the marketplace that allows it to flourish (see Section 3.1). This section describes processes and capabilities that are currently used to investigate and prosecute traffickers, and provides findings and recommendations for technologies that can enhance these processes. A wide range of approaches are used to investigate and prosecute human trafficking and related crimes, which cannot be holistically accounted for in this document. However, during interviews with law enforcement and prosecutorial organizations (see Appendix A), several key attributes of these approaches emerged which can be significant factors regarding technology employment, adoption, and impact, as follows:

Personnel and resources: There is a wide disparity in investigative and prosecutorial resources across the federal government, states, cities, and towns. For example, some states have dedicated human trafficking response teams at the attorney general or district attorney-level that both investigate and prosecute trafficking, while other areas may only have a handful of state or local police officers assigned to cover investigations across an entire state or city. Several states also have regional human trafficking task forces; grant-funded multidisciplinary anti-trafficking teams that pool prosecutors, officers, victim service providers, and other groups into a coordinated, collaborative organization. This disparity of resources hinders development of a more uniform national-level response to countering human trafficking.

Targeted approaches: Some states use a "victim-centered" or "victim-informed" approach that emphases identifying victims and establishing rapport early-on in order to have more long-term success at trial. Other states are more "facilitator-centric," using tactical counter-supply operations to yield lower-level trafficking-related sentences and get more facilitators incarcerated, even if the lower-level offense results in a shorter sentence than could be achieved from a full trafficking prosecution. Some states take a more "buyer-centric" approach by targeting commercial sex purchasers and by investing in robust deterrence and compulsory john school programs. Strategic approach decisions vary greatly by region: some elected officials run for office on platforms that address sexual violence and exploitation while others seek to de-fund investigative units in response to anti-police sentiment.

Jurisdiction and statutes: While the federal trafficking statute of *force, fraud, or coercion* is clear and well known, the evidentiary burden at trial can be quite high and the timeline long to yield a successful federal prosecution. A patchwork of state statutes also addresses trafficking and related commercial sex crimes [42]. Because state statutes can be more efficient and can yield longer sentences with a lower evidentiary burden, state-level prosecutors often pursue cases under local statues rather elevating to a federal case. Additionally, many prosecutors pursue commercial sex-related statutes such as *pimping*, *pandering*, or *material support to prostitution*, instead of a trafficking statute, as they can be more efficient, require less evidence, and have robust minimum sentences and compulsory sex offender registration. For example, in California a person can be convicted of *pandering* by encouraging another person to engage in prostitution. There is also a wide range of vacature and forfeiture clauses related to trafficking that can significantly affect the ability of survivors to advance their lives post-victimization.

Access to technology: There is a wide disparity in the availability of purpose-built technology to assist trafficking investigations and prosecutions. Some states have advanced capabilities such as electronic forensic exploitation teams, dedicated financial analysis cells, and in-house software developers. In other states, investigators and prosecutors have few purpose-built capabilities beyond NGO-developed tools and standard spreadsheet software. There is also a wide range in the cost of technology, with some NGO-developed tools provided free to law enforcement, while some electronic forensics tools cost tens of thousands of dollars in annual licensing and maintenance fees. High cost can be a significant to insurmountable burden for smaller agencies to access relevant technology to assist in investigations. Lastly, across this spectrum, investigating and prosecuting trafficking cases is significantly time- and human-intensive, often requiring over a year of work to complete single cases.

3.2.1 Overview of the Criminal Justice Process for Trafficking Investigations

Across these disparate regional, organizational, and technological approaches, the following set of common trafficking investigation and prosecution processes, stages, and techniques were observed:

i. **Observation and investigation generation:** Organizations use either a proactive or reactive process to start investigations. Proactive organizations monitor data sources to find activity or vulnerability. This can include monitoring classified sex advertisements in a geographic area and cross-checking contact phone numbers in the advertisements with prostitution arrests, domestic violence reports, or other indicators. Reactive organizations begin the investigative process after receiving a tip from the federal government, a regional law enforcement organization, or other public organizations. Once a tip is received, it is adjudicated using government and open-source databases to verify the identities of individuals and to determine whether the indicators found warrant opening a case. For example, a phone number in a domestic violence report that is also used to advertise sex services may indicate that further investigation is warranted.

ii. **Evidence triage and initial court preparation:** After an investigation is opened the primary evidence collection process begins, typically with subpoenas to obtain information about the individuals and organizations identified in the previous step. Subpoenaed information can include call detail records from telecommunication providers; account or subscriber information from social media companies, banks, or internet service providers; advertisement data from classified sites; and other information. From these records, analysis is performed to establish connections between actors, locations, transactions, and activities to determine if illicit activity may be present. For example, a call frequency analysis may be performed on call detail records to better understand the interactions of suspected facilitators and providers. Individuals in the investigation may also be interviewed by a combination law enforcement and victim services organizations, to gather first-hand information about potential or suspected activities. The objective of this phase of the investigation is to collect sufficient evidence to establish probable cause and obtain search warrants from a judge.

iii. **Main evidence analysis and grand jury preparation:** Once search warrants are obtained, the evidence collection grows substantially as the compelled information returned becomes more diverse and complex. For example, instead of general account information for a social medial profile obtained via subpoena, the records returned for a warrant can contain the entire profile including all images, videos, messages, and other data the user uploaded or interacted with. Triaging and analyzing such large and complex data sets is time-intensive and can represent the lion's share of work involved in developing a case. Interviews with individuals associated with the investigation also continue in this phase, often in response to information uncovered from compelled records. Investigators and prosecutors perform detailed financial, communication, and geospatial analysis to demonstrate violation of the law, such as through force, fraud, or coercion, in order to secure an indictment, often from a grand jury.

iv. **Final analysis, incarceration monitoring, and trial preparation:** Once the indicted individuals are arrested, additional evidence collection takes place, often using search warrants to seize mobile phones, computers, vehicle embedded systems, and other electronic devices during an authorized search of the premises of suspects. Forensic extractions and examinations are performed on these devices, generating large volumes of complex data requiring specialized analysis. While suspects are incarcerated awaiting trial, they often use

jail phone systems to witness-tamper and to run illicit operations. These communications must be monitored, transcribed, and analyzed as the case builds towards trial. After the final sources of evidence and interviews are analyzed and corroborated, the entire corpora must be laid out in thorough presentations for trial, which can take weeks to months to produce.

3.2.2 Evidence types and attributes

During analysis of stakeholder investigative processes and workflows, four major categories of evidence data emerged, as follows:

Advertisement sites and review forums: One major reason sex trafficking investigations are tractable is that supply and demand need a marketplace in which to connect. Over the last decade, that marketplace has largely been facilitated by the internet. The ability to find and characterize advertisements that connect providers and buyers is critical to the investigative process. Currently there are three major platforms for searching and analyzing advertisements and user posts from commercial sex classified websites and buyer review forums. Together these platforms provide access to hundreds of active and inactive websites comprised of several hundred million unique archived web pages of advertisements, provider reviews, and buyer discussion threads. Web-based interfaces allow authorized law enforcement users to query and filter ads based on geographic region, keyword, phone number, and other identifiers. Users can upload photos of investigative interest, detect faces, and find advertisements with similar faces as the seed image. Additional advanced capabilities for network building, classification of advertisement text, and other features are available on a subset of platforms.

Commercial and open-source intelligence: As described in Section 3.2.1, openly available information is commonly used to triage leads and reduce the ambiguity of personally identifying information, often before seeking judicially compelled records for an individual. This information comes in two major forms: information openly available on the internet that can be accessed for free from the major commercial search engines, and information purchased from commercial data brokers. Data brokers can provide information such as addresses, phone numbers, email addresses, social media handles, places of employment, criminal records, bankruptcies, liens, assets, known relatives, vehicle registrations, drivers license information, license plates, and more. Other source-specific data brokers provide information including mobile phone geolocation histories from partnering phone applications and other methods, and vehicle movement history from networks of license plate readers. These data broker services can be expensive to use on a regular basis but can also reduce the time and uncertainty of investigations.

Judicially compelled records: A significant portion of the investigative process involves analyzing evidence that is derived by compelled judicial orders, as described in Section 3.2.1. These consist of several different classes of information, including: business records, such as bank statements, hotel receipts, wire transfers, ride-share receipts, or travel tickets; communications data, such as phone call records, text messages, social media messages, or wiretaps; internet data, such as social media profiles or web access histories; and geolocation information, such as mobile phone tower associations, social media check-ins, or other identifying

21

information. Responses to authorized requests for these data can take days to months to be returned, and come in a wide variety of data formats, from paper and scanned documents to large complex multimedia archives. The analysis of these varied information sources enables investigators to piece together the puzzle of illicit and criminal activities.

Forensic device extractions: Many investigations have some component of evidence derived from forensic extractions of data from cell phones, tablets, computers, and other electronic devices. Extractions begin with hardware that physically attaches to the device and use various methods to subvert access controls and expose data, such as emulating known backup protocols to extract specific files, or other sensitive methods that extract full file systems or copies of a device's memory. These data are stored in proprietary "image" file containers on a computer for subsequent analysis. Specialized software is used to read and decode the device image into known file and database formats that are subsequently parsed and exploited by analysis software. Some tools focus on messaging and call analysis, image and video exploitation, file and database formats of third-party phone applications, or retrieval of cloud-based information. Some tools are designed for single users and some are designed for collaborative workflows, including case management components with the ability to log and maintain the chain of custody across an organization. Many of these tools are intended for law enforcement only and have sensitive capabilities, and almost all of them require judicial order or consent in order to be used. Together the many tools in this space provide a robust set of options to assist an investigation.

3.2.3 Recommendations

Recommendation I-1: Create trafficking signature repository

Finding: Investigators are often looking for known indicators of sex trafficking, but there exists no repository for these text or visual-based signatures across law enforcement organizations.

Discussion: Indicators of trafficking or illicit commercial sex operations; which can appear on sex advertisement sites, in social media accounts, or within communications between individuals; can be important cues for investigators. These indicators, or signatures, can be based on specific language usage in speech or text, or based on visual appearance, such as specific objects or scenes in imagery or video. Text-based signatures are terms or phrases associated with specific trafficking-related activities, including: advertisement terms, such as *"new in town"*; buyer terms, such as *"GFE"*; pimping-related terms, such as *"choose up"*; or incarceration-related terms, such as *"automatic"*. Visual signatures can also contain useful trafficking-related indicators such as pictures or video of cash, drugs, hotel rooms, lingerie, or specific tattoos. Currently there is limited tooling to identify these indicators in a robust way within an investigative setting.

Suggested Actions: Develop a repository of known trafficking indicators and signatures accessible to federal, state, and local law enforcement organizations. Allow authorized organizations to search and retrieve known signatures and label and submit new ones. Taxonomies

should be developed for each class of data and updated over time as the vernacular of the illicit market changes. The signature repository should also be shared with tool developers, such as those who make sex advertising search platforms or electronic device forensic software, so that text and visual classification capabilities can be integrated into their tools.

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Narrow	High	Short

Recommendation I-2: Monitor online vulnerability indicators

Finding: Investigators are often monitoring social media profiles for potential indicators of vulnerability, but in manual and ad-hoc ways.

Discussion: Traffickers regularly use social media to look for susceptible or vulnerable individuals to contact and groom for later exploitation. Vulnerability indicators include sexually suggestive photos or post content and hash-tags with themes associated with familial trouble, truancy, drug use, running away, and depression. Recently, some law enforcement organizations, in consultation with victim service organizations, have started to monitor for such indicators on social media accounts suspected of being located within their jurisdictions. Law enforcement and service organizations would like to be tipped to accounts expressing these indicators, in order to develop intervention strategies, reach out with services, or alert family members.

Suggested Actions: Develop a capability to monitor publicly available social media profiles in a given region and provide warnings to law enforcement and social service agencies when a profile begins to express indicators associated with vulnerability, interest in the commercial sex industry, or potential trafficking. Many social media services provide Application Programming Interfaces (APIs) to access their public profiles, which could be used to screen accounts, monitor for known signatures, and infer locality.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Medium	Medium	Medium

Recommendation I-3: Develop advertisement platform federated search

Finding: Multiple sex advertisement search tools gather different information from a wide range of websites. Most law enforcement agencies use only one of these tools, leading to partial information and potentially missed victims or cases.

Discussion: Currently there are three major indexing and search platforms in regular use within law enforcement (see Appendix B). These platforms collect advertisement and buyer reviews from a range of different websites, each platform having a slightly different geographic or market focus, with a super-set of websites common across all three platforms. Because no one platform indexes all pertinent websites, investigators may have to visit multiple search platforms and run the same queries, in order to have a complete understanding of the market in their area of interest. Running identical searches manually in multiple tools is both cumbersome and time-consuming.

Suggested Actions: Work with the sex advertisement search platforms to develop concurrent machine-to-machine federated search methods across their platforms. Explore new revenue models to compensate platforms based on the utilization of information they provide in search returns, in order to sustain the benefits of federated information sharing. Additionally, the machine-to-machine search API should be designed to allow queries from other external applications. For example, if a domestic violence report is entered into a law enforcement database with a known phone number, software could be developed to automatically query that number against the federated search API to see if it has appeared in advertisements across all three platforms, and automatically alert investigators to an advertisement if a match is found.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Wide	High	Medium

Recommendation I-4: Develop case knowledge management system

Finding: Many law enforcement agencies have limited case management technology and often employ ad-hoc analysis logs and spreadsheets to keep track of discovered information throughout an investigation.

Discussion: Trafficking cases can take months to years to fully develop; over their lifetime they may be worked on by several investigative personnel and prosecutors. Additionally, compelling records, waiting for their return, and their subsequent analysis is a cyclical process that occurs concurrently across multiple investigations over timescales of days to months. Lastly,

many prosecutorial agencies have parallel caseloads that require personnel to frequently rotate their effort between cases. All of these factors complicate the chain of custody and analysis of evidence; institutional knowledge of what information was gleaned from which piece of evidence and by whom; and frequent context switching that can lead to missed connections and lost knowledge.

Suggested Actions: Develop or acquire a purpose-built integrated case and knowledge management platform for collaborative trafficking investigations across multiple law enforcement and prosecutorial organizations. The capability should include natural-language processing to extract and structure pertinent metadata from primary evidence stored in the system, and should provide full-text search. Access to data, information viewed, and user analysis history and notes should be automatically logged during the investigative process to ensure chain of custody; reporting and auditing; and knowledge management best practices.

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Wide	Mature	Long

Recommendation I-5: Establish data template repository

Finding: Businesses often use unique data and document formats when responding to judicially-compelled record requests. Several law enforcement organizations are building capabilities to read and parse these formats, incurring collective duplication of effort.

Discussion: Judicially compelled evidence is the lifeblood of many investigations, but often record returns come back in unique and difficult-to-use formats that are specific to each data provider. For example, mobile phone call detail records are produced in various spreadsheet formats, with each telecommunication provider having a different document structure and data ontology. Similarly, some social media full account archives can be returned as a large PDF document or as a complex nested file and folder structure. Many of the organizations interviewed for this work have built internal tools to parse, structure, and display these different formats for their own investigative purposes. However, this duplication of effort in producing internal-facing tool suites that are not shared across organizations, results in the reinvention of siloed capabilities across the counter-human trafficking community.

Suggested Actions: Establish a working group of commercial tool vendors, law enforcement agency software teams, and other stakeholders to develop templates for common judicially-compelled record formats and forensic device reports, and the structured extraction software libraries for each format. Templates and libraries could be stored in a restricted repository where approved vendors and agencies can access them to integrate into their own tools, increasing the efficiency of scarce software development personnel and resources.

Impact	Scope	Maturity	Timescale
Significant	Wide	High	Short

Recommendation I-6: Develop data enrichment capabilities

Finding: Investigators often have to work with complex data formats that are difficult to use efficiently, such as business records scanned as images, full social media profile archives, and electronic device forensic extractions.

Discussion: A fundamental requirement of evidence analysis is knowing what information is in your holdings. This becomes extremely important at trial, when the defense examines a similar set of evidence and when the mishandling of evidence can lead to mistrial. Currently, many forms of evidence analysis involve manual review of a wide range of document formats and other complex data, requiring investigators to manually cut and paste relevant pieces of information into other formats for later analysis. Also, several forms of evidence are inherently difficult to work with, such as scanned documents; photos or screen-shots of mobile phone screens; social media profile archives; and voluminous mobile phone reports generated from forensic extraction software. These challenging data formats contain rich information that is not easy to search in their native forms, leading to manual, time-consuming, and ineffective analysis processes.

Suggested Actions: Develop an integrated software capability that leverages natural language processing, computer vision, machine learning, and other techniques to extract, structure, and enrich information contained in primary evidence. Textual information, such as names, locations, organizations, phone numbers, IP addresses, and other information should be extracted automatically using statistical and rule-based methods. Faces, objects, and text overlays should be recognized and extracted from scanned documents, images, and videos, and used to enrich the source media with the enriched metadata. Social network analysis techniques should be employed to find connections between information within and across documents and data, in order to automatically build relations between pieces of evidence to accelerate triage, analysis, and confirmation. Original evidence, extracted information, and enriched metadata should be stored in a distributed full-text search engine as opposed to a traditional database, to provide increased discoverability for investigators and prosecutors.

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Wide	Medium	Medium

Recommendation I-7: Use geospatial information effectively

Finding: Investigators lack capabilities to effectively discover and extract geospatial information found in evidence, limiting their ability to make best use of geospatial attribution and analysis during investigations.

Discussion: One of the tenets of trafficking investigations is piecing together a timeline of when and where events occurred and which individuals in the investigation are associated with those events. Many of these events are referenced in location-based information in various types of primary evidence. This information may be explicit, consisting of text mentions of exact latitude and longitude coordinates; or it may be implicit, such as referencing a specific facility whose latitude and longitude coordinates can be determined with auxiliary information. Examples of evidence containing both explicit and implicit mentions include; business records containing IP addresses of user login histories, cell phone records containing cell tower associations, digital photographs containing geospatial EXIF metadata, license plate reader records, and more. Much of this evidence also has temporal information, enabling joint spatiotemporal analysis and development of *patterns of life* for individuals in the investigation. For example, when a facilitator and a provider travel to an out-call at a hotel with a buyer, this meeting event can have rich spatiotemporal information, such as the listed service area of the sex advertisement, address of the hotel, the date and time on the hotel stay, GPS location records of the facilitator's cell phone en route to the hotel, and more. However, investigators have few capabilities that can robustly understand and make accessible explicit and implicit geospatial information contained within evidence.

Suggested Actions: Develop a software capability that automatically extracts from various evidence formats explicit geospatial and temporal mentions, and can also recognize and georeference implicit geospatial information. Extracted events should be stored in a geospatial datastore and integrated with a user interface that combines timeline and map visualization tools. This environment can help investigators better understand spatiotemporal information throughout an investigation, and can aid in critical areas like victim cooroboration and more effective presentation of evidence in court.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Narrow	High	Medium

Recommendation I-8: Improve call record analysis

Finding: Phone call records play vital roles in many trafficking investigations, yet investigators often do not have proper tooling to make best use of call record data.

Discussion: Information derived from mobile, landline, and calling card phone records is a cornerstone of trafficking investigations. Investigators often receive months to years of call records from telecommunication companies in response to subpoenas and warrants, and can be augmented with call histories extracted from seized mobile phones. These records are often processed and analyzed by individual source phone number with basic analysis techniques, such as ranked call frequency, using common spreadsheet software. This approach does not scale to the large amount of target phone numbers and records often associated with complex cases, and the lack of joint processing limits the ability to make connections across call record datasets. These limitations may lead to missed discovery of coordinated activities or obfuscation techniques, and missed identification of additional suspects and victims.

Suggested Actions: Develop a capability to extract, structure, and store call records from any relevant investigative source and format. Apply social network analysis techniques to enable in-depth analysis of phone call behaviors across the collective data within an investigation. Augment the deployment of the capability with personnel training in common social network analysis techniques and best practices.

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Narrow	High	Short

Recommendation I-9: Enhance financial record analysis

Finding: Investigations often have a financial component, but agencies lack the access to relevant data, tooling, and expertise to best integrate financial analysis into investigations.

Discussion: Illicit commercial sex operations use complex financial processes in order to conceal the generation, transfer, and storage of their proceeds, and uncovering these activities can be critical to successful trafficking investigations. These processes can range from small-scale all-cash schemes to large-scale complex international money laundering operations using the global financial system. Local trafficking operations often leverage the retail banking system, where accounts will be opened in a victim's name and are managed by the trafficker using various techniques [43]. Similarly, commercial banking is often used in conjunction with illicit massage parlors, and gaining access to merchant service provider records can be critical for characterizing an operation's inner-workings and for identifying buyers. [44].

International trafficking operations often involve money transmitters, such as Western Union or MoneyGram, to wire proceeds to beneficiaries outside the U.S. [44]. Ad-hoc financial methods are also used, such as the laundering of money through gift cards and prepaid debit cards, and the use of cryptocurrencies, such as Bitcoin, to obfuscate the sender and recipient of funds [45].

Several organizations associated with the global banking system develop and share financial-related indicators and behaviors associated with human trafficking, to help flag suspected activities. For example, the Financial Action Task Force (FATF), a policy-making body which works on an inter-governmental level to combat illicit money laundering, develops comprehensive lists of such indicators [43]. Similarly, the Financial Crimes Enforcement Network (FinCEN), an enforcement agency under the U.S. Department of Treasury, develops advisory criteria to help financial institutions codify suspected human trafficking using Suspicious Activity Reports (SAR) [46]. NGOs also play a role in helping to coordinate analysis of indicators between financial institutions that cannot be easily undertaken at the federal level due to authority or policy limitations, such as the financial intelligence analysts at the Polaris Project. Following sources and flows of money can be critical to gaining an understanding of the scope of an illicit operation, characterizing the involvement of individuals, and identifying accounts and funds for potential seizure and forfeiture.

Suggested Actions: Develop new capabilities to process, store, and analyze financial data relevant to human trafficking investigations. Focus areas should include capabilities to process large amounts of bank transaction and money transmitter records, to more easily build integrated spatiotemporal networks of activities. Additionally, automated detection and alerting capabilities should be developed for known indicators and coordinated behaviors, such those from the FATF, FinCEN, and the broader anti-money laundering community. Next, make better use of financial data and resources not derived from compelled primary evidence, including incorporation of FinCEN SAR information and working with data from NGOs, such as Polaris and other financial institution partnerships. Remain cognizant of abilities to monitor and analyze various cryptocurrencies, and incorporate into investigations as required. Lastly, seek out training and best practices for personnel for how to better leverage financial analysis techniques within the context of investigations, and maintain awareness of ongoing federal policy and collaboration efforts within the U.S. Department of the Treasury, specifically within the Office of Terrorist Financing and Financial Crimes (TFFC) and the Office of Foreign Assets Control (OFAC).

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Wide	Medium	Short

Recommendation I-10: Utilize surveillance video analytics

Finding: Investigations often have a video evidence examination component; however, current analysis methods are inefficient and human-intensive, with limited technology to assist.

Discussion: Many human trafficking cases often involve obtaining closed-circuit video surveillance footage from hotels, stores, banks, and other businesses that are associated with trafficking operations. Some undercover operations also use covert video surveillance equipment, such as cameras strapped to light or telephone poles outside suspected illicit facilities. Organizations interviewed for this study suggested that upwards of 60% of their cases involve monitoring and analyzing video footage in some form, which entails personnel watching hours of video in real-time, waiting for persons of interest to appear in scenes and then characterizing their movements. While some vendors provide hardware and software systems that assist with the forensic extraction of video data from proprietary video surveillance systems, there are few analysis tools for the triage and analysis of the extracted video data.

Suggested Actions: Employ video analytic capabilities to reduce the time- and human-intensive elements of surveillance video analysis. Technologies to consider include attribute-based search; which allows a user to find video segments based on attributes of interest, such as of a person wearing an article of clothing of a specific color; and time compression or video summarization; which allows segments of video with little or no activity to be removed, focusing analysis on scenes with significant activity. Additionally, facial recognition and object detection techniques could be employed on video data, if the video resolution is high and the camera geometries suitable for common image classification techniques.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Wide	Medium	Medium

Recommendation I-11: Employ speech transcription tools

Finding: Investigators and prosecutors spend large amounts of time listening to, transcribing, and analyzing phone call speech content.

Discussion: As noted in 3.2.1, an important time to gather additional evidence on indicted suspects is when they are incarcerated awaiting trial. Prison and jail facilities use specialized phone systems to record calls from inmates, who are given a Personal Identification Number (PIN) to identify themselves when they place calls. Aware they are being monitored, inmates often employ tactics to evade attribution by swapping PINs with other inmates, or by calling intermediaries to setup three-way calls that obfuscate one or more call participants. Once

calls from the correctional facility are recorded, they are provided via court order to authorized personnel, often specialized linguists, for analysis and transcription. Many prosecutorial agencies interviewed for this study noted that call transcriptions are an element of most of their human trafficking cases, and stressed the large amount of time and personnel resources expended to manually transcribe calls. Compelled call recordings are also delivered in bulk, and personnel often have no a priori knowledge of which calls might have the highest investigative value and should be prioritized for transcription. Complicating this task further, many calls contain foreign languages, frequently Spanish, Mandarin, and Russian.

Suggested Actions: Speech processing technology offers the potential for reducing the time- and human-intensive nature of jail and prison call analysis. First, speaker biometrics should be employed help keep track of unique parties on calls and to mitigate against PIN number swapping. Language identification processing techniques should used to survey a call dataset and determine what linguist personnel would be required for a given case. Next, automatic speech recognition should be used to convert call audio into partial text transcripts. While these text transcripts are only partially complete and have varying degrees of accuracy, they can be helpful to assist in the triage of large volumes of calls. Several analytic methods can be used to support this triage process, including word frequency analysis, topic modeling, and text summarization; language identification and text-to-text machine translation; keyword spotting; and other techniques. While these technologies will not replace the need for manual transcription of calls for evidentiary or trial purposes, they can serve as critical tools for directing and prioritizing scarce manpower to the highest calls of interest. Lastly, many of these component technologies could be combined to create a machine-assisted transcription system to increase the overall efficiency when full transcription is required.

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Narrow	Medium	Medium

3.3 INFORMATION SHARING

In interviews conducted for this study, practitioners often expressed an interest in improving on current quantitative, data-driven approaches to understanding trafficking at local, state, and national levels. Making such improvements involves more comprehensive collection and sharing of data, addressed in this section, and new methods for measuring relevant quantities in that data, discussed in Section 3.4. This section begins by characterizing the benefits and challenges of sharing human trafficking-related data, and then reviews existing data sharing systems and standards relevant to the counter-human trafficking community, followed by a set of findings and recommendations for technologies that can help to address concerns about data sharing.

3.3.1 Benefits and Challenges of Sharing Data

There are potential benefits of increased data sharing [47, 48], including:

i. Enabling discovery of victims and potential victims by aggregating weak signals across multiple data sources.

ii. Increasing prosecution of facilitators, traffickers, and buyers by aggregating weak signals across multiple data sources.

iii. Providing a more robust picture of the scale and scope of trafficking at the local, state, regional, and national levels in order to raise awareness and improve resource allocation.

iv. Improving the accuracy of trafficking statistics by de-duplicating data reported by multiple entities.

v. Guiding public policy by enabling quantitative evaluation of policy effectiveness.

vi. Facilitating the provisioning of victim services.

Despite these clear benefits, broad sharing of data within the counter-human trafficking community remains elusive. In interviews for this study, practitioners indicated that they place a high value on quantitative, data-driven analyses but lamented that they frequently do not have access to the data needed for these analyses. Data inaccessibility has a number of root causes, such as:

i. Concerns about traumatizing victims if their personally identifiable information (PII) becomes publicly available.

ii. Privacy concerns for victims and other individuals who may or may not be associated with trafficking.

iii. Challenges with sharing law enforcement sensitive data with non-law enforcement organizations.

iv. In the medical community, concerns about violating the Health Insurance Portability and Accountability Act (HIPAA).

v. Loss of control over proprietary data sets that have taken significant time and resources to construct.

3.3.2 Existing Data Sharing Systems and Standards

At the national level, there are two general types of data sharing mechanisms associated with counter-human trafficking. The first consists of databases and agreements that focus specifically on the sharing of human trafficking data. These include the Banks Alliance Against Trafficking (BAAT), the Human Trafficking Reporting System (HTRS), and Trafficking Information Management System (TIMS). However, this list excludes databases with no sharing mechanism, such as

The National Human Trafficking Resource Center (NHTRC) database of National Human Trafficking Hotline calls, and global sharing mechanisms such as the IOM Global Human Trafficking Database. These sharing mechanisms are limited in scope, access, or both. For example, BAAT is limited in scope to financial information, and its participants are limited to financial institutions, financial crime compliance standard-setting bodies, national Financial Intelligence Units, some law enforcement, and a small number of counter-human trafficking NGOs. Similarly, TIMS is limited in scope to information associated with Office for Victims of Crime (OVC) grants, and access to TIMS is limited to current grant recipients.

The second category consists of data sharing mechanisms accessible primarily by law enforcement, such as the Law Enforcement Enterprise Portal (LEEP), National Data Exchange (N-DEX), and Regional Information Sharing System (RISS). Together, these systems enable sharing of data such as incident and case reports; arrests and incarcerations; and parole and probation information across local, state, regional, federal, and tribal law enforcement agencies; public safety agencies; and task forces. The databases do not focus specifically on human trafficking and are not accessible to the broader counter-human trafficking community. Moreover, trafficking data within these systems can suffer from reporting inconsistencies across different organizations, as discussed further in Section 3.4. Nevertheless, they can enable law enforcement to gather multi-jurisdictional evidence of trafficking activity. More details of these systems appear in Appendix C.

Of these systems, RISS, a network of six regional centers that enables information sharing among law enforcement agencies, offers several capabilities and features that may be relevant to the counter-human trafficking community. The first capability is RISSIntel, which provides the ability to share information and to search over 50 federated databases simultaneously, and also provides analytical tools and support. The second capability is the RISS Automated Trusted Information Exchange (ATIX), which enables collaboration among federal and local agencies, as well as, relevant private entities, such as hotels, during and after disaster or terror incidents. Lastly, the RISS Master Telephone Index (MTI), is a database system that can compare and match common telephone numbers in law enforcement investigations nationwide.

The Western States Information Network (WSIN), discussed during the stakeholder interviews, is one of the six RISS regional centers and leverages RISSIntel in human trafficking investigations to provide intelligence products, such as [49]:

  i. Organizational or link charts that identify relationships among victims and traffickers;

 ii. Analyses of communications among traffickers, victims and buyers including cell towers, locations, and emails;

iii. Analyses of money flows associated with trafficking;

 iv. Timeline visualization of events associated with trafficking cases.

RISS ATIX, while not currently focused on human trafficking, provides a model for how the RISS system can be used to include non-law enforcement entities, such as hotels.

In addition to these systems, a number of national standards exist for sharing data in general and law enforcement data specifically, including the National Information Exchange Model (NIEM), the Global Justice XML Data Model (GJXDM), and Logical Entity Exchange Specifications (LEXS). NIEM, which serves as the foundation for RISS, is an Extensible Markup Language-based information exchange framework for sharing data among agencies and organizations across all levels of government (federal, state, and local) and with private industry. NIEM has models for 14 domains, such as human services, justice, emergency management, biometrics, and agriculture. Standards within these domains are developed within Communities of Interest (COIs), collaborative groups of users who desire to exchange information in pursuit of shared goals, interests, missions, or business processes and who therefore must have a shared vocabulary for the information they exchange. COIs capture domain-specific data structures in Information Exchange Packet Documentation (IEPD). The Administration for Children and Families (ACF) Office on Trafficking in Persons (OTIP) is currently a member of the COI that creates standards for the NIEM human services domain.

3.3.3 Recommendations

Recommendation S-1: Leverage existing sharing mechanisms

Finding: Data available within existing sharing mechanisms is not collected consistently and is not available beyond law enforcement.

Discussion: While existing systems such as RISS are effective for enabling law enforcement to discover and link evidence associated with traffickers, currently they lack consistent data standards and do not allow for sharing data with the broader counter-human trafficking community beyond law enforcement. At the same time, RISS provides a proven operational infrastructure with features including analytics, database federation, and network security. Moreover, the RISS ATIX public-private partnership may serve as a model for engaging with non-law enforcment counter-human trafficking entities.

Suggested Actions: Investigate existing data sharing mechanisms for their ability to address counter-human trafficking community needs. Coordinate with efforts under Recommendation M-3 to develop and disseminate data sharing standards within one or more of these mechanisms. If feasible, work with an appropriate sharing mechanism to implement standards and improve sharing. Specific possibilities include leveraging RISS to improve counter-human trafficking data sharing not only within law enforcement but also between law enforcement and non-law enforcement entities, and working within NIEM to develop standards for sharing human trafficking data.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Wide	High	Medium

Recommendation S-2: Encourage novel data sharing

Finding: Many data sources relevant to counter-human trafficking efforts are not accessible using current sharing mechanisms.

Discussion: A number of community members interviewed for this study identified relevant but inaccessible data sources that either they or others in the community own. Examples include social media of traffickers, victims, potential victims, and buyers; hospital and other medical records; phone records associated with buyer stings and counter-demand operations; data from law enforcement license plate reading systems in the vicinity of known trafficking circuits; buyer hobby board reviews; and other closely-held data sets collected and managed by counter-human trafficking NGOs. (Appendix D lists a number of additional data sources of interest). In the long run, wide-scale sharing of such data will likely require advanced technologies such as those discussed in Finding S-4 below. In the interim, it may be possible for small groups of organizations sharing mutual interests to collaborate in sharing data among themselves as a pathfinder toward larger, community-wide sharing, similar to the one described in Appendix E.

Suggested Actions: Pursue a bottom-up approach to building collaborative data sharing agreements among small groups of counter-human trafficking organizations on a case-by-case basis in order to encourage sharing of novel data sources. Elements of this approach include identifying novel data sources and organizations twilling to share them, implementing a pilot sharing program, and extending this program to include other sources and organizations.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Narrow	High	Medium

Recommendation S-3: Encourage local data sharing

Finding: Differences in trafficking patterns, law enforcement human trafficking training and awareness, and legal statutes at the local, state, and regional levels all hinder data sharing.

Discussion: A common theme in interviews for this study was that trafficking presents very differently in disparate communities, sometimes only a few miles apart. An implication is that, even if a standards-based approach is adopted as described in Recommendation S-1, these differences may lead to discrepancies in what data is collected and even what data is of interest among collaborators. For example, specific practices in the state of Oregon include

using state Department of Human Service Form 307 as an information source to initiate cases and using the state's "promoting prostitution" statute as an effective way to bring traffickers to justice. Data associated with these practices are likely to be of interest within Oregon but of less interest to organizations in other states that have no equivalent standard forms or statutes.

Suggested Actions: Pursue a local approach to building collaborative data sharing agreements among small groups of counter-human trafficking organizations that serve a local community or region. Use an approach that is similar to the one described for Recommendation S-2: identify a locality with proactive, forward-leaning counter-human trafficking organizations interested in sharing data, implement a pilot sharing program among these organizations, and seek to extend to include other organizations and/or localities.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Narrow	High	Medium

Recommendation S-4: Leverage secure computing

Finding: Traditional data security approaches may not be sufficient for sharing human trafficking data.

Discussion: As noted in Section 3.3.1, data sharing efforts are limited by concerns such as traumatizing victims, revealing PII of victims, violating HIPAA, and legal and proprietary issues. Current data sharing mechanisms use traditional security mechanisms to protect information shared on their systems: RISS, for example, uses encryption, user authentication, Internet Protocol security standards, and firewalls to prevent unauthorized access. These technologies enable authorized users to interact electronically with one another in a secure environment; however, their only mechanism for addressing the above concerns is simply to deny access.

Several technology solutions are available that may address data sharing concerns while also providing access to the full counter-human trafficking community. The simplest is data anonymization, which involves encrypting or removing PII from data sets so that the people whom the data describe remain anonymous. A number of data anonymization tools are commercially available today. While this technology is mature, it suffers from two drawbacks. First, it is vulnerable to de-anonymization attacks, in which anonymized data may be correlated with PII or other auxiliary data to reveal the identity of the individual associated with the data. Second, it hinders correlation with other data sets, making it difficult to achieve some of the key goals of data sharing, such as enabling discovery of victims, potential victims, and traffickers by aggregating weak signals across multiple data sources.

More sophisticated technologies currently under development for commercial data sharing applications have the potential to address privacy and other concerns while also enabling full data access across the counter-human trafficking community. These technologies, which include secure multi-party computation and homomorphic encryption, enable users to query and perform computation on encrypted data without first decrypting it. They are currently being investigated in the DHS S&T Data Privacy project [50] and other DHS S&T cybersecurity efforts which address secure cloud computing. In addition, these techniques have been proposed for applications including healthcare record sharing, biometric data matching, and smart energy grids [51, 52]. Limitations of these approaches include low maturity, computational latency, and restrictions on the types of functions that can be computed on encrypted data [53].

Suggested Actions: Monitor advances in secure computing, including those currently being investigated within DHS S&T. Consider one or more pilot data sharing programs that leverage these capabilities after they have been proven in the commercial sector. Pilot programs could, for example, involve sharing among law enforcement, other government entities, healthcare providers, and other NGOs.

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Wide	Low	Long

Recommendation S-5: Implement peer-to-peer marketplaces for victim services

Finding: Insufficient availability of services and options for victims impedes their ability to escape exploitation.

Discussion: A number of law enforcement agencies and victim services NGOs interviewed for this study indicated that the largest impediment to removing victims from exploitative circumstances is their lack of options and available services, especially immediate bed space after encounters with social workers, victim advocates, and law enforcement. Victims encountered by the organizations often return to "the life" simply because they have no other choice. Matching of victims to the services they need is typically done at the local level, based on local knowledge of specific victim services NGOs that operate within a region or jurisdiction. Traditionally, the staff group home model has been used to care for victims, and more recently specialized therapeutic foster homes are being used with individual victims [54].

Peer-to-peer (P2P) marketplaces may offer a partial solution to these challenges of aligning need with availability. P2P marketplaces have become commonplace commercially in recent years because of their ability to efficiently match providers of products and services with

individuals who want them. Examples include marketplaces for ride-sharing (Uber, Lyft), vacation rentals (VRBO, Airbnb), and home maintenance and repair services (TaskRabbit). Commercial software packages that enable simple, rapid deployment of robust P2P marketplaces are available today. While the needs of trafficking victims are significantly more complex than those of participants in these ad-hoc markets, there may be a role for P2P technologies to address at least some of them. For example, a P2P solution could be used to provide visibility into available bed capacity across a region to better help victim service organizations have better awareness of potential options. Additionally, P2P capabilities could potentially be used to help identify and match candidates for the individualized foster home programs.

Suggested Actions: Investigate implementing P2P marketplaces for victim services. Specific tasks include working with NGOs and law enforcement to gauge feasibility of and interest in providing victim services through a P2P marketplace, and implementing one or more local pilot P2P programs.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Narrow	High	Short

3.4   MEASUREMENT OF TRAFFICKING

As noted in Section 3.3, there is significant interest in the counter-human trafficking community for quantitative, data-driven approaches to understanding trafficking at local, tribal, state, and national levels. The previous section focused on enabling this goal through technologies that facilitate data sharing among community members. This section addresses the types of studies that should be conducted with this shared data, the data sets that should be developed to support these studies, and the underlying technical approaches needed to address data limitations. Appendix D contains a much more detailed discussion of the potential utility of a wide range of data sources for measuring and understanding trafficking, and Appendix E describes a preliminary statistical modeling effort using such data.

3.4.1   Benefits and Challenges of Measuring Trafficking

The benefits of quantitative analyses have been noted in numerous human trafficking studies. A robust, quantitative understanding of the scale and scope of trafficking can be used to raise awareness, improve resource allocation, guide public policy by enabling quantitative evaluation of policy effectiveness, and ultimately reduce the prevalence of trafficking. Building such an understanding is challenging. Some of these challenges arise from the process of gathering and sharing

the necessary data; these are discussed in Section 3.3.1. However, significant analysis challenges remain once data has been aggregated [7, 21, 22, 55–58]. These challenges include:

Incompleteness: Victims are often hidden and may not come into contact with officials or organizations who collect data [22, 55].

Inconsistency in identifying trafficking: Even when contact is made, victims may not self-identify or may not be identified by those making contact [22].

Inconsistency in recording trafficking data: Collection and reporting procedures are often inconsistent at local, state, and national levels. Such inconsistencies arise for many reasons. For example, differences in legal statutes can affect how trafficking-related activities are prosecuted; often traffickers may be prosecuted for other crimes that are easier to prove than trafficking. Inconsistencies can also arise from how trafficking activities are recorded geospatially and temporally. For example, some organizations may aggregate information annually or at the state level, while others may use finer granularity; or organizations may use different procedures for recording trafficking activity that occurs over wide geographic areas or for long durations [7].

Appendix E discusses these issues in more detail in the context of a specific set of heterogeneous, nation-wide data sets associated with human trafficking.

3.4.2 Quantitative Study Examples

Numerous quantitative studies of trafficking have been conducted; many are reviewed in the companion report [7]. Rather than providing an exhaustive review, this section highlights reports that are relevant to the findings and recommendations below.

Researchers have used a variety of methods to address the challenges discussed in Section 3.4.1. An approach known as Multiple Systems Estimation (MSE) has been used to address data set incompleteness in computing victim prevalence estimates. MSE works by matching victim identities across multiple data sets, determining the number of individuals appearing in multiple data sets, and using statistical techniques, combined with assumptions of data set independence, to extrapolate the estimated true prevalence [57]. Another approach is to correlate incomplete trafficking data with more readily available proxy data sets, such as data for underground drug and weapons markets, and to extrapolate to the true estimated prevalence using the derived correlation model [59]. Other researchers have analyzed online data sources, such as sex advertising sites and hobby boards, to measure supply and demand in the commercial sex economy [35]. Online sources can provide a richer view into marketplace dynamics than conventional data sources, but they introduce new challenges such as duplication and fake data mitigation, and association of online personas to real-world individuals. Finally, some researchers have addressed data inconsistency by focusing on small geographic regions across which data is reasonably consistent [60–62].

3.4.3   Recommendations

Recommendation M-1: Quantify impact of actions

Finding: There is a need for quantitative measurement of the impact of counter-human trafficking efforts over time.

Discussion: Many practitioners interviewed for this study expressed the desire to quantitatively understand the effect of their efforts on trafficking. For example, law enforcement agencies noted that, while they have the ability to perform buyer stings almost daily, arresting tens of buyers per operation, they were unsure whether these stings had a deterrence effect and whether they were an effective use of their resources. Similarly, prosecutors were interested in determining whether requiring buyers to attend john schools in exchange for expunction of charges results in lower recidivism. Outside of law enforcement, there is interest in understanding the impact of counter-human trafficking statutes and counter-demand and other awareness campaigns. Second- and higher-order effects are also of interest, such as the effect of counter-demand campaigns on suppliers and ultimately on the number of individuals trafficked.

Causal inference techniques, which seek to rigorously determine causal relationships between events, have been applied in diverse fields including marketing, epidemiology, biomedical research, and pharmaceutical testing [63, 64]. While some work has been done in applying these techniques to human trafficking [65, 66], more research needs to be done in this area.

Quantifying the impact of a given action requires measuring relevant variables both before and after the action takes place, which in turn requires anticipating data that will be needed well in advance of the action. Identifying this data requires collaboration among stakeholders from across the counter-human trafficking community. This need is addressed in Recommendation M-3.

Suggested Actions: Convene a workshop that includes practitioners from other fields, such as health and biomedicine, to identify relevant techniques for measuring impact in the human trafficking domain. Initiate a pilot program that applies these techniques to measure the impact of a specified action. Through the working group discussed in Recommendation M-3, identify the appropriate data sets needed to conduct such measurements. In the long term, use these techniques and data sets to identify successful counter-human trafficking strategies and to target them for additional funding.

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Medium	High	Long

Recommendation M-2: Develop geospatially hierarchical methods

Finding: There is not a unified understanding of trafficking across local, state, regional, and national scales.

Discussion: As noted in Finding S-3, trafficking patterns and characteristics can vary substantially across local, state, and regional levels. These differences can have a substantial impact on the outcome of quantitative studies. Often, studies focus on one of these levels without connecting their results to those at lower or higher levels.

Many kinds of observational data—including data collected in the social and biological sciences—have a hierarchical or clustered structure. For example, in data associated with education, the academic performance of two children from the same family is more likely to be similar than that of two students chosen at random from their school population. At the same time, the performance of two students from the same school is more likely to be similar than that of two students chosen at random from the same state.

Techniques that address data analysis and modeling in such multilevel or hierarchical systems have been developed for education, sociology, biomedical sciences, and other fields [67–70]. These techniques take into account the existence of data hierarchies by allowing for residual components at each level in the hierarchy. For example, a two-level model which allows for grouping of child outcomes within schools would include residuals at the child and school level. Thus the residual variance is partitioned into a between-school component (the variance of the school-level residuals) and a within-school component (the variance of the child-level residuals) [71, 72].

Generally, multilevel methods are relevant when [70]

- Observations are correlated or clustered along spatial, non-spatial, or temporal dimensions;
- Causal processes operate simultaneously at more than one level; and
- Describing population variability is of interest.

As the first bullet suggests, the techniques are not limited to data that is spatially hierarchical. They can also be used to perform meta-studies that aggregate results from existing research while adjusting for differences in when, how, and what data was collected [73]. In this context, the residual variance is partitioned into between-study and within-study components. Similarly, they can be applied to longitudinal studies by treating time as another level. In the trafficking domain, generalized linear mixed models, a form of multilevel modeling, have been used to combine data from multiple sources and studies at the city and state levels in the United States [73], but more work needs to be done in this area.

Suggested Actions: As for Recommendation M-1, convene a workshop that includes practitioners from other fields to identify appropriate multilevel modeling techniques. Initiate a pilot program that applies these techniques to a specific region known to exhibit geospatial differences in trafficking patterns at multiple levels. Through the working group discussed in

Recommendation M-3, identify the appropriate data sets needed to conduct such measurements. In the long term, leverage multilevel techniques to integrate understanding of the trafficking marketplace over local, state, regional, and national levels.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Medium	High	Medium

Recommendation M-3: Establish data standards working group

Finding: Incomplete and inconsistent data sets continue to limit the ability to measure trafficking.

Discussion: As discussed in Section 3.4.2, methods such as multiple systems estimation and correlation with proxy data sets have successfully addressed incompleteness and inconsistency in estimating trafficking prevalence [57, 59]. Techniques developed for handling missing data have also been developed in biomedical and other research areas, and these techniques have been applied to both longitudinal studies [74, 75] and hierarchical studies [76].

While these techniques are powerful and should be applied in the trafficking domain, they are not a substitute for high quality data. The longitudinal and multilevel studies recommended in Findings M-1 and M-2, respectively, will benefit from the establishment of a stakeholder group that seeks to establish uniform data collection standards within the counter-human trafficking community. This group will have two objectives. The first is to establish data standards in order to improve the uniformity of data reporting among stakeholders at the local, state, and national levels. Efforts to achieve this objective should include engaging with standards-setting organizations as discussed in Finding S-1. The second is to identify data that should be collected in support of the recommendations for longitudinal and multilevel analyses. This objective is especially important in the case of longitudinal studies, which require identifying appropriate data to be collected in anticipation of future events or actions.

As noted in Section 3.3.2, a number of mechanisms exist for developing standards, most notably NIEM, and organizations such as ACF OTIP have or are currently engaged with these mechanisms. Efforts associated with this recommendation should build on existing mechanisms and standards when possible.

Suggested Actions: Identify and reach out to previous and existing efforts to establish data collection standards for human trafficking. Use lessons learned from these to establish a data analysis and standards working group of counter-human trafficking stakeholders. Through this working group, develop and disseminate standards, initially through a pilot program that targets a specific organization or information sharing mechanism. Identify data collection activities to support other recommendations, such as M-1 and M-2.

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Wide	High	Short

Recommendation M-4: Build victim prevalence models

Finding: The counter-human trafficking community needs better victim prevalence estimates.

Discussion: Accurate victim prevalence estimates are crucially important to the counter-human trafficking community because they drive awareness of the problem and allocation of resources needed to address it. As discussed in Section 3.4.2 as well as in the companion report [7], numerous studies have reported estimates using a variety of data sources and methodologies. This research needs to continue, leveraging recommendations made in this section, in Section 3.3, and in Appendix D. As discussed in Appendix D, potential data sources include law enforcement reporting and arrest data, judicial data, case work, trafficking tip line reporting, hospital and medical services records including public health human trafficking reports, and victim services provider data. Relevant methodologies include the longitudinal and multilevel techniques recommended in Findings M1 and M2. Minimization of data sharing concerns regarding PII, HIPAA, and proprietary considerations should be addressed using recommendations made in Section 3.3.3. Objectives include understanding the impact of policy changes and other actions on victim prevalence.

Suggested Actions: Conduct longitudinal, multilevel, and other studies of victim prevalence using a broad array of data sources at local, state, regional, and national levels with the goal of understanding the scope of victimization and how to allocate resources to minimize it.

Assessment:

Impact	Scope	Maturity	Timescale
Signficant	Medium	Medium	Long

Recommendation M-5: Build buyer demand models

Finding: The counter-human trafficking community needs a better understanding of buyer demand.

Discussion: As a subject for quantitative analysis, the demand for commercial sex has received less attention than victim prevalence. Studies often use surveys to attempt to understand buyer attitudes and motivations [31, 77]. One such study uses survey responses of over 8,000 men across the United States to provide a detailed picture of buyers nationally at a specific point in time [31]. There is a need for buyer demand studies that expand on this work by leveraging new data sources and methodologies to build a more complete understanding of demand over time and at local, state, and national levels. As discussed in Appendix D, relevant data sources include hobby boards, data collected by buyer deterrence platforms, law enforcement reporting, buyer stings, and john school enrollment. For example, some demand deterrence platforms can post decoy advertisements to multiple commercial sex classified websites, and, by using software bots to chat with buyers who respond to ads to negotiate service levels and price, are able to automatically sample and directly measure buyer demand and service economics in regional areas [35, 36]. Relevant methodologies include the longitudinal and multilevel techniques recommended in Findings M1 and M2. Objectives include understanding the impact of policy changes and of actions such as buyer stings and deterrence campaigns on demand and ultimately on victim prevalence.

Suggested Actions: Leverage automation to measure marketplace attributes from sex advertising forums and buyer review sites. Develop methods for continuous measurement of buyer demand. Apply analysis to understand impact of policy changes and counter-demand activities.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Narrow	Medium	Medium

Recommendation M-6: Investigate remote sensing

Finding: Remote sensing is the science of obtaining information about objects or areas from a distance, typically from aircraft or satellites. Remote sensing has diverse applications including assessing agricultural crop health, estimating economic activity, characterizing land cover and land use, monitoring ecological change, and collecting military intelligence. With commercial satellites now imaging the Earth's entire land mass daily [78], vast areas that are otherwise inaccessible can be monitored, enabling detection of illicit activity around the globe.

Recently researchers have used satellite imagery to detect and monitor activities thought to be associated with labor trafficking such as deforestation near fisheries and construction of brick kilns in south Asia [79–81]. More generally, there may be a wide range of technologies that can be used to collect unconventional data related to both sex and labor trafficking. For

sex trafficking, possibilities include DNA collection for identifying buyers and pole cameras set up a locations such as hotels or known local trafficking "tracks."

Suggested Actions: Conduct systems analysis study to identify remote sensing techniques that show promise for detecting labor and sex trafficking. Based on the results of this study, initiate a pilot project to demonstrate the utility of promising remote sensing capabilities.

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Wide	Low	Long

3.5 TRAINING

3.5.1 Training Needs

In interviews for this report, a frequent topic of discussion was the need for better training for law enforcement and others involved in counter-trafficking efforts. Observations in our interviews included the following:

i. In some local law enforcement agencies, the biggest problem can be simply getting officers and their superiors to understand that there may be trafficking activity in their jurisdictions.

ii. Even when individuals in law enforcement are aware of the problem, some may treat victims as criminals, while others may not know how to proceed.

iii. There is a need for training to understand the specific implications of state statutes on how police and prosecutors conduct their work.

iv. Law enforcement may need training to understand certain technology aspects of trafficking, such as how to detect victims online and how to discover useful evidence in mobile device extraction data.

v. Beyond law enforcement, there is a need for training of others who may come in contact with victims or with trafficking activities, such as healthcare providers and hotel employees.

Some of these findings have been reported in other studies [82], which have also noted the need for training in

i. Best practices for acquiring and utilizing corroborating evidence, including improving interviews with human trafficking victims.

ii. The impact of trauma and violence on victim behavior (including criminal behavior) and testimony.

iii. Techniques for presenting evidence at trial with a victim who may be perceived as less credible.

These lists include training needs that, because they do not involve technology, are beyond the scope of this study. The findings and recommendations below focus on two ways that the worlds of training and technology intersect: technologies that can be used to improve counter-trafficking training, and training about technology aspects of counter-trafficking efforts.

3.5.2 Recommendations

Recommendation T-1: Train using serious games

> **Finding:** Individuals across the counter-trafficking community as well as others who may come into contact with victims often lack the training they need to respond appropriately.
>
> **Discussion:** Effective response to human trafficking is a complex process that involves not only law enforcement, other government agencies, and NGOs but also individuals in healthcare, hospitality, and other sectors who may come into contact with victims or with trafficking activity. Many of these individuals lack the training they need to identify trafficking or to respond appropriately and effectively. Training needs include the following:
>
> – Local law enforcement agencies may not be aware of trafficking in general and may not recognize that it is likely to occur in their jurisdictions.
> – Patrol officers, emergency room personnel, hotel workers, and others may not have sufficient training to spot indicators of involvement in trafficking or to respond appropriately when they do.
> – Law enforcement may need specialized training that addresses gathering evidence for and building effective counter-trafficking cases.
> – Individuals may need training to understand how to interact with and help victims who have been profoundly affected by trauma and violence.
>
> Challenges in meeting these needs include diversity of objectives, diversity of trainees, and the need for customization to address particular state statutes and specific ways that human trafficking presents in different localities.
>
> Serious games are a promising approach for providing this type of training. Serious games are a type of game, typically computer-based, whose primary purpose is learning, training, or problem-solving rather rather than entertainment [83,84]. Serious games have been applied in fields such as law enforcement, defense, education, medicine, and business. They are especially useful in allowing participants to undertake tasks and experience situations that are impossible or undesirable to due to cost, time, logistical, or safety considerations [85, 86]. Example training applications include helping doctors to learn surgical skills, training first responders to triage to mass casualty or hazardous materials incidents, and training intelligence analysts and operators to work with complex data sets or under challenging conditions [85, 87–89] .

Suggested Actions: Use serious games to address the diverse training needs of the counter-trafficking community. Start with a pilot program that addresses a specific need at the local or state level. Ensure that the approach is sufficiently flexible to eventually offer appropriate training at the local, state, and national levels to a wide array of individuals including LE, prosecutors, healthcare providers, NGOs, and hospitality employees.

Assessment:

Impact	Scope	Maturity	Timescale
Significant	Wide	High	Long

Recommendation T-2: Implement training for online investigations

Finding: Investigators need training for conducting online investigations efficiently and safely.

Discussion: The internet is conducive to illicit activities because it allows criminals to gain access to targeted markets worldwide while remaining largely anonymous [90]. As traffickers increasingly exploit the internet for advertisement, communications, and other aspects of their operations [91], law enforcement and prosecutors are challenged to conduct increasing amounts of their counter-trafficking investigations online. These investigations can include

- Inspection of commercial sex advertisements to identify potential victims;
- Inspection of social media posts to uncover trafficking activity;
- Use of search engines, white pages, reverse phone lookup, and other online tools to discover PII of traffickers and victims;
- Engagement on social media as part of sting operations.

The first three tasks on this list involve extracting information from online open sources, which is a complex process that requires significant training [90, 92]. Without this training, investigators may miss relevant information or may spend unnecessarily large amounts of time conducting time-consuming manual searches. The military and intelligence communities have developed extensive training programs for teaching open source intelligence tradecraft; it may be possible for law enforcement to leverage or at least emulate these.

The fourth task on the list, which involves actively engaging with others online, has potential to expose investigators to considerable personal risk if not done properly. For example, use of personal social media accounts to conduct stings or otherwise engage online with individuals associated with the commercial sex industry, can put investigators themselves at risk of being "doxxed" or investigated by traffickers, with potential for danger to themselves or others. Again, the military and intelligence communities have adopted tools and tradecraft

for managed attribution of online activities and personas that law enforcement may be able to leverage.

Suggested Actions: Develop training in the tools, techniques, and tradecraft for performing open source intelligence analysis, including management of undercover online personas. Where possible, seek to leverage or emulate capabilities developed in the military and intelligence communities.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Narrow	High	Short

Recommendation T-3: Implement training for technical evidence gathering

Finding: Investigators lack the technical training to exploit evidence extracted from mobile phones and other devices.

Discussion: A number of interviews conducted for this study highlighted the importance of data extracted from mobile phones in trafficking cases. Other studies have reported similar observations [17]. Given challenges associated with asking traumatized victims to testify, mobile phone data can often be the primary source of corroborating evidence of victim stories.

At the same time, mobile device extractions are complex, and investigators typically do not have sufficient technical background to discover relevant evidence in raw extraction data. Conversely, forensic units that provide extracted device images to law enforcement may not have knowledge of their needs. The result can be missed opportunities to extract relevant evidence, potentially leading to mistrial if the defense finds information not included in discovery.

Suggested Actions: Develop training that addresses the technology aspects of acquiring and utilizing corroborating evidence, such as better usage of data sets derived from forensic device extractions.

Assessment:

Impact	Scope	Maturity	Timescale
Incremental	Narrow	High	Short

4. IMPLEMENTATION STRATEGY AND FUTURE WORK

The previous section enumerated a set of recommendations to address needs in five focus areas: marketplace interdiction operations, investigations and prosecution, information sharing, measurement of trafficking, and training. This section proposes a strategy for implementing these recommendations. Section 4.1 introduces a general implementation framework for maturing new technologies from initial concept development to operational deployment. The framework also describes organizational roles and structures for this process. Section 4.2 then applies this framework to the specific recommendations made in Section 3. A key result of Section 4.2 is a prioritized ranking of the recommendations, ranging from "quick wins" that implement mature technologies to "moonshots" that require substantial technology development. This ranking provides a flexible plan of action depending on available resources and appetite for risk.

4.1 IMPLEMENTATION FRAMEWORK

While each of the recommendations of this study ultimately requires its own unique implementation strategy, there will be some commonality to these strategies. This section describes common processes and organizational structures for technology development associated with the recommendations of Section 3.

Implementing a technology-based capability generally requires some or all of the following five stages:

Concept development: Identify specific needs or gaps, technologies that address them, and potential risks of those technologies; choose an approach that best fills the needs or gaps while minimizing risks; identify collaborators and data for prototyping; identify a technology transition partner.

Capability prototyping: Implement a proof-of-concept demonstration with the technologies, collaborators, and data identified during concept development.

System development: Work with the technology transition partner to extend the prototype by improving its scope, functionality, usability, and/or robustness.

Technology transition: Harden the system and deploy into operational use.

Test and evaluation: Measure the utility of the deployed system to end users in order to modify or improve its capabilities.

The relative importance of each of these stages varies across the recommendations depending on the maturity of the underlying technology. For recommendations based on mature technologies, capability prototyping and even system development may be unnecessary. On the other hand, for recommendations based on novel technologies, the appropriate strategy may be simply to monitor progress until sufficient maturity is reached to begin the development process.

In addition to DHS S&T and alike federal agencies, organizations that participate in this process may include federally funded research and development centers (FFRDCs), operational transition partners, NGOs, industry, and academia. Participants may contribute at some or all stages; for example, academia is likely to participate in the initial concept development and prototyping, while industry may contribute at later stages when the requirements are well defined and the system is more robust. A summary of anticipated participation at each stage appears in Figure 3.

Development Stage	Academia	FFRDCs	NGOs	Industry	Operational Partners
Concept Development	✓	✓	✓		✓
Capability Prototyping	✓	✓	✓	✓	✓
System Development		✓	✓	✓	✓
Technology Transition		✓	✓	✓	✓
Test and Evaluation		✓			✓

Figure 3. Participants in the development process. Checks denote participation in a given development stage.

To achieve the collaboration implied by Figure 3, organizational structures that have been successfully deployed on Department of Defense (DoD) projects may prove useful. One is the working group or community of practice model, which brings together individuals from disparate organizations who seek to achieve a common goal or goals. This model has been used in the DoD to bring together sensor manufacturers, software developers, and intelligence analysts in order to establish standards for collecting and processing certain types of sensor data [93]. It is relevant to many recommendations of this study, especially those pertaining to data sharing and measuring of trafficking, and it arises specifically in Recommendations S-1 and M-3. Another model is the industry consortium, used by the Defense Advanced Research Projects Agency (DARPA) Memex program and others to develop tools for mining the dark web for counter-human trafficking and other applications [94]. In this model, a government agency provides funding, direction, and oversight; a number of commercial entities or performers provide technology components; and an FFRDC conducts integration and/or test and evaluation functions. The Memex team worked closely with the New York County District Attorneys Office and other law enforcement agencies to ensure that program efforts aligned with operational needs. Quarterly hackathons were used to focus performers on collaborating to solve specific challenge problems. Software developed under the Memex program is freely available online [95] and was incorporated into operational tools by NGOs and other organizations.

Pulling these pieces together, a generic development program might work as follows. In the concept development stage, FFRDCs team with academia, NGOs, and an operational partner,

such as a prosecutorial agency, to scope the problem and identify potential technical solutions. This process may take 3-6 months. Next, in the prototyping stage, multiple performers are funded simultaneously to yield a diversity of solutions. At the same time, FFRDCs and academia continue to establish the technology vision and perform test and evaluation to ensure go/no-go criteria are met before additional investment. Continued partnership with the operational agency, along with discussions with other agencies, ensures that the design meets operational needs. This stage may take 1-2 years. Next, during the system development stage, industry takes a leadership role to develop the prototype into robust commercial software, while FFRDCs continue to perform test and evaluation. A wide-range of stakeholders should be included at this stage to assess long term applicability and utility across a diverse mission set. Technology transition agreements should be in place so that product funding can originate largely from vendors and stakeholders, with federal funding of test and evaluation and program management only. This stage may take 1-3 years. Finally, the hardened system is transitioned into operational use, with continued funding for test and evaluation and operations and maintenance to ensure long term reliability of the capability.

4.2 IMPLEMENTATION ROADMAP

This section presents an implementation roadmap that consists of a prioritized ordering of the 29 recommendations of Section 3 and a phased set of actions for the 19 highest-priority recommendations. The prioritized ordering combines the impact, scope, maturity, and timescale assessments from Section 3 with feedback from the DHS Human Trafficking Advisory Group. The phased set of actions applies the framework of Section 4.1 to each of the highest-priority recommendations.

4.2.1 DHS Human Trafficking Advisory Group Feedback

The DHS Human Trafficking Advisory Group met July 12, 2019 in Washington, DC to review and discuss the recommendations of this report. Advisory Group members received a briefing that summarized the roadmap objective and approach and presented each of the recommendations of Section 3 in detail. Feedback was elicited from the group through a round-table discussion and through handouts that provided the opportunity to rank or highlight the recommendations attendees assessed to be most important.

In general, the Advisory Group was most interested in recommendations in the investigations and prosecution category, especially technologies with potential to save time for investigators or agents so that they can focus on the complexity of cases. Within this category, the group was particularly interested in establishing a data template repository (I5) and in technologies for financial record analysis (I9); some interest was also shown for monitoring online vulnerability (I2) and developing data enrichment capabilities (I6). Across all categories, the group tended to value ease of implementation over technologies with high risk and reward. Specifically, they favored recommendations with high maturity and wide scope (such as I5, S1, and M3), recommendations that leverage existing work and capabilities (such as S1 and M3), and recommendations that involve establishing working groups (including I5, M1, and M3). Other recommendations that received high prioritization were encouraging novel data sharing (S2) and training for technical evidence gathering (T3).

4.2.2   Phased Implementation Approach

In keeping with the DHS S&T guidance to make recommendations both for "short-term, small-scale technology for quick implementation, as well as long-term, large-scale nationally disruptive technology" [6], this section provides a prioritization of the 29 recommendations presented in Section 3. In addition, for each 19 of the highest priority recommendations, a phased implementation approach is provided.

As shown in Figure 4, we prioritize the 29 recommendations into four categories: "immediate action," "longer-term focus," "moonshots," and "additional efforts." Within each category, recommendations appear in the same unprioritized order as in Section 3. Generally, immediate action recommendations scored high in our assessment, generated substantial advisory group interest, and are relatively mature and/or can be implemented relatively quickly. They are divided into two subcategories: "quick wins," which have significant impact, wide scope, high maturity, and a short timeline; and "some development," which have significant impact and wide scope but may require time for development. Recommendations in the the immediate action category should receive highest priority initially from DHS S&T. The longer-term focus category is a mixed group of recommendations that scored relatively high in two or more categories of our assessment and received some advisory group interest. If time and resources permit, DHS S&T should consider funding efforts in these areas as well. Moonshots are high-risk, high-reward recommendations: they have significant impact and wide scope but require significant development time and effort. DHS S&T should at least monitor activity in these areas and should consider supporting research efforts. The final category collects recommendations that did not fit into the other three categories. DHS S&T should place a lower priority on these but should consider funding them if opportunities and interest exist and resources are available.

Immediate Action	Longer-Term Focus	Moonshots	Additional Efforts
Quick Wins	I2 Monitor online vulnerability indicators	S4 Leverage secure computing	O1 Improve online persona management
O3 Establish buyer repository	I4 Develop case knowledge management system	M6 Investigate remote sensing	O4 Improve sting operations
I5 Establish data template repository	I11 Employ speech transcription tools	O2 Determine buyer identities	I3 Develop advertisement platform federated search
I9 Enhance financial record analysis	S2 Encourage novel data sharing		I7 Use geospatial information effectively
M3 Establish data standards working group	S3 Encourage local data sharing		I10 Utilize surveillance video analytics
S1 Leverage existing sharing mechanisms	M1 Quantify impact of actions		S5 Implement P2P markets for victim services
Some Development	T3 Implement training for technical evidence gathering		M2 Develop geospatially hierarchical methods
I1 Create trafficking signature repository			M4 Build victim prevalence models
I6 Develop data enrichment capabilities			M5 Build buyer demand models
I8 Improve call record analysis			T2 Implement training for online investigations
T1 Train using serious games			

Figure 4. Prioritization of recommendations.

For the 19 recommendations in the immediate action, longer-term focus, and moonshots categories, Figure 5 provides a phased set of actions or next steps. The timing of the phases is independent for each recommendation; it depends on the difficulty of implementation and on the available resources. For example, each phase of the quick win recommendations may take months, while phases of the moonshot recommendations may take years. For recommendations that involve developing a new capability (such as I4, I6, I8, I11, and O2), the phased approach follows the framework introduced in Section 4.1. In contrast, actions for recommendations associated with existing or mature capabilities, such as many of the quick wins, involve reaching out to existing efforts or capabilities and determining how they can support counter-human trafficking efforts.

Priority			Recommendation	Phase I	Phase II	Phase III
Immediate Action	Quick Wins	O3	Establish buyer repository	Identify stakeholders	Form working group	Establish repository
		I5	Establish data template repository	Identify stakeholders	Form working group	Establish repository
		I9	Enhance financial record analysis	Identify stakeholders	Enable data access	Develop tools and training for financial analysis
		M3	Establish data standards working group	Reach out to previous and existing standards efforts	Form working group	Initiate pilot program to develop and disseminate standards
		S1	Leverage existing sharing mechanisms	Investigate utility of existing mechanisms such as RISS	Coordinate with standards working group (M3)	Improve sharing within LE and implement sharing with non-LE
	Some Development	I1	Create trafficking signature repository	Identify existing signature databases within community	Create centralized repository	Provide repository access and updating mechanisms
		I6	Develop data enrichment capabilities	Identify stakeholders and develop concept	Build prototype leveraging existing language & image processing tools	Harden prototype and transition to operational use
		I8	Improve call record analysis	Identify stakeholders and develop concept	Build prototype	Harden prototype and transition to operational use
		T1	Train using serious games	Develop game concept and prototype	Implement pilot program addressing specific training need	Expand pilot to address multiple needs at local/state/national levels
Longer-Term Focus		I2	Monitor online vulnerability indicators	Assess legal and privacy aspects	Identify stakeholders and develop concept	Leverage signature repository (I1) to develop prototype
		I4	Develop case knowledge management system	Identify stakeholders and develop concept	Build prototype	Harden prototype and transition to operational use
		I11	Employ speech transcription tools	Identify stakeholders and develop concept	Build prototype	Harden prototype and transition to operational use
		S2	Encourage novel data sharing	Identify data sources and organizations willing to share them	Implement pilot data sharing program	Extend pilot to other sources or organizations
		S3	Encourage local data sharing	Identify locality with organizations interested in data sharing	Implement pilot data sharing program	Extend pilot to other sources, organizations, or localities
		M1	Quantify impact of actions	Convene multidisciplinary workshop	Implement pilot program	Use results to identify and fund successful programs and strategies
		T3	Implement training for technical evidence gathering	Identify stakeholders, training needs, and approach	Partner with LE organization to implement pilot training program	Integrate into serious games (T1)
Moonshots		S4	Leverage secure computing	Monitor advances in DHS S&T and commercial sector	Identify stakeholders and secure computing approach	Implement pilot program
		M6	Investigate remote sensing	Conduct systems analysis to identify approach	Initiate pilot program for most promising approach	Extend to other sensing modalities
		O2	Determine buyer identities	Identify stakeholders and develop concept	Build prototype	Harden prototype and transition to operational use

Figure 5. Phased implementation of recommendations.

4.3 FUTURE WORK

As noted in Section 1.1, human trafficking is global in scope, encompassing sex and labor exploitation as well as domestic servitude, forced marriage, forced criminality, and other types of exploitation. This roadmap focuses on domestic sex trafficking, with an emphasis on the interplay between federal and state organizations. Despite this focus, many of our recommendations are likely to be applicable to other types of trafficking. For example, many recommendations for improving investigations and prosecution (Section 3.2) are relevant to any complex casework. Similarly, those for improving measurement of trafficking (Section 3.4) apply to all types of trafficking; and some, such as the use of remote sensing, may in fact provide more utility for international labor and other trafficking than for domestic sex trafficking.

At the same time, a detailed technical analysis of the broader global trafficking problem would likely identify other recommendations beyond those in this report. Some extensions of this work that may be of interest to DHS S&T include analyses of domestic labor trafficking, trafficking associated with immigration into the U.S., and international trafficking networks operating in multiple countries including the U.S.

This page intentionally left blank.

GLOSSARY

API	Application Programming Interface
ATIX	Automated Trusted Information Exchange
BAAT	Banks Alliance Against Trafficking
C-HT	Counter-Human Trafficking
CDC	Centers for Disease Control and Prevention
CDR	Call Detail Record
CDTC	Counter Trafficking Data Collaborative
CEASE	Cities Empowered Against Sexual Exploitation
COI	Communities of Interest
CTR	Currency Transaction Report
DARPA	Defense Advanced Research Projects Agency
DHS	Department of Homeland Security
DOD	Department of Defense
DOJ	Department of Justice
DOS	Department of State
ECCIE	Escort Client Community Information Exchange
FATF	The Financial Action Task Force
FBI	Federal Bureau of Investigation
FFRDC	Federally Funded Research and Development Center
FinCEN	Financial Crimes Enforcement Network
FIU	Financial Intelligence Units
FOSTA	Allow States and Victims to Fight Online Sex Trafficking Act
GAN	Generative Adversarial Networks
GJXDM	Global Justice XML Data Model
HIPAA	Health Insurance Portability and Accountability Act
HIV	Human Immunodeficiency Virus
HHS	Health and Human Services
HSI	Homeland Security Investigations
HSTC	Human Smuggling and Trafficking Center

GLOSSARY
(Continued)

HT	Human Trafficking
HTAM	Human Trafficking Analysis Model
HTRS	Human Trafficking Reporting System
ICAC	Internet Crimes Against Children
ICD	International Statistical Classification of Diseases
IEPD	Information Exchange Packet Documentation
LE	Law Enforcement
LEEP	Law Enforcement Enterprise Portal
LEXS	Logical Entity Exchange Specifications
MIT LL	Massachusetts Institute of Technology Lincoln Laboratory
MSE	Multiple Systems Estimation
N-DEX	National Data Exchange
NAICS	North American Industry Classification System
NGO	Non-Governmental Organization
NICS	Next Generation Incident Command System
NIEM	National Information Exchange Model
OFAC	Office of Foreign Assets Control
OVC	Office for Victims of Crime
OTIP	Office on Trafficking in Persons
P2P	Peer-to-peer
PII	Personally Identifiable Information
PNNL	Pacific Northwest National Laboratory
RISS	Regional Information Sharing System
S&T	Science and Technology

GLOSSARY
(Continued)

SAR	Suspicious Activity Report
SESTA	Stop Enabling Sex Traffickers Act
SUSB	Statistics of U.S. Business
STD	Sexually Transmitted Disease
SVI	Social Vulnerability Index
TFFC	The Office of Terrorist Financing and Financial Crimes
TIGER	Topologically Integrated Geographic Encoding and Referencing
TIMS	Trafficking Information Management System
TIP	Trafficking in Persons
TVPA	Trafficking Victim Protection Act
UCR	Unified Crime Reporting
U.S.	United States
VSP	Victim Service Provider
WiGLE	The Wireless Geographic Logging Engine
WSIN	Western States Information Network

This page intentionally left blank.

REFERENCES

[1] K. Kangaspunta, F. Sarrica, R. Johansen, J. Samson, A. Rybarska, and K. Whelan, *2018 Global Report on Trafficking in Persons*, United Nations Office on Drugs and Crime (2018).

[2] The Department of State, "Trafficking in persons report," URL https://www.state.gov/wp-content/uploads/2019/06/2019-Trafficking-in-Persons-Report.pdf, accessed June 16, 2019.

[3] The Presidents Interagency Task Force to Monitor and Combat Trafficking in Persons, *Report on U.S. Government Efforts To Combat Trafficking In Persons*, U.S. State Department (2017).

[4] Administration for Children and Families, "Federal government efforts to combat human trafficking," (2019), URL https://www.acf.hhs.gov/otip/resources/federal-efforts, accessed June 11, 2019.

[5] The Department of Homeland Security, "Homeland security enterprise," URL https://www.dhs.gov/topic/homeland-security-enterprise, accessed June 22, 2019.

[6] The Department of Homeland Security, "S&T combatting human trafficking using social science," (2019), URL https://www.dhs.gov/science-and-technology/news/2019/01/30/st-combatting-human-trafficking-using-social-science, accessed June 5, 2019.

[7] H.J.D. Reynolds, C.D. Engholm, J.A. Sloboda, and C. Heine, "Human trafficking systems analysis," MIT Lincoln Laboratory, Lexington, MA (United States), Technical rep. (2019).

[8] Polaris Project, *The Typology of Modern Slavery: Defining Sex and Labor Trafficking in the United States* (2017).

[9] P. Szekely, C.A. Knoblock, J. Slepicka, A. Philpot, A. Singh, C. Yin, D. Kapoor, P. Natarajan, D. Marcu, K. Knight, et al., "Building and using a knowledge graph to combat human trafficking," in *International Semantic Web Conference*, Springer (2015), pp. 205–221.

[10] H. Alvari, P. Shakarian, and J.K. Snyder, "Semi-supervised learning for detecting human trafficking," *Security Informatics* 6(1), 1 (2017).

[11] L. Li, O. Simek, A. Lai, M. Daggett, C.K. Dagli, and C. Jones, "Detection and characterization of human trafficking networks using unsupervised scalable text template matching," in *2018 IEEE International Conference on Big Data (Big Data)*, IEEE (2018), pp. 3111–3120.

[12] E. Tong, A. Zadeh, C. Jones, and L.P. Morency, "Combating human trafficking with deep multimodal models," *arXiv preprint arXiv:1705.02735* (2017).

[13] R.A. Konrad, A.C. Trapp, T.M. Palmbach, and J.S. Blom, "Overcoming human trafficking via operations research and analytics: Opportunities for methods, models, and applications," *European Journal of Operational Research* 259(2), 733–745 (2017).

[14] J. McKenzie, *Identifying and Mitigating Human Trafficking Risk through the Use of Financial Institutions*, Ph.D. thesis, Utica College (2019).

[15] S.J. Kreyling, C.L. West, and J. Olson, "Technology and research requirements for combating human trafficking: Enhancing communication, analysis, reporting, and information sharing," Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), Technical rep. (2011).

[16] C.L. West and S.J. Kreyling, "Law enforcement technology roadmap: Lessons to date from the Northwest technology desk and the Northwest FADE pilots," Pacific Northwest National Laboratory (PNNL), Richland, WA (United States), Technical rep. (2011).

[17] M. Latonero, J. Musto, Z. Boyd, E. Boyle, A. Bissel, K. Gibson, and J. Kim, *The rise of mobile and the diffusion of technology-facilitated trafficking*, University of Southern California, Center on Communication Leadership (2012).

[18] D. Boyd, H. Casteel, M. Thakor, and R. Johnson, "Human trafficking and technology: A framework for understanding the role of technology in the commercial sexual exploitation of children in the us," *Microsoft Research, Cambridge, MA* (2011).

[19] M. Morris, *The Impact of Advancing Technologies Upon Global Human Trafficking and Sexual Exploitation in Society Today*, Ph.D. thesis, Utica College (2017).

[20] C. Aghazarm, F. Laczko, et al., "Human trafficking: new directions for research," *Geneva: International Organization for Migration* (2008).

[21] R. Weitzer, "New directions in research on human trafficking," *The ANNALS of the American Academy of Political and Social Science* 653(1), 6–24 (2014).

[22] A. Farrell and I. de Vries, "Measuring the nature and prevalence of human trafficking," *The Palgrave International Handbook of Human Trafficking* pp. 1–16 (2019).

[23] International Labour Organization, *Global estimates of modern slavery: Forced labour and forced marriage*, International Labour Office, Geneva, Switzerland (2017).

[24] International Labour Organization, *Profits of Poverty: The Economics of Forced Labour*, IInternational Labour Office, Geneva, Switzerland (2014).

[25] N. Kulish, F. Robles, and P. Mazzei, "Behind illicit massage parlors lie a vast crime network and modern indentured servitude," (2019), URL https://www.nytimes.com/2019/03/02/us/massage-parlors-human-trafficking.html, accessed May 15, 2019.

[26] D. Barry and J.E. Singer, "The case of Jane Doe Ponytail: An epic tragedy on a small block in Queens," (2018), URL https://www.nytimes.com/interactive/2018/10/11/nyregion/sex-workers-massage-parlor.html, accessed May 30, 2019.

[27] T. Rosenberg, "Fighting sex trafficking at the truck stop," (2019), URL https://www.nytimes.com/2019/04/02/opinion/fighting-sex-trafficking-at-the-truck-stop.html, accessed May 30, 2019.

[28] V. Bouch, B. Carr, T. Grover, and U. Roy, "Human trafficking data (HTD) open-source database of federally prosecuted human trafficking cases in the United States," (2019), URL http://www.humantraffickingdata.org, accessed April, 2019.

[29] Polaris Project, "2017 statistics from the National Human Trafficking Hotline and BeFree Textline," Technical rep. (2017).

[30] Child Welfare Information Gateway, *Human Trafficking and Child Welfare: A Guide for Caseworkers*, Childrens Bureau, Administration for Children and Families, U.S. Department of Health and Human Services (2017).

[31] Demand Abolition, "Who buys sex? Understanding and disrupting illicit market demand," Technical rep. (2018).

[32] Manhattan District Attorney's Office, "Sex trafficker who targeted homeless and runaway teens convicted at trial," (2017), URL https://www.manhattanda.org/da-vance-sex-trafficker-who-targeted-homeless-and-runaway-teens-convicted-trial/, accessed June 14, 2019.

[33] Manhattan District Attorney's Office, "Froilan rosado sentenced to 7-to-14 years for prostituting teenage girls," (2015), URL https://www.manhattanda.org/da-vance-froilan-rosado-sentenced-7-14-years-prostituting-teenage-girls/, accessed June 28, 2019.

[34] National Center for Missing and Exploited Children, "National Center for Missing and Exploited Children amicus curiae supporting respondents, J.S., S.L., and L.C. v. Village Voice Media Holdings, L.L.C., d/b/a Backpage.com; Backpage.com L.L.C." (2014).

[35] R. Spectre, "Beyond Backpage: Buying and selling sex in the United States one year later," childsafe.ai, Technical rep. (2019).

[36] T. Rosenberg, "A.I. joins the campaign against sex trafficking," *The New York Times* (2019), URL https://www.nytimes.com/2019/04/09/opinion/ai-joins-the-campaign-against-sex-trafficking.html, accessed June 2, 2019.

[37] S. Downing, "How to Scare a Predator: Neutering the Sex Industry One Buyer at a Time," (2018), URL https://chronicleofsocialchange.org/analysis/how-to-catch-a-predator-neutering-the-sex-industry/29426, accessed July 2019.

[38] L. Mucciolo, "Frontline: Sex trafficking in America," (2019), URL https://www.pbs.org/wgbh/frontline/film/sex-trafficking-in-america/, accessed June 7, 2019.

[39] K. Hill, "People you may know: A controversial Facebook feature's 10-year history," (2018), URL https://gizmodo.com/people-you-may-know-a-controversial-facebook-features-1827981959, accessed June 12, 2019.

[40] A. Abbasi and H. Chen, "Applying authorship analysis to extremist-group web forum messages," *IEEE Intelligent Systems* 20(5), 67–75 (2005).

[41] L. Li and W.M. Campbell, "Matching community structure across online social networks," *arXiv preprint arXiv:1608.01373* (2016).

[42] Shared Hope International, "The protected innocence challenge legislative framework and methodology 2018 toolkit," URL https://sharedhope.org/wp-content/uploads/2018/11/2018ProtectedInnocenceChallengeToolkit.pdf, accessed June 3, 2019.

[43] Financial Action Task Force, "Financial Flows from Human Trafficking," Asia/Pacific Group on Money Laundering, Technical rep. (2018).

[44] B. Anthony, "On-ramps, intersections, and exit routes: a roadmap for systems and industries to prevent and disrupt human trafficking," Polaris Project, Technical rep. (2018).

[45] R.S. Portnoff, D.Y. Huang, P. Doerfler, S. Afroz, and D. McCoy, "Backpage and bitcoin: Uncovering human traffickers," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2017), pp. 1595–1604.

[46] F.C.E. Network, "Guidance on recognizing activity that may be associated with human smuggling and human trafficking financial red flags," Technical rep. (2014).

[47] J. Brunner, "Getting to good human trafficking data," WSD Handa Center for Human Rights and International Justice, Stanford University, Palo Alto, CA (United States), Technical rep. (2018).

[48] M. Coakley, "Massachusetts Interagency Human Trafficking Policy Task Force: Findings and recommendations," Massachusetts Office of Attorney General, Boston, MA (United States), Technical rep. (2013).

[49] J. Chua, "WSIN offers agencies assistance on human trafficking cases," (2019), URL https://inpublicsafety.com/2019/01/wsin-offers-agencies-assistance-on-human-trafficking-cases/, accessed June 21, 2019.

[50] Department of Homeland Security Science and Technology Directorate, "Data privacy," (2019), URL https://www.dhs.gov/science-and-technology/cybersecurity-privacy, accessed June 13, 2019.

[51] D.W. Archer, D. Bogdanov, Y. Lindell, L. Kamm, K. Nielsen, J.I. Pagter, N.P. Smart, and R.N. Wright, "From keys to databases: Real-world applications of secure multi-party computation," *The Computer Journal* 61(12), 1749–1771 (2018).

[52] R.L. Lagendijk, Z. Erkin, and M. Barni, "Encrypted signal processing for privacy protection: Conveying the utility of homomorphic encryption and multiparty computation," *IEEE Signal Processing Magazine* 30(1), 82–105 (2012).

[53] S. Yakoubov, V. Gadepally, N. Schear, E. Shen, and A. Yerukhimovich, "A survey of cryptographic approaches to securing big-data analytics in the cloud," in *2014 IEEE High Performance Extreme Computing Conference (HPEC)*, IEEE (2014), pp. 1–6.

[54] R.S. Rani, "Homes to heal trafficked children," *The New York Times* (2019), URL https://www.nytimes.com/2019/06/19/opinion/foster-child-trafficking.html, accessed June 26, 2019.

[55] H.J. Clawson, M. Layne, and K. Small, *Estimating human trafficking into the United States: Development of a methodology*, Caliber, Fairfax, VA (United States) (2006).

[56] E.U. Savona and S. Stefanizzi, *Measuring human trafficking: Complexities and pitfalls*, Springer (2007).

[57] M. Cruyff, J. van Dijk, and P.G. van der Heijden, "The challenge of counting victims of human trafficking: Not on the record: A multiple systems estimation of the numbers of human trafficking victims in the netherlands in 2010–2015 by year, age, gender, and type of exploitation," *Chance* 30(3), 41–49 (2017).

[58] D.P. Durgana and P.L. Zador, "Fighting slavery through statistics: A discussion of five promising methods to estimate prevalence in the United States," *Chance* 30(3), 50–53 (2017).

[59] M.L. Dank, B. Khan, P.M. Downey, C. Kotonias, D. Mayer, C. Owens, L. Pacifici, and L. Yu, "Estimating the size and structure of the underground commercial sex economy in eight major U.S. cities," Urban Institute, Washington, DC (United States), Technical rep. (2014).

[60] J.M. Wilson and E. Dalton, *Human trafficking in Ohio: Markets, responses, and considerations*, Rand Corporation (2007).

[61] M. Ibanez and D.D. Suthers, "Detection of domestic human trafficking indicators and movement trends using content available on open internet sources," in *2014 47th Hawaii International Conference on System Sciences*, IEEE (2014), pp. 1556–1565.

[62] N.A. Giacobe, J.B. Altmire, A.E. Forster, A.C. Jackson, E.W. Raibick, J.A. Reep, R.Y. Tsang, and P.K. Forster, "Characterizing sex trafficking in Pennsylvania for law enforcement," in *2016 IEEE Symposium on Technologies for Homeland Security (HST)*, IEEE (2016), pp. 1–5.

[63] K.H. Brodersen, F. Gallusser, J. Koehler, N. Remy, S.L. Scott, et al., "Inferring causal impact using Bayesian structural time-series models," *The Annals of Applied Statistics* 9(1), 247–274 (2015).

[64] G.W. Imbens and D.B. Rubin, *Causal inference in statistics, social, and biomedical sciences*, Cambridge University Press (2015).

[65] N. Jakobsson and A. Kotsadam, "The law and economics of international sex slavery: prostitution laws and trafficking for sexual exploitation," *European Journal of Law and Economics* 35(1), 87–107 (2013).

[66] M.L. Antonio, *Quantitative Study of the Cost and Benefit Analysis of Human Trafficking*, Ph.D. thesis, Northcentral University (2019).

[67] J.J. Hox, M. Moerbeek, and R. Van de Schoot, *Multilevel analysis: Techniques and applications*, Routledge (2017).

[68] J.W. Graham, "Multiple imputation and analysis with multilevel (cluster) data," in *Missing Data*, Springer, pp. 133–150 (2012).

[69] M. Moerbeek and S. Teerenstra, *Power analysis of trials with multilevel data*, Chapman and Hall/CRC (2015).

[70] S.V. Subramanian, "Multilevel modeling," Office of Behavioral and Social Sciences Research, National Institutes of Health,Bethesda, MD (United States), Technical rep. (2019).

[71] University of Bristol Centre for Multilevel Modeling, "What are multilevel models and why should I use them?" (2019), URL http://www.bristol.ac.uk/cmm/learning/multilevel-models/what-why.html, accessed June 16, 2019.

[72] J. De Leeuw, E. Meijer, and H. Goldstein, *Handbook of multilevel analysis*, Springer (2008).

[73] D.J. Corliss and H.M. Hill, "Meta-analysis of human trafficking in the United States," Peace-Work, Plymouth, MI (United States), Technical rep. (2016).

[74] M.J. Daniels and J.W. Hogan, *Missing data in longitudinal studies: Strategies for Bayesian modeling and sensitivity analysis*, Chapman and Hall/CRC Press (2008).

[75] J.G. Ibrahim and G. Molenberghs, "Missing data methods in longitudinal studies: a review," *Test* 18(1), 1–43 (2009).

[76] H. Goldstein and G. Woodhouse, "Multilevel models with missing data," in *Eleventh International Workshop on Statistical Modelling, Orvieto, Italy* (1996).

[77] M. Farley, E. Schuckman, J.M. Golding, K. Houser, L. Jarrett, P. Qualliotine, and M. Decker, "Comparing sex buyers with men who don't buy sex," *San Francisco, CA: Prostitution Research & Education* (2011).

[78] Tim Fernholz, "The company photographing every spot of land on earth, every single day," (2017), URL https://www.state.gov/wp-content/uploads/2019/06/2019-Trafficking-in-Persons-Report.pdf, accessed July 23, 2019.

[79] C. McGoogan and M. Rashid, "Satellites reveal 'child slave camps' in UNESCO-protected park in Bangladesh," (2016), URL https://www.telegraph.co.uk/technology/2016/10/23/satellites-reveal-child-slave-camps-in-unesco-protected-park-in/, accessed July 15, 2019.

[80] B. Jackson, K. Bales, S. Owen, J. Wardlaw, and D. Boyd, "Analysing slavery through satellite technology: How remote sensing could revolutionise data collection to help end modern slavery," *Journal of Modern Slavery* 4(2), 169–199 (2018).

[81] D.S. Boyd, B. Jackson, J. Wardlaw, G.M. Foody, S. Marsh, and K. Bales, "Slavery from space: Demonstrating the role for satellite remote sensing to inform evidence-based action related to UN SDG number 8," *ISPRS journal of photogrammetry and remote sensing* 142, 380–388 (2018).

[82] A. Farrell, J. McDevitt, R. Pfeffer, S. Fahy, C. Owens, M. Dank, and W. Adams, "Identifying challenges to improve the investigation and prosecution of state and local human trafficking cases," (2012).

[83] D.R. Michael and S.L. Chen, *Serious games: Games that educate, train, and inform*, Muska & Lipman/Premier-Trade (2005).

[84] T. Susi, M. Johannesson, and P. Backlund, "Serious games: An overview," (2007).

[85] M.P. Daggett, T.J. Dasey, A.S. Norige, and R.M. Seater, "Game-based human-system analysis for national security R&D," *Lincoln Laboratory Journal* 23(1), 4–7 (2019).

[86] K. Corti, "Games-based learning; a serious business application," *Informe de PixelLearning* 34(6), 1–20 (2006).

[87] M.P. Daggett, D.J. Hannon, M.B. Hurley, and J.O. Nwagbaraocha, "Serious games for collaborative dark network discovery," *MIT Lincoln Laboratory Journal* 23(1), 81–107 (2019).

[88] G.M. Jones, M.C. Gombolay, R.E. Jensen, and S.L. Nelson, "Strike group defender," *MIT Lincoln Laboratory Journal* 23.

[89] M. Graafland, J.M. Schraagen, and M.P. Schijven, "Systematic review of serious games for medical education and surgical skills training," *British Journal of Surgery* 99(10), 1322–1330 (2012).

[90] D.L. Carter, "Law enforcement intelligence: A guide for state, local, and tribal law enforcement agencies," Office of Community Oriented Policing Services, U. S. Department of Justice, Washington, DC (United States), Technical rep. (2009).

[91] M. Latonero, *Human trafficking online: The role of social networking sites and online classifieds*, University of Southern California, Center on Communication Leadership (2011).

[92] M. Bazzell, *Open source intelligence techniques: resources for searching and analyzing online information*, CreateSpace Independent Publishing Platform (2016).

[93] M.S. Lee, "Innovative collaborative community helps warfighters connect the dots," (2010), URL https://www.mitre.org/publications/project-stories/innovative-collaborative-community-helps-warfighters-connect-the-dots, accessed June 22, 2019.

[94] C. Pellerin, "DARPA program helps to fight human trafficking," (2017), URL https://dod.defense.gov/News/Article/Article/1041509/darpa-program-helps-to-fight-human-trafficking/, accessed June 22, 2019.

[95] Memex Program, "Memex tools and components," (2019), URL https://github.com/darpa-i2o/memex-program-index, accessed June 22, 2019.

67

This page intentionally left blank.

APPENDIX A: LIST OF ORGANIZATIONS INTERVIEWED

This appendix details some the 35+ organizations that were interviewed during the creation of this document.

A.1 FEDERAL ORGANIZATIONS

Department of Homeland Security, Blue Campaign — A national public awareness campaign, designed to educate the public, law enforcement and other industry partners on how to recognize the indicators of human trafficking and respond to potential cases. Works with other agencies to create general awareness training and materials for law enforcement to better identify victims.

Department of Homeland Security, Homeland Security Investigations, Child Exploitation Investigations Unit — The unit uses investigative techniques to identify consumers, producers, and distributors of child pornography and those engaging in child sex tourism. Collaborates with law enforcement partners around the world including Interpol and the National Center for Missing & Exploited Children and participates on all 61 Internet Crimes Against Children (ICAC) Task Forces in the U.S.

Department of Homeland Security, Homeland Security Investigations — An investigative arm of the Department of Homeland Security combating criminal organizations that illegally exploit travel, trade, financial, and immigration systems. Has a broad legal authority to enforce a diverse array of federal statutes, including investigation of human smuggling and trafficking.

Department of Homeland Security, Human Smuggling and Trafficking Center — Serves as an intelligence clearinghouse for information related to human smuggling and trafficking within the U.S. Government. Primary functions include facilitating dissemination of all source information, preparing strategic assessments, identifying issues for inter-agency attention, coordinating select operational initiatives, and working with allied foreign governments and organizations.

Federal Bureau of Investigation, Child Exploitation Task Force — Task forces which combine the resources of the FBI with other federal, state, and local law enforcement agencies across each of the FBIs 56 field offices. Investigative priorities include child abductions, child sexual exploitation enterprises, contact offenses against children, trafficking of child pornography, and other crimes against children.

U.S. Department of State, Bureau of Intelligence and Research — The primary mission of the Bureau of Intelligence and Research is to harness all-source intelligence to serve U.S. diplomacy and provide value-added independent analysis of events to U.S. State Department policymakers.

U.S. Department of State, Office to Monitor and Combat Trafficking in Persons — Partners with foreign governments, international organizations, other federal agencies,

civil society, the private sector, and survivors of human trafficking to develop and implement effective strategies to confront modern slavery. The office is responsible for bilateral and multilateral diplomacy, targeted foreign assistance, and public engagement on trafficking in persons.

A.2 STATE AND LOCAL ORGANIZATIONS

Clackamas County District Attorneys Office — Works through the Clackamas County Sex Trafficking Multidisciplinary Team to prosecute traffickers in Clackamas County, Oregon using a victim-centered approach; educates local law enforcement on methods for detecting and disrupting trafficking in their communities.

Connecticut Intelligence Center — A U.S. Department of Homeland Security intelligence fusion center, which serves as the focal point within the state of Connecticut for the receipt, analysis, gathering, and sharing of threat-related information among federal, state, and local partners.

Internet Crimes Against Children Task Force — A national network of 61 coordinated task forces representing over 4,500 federal, state, and local law enforcement and prosecutorial agencies. These agencies are continually engaged in proactive and reactive investigations and prosecutions of persons involved in child abuse and exploitation involving the Internet.

Massachusetts State Police, Cyber Crime Unit — Investigates and does forensic examination for complex criminal cases involving digital evidence; consults on criminal matters involving technology.

Massachusetts State Police, High Risk Victim Unit — Investigates human trafficking cases across the commonwealth of Massachusetts. Co-located with the Massachusetts Commonwealth Fusion Center.

Multnomah County District Attorneys Office — Addresses trafficking in Multnomah County, OR through a Human Trafficking Team, which works to prosecute traffickers, reduce demand for exploitation, and ensure protection and support for victims.

Multnomah County Sheriffs Office, Jail Investigations — Collects and analyzes jail intelligence information in support of human trafficking investigations.

New Hampshire Information & Analysis Center — A U.S. Department of Homeland Security intelligence fusion center, which serves as the focal point within the state of New Hampshire for the receipt, analysis, gathering, and sharing of threat-related information among federal, state, and local partners.

New York County District Attorney's Office, Human Trafficking Response Unit — The Manhattan District Attorney's Office Human Trafficking Response Unit was established in 2014 and is housed within the Offices Special Victims Bureau in the Manhattan Family Justice Center. The unit investigates and prosecutes sex and labor trafficking cases, and

provides support for victims and their families. The Unit also works to identify victims in cases that may involve elements of trafficking from hundreds of street arrests that are screened and prosecuted by other areas of the Office, such as the Domestic Violence Unit.

Portland Police Bureau — Investigates trafficking cases in Portland, OR using a victim-centered approach; coordinates with local government agencies including the Oregon Department of Human Services, the Multnomah County District Attorneys Office, the FBI Child Exploitation Task Force, and police departments in neighboring communities; provides referrals to immediate and long-term victim services offered by local NGOs.

San Diego Computer and Technology Crime High-Tech Task Force — A multi-agency task force formed in 2000 by combining local, state and federal law enforcement agencies from San Diego, Riverside, and Imperial counties, to apprehend and prosecute criminals who use technology to prey on citizens. Provides digital forensics examinations in support of the District Attorney's Office.

San Diego Human Trafficking Task Force — A cooperative effort between 10 state and local agencies with support from 14 regional, state, and federal agencies, led by the California Department of Justice Attorney General's Office. The Task Force was established in 2015, and is currently comprised of positions from two San Diego vice teams, members from the County Sheriff's Department, members from the San Diego Police Department, multiple Federal Bureau of Investigation staff, and liaisons to the San Diego District Attorney's Office. An executive board oversees the operations of the Task Force and meets quarterly to review activities.

South Carolina Human Trafficking Task Force — Supports the efforts of law enforcement and prosecutors through specialized training, victim support, and legislation advocacy.

A.3   NON-GOVERNMENTAL ORGANIZATIONS

childsafe.ai — A New York based software company that deploys machine learning and automated collection capabilities to collect data on facilitators and consumers of commercial sex services from within the surface, deep, and dark web marketplaces. For law enforcement, childsafe delivers a Demand Deterrence Platform serving human trafficking units around the country to disrupt local human trafficking economies.

Guardian Group — Aims to prevent and disrupt the sex trafficking of women and children while enabling partners to identify victims and predators in the United States; identifies victims of sex trafficking to law enforcement; researches trafficking leads received from law enforcement and the public; provides counter-trafficking training for law enforcement and employees of hotels and other organizations.

Marinus Analytics — Devloper of the Traffic Jam tool, which scrapes online sex advertisement and buyer review boards to help law enforcement identify and investigate commercial sex providers involved in human trafficking.

National Center for Missing & Exploited Children — A private nonprofit that provides assistance to victims, families, law enforcement, social service agencies, mental health agencies and other organizations when they need help in dealing with cases of missing, exploited, or recovered children. The Center also operates the CyberTipline, a nationwide central reporting system that works with the public and electronic service providers to report suspected online exploitation of children.

Polaris — Polaris is a nonprofit, NGO that works to prevent human trafficking; for the last decade, Polaris has run the National Human Trafficking Hotline and Textline. Polaris is also involved in data-driven research on human trafficking data and associated awareness campaigns.

Project Concern International — A non-profit, humanitarian NGO based in San Diego, California, which supports domestic early educational public school programs for teaching gender dynamics to children; runs a group mentoring program designed to promote self-esteem, develop life skills, and inspire positive motivation in at-risk youth aged 8-15 years; and coordinates the CEASE San Diego Network, which is dedicated to developing strategies to deter people from buying sex.

Safety Compass — A non-profit that provides advocacy services for survivors of sex trafficking and commercial sexual exploitation in Clackamas, Marion, and Washington Counties, OR; provides victim advocates 24/7 to meet with survivors beginning with initial contact, through the interview process, during trial and grand jury preparations, and afterward; coordinates with law enforcement, the Oregon Department of Human Services, the juvenile justice system, and other government and nonprofit social services organizations.

The EPIK Project — The EPIK Project is a nonprofit organization that mobilizes men to disrupt the commercial sex market at the point of sale, equipping them to combat the roots of exploitation and to collaborate effectively with the wider anti-trafficking movement.

Thorn — Technology NGO that develops the sex advertisement analysis platform, Spotlight.

Uncharted — Commercial software company that develops the sex advertisement analysis platform Tellfinder.

APPENDIX B: TECHNOLOGY CAPABILITY INVENTORY

This Appendix catalogs known capabilities that were identified during the stakeholder engagment process, and used in human trafficking investigations, prosecutions, and associated law enforcement operations.

B.1 DATA AGGREGATORS AND WEB-BASED SERVICES

Accurint / LexisNexis — Data broker with a comprehensive database of public records geared towards uncovering connections between people, businesses, assets, and locations. Often used to expedite identification of individuals and their relatives.

Apify / Apify Technologies — A web scraping and automation platform that enables extraction of data or execution of automated actions on websites. Jobs can be automatically scheduled on intervals and downloaded or sent to other cloud-based services. The capability is frequently used to scrape sex advertisement websites.

Bolo Mobile / Bolo Mobile — "Be On the Look Out" is a cell phone based photo sharing application for law enforcement only.

Callyo / Callyo — A call and text platform designed to assist law enforcement with investigations. Allows the ability to generate numbers in an area code of choice and have those texts and calls forwarded to an officer's own phone number. Used in trafficking investigations where throw-away numbers are needed for decoy sex advertisements or undercover operations. Call and text logs can be downloaded from a website for evidence preservation.

CLEAR / Thomson Reuters — A data broker with a comprehensive database of public records and other 3rd party data geared towards law enforcement needs. Large amounts of personally identifiable information from public and other business records, augmented with law enforcement specific data feeds, such as cell, landline, and VoIP phone records; licence plate scans; arrest records and intake photos from thousands of facilities.

Demand Deterrence Platform / Childsafe.ai — An automated system to measure sex buyer activities and behaviors on commercial sex advertisement sites, sugar daddy sites, and review sites.

Fake Name Generator / Corban Works — An online service that can create automated fake personas by leveraging databases of known names addresses, combined with an email forwarding service.

IntelTechniques OSINT Portal / Michael Bazzell — Custom search internet tools, used by law enforcement. Search facets include email addresses; Facebook, Twitter, and Instagram profiles; real name, username, telephone number, and other attributes.

Law Enforcement Archival Reporting Network (LEARN) / Vigilant Solutions — A national vehicle locator service using license plate recognition scans submitted by various law enforcement organizations.

Nfusion / Ntrepid — Virtual Desktop Infrastructure service for providing managed attribution web-browsing services to increase privacy and security when performing investigations. Used primarily for disassociating undercover social media accounts and persona from law enforcement information technology resources and identifying network information.

OneReach / Twilio — A front-end and application programming interface for integrating voice call and text messaging into 3rd-party applications. The back-end that is used for some demand deterrence platforms.

Pen-Proxy / PenLink — Service for performing intelligence gathering on phone numbers. Automates the queries of collected phone numbers and location codes against third-party data services; can query third-party data brokers and providers to find a subscriber, porting history, and current carrier information for a given number.

Rekognition / Amazon Web Services — Software service to add image and video analysis to customer applications. Users provide images or video to the Rekognition Application Programming Interface, and the service identifies objects, people, text, scenes, and activities in the image or video. Also provides facial analysis and facial recognition, which is used as the back-end of multiple sex advertisement search tools, rather than custom proprietary implementations.

Scrapinghub / Scrapinghub — An automated web scraping and platform that extracts data from websites and provides capabilities to build custom web crawlers. The capability is frequently used to scrape sex advertisement websites and buyer review websites.

Spotlight / Thorn — Web application that scrapes sex advertisements from commercial sex classified sites and buyer review forums, and makes them searchable by users. Allows users to query and filter ads based on geographic region, keyword, phone number, and other identifiers. Users can upload photos of interest, detect faces, and potentially find advertisements with similar faces in their images. Includes additional capabilities for network building, text classification of advertisement copy, and other features. Uses well-known cloud-based services for the application back-end.

Tellfinder / Uncharted Software — Web application that scrapes sex advertisements from commercial sex classified sites and buyer review forums, and makes them searchable by users. Currently includes 86 active web sites and 244 historical web sites for search and analysis, with over 180 million individual web pages and 110 million unique collected images since 2001. Allows users to query and filter ads based on geographic region, keyword, phone number, and other identifiers. Users can upload photos of interest, detect faces, and potentially find advertisements with similar faces in their images. Additional capabilities for network building, text classification of advertisement copy, and other features.

TinEye / TinEye — An image search and recognition service, known mostly for its reverse image search capability.

This Person Does Not Exist / Phil Wang — A website that uses a technology called GANs (generative adversarial networks) to produce synthetic images of individuals that are unique. Used for profile pictures in undercover social media profiles.

TLOxp / Trans Union — Commercial data broker service with data on an estimated 95% of the U.S. population. Data sources include a wide range of personally identifying information, including social security number, address, employment, and phone records; license plate recognition and drivers license databases; and mobile phone geolocation services.

Traffic Jam / Marinus Analytics — Web application that scrapes sex advertisements from commercial sex classified sites and buyer review forums, and makes them searchable by users. Allows users to query and filter ads based on geographic region, keyword, phone number, and other identifiers. Users can upload photos of interest, detect faces, and potentially find advertisements with similar faces in their images. Additional capabilities for network building, text classification of advertisement copy, and other features. Uses well-known cloud-based services for the application back-end.

Whooster LE / Whooster — Mobile phone based data as a service platform for delivering business intelligence information and personally identifying information.

WiGLE / WiGLE — The Wireless Geographic Logging Engine (WiGLE), is a website for collecting information about the different wireless hotspots around the world. Users can upload hotspot data like GPS coordinates, SSID, MAC address, and encryption used, on the hotspots discovered. Additionally, cell tower data is uploaded and also displayed.

B.2 INVESTIGATION TOOLS

Cell Hawk / Hawk Analytics — Call Detail Record analysis software. Allows users to upload phone records received in response to court orders from all major U.S.-based cell phone telecommunication providers. Enables graphical representations of individual and aggregate calls on a map, animations of call histories, and visualization of cell tower locations.

Gotham / Palantir — Distributed software platform for structuring network-based (relational) data enterprise-wide and performing social network analysis for investigative purposes. Often requires human-intensive tagging of unstructured documents and data into a predefined ontology, which can be a significant burden on manpower-limited investigative organizations. Once data is structured in the system, users can perform a variety of collaborative all-source and graph-based analysis to uncover connections in their data.

Hunchly / Hunchly — Web-browser plugin that automatically tracks the URL, timestamps, and hashes of every web-page visited during an open-source investigation. Allows investigators to go off on a variety of investigative tangents looking for leads, without having to worry about bookmarking or taking screenshots of what they have already discovered. Browsing histories and associated image files are stored in a cloud-based service, which allows for easy download by investigators.

i2 Analysts Notebook / IBM — A visual analysis tool to help structure data into intelligence, primarily by providing a mechanism to label textual data in order to perform network visualization and social network analysis.

Investigator Pro / Securus Technologies — Voice biometrics software for inmates who try to obscure their identities and commit crimes using the inmate telephone system. Privdes the ability to search call databases by inmate voice biomarkers.

NetViz / PenLink — Social network visualization software for use with other PenLink products.

OmniPage / Nuance — Commercial software for optical character recognition of scanned documents for document conversion to electronic format. Mixed performance on judicially compelled records, which limits overall utility.

TRAX / ZetX — Automates the mapping process of complex phone records into time-lapse, movie like presentations. Can support input data from call detail records, war-drive tests, cell tower dumps, exigent pings, wire taps, and pen registers.

B.3 DIGITAL FORENSICS TOOLS

Most investigations have some component of evidence derived from forensic extractions of data from cell phones, tablets, computers, and other electronic devices. Extractions begin with hardware that physically attaches to the device and then uses various methods to extract data off, such as emulating a known backup protocol or using very closely held proprietary methods to extract full file systems or memory copies of the devices, which are stored in an "image" file on a computer for subsequent analysis. Next, specialized software is used to read and decode the device image into file and database formats that can be parsed and exploited by analysis software. There is a wide breadth and depth of the capability of the varied analysis software packages, and some tools focus more on specific areas of need, such as messaging and call analysis; image and video exploitation; file and database formats of 3rd party phone applications, and the ability to use cached credentials to retrieve cloud-based information. Some tools are made for single users and some are designed for collaborative workflows, including some case management components and the ability to log and maintain the chain of custody across organizations. Many of these tools are intended for law enforcement only and have sensitive capabilities, and almost all of them require judicial order or consent, in order to be used. There are many vendors and solutions in this space, and for the purposes of simplicity, they are grouped into a simple list without explicit details on the features and differences of each capability.

Analyze DI Pro / Griffeye Technologies

ATLAS / Magnet Forensics

AXIOM / Magnet Forensics

Blacklight / BlackBag Technologies

Cellebrite Reader / Cellebrite

Detective / Oxygen Forensic

Digital Evidence Investigator / Advanced Digital Forensic Solutions

E3 / Paraben Corporation

EnCase Forensic / opentext

Forensic Toolkit / AccessData

Graykey / Grayshift

MacQuisition / BlackBag Technologies

Passware Kit Forensic / Passware

Recon / Sumuri

Triage / Advanced Digital Forensic Solutions

Truxton Forensics / Probity

UFED Ultimate / Cellebrite

X-Ways Forensics / X-Ways Software Technology

XAMN / MSAB

XRY / MSAB

Other forensic-centric tools not in the aforementioned group include:

Burner Breaker / Susteen — Robotic arm for physically brute forcing mobile phone pin codes.

DVR Examiner / DME Forensics – Software solution for the recovery of video and metadata from Digital Video Recorder surveillance systems, in a forensically sound manner.

Project-A-Phone / Paraben Corporation — A device to take pictures of phone screens, for when logical forensic extraction is unavailable.

VideoFOCUS / Salient Sciences — A family of digital multimedia analysis and clarification software tools for audio, video, and images, providing workflows for forensic video analysts. Can import more than 200 standard multimedia and numerous proprietary CCTV file formats.

This page intentionally left blank.

APPENDIX C: DATA SHARING AGREEMENTS

This Appendix catalogs known data sharing agreements for data associated with human trafficking at the regional and national levels. This list includes national databases that focus on human trafficking:

Banks Alliance Against Trafficking / Thomson Reuters Foundation) — Information-sharing agreement that aims to to identify money flows associated with trafficking. Participants include financial institutions, counter-trafficking NGOs, financial crime compliance standard-setting bodies, national Financial Intelligence Units (FIUs), and law enforcement.

Human Trafficking Reporting System (HTRS) / Bureau of Justice Assistance (managed by Northeastern University) — Tracks investigations of alleged human trafficking reported by federally funded human trafficking task forces. Data reported into the HTRS are used to track task force performance, to characterize the trafficking cases that task forces investigate, and to characterize trafficking victims who are identified by agencies participating on these task forces.

Trafficking Information Management System (TIMS) / Department of Justice — Helps Office for Victims of Crime (OVC) grantees to gather, record, analyze, and report required performance measures; serves as a central repository of information related to activities under the OVC human trafficking grant initiative; enables OVC grantees to organize, standardize, and centralize data collection efforts on a variety of performance measures.

This list includes law enforcement mechanisms for sharing general investigative information at the regional and national levels:

Law Enforcement Enterprise Portal (LEEP) / Federal Bureau of Investigation — Provides law enforcement, intelligence, and criminal justice entities with centralized access to resources in order to both aid case development for investigators and enhance information sharing between agencies.

National Data Exchange (N-DEx) / Federal Bureau of Investigation — Supports law enforcement agencies, criminal justice agencies and multi-jurisdictional task forces by bringing together investigative data from across the United States, including incident and case reports, booking and incarceration data, and parole/probation information. Enhances national information sharing across federal, state, regional, local, and tribal investigative agencies and task forces. Provides data exploration tools to identify relationships and correlations between personal, vehicle/property, location, and crime characteristics.

Regional Information Sharing Systems (RISS) / Department of Justice — Supports law enforcement efforts nationwide to promote officer safety and to combat illegal drug trafficking, identity theft, human trafficking, violent crime and terrorist activity. Comprised of six multi-state centers that operate on a regional basis and more than 9,000 local, state, federal, and tribal law enforcement and public safety member agencies.

This page intentionally left blank.

APPENDIX D: MEASUREMENT DATA SOURCES

This appendix catalogs major classes of measurement of trafficking and detail existing and potential sources of data to produce those measurements.

D.1 INCIDENCE AND PREVALENCE

Knowledge of the prevalence of human trafficking would establish a baseline level of situational awareness concerning the crime and enable data-driven decisions regarding resource allocation, quantification of impact, and open research questions such as vulnerability modeling and prevalence forecasting. Information on the time and location of trafficking-related activity, number and attributes of persons involved (spanning victims, traffickers, and buyers), and attributes of the crime itself are of utility. While it is impossible to perfectly measure prevalence, incidence of human trafficking—with varying levels of validity relative to proof of lawful conviction—can be measured from the following types of data:

Prosecuted trafficking cases: Prosecuted case records contain information about victim, defendant, and crime attributes as well as the number of victims and defendants. While cases resulting in trafficking convictions have the highest level of proof, the process between identification of potential trafficking activity and conviction can take years.

Visas granted to trafficking victims: T- and U-visa records provide counts of the number of non-citizen victims in U.S. jurisdictions (and immediate family members) cooperating with law enforcement in the investigation and prosecution of human trafficking crimes.

Law enforcement cases and investigations: Law enforcement human trafficking cases provide information about victim, perpetrator, and crime attributes; the number of victims and perpetrators; case status (i.e., open, prosecution declined, etc.); and agencies or type of agencies involved (i.e., local, state, federal, or foreign).

Healthcare provider and public health human trafficking reports: ICD-10 reporting provides counts of suspected and confirmed forced sexual or labor exploitation for adults and children as identified by medical service providers and reported through public health channels. Counts of individuals with a personal history of exploitation and multiple perpetrators of maltreatment/neglect are also recorded. ICD-10 codes for human trafficking went into effect in October 2018.

Victim services provider data: Data on utilization of victims services, particularly the number of unique victims served, can inform victim-based prevalence estimates and uniquely capture victims who have not had contact with organizations reporting to the government (e.g., law enforcement or health services).

Tip line reports: Reports of possible incidents of human trafficking, through phone or cyber tip lines, primarily provide information on victim and crime attributes. Incident quantity may be recorded as the number of incoming tips or reported victims, or resolved to a number

of cases (if incoming reports are consolidated in such a way). Of the trafficking incidence measures listed, tips are the least robust but have the smallest delay between time of crime and time of reporting and may have significantly more data points than sources that require more manpower for investigation and higher burdens of proof.

Other data sources—such as missing persons reports, arrests for crimes frequently made before trafficking can be proven (truancy, prostitution), and ad postings of confirmed victims—may be utilized as proxies or noisier measures of human trafficking activity, but require a notion of "ground truth" trafficking activity data, against which the validity of such techniques can be proven. It is also worth noting that the legal definition of trafficking, which may be seen as the gold standard for validity of a trafficking incident, can vary between jurisdictions. Additionally, producing a unified incidence or prevalence measure requires identification of victims or cases across sources to avoid duplicate counting, which may be problematic or impossible to resolve.

TABLE D.1

List of potential data sources for measuring incidence of human trafficking.

Organization	Data Source	Description
Polaris Project	Case files from the National Trafficking Hotline	Cases, signals, and victims reported to the National Trafficking Hotline.
National Center for Missing and Exploited Children	NCMEC Hotline and CyberTipline reports	Trafficking-involved missing or exploited children calls and tips.
Center for Disease Control	National Notifiable Diseases Surveillance System public health database	ICD-10 human trafficking code reports.
DHS Homeland Security Investigations	Human trafficking case files	Human trafficking arrests, opened cases, and associated statistics.
U.S. Citizenship and Immigration Services	Trafficking victim visas granted	T-visas, U-visas granted to trafficking victims.
Human Trafficking Data	Human trafficking prosecutions	Corpus of federally prosecuted human trafficking cases in the U.S., including crime locations, victim and defendant information.

D.2 COMMERCIAL SEX MARKETPLACE DYNAMICS

Estimates of commercial sex provider supply and buyer demand in a geographic region can provide intelligence on a marketplace that exists irrespective of provider legality or trafficking status. Potential sources of data on commercial sex marketplace dynamics include:

Online advertisements: Advertisements for commercial sex services can be used to estimate the supply in a given geographic area and serve as an indication of the unique number of providers or operations. A subset of this data could also be used to estimate the number of potential victims. Example sites include: Adult Look, Adult Search, Bedpage, City X Guide, Eros, Listcrawler, Megapersonals, One BackPage, Plenty of Fish, Skip The Games, Switter, and others.

Buyer forums: Data collected from "hobby boards" and "john forums," including pricing information, can be used to characterize high-frequency buyers. Example sites include: ECCIE, Erotic Monkey, The Erotic Review, Rubmaps, TNA Board, USA Sex Guide, and others.

Demand deterrence operations: Data from NGO demand deterrence platforms and law enforcement sting operations, including information on the diversity of responses to decoy ads, frequency of response from specific phone numbers (buyers), types of services requested or responded to, and pricing information, can be used estimate the buyer demand in a geographic region.

Web statistics: General estimates of the relative popularity or market-share of classified advertising sites, sugar daddy websites, hobby boards, and other online sites associated with the commercial sex industry.

TABLE D.2

List of potential data sources for measuring commercial sex marketplace dynamics.

Organization	Data Source	Description
Thorn	Spotlight	Commercial sex advertisements and buyer forum reviews. See Appendix B.
Marinus Analytics	Traffic Jam	Commercial sex advertisements and buyer forum reviews. See Appendix B.
Uncharted Software	Tellfinder	Commercial sex advertisements and buyer forum reviews. See Appendix B.

Organization	Data Source	Description
Childsafe.ai	Demand Deterrence Platform	Data derived from responses to decoy advertisements. See Appendix B.
Epik Project	Cyber Patrol Deterrence Platform	Data derived from responses to decoy advertisements. See Appendix B.
Various	Law enforcement operations	Data derived from responses to and from advertisements during counter-demand and counter-supply operations. See Section 3.1.
Alexa	Website rankings	Commercial web traffic data and analytics

D.3 LOW-WAGE AND FORCED LABOR MARKETPLACE DYNAMICS

Estimates of supply and demand for below-living-wage labor and related market dynamics in a given geographic region or industry would provide intelligence on a broader marketplace involving legal low-wage labor, illegal but non-trafficking labor abuse, and trafficked laborers. Potential sources of data on low-wage and forced labor marketplace dynamics include:

Employer and recruiter reviews: Online postings reviewing employers and recruiters could be used to inform estimates of demand for forced and/or low-wage labor.

Online advertisements: Number of posted positions (e.g. for household or food service workers) could inform demand estimates for low-wage workers in a region. Salary information in postings, or from aggregation sites, could be used to inform industry/venue-specific wage estimates in a region.

Wage data: Wage data by region and occupation can be used identify the lowest-wage jobs in a region. Numbers reported through official channels are less likely to capture information on wages for trafficked individuals or wage abuse.

Market-share/pricing data: Measures of price and purchaser-driven price control, such as percent market share determined by a small number of purchasers or pricing for a good, could inform low/forced labor risk estimates. Producers are driven to reduce costs and therefore wages in order to meet prices dictated by large market-share purchasers. For example, large fast food and grocery corporations exert this type of influence on particular crop prices.

Workplace Inspection, Labor Investigation, Licensing, and Enforcement: Data on the number of Department of Labor investigations or investigators in a geographic region and on licensing and inspection requirements and enforcement by industry/region could be used to measure demand-side deterrence for illegal/forced labor.

Remote sensing data: Satellite and other remote sensor data can provide regional agricultural and economic assessments and potentially provide measurements of activity associated with human trafficking (such as deforestation in protected regions or near fisheries and construction of man-made structures).

Measurements of consumer/investor awareness of and adherence to slavery-free supply chains: Such data could be used to inform illegal labor demand estimates or measure the deterrent influence consumer/investor behavior exerts on supplier-side behavior.

TABLE D.3

List of potential data sources for measuring low-wage and forced labor marketplace dynamics.

Organization	Data Source	Description
Centro de los Derechos del Migrante, Inc.	Employer and recruiter reviews	Yelp-like, user-driven information sharing platform hosting reviews of recruiters and employers in low-wage labor recruitment involving the Mexico-U.S. migrant stream. (see contratados.org)
Care.com	Job postings	Domestic work demand (based on number of postings) and salary information listed in postings.
Bureau of Labor Statistics	*not identified*	Survey-based wage data by area and occupation for national, regional, state, and many metropolitan/non-metropolitan areas.
USDA National Agriculture Statistics Service	*not identified*	Crop prices, size of agribusiness companies (e.g., corporation vs. family-owned).
not identified	*not identified*	Percentage of crop market or pricing dictated by large corporate buyers.

D.4 INDICATORS OF VULNERABILITY

Knowledge of the degree of vulnerability of an individual, population, or geographic region to human trafficking, apart from whether or not exploitation is taking place, would aid in the deployment of preventative actions to eliminate victimization before it begins. Populations especially vulnerable to human trafficking include children in the foster care system, homeless/displaced populations and those at risk for becoming homeless, drug users, recent immigrants, those in poverty, and survivors without viable economic opportunity. Factors such as legislative/policy protections, gender culture, access to basic necessities, and political/physical safety and stability impact vul-

nerability to human trafficking. Research in identifying and understanding these factors is ongoing. The size of these populations and severity of vulnerability indicators could be measured from the following types of data:

Child welfare and foster care system data: Data on the number of children in the foster care system and number of children reported to have experienced abuse or neglect in or outside of foster care informs estimates of vulnerable minor populations.

Federal homeless and displaced population estimates: Estimates of the size of homeless populations and number of persons displaced due to natural disaster provides information on populations at risk for exploitation directly or indirectly related to residential instability to the point of homelessness.

Shelter provider data: Data directly from shelter providers can also provide estimates of populations at risk for exploitation related to residential instability, with a more frequent update rate but less standardized/aggregate scope than federal estimates.

Social and demographic data: Census-based social/demographic data can provide estimates of current or projected vulnerable population sizes, stratified by demographic attributes, as well as measures of the prevalence or severity of environmental factors impacting vulnerability of populations and/or regions. These encompass, but are not limited to: household finance data such as income, employment, population in poverty, and economic opportunity (including access to credit, trade skills, and education level); housing cost, availability, and stability; and demographics related to age, gender, population density, and migrant and non-English speaking populations (who may not be informed of rights and hiring practice or salary norms and are typically over-represented in low-wage jobs). Such data can be applied to measuring populations or regions currently experiencing or at risk of experiencing vulnerability indicators.

Public health data: Measures of adult risk behavior (drinking, drug use, sexual activity), a community's resilience to human suffering and financial loss in a disaster, and infant mortality rate (considered a sensitive indicator of overall societal health) as reported through public health channels may be indicative of the level of vulnerability to trafficking in a community.

Crime data: Prevalence of crimes related to trafficking and individual physical safety—violent crimes, sex offenses, runaways, offenses against family and children, and drug use—may be indicative of trafficking risk for individuals or populations within a jurisdiction. Certain charges known to be used before trafficking can be proven or when trafficking is mis- or unidentified (perhaps due to lack of training) may contain information on prevalence in addition to vulnerability.

Missing persons reports: Data on missing persons, especially runaways and cases with suspected links to trafficking, could inform estimates on the number of persons with elevated vulnerability to human trafficking.

Cultural and societal measures: Broad assessments of gender culture—via measures such as equality and protection in legislation, workforce representation statistics, and prevalence

of domestic violence and sex or discrimination-based crimes—and social or political conflict, corruption, or unrest—via measures such as public opinion surveys and public corruption convictions—could inform regional vulnerability estimates at sufficiently large scales. Sentiment analysis of news and social media is another potential source of region-based population-level estimates on these topics.

Social media: Text/emoji or image data from user profiles on social media sources, such as Twitter and Instagram, consistent with common trafficking indicators could be used to identify individuals at-risk for recruitment, marketing, or exploitation stages of trafficking. Counts of at-risk individuals or other usage metrics could be used to inform population or region-based vulnerability estimates.

Legislation: Measures of the prevalence, coverage, and strength of anti-trafficking legislation may be indicative of governmental and/or societal awareness of human trafficking issues as well as protections provided via legal deterrence and, thus, vulnerability of individuals or populations within that jurisdiction.

It should be noted that "hidden" populations, which are not covered by traditional statistical survey methods like the census (i.e., those without a household like homeless, orphan, and institutionalized populations), pose a challenge in population-level vulnerability measurement. Additionally, vulnerability indicators can be expanded to factors concerning trafficker/facilitator and buyer/consumer populations, but current research focuses primarily on victim-side vulnerability.

TABLE D.4

List of potential data sources for measuring indicators of vulnerability

Organization	Data Source	Description
Department of Health and Human Services, Administration for Children and Families Children's Bureau	Adoption and Foster Care Analysis and Reporting System	Number of children in foster care.
Department of Health and Human Services, Children's Bureau	National Child Abuse and Neglect Data System	Reported cases of child abuse or neglect by region; a voluntary reporting system.
National Center for Missing and Exploited Children	Hotline and CyberTipline reports	Calls or cases involving missing or exploited children.

Organization(s)	Data Source	Description
Department of Housing and Urban Development	Point-in-Time homelessness estimates and Housing Inventory Count	Overall homelessness counts, sheltered and unsheltered, reported by Continuum of Care.
Federal Emergency Management Agency	Housing displacement	Number of persons or households displaced due to natural disaster.
not identified	*not identified*	Shelter capacity and utilization data from service providers.
U.S. Census	American FactFinder	Counts by region of residential stability, total housing units, vacant housing units, owner / renter occupied housing, median owner-occupied value. Rent as percentage of household income.
U.S. Census	American FactFinder	Counts by region of employed / unemployed, males / females in labor force, persons employed in service occupations, households with both parents in workforce, persons without health insurance coverage.
U.S. Census	American FactFinder	Median household income, median family income, average per capita income, percentage of under-18 population below poverty, percentage of all families below poverty, percentage of female-headed households.
not identified	*not identified*	Access to banking/credit by region.
Centers for Disease Control and Prevention	Behavioral Risk Factor Surveillance System data	Adult risk behaviors: drinking, drug use, sexual activity.
Centers for Disease Control and Prevention	Wide-ranging Online Data for Epidemiologic Research	Number of opiod and non-opiod overdose deaths; infant mortality rate (indicator of societal health).

Organization(s)	Data Source	Description
Department of Health and Human Services Agency for Toxic Substances and Disease, Centers for Disease Control and Prevention	Social Vulnerability Index (SVI)	A measure of social vulnerability—a community's ability to prevent human suffering and financial loss in a disaster, based on 5 social factors, including poverty, lack of vehicle access, and crowded housing—for every census tract.
Federal Bureau of Investigation	Unified Crime Reporting (UCR) statistics	Number of arrests and offenses—particularly violent crimes, sex offenses, runaways, offenses against family and children, and drug use—per region as tracked by the UCR program.
not identified	*not identified*	Measure of cultural valuation of women.
Bureau of Justice Statistics	2012 Census of Problem-Solving Courts	Number of problem solving courts, by type, per region.
Office of Juvenile Justice and Delinquency Prevention	Juvenile Residential Facility Census	Number and populations of juvenile detention facilities.
Office of Juvenile Justice and Delinquency Prevention	Census of Juveniles in Residential Placement	Number of juveniles in a residential facility by race and gender and in total.
not identified	Juvenile Court data	Delinquency and truancy.
Gallup	Annual poll on crime	Percentage of population with favorable opinion regarding law enforcement and regarding courts/judges/judiciary.
Gallup	Annual governance poll	Measure of public opinion on trust and confidence in state and local governments when it comes to handling state/local problems.
Department of Justice Public Integrity Section	*not identified*	Statistics on federal public corruption convictions by judicial district.

Organization(s)	Data Source	Description
Twitter, Instagram	User feeds	Measure of usage of common trafficking indicators in public profiles/postings by user location.
not identified	*not identified*	Measure of conflict and social unrest.
not identified	*not identified*	Measure of sophistication of sex trafficking legislation in a region/jurisdiction.
not identified	*not identified*	Measure of sophistication of labor trafficking legislation in region/jurisdiction.
not identified	*not identified*	Measure of sophistication of labor and farm workers protections in region/jurisdiction.

D.5 CONTEXTUAL AND ENVIRONMENTAL ATTRIBUTES

Understanding human trafficking activity and interdiction requires an understanding of the broader environment in which the crime takes place. Attributes characterizing individuals (i.e., victims, traffickers, facilitators, buyers, consumers), populations, geographic regions (e.g., rural and urban locations), and crimes involved provide contextual information for analyses with finer scopes than national-level analyses of sex or labor trafficking. Data types measuring such environmental and contextual attributes include many of the data sources previously mentioned for trafficking incidence, vulnerability, and marketplace dynamics (which will not be repeated in table D.5), and fall broadly into the following categories:

Crime: Data on arrests, offenses, and gang activity in a region measure human trafficking and possibly related crimes.

Law Enforcement: Counts of sworn law enforcement employees, assessment of human trafficking training/awareness/expertise among law enforcement, and existence of anti-human trafficking initiatives (e.g., task forces) help characterize law enforcement presence and counter-human trafficking response.

Health: Factors such as adult risk behavior prevalence, HIV and sexually transmitted disease prevalence and new diagnosis rates, infant mortality rates, and community resiliency to natural disaster help characterize population-level health impacting human trafficking prevalence and vulnerability.

Demographics: Data characterizing populations and sub-groups within populations at multiple geospatial scales is required to perform population- and region-level analyses.

Housing: Housing data is relevant to contextualizing HT crime in several different ways: homelessness, risk of homelessness, and residential instability indicate issues with access to life necessities and increased vulnerability to human trafficking; shelter providers are a point of contact between victims, survivors, or at-risk individuals and service providers; victims may require housing assistance to exit the trafficking cycle; individuals may be targeted at shelters or by persons with power over their housing situation (e.g., landlords); and traffickers may provide victims with housing to sustain the trafficking cycle. Generally speaking, census and shelter provider data may be used to inform victim prevalence and vulnerability estimates in a region.

Transportation: Transportation data captures information about human movement—including that of human trafficking victims, traffickers, buyers, and knowing or unknowing accomplices (e.g., taxi drivers)—as enabled by transportation systems. Data on transportation route availability and use by region (via road, rail, air, and other means) may relate to a region's vulnerability to or prevalence of trafficking. Additionally, transit stations may be used by traffickers for recruitment, and transportation hubs are key points for victim identification and public awareness, thus it would be useful for vulnerability modeling and interdiction efforts to measure exposure at these locations.

Finance: Access to financial resources plays a key role in individuals' vulnerability to trafficking and ability to exit the trafficking cycle. In the case of sex trafficking or commercial sex, the financial resources of buyers in a region may also be relevant. Thus, data such as household income, employment, creditworthiness, and access to credit are of utility for understanding individual and population-level vulnerability or prevalence.

Education: Viable economic opportunity also plays a key role in whether a person enters or can exit the trafficking cycle. Measures of education level (including vocational education) are therefore of utility in characterizing individual and population-level vulnerability to and prevalence of human trafficking in a region.

Business: Data on the types and sizes of industries in a region and location and financial information of particular establishments with higher probability of association with trafficking (e.g., massage parlors) could inform supply-side prevalence and vulnerability estimates in a region as well as demand-side estimates via the potential to identify increased likelihood of buyer/consumer activity in a region. Estimates of the size of illegal economic activity in a region may also be a relevant indicator. Additionally, sub-types of human trafficking (such as those identified in Polaris's typology of modern slavery) are typically defined based on the industry they related to (e.g., forestry, manufacturing, or hotels and hospitality) as each are associated with different characteristics of victims, offenders, offenses, recruitment, transportation, and exploitation.

List of data sources for measuring environment attributes

Organization	Data Source	Description
Federal Bureau of Investigation	Unified Crime Reporting statistics	Number of arrests and offenses, by type and region as tracked by the UCR program.
National Gang Center, National Drug Intelligence Center	National Youth Gang Survey, 2011 National Gang Threat Assessment	Gang-related activity by region: number of gangs, gang membership, prevalence of gang-related problems.
Bureau of Justice Statistics	National Sources of Law Enforcement Employment Data	Count of law enforcement officers, law enforcement agencies, and sworn employees (civilian and non-civilian) by region. (This data source is a centralized source 3 others: FBI UCR, Census Annual Survey of Public Employment and Payroll, and the Bureau of Justice Statistics' Census of State and Local Law Enforcement Agencies)
Department of Justice, Bureau of Justice Assistance	Anti-Human Trafficking Task Force Initiative	Binary flag for existence of HT task force in county or state.
not identified	*not identified*	Measure of law enforcement training and awareness regarding human trafficking.
Centers for Disease Control and Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention	Atlas	Prevalence of HIV and STDs and new cases diagnosed.

Organization	Data Source	Description
U.S. Census	American FactFinder	Counts by region of population (total and stratified by age), race, ethnicity, gender, female householder/no spouse, households with children, child birthrate, grandparents responsible for children, foreign-born persons, non-citizens, foreign-born persons entering the U.S. 2010 or later, region where foreign-born persons were born, languages spoken other than English, population density, percentage of population living in an urban area.
Gallup	*not identified*	Measure of religiosity (percentage): attends church regularly, professes faith in God.
not identified	*not identified*	Measure of cultural valuation of women.
U.S. Census	American FactFinder, Annual Survey of State and Local Government Finances	State and local government expenditures: education, public welfare, higher education, health, police/corrections, transportation.
U.S. Census	American Fact Finder	Counts by region of high school graduates, college graduates, and persons who did not graduate high school (in population 18-24 and 25+).
U.S. Census, Department of Commerce	Statistics of U.S. Business (SUSB)	Types of industries in a region—number of firms, number of establishments, employment, and annual payroll for most U.S. business establishments.
U.S. Census	American Fact Finder	Number of drinking, strip club, massage parlor establishments by region as classified by NAICS code.
not identified	*not identified*	Energy consumption by region. Studies which have compared forecasted energy consumption based on economic activity to measured energy consumption in a region and used it to estimate the size of illegal economic activity.
not identified	*not identified*	Food production index by region.

Organization	Data Source	Description
not identified	*not identified*	Number of military bases/installations in region.
U.S. Census	American FactFinder	Number (or percentage) of workers 16+ commuting on public transportation.
Federal Aviation Administration	Enhanced Traffic Management System	Count of airports by region, measure of airport activity (count of ingoing/outgoing flights or other network characteristics).
U.S. Census	TIGER / LINE shapefiles	Count of highway, interstate, and railroad intersections by region.
not identified	*not identified*	Count of train stations, bus stations, and interstate or international bus stations by region.
not identified	*not identified*	Event-related temporary population influx: e.g., major sporting events, political party conventions, technology conferences, etc.

D.6 INTERVENTION IMPACTS AND OUTCOMES

A need for quantitative impact and outcome assessment of interventions across all stages of the human trafficking process has been widely expressed by the counter-human trafficking community. Such assessments require measurement of the magnitude and scope of an intervention paired with measures of prevalence, vulnerability, or more targeted measures before and after the time of intervention. Data types for measuring interventions and outcomes include:

John school enrollment and impacts: Counts of offenders attending john school, alone and relative to recidivism of offenders and local changes in prevalence, can be used to understand the intervention's impact on individuals and the propagation of that impact through the larger community.

LE sting impacts: Data on the impact of performing both counter-supply and counter-demand sting operations, such as recidivism rate for buyers, identification of unique buyers across multiple geographic areas, and other information.

Law and judiciary process: Data on arrest, prosecution, and conviction rates as well as accompanying reasons why a case did or did not progress to the following stage towards conviction enable analysis of case flow through law enforcement and judiciary processes, identification of strengths or limitations, and quantification of changes in effectiveness in response to policy, leadership, or process changes.

Recidivism: Recidivism rates for individuals charged with trafficking crimes could be used to identify regions where deterrence and rehabilitation efforts are ineffective and to understand associations between effectiveness of deterrence and rehabilitation efforts and characteristics of crimes or individuals in order to inform improvements in interdiction efforts.

Victim benefits and services rendered: Counts, utilization, and effectiveness measures of survivor support services (e.g., shelters), networks, and other resources in a region aid in identifying strengths or areas of improvement for individual service providers as well as the coverage of collective efforts in a region, which may be used to inform resource allocation decisions.

Victim outcomes: Counts of victims who exit or re-enter the trafficking cycle would aid in quantifying effectiveness of interdictions and support services.

Training and awareness: Data such as the percentage of a workforce which has received training and post-training performance/retainment metric for law enforcement personnel and others in contact with victims (e.g., health care professionals, shelter providers, credit providers) would quantify the extent and quality of C-HT personnel training and aid in decision-making on training techniques and requirements in a region. Similarly, sample-based or language usage-based metrics of awareness in law enforcement, policy-makers, and the general public, on a rolling basis and around awareness campaigns, would help quantify the effectiveness of such interdictions and identify gaps at a population or region level.

Intervention funding: Total federal and state funding of counter-human trafficking interventions in a region would inform estimates of resources deployed against human trafficking.

TABLE D.6

List of data sources for measuring intervention impacts and outcomes

Organization	Data Source	Description
U.S. Department of Justice Bureau of Justice Assistance	Anti-Human Trafficking Task Force Initiative	Binary flag for existence of HT task force in county or state.
DHS Blue Campaign	*not identified*	Count and measure of effectiveness of human trafficking public awareness campaigns by region.

This page intentionally left blank.

APPENDIX E: TRAFFICKING ANALYTICAL MODEL

One of the key missing pieces in the battle against human trafficking is a national view of the situation that provides a collective awareness that could foster multi-agency collaboration. A real-time virtual display of human trafficking-related activity, patterns, and forecasting must address each agency's data protection needs as well as the need for all agencies to contribute to a national assessment of prevalence and/or vulnerability. Without this view, allocating resources effectively or crafting optimal policies is extremely difficult. This kind of tool could enable agencies combating trafficking crime to be proactive instead of reactive. Resource allocation, including training, victims services and law enforcement personnel, would benefit greatly from knowledge of where trafficking activity is occurring which is not showing up in current reports, and where activity is most likely to grow. This requires a model that can use currently measured factors to predict unknown or future trafficking.

E.1 EFFORT SCOPE

The Massachusetts Institute of Technology's Lincoln Laboratory (MIT LL) was tasked with early development of a prototype national display and model to address these gaps. The prototype model had three goals: (1) Identify trafficking that may be occurring but is as yet unknown, (2) Identify regions with high likelihood of experiencing trafficking in the near future and (3) Identify factors associated with trafficking levels, and their relative priority. This early development was focused on child sex trafficking at the county and yearly level in the contiguous U.S.

Display capability was established inside MIT LL's open-source Next Generation Incident Command System (NICS) environment. NICS is a web-based communication and collaboration platform, designed with the data protection capabilities and multi-agency big picture access desired for this type of task: with this system, no agency need have the details of every source, but the combination of sources can be available to all agencies in order to provide each with additional perspective. Figure E.1 shows an example of combining two publicly available national level datasets, capturing airline activity and Polaris Hotline activity.

The larger effort was the underlying predictive modeling. Modeling involves leveraging the relationship between a set of input "independent" variables to predict a desired output, or "dependent" variable, such as trafficking prevalence. The relationship is usually unknown, and must be uncovered by relating the input data to an instance of the desired output, i.e. a truth dataset. Success of this process requires relationships that are strong enough to make the prediction and data which is sufficient and accurate. These two aspects are described in the following sections.

E.2 DATA COLLECTION

Building a model requires significant amounts of data, often bringing with that the challenge of how to combine widely varying types of data. Development of this model was done in cooperation with the Human Trafficking Activity Model (HTAM) team: a collective of advisory, stakeholder and data contributor agencies across government intelligence, counter-trafficking NGOs, and academia.
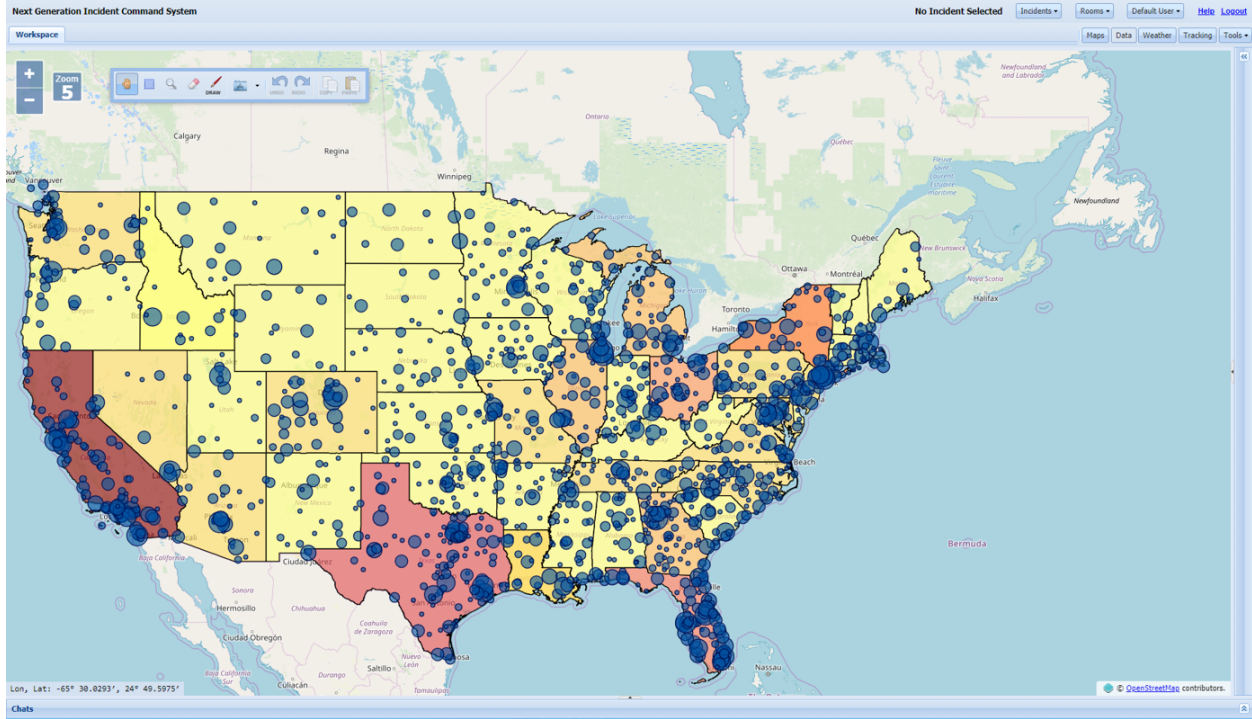
Figure E.1. Overlay of hotline and air travel activity using NICS. States are shaded to indicate total Polaris cases (color-coded from low (yellow) to high (dark red) activity); blue circle overlay displays airline flight counts, with size commensurate to airport activity.

HTAM was able to assemble the bulk of the data, much of it from agencies participating in the team. This was a difficult feat. Privacy is a critical issue when working with human trafficking data: often, the data contain personally identifiable information that must be protected in order to ensure the privacy of both victims and those identified as potential traffickers. The collector usually ensures privacy by choosing to clean the data, aggregate the data or simply not share the data. The sensitivity of law enforcement cases and operations is another potentially prohibitive privacy concern. HTAM took on the role of arranging terms for sharing of data. To this data, MIT LL added data that were already in-house or publicly available. The result was one of the largest datasets ever used for human trafficking activity modeling.

E.3 MODEL SCOPE

An HTAM sub-group – consisting of MIT LL staff, criminologists, and domain-expert statisticians – assembled for the purpose of planning the details of the initial modeling efforts, including prioritization of analyses and choice of initial methods. The initial goal was to build a model specifically for sex trafficking of minors at the county level. Although most past efforts have been at the state and national level, responses to trafficking include those from a lower governmental level,

so the county level was a step in this direction. The team prioritized use of data based on ease of access, coverage and expected predictive capability. This resulted in a synthesized data extract which aimed to cover concepts of trafficking incidence, population vulnerability, law enforcement presence, advertisement prevalence, related crime and transportation networks. The analysis sub-group advocated for linear regression models, based on their past experience. MIT LL was tasked with preparing its data contribution, preparing the synthesized data extract for analysis and initial model development.

Trafficking incidence (dependent variable) data was provided from the Polaris National Human Trafficking Hotline and a second non-publicly-identifiable source.

The following sections describe specifics of the modeling effort and results. Discussion surrounding the issues faced and strategies employed to deal with them speak to the more general task of predictive modeling for counter human trafficking purposes. It will be noted when other decisions might be more appropriate for this general case or other targeted models.

E.4 DATA CLEANING, STANDARDIZATION, AND SYNTHESIS

HTAM accomplished much of the aggregation and cleaning of the data they provided. Data was aggregated to yearly and county levels and limited in scope to sex trafficking of minors whenever possible. Time constraints prevented preparation of all the data collected, so prioritization was made according to the expertise of the HTAM analysis sub-group, informed by an estimate of the level of effort required to prepare the data. For the prioritized data, it was desirable to choose consistent measures whenever possible. Standardization of collection and reporting is a major issue. Since collectors work independently, they often have their own definition of things such as case or signal and their own way of assigning a time and/or a location. One common strategy is to report a case in all impacted counties, another is to report it only in the most relevant county for the collectors interests. This creates disparate datasets that must be processed and massaged individually before a model can be built.

Time and location are two key factors in the measurement and modeling of human trafficking. These factors were handled in the most straightforward way for the initial model, but future work may benefit from revisiting this issue to see if a more complex handling would result in increased fidelity. For this initial effort, the time tag associated with the data was applied as the year of activity. For example, a Polaris Hotline call received in 2018 was counted in 2018. An arrest occurring in 2018 was also counted in 2018, recognizing that it was possible that the trafficking itself occurred in 2017 or even earlier.

Only one year of data was present for several sources. Ground transportation data was initially acquired for 2018 because a census update was done in 2018 which improved the quality of the data. It was assumed that these data did not significantly change between 2015 and 2018, such that they could be applied to any of those years. Airline data was readily available for 2018 as well, and the same assumption was applied for the initial model. Should these data prove significant as predictors, data for prior years can be acquired. The most recent Rural/Urban designation available was from the 2010 census. This was used for initial model development but, should

it prove significant, a surrogate based on data updated more frequently may be desirable. Two variables, from foster care and crime and law enforcement datasets, were only recorded or shared for a single year.

Another issue is that of latency – the delay between the measurement and ingestion that will limit the model's ability to respond to short term changes. Examples here might be the uptick in trafficking associated with major sporting events or following a natural disaster. Prediction of changes on this time scale requires ingestion of data responsive on a similar time scale (such as arrests or online advertisement postings). Lack of this type of data led the team to decide to ignore these fluctuations and work on a yearly basis.

The second key factor was geographic location. The desired unit of analysis was U.S. county. The mapping of data to a county was often done by the collecting agency and/or added as part of the HTAM collection process. For data not assigned a county by HTAM or the collecting agency, MIT LL took on this task. Boundaries of US counties are available from the Census Bureau. Because of an improvement in data quality beginning that year, 2018 boundaries were obtained first. These 2018 boundaries were then used throughout the model effort (even for data from prior years) to provide a consistent location mapping. It is expected that this choice has minimal impact, but that has yet to be verified by analysis. For most datasets not assigned a county, an address (or, in the case of transportation data, latitude and longitude coordinates) were provided and the county assignment was made on this basis. Some of the online advertisement measures collected did not have an easily resolvable county (e.g., "Northern Wisconsin") and were removed from the initial dataset. This approach could be improved if these data prove to be a significant predictor.

Note that the location given for human trafficking activity is not always consistent across datasets. Arrest locations may or may not be in the same county as the offense. Ad postings may or may not be in the same county where victims are housed and/or exploited.

E.5 MODELING APPROACH

Past efforts at modeling human trafficking have shown that strong indicators are hard to find, if they exist at all. Goals for the initial modeling effort were to try and distinguish which independent variables from the initial set were most correlated to the dependent truth data and to consider what could realistically be predicted with this dataset. Building a single model – utilizing every independent variable in the dataset to predict a dependent variable describing the level of human trafficking in each county – may seem like an appealingly simplistic initial approach, but it was not feasible. Various constraints – relating to data sparsity, ambiguity in interpreting lack of trafficking reports, insufficient information to resolve cases across dependent variable sources, difficulties in using data across years and differences in fundamental characteristics and data reported in different geographic regions – led to the development of a set of models, rather than just one. These models varied in five attributes: the dependent variable, the set of independent variables, the mathematical relationship between these two, temporal coverage and geographic regions included. Following is a description of issues encountered and how each constrained model design. Variable treatments are also discussed.

In order to develop a predictive model, a dependent variable, or "truth variable," is required. The richest available data source for this is from the Polaris National Hotline. However, even this source has no sex trafficking cases involving minors associated with the vast majority of the 3142 U.S. counties. Such sparsity appears throughout the HTAM dataset: as shown in Figure E.2, every data source is limited in temporal coverage, spatial coverage, or both. This sparsity raises questions about how to work with the data. When a region has no reports, should that be interpreted to indicate zero activity present or incomplete data? Does this data accurately represent trafficking prevalence at this scale? These are pivotal questions, as they reflect the ability to build a model at the county level, and become even more important when considering the additional dependent variable which is even more sparse. Non-null trafficking incident data were present in only a small percentage of counties from 2015 to 2017. This sparsity proved to be a limiting factor of what analyses could be performed.

Data sparsity was also an issue for model inputs (independent variables). When coverage is low, the amount of coincident data becomes critical - i.e. how many counties have sufficient number of the input variables to be viable candidates for the model? This proved to be another limiting factor in the model effort.

Data Source	2018	2017	2016	2015	2014	2013	2012	2011	2010
Truth 1	⌄	⌄	⌄	⌄	⌄	⌄	⌄		
Truth 2	⌄⌄	⌄⌄	⌄⌄	⌄⌄					
Social Vulnerability Index			⌃⌃		⌃⌃				
Airport Activity	—								
Highway Intersections	—								
Railroad Intersections	^								
Rural/Urban									⌃⌃
Population		⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃
Ad Postings	⌄	⌄	⌄	⌄	⌄	⌄			
Foster Care 1		⌄⌄	⌄⌄	⌄⌄	⌄⌄	⌄⌄			
Foster Care 2			⌄						
Crime / Law Enforcement 1		⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃
Crime / Law Enforcement 2			⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃	⌃⌃
Crime / Law Enforcement 3			^						
Crime / Law Enforcement 4		⌄⌄	⌄⌄	⌄⌄					

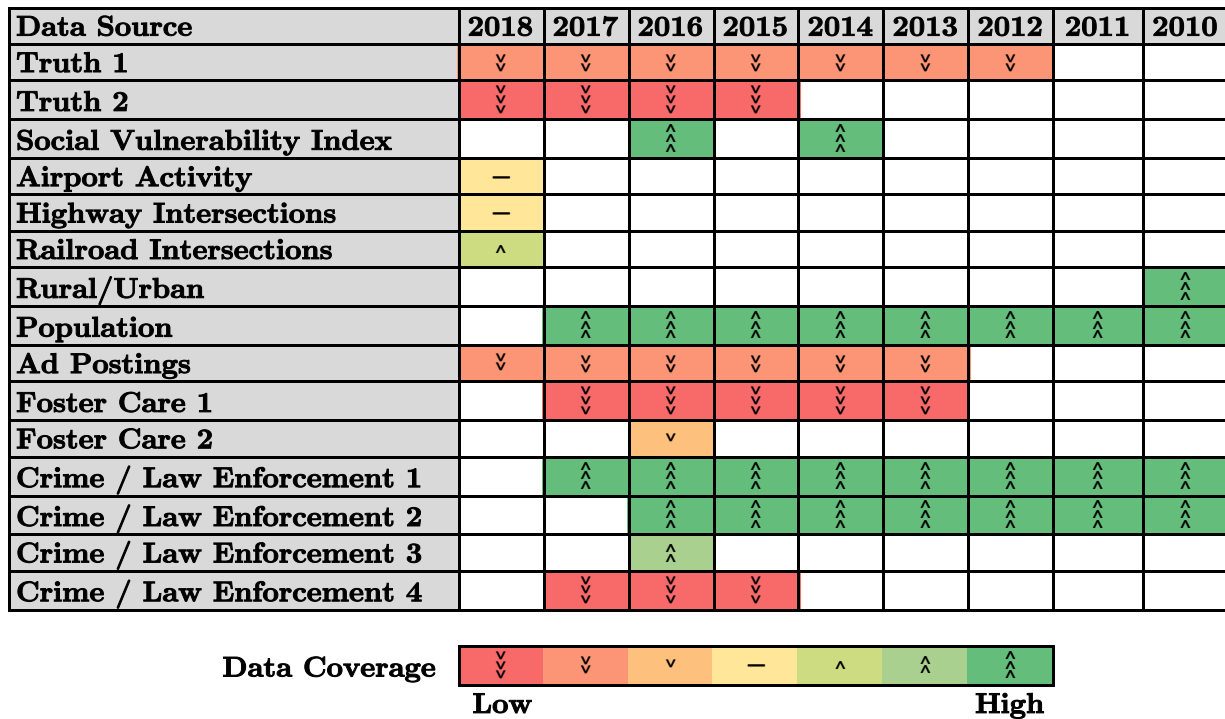| Data Coverage	⌄⌄	⌄	⌄	—	^	⌃	⌃⌃
	Low						High

Figure E.2. Number of counties with non-zero data; coverage color-coded from high (green) to low (red).

Due to the sparsity of truth data and ambiguity in interpreting reports of zero activity, there was no true measure of "zero trafficking activity" and no accurate way to represent the level of trafficking activity at a scale beginning with zero. Thus, a two-part approach was adopted to separately model presence and level of human trafficking activity. Logistic regression was used to address what factors might predict whether or not trafficking occurs in a particular county and year. The dependent variable in these models was a binary indicator of trafficking presence (0 if the county had any cases of minor sex trafficking, 1 otherwise). Linear regression was used to address what factors might predict how much trafficking occurs, in counties that show signs of trafficking. The dependent variable in these models was a count of the total number of minor sex trafficking cases in a county, always greater than zero, as the data was restricted to the set of counties with at least one case.

The dependent variables for these models consisted of information from a single trafficking incident data source, as there was not enough information provided in the data to link cases across sources and differences in units, timing and patterns of activity were unable to be resolved. It is important to note that human trafficking incidence reports are considered to be convenience samples: a sample of data drawn without any underlying probability-based selection mechanism. It is impossible to quantify the relationship between a single convenience sample and the underlying population, which means these models only capture relations between the independent variables and the chosen dependent variable (i.e., Polaris Hotline cases), not human trafficking activity in a broader sense.

Models used a single year of data between 2015 and 2017, as these were the years with the best data coverage and consistent data recording practices in contributing organizations. Data pre- and post-2015 were not comparable from one data source due to a change in the organization's recording practices. Independent variables were chosen from the same year as dependent variables, wherever possible.

Several variable treatments were applied: (1) population effects were mitigated by normalizing dependent and count-based independent variables by county population, (2) a combination of zero-imputation and multiple imputation by chained equations were used to impute sporadic missing values, (3) a non-linear relation between the independent variables and dependent variables was mitigated by applying a logarithmic transformation to dependent variables in linear regression models. Note that the population normalization approach introduces an issue – trafficking rates in low-population areas can exceed possible rates in high-population areas – which must be taken into account when interpreting model results. Future work could explore other approaches to normalization, perhaps incorporating land area.

The first set of models to be discussed are the linear regression models. The first model built considered all counties with at least one Polaris case. Lessons learned were: some data (for example, a Human Trafficking Task Force indicator) had to be completely dropped due to sparsity and the Rural/Urban variable was of relatively high significance for all years. This drove the decision to divide the dataset into urban and rural counties – separate models were run on the two groups to see if the relationship between the independent variables and trafficking activity was different in fundamentally dissimilar types of places. The "Urban" model was limited in scope to only those

counties with 95% or more of the population located in an urban setting. The "Rural" model was limited to only those counties where less than 80% of the population live in an urban setting. For these models, the Foster Care data was too sparse to be included; thus, a separate model was built to investigate only those counties for which there was Foster Care data. These counties were mostly urban (on average, around 94% of their residents live in an urban area) and may be centers of foster care reporting for larger surrounding areas.

The second set of models were logistic regression models which address the binary categorization of whether or not human trafficking activity took place in a given county. The first three models built had Polaris data as the dependent variable (here, in the form of a 0/1 indicator representing no trafficking cases/at least one trafficking case), where geographic scope was again varied to provide separate models for all, urban and rural counties. The final model was a logistic regression using all data, but with the second incidence data source as the dependent variable. This was only performed for 2016, as the sparsity led to questions about how best to continue with the analysis.

The combined set of models – looking at presence and prevalence of trafficking activity in national, urban, rural, and foster-care-reporting regions, per incident data source and per year between 2015 and 2017 – were analyzed and compared. The percentage of variance in data explained by the model (on the full set of counties, as well as on a random set of counties held out for validation), as measured by the coefficient of determination, and regression coefficients and significance levels (p-values) for each variable were analyzed to provide additional insights.

E.6 KEY RESULTS

Key results stemmed from both the development process and interpretation of the resulting models. Insights may be drawn across model-, variable-, county-, and domain-level perspectives.

Model-level:

- The coefficient of determination for these relatively naive models showed they held some explanatory power.

- Sub-national models were important: regression coefficients and significance values associated with each variable differed across national, urban, and non-urban models; the percent urban variable was the most consistently significant variable; data reporting differed by region thus sub-national models allowed for incorporation of datasets otherwise too sparse for inclusion (e.g., foster care data).

- Modeling capability was limited to individual trafficking measures; statistical inferences about human trafficking activity in a broad sense cannot be drawn from the data collected.

- There is a respectable amount of heterogeneity in the models between years which does not bode well for prediction of future activity using the current data extract.

Variable-level:

- The number of law enforcement employees, arrests, and violent offenses are generally consistently significant variables and more helpful in predicting presence of activity than volume of activity.

- Online advertisement data is specifically helpful in predicting volume (not presence) in urban (not rural) counties.

- The percentage of a population living in an urban area was highly significant and positively correlated with trafficking presence, but negatively correlated with trafficking prevalence.

- The impact of the data interpretation and choices around its combination and use will need to be considered before any conclusions can be drawn. What can be noted is the value of data on law enforcement efforts, urban/rural status, and advertisements in the modeling effort. This can inform the next data choices for future models.

- With more robust models, analysis of regression coefficients could be used to generate hypotheses for future trafficking research and consistently highly significant variables could be used to identify risk-factors associated with reported trafficking activity.

County-level:

- Model residuals can be used to identify outlier counties, for which model-predicted cases per number-of-people are higher than the observed cases. With a more robust model, this could be used to identify counties with undetected trafficking or successful counter-trafficking strategies, as well as counties which are at risk for future trafficking growth based on current projections.

Domain-level:

- The nature of human trafficking crime makes truth data difficult to establish: aspects of the crime and response can occur at multiple times and locations over a variable-length range of time; there is a large amount of undetected activity, which is problematic for establishing "no-trafficking" observations; and current trafficking measures are convenience samples, which poses limitations on the inferences that can be made.

- Existing data is difficult to collect, due to the disparate datasets involved, barriers to data sharing, data loss due to privacy concerns, and lack of persistent collection mechanisms.

- The desired data for modeling, in a usable format, may not exist: there is a gap between data collected for individual organizations' needs and that which is needed to address field-wide needs and research questions.

- Data fusion can be difficult due to geospatial and temporal alignment considerations and a lack of recorded information necessary to fuse across sources.

- Data sparsity is a significant hurdle in modeling human trafficking activity.

- Technology solutions less sophisticated than modeling (e.g., centralized data display, comparison and trend analysis) could make large steps towards providing national situational awareness of human trafficking crime.

E.7 CONCLUSION

The need for a national view of human trafficking so that agencies can collaboratively wage a battle is clear. The capabilities of NICS, serving as a platform for this collaboration, were demonstrated so that domain experts could better refine their needs in this area. One such need is a predictive model to inform resource allocation and policy making. To our knowledge, a model at the county level has never been built before. A fundamental issue is understanding the strengths and weaknesses of our current level of data collection. This modeling effort is helping to tease out how much data and of what type is required, but remains incomplete. Foster Care and Law Enforcement data are emerging as key inputs, which today are not available for all counties or in real-time. The role of the 2010 Census variable measuring the percentage of population living in an urban area needs to be understood to indicate whether this variable itself is of value, or something related to those environments needs to be measured. It is recommended that efforts continue in order to uncover the best current data available and ascertain its current utility to operational partners and researchers, while defining and refining requirements for potential new datasets, data infrastructure, and intelligence tools that could aid the fight against human trafficking.