Code 8.7: Applying AI in the Fight Against Human Trafficking

Lehigh Back to the Classroom
April 26, 2019

Professor Daniel Lopresti
Chair, Computer Science and Engineering
Director, Data X Initiative
Human Trafficking in the News

But it may be more complicated than this ...
What is Human Trafficking?

“The Protocol to Prevent, Suppress and Punish Trafficking in Persons defines Trafficking in Persons as ...

- the recruitment, transportation, transfer, harbouring or receipt of persons,
- by means of the threat or use of force or other forms of coercion, of abduction, of fraud, of deception, of the abuse of power or of a position of vulnerability or of the giving or receiving of payments or benefits to achieve the consent of a person having control over another person,
- for the purpose of exploitation.”

What is Human Trafficking?

“Exploitation shall include, at a minimum, the exploitation of the **prostitution** of others or other forms of sexual exploitation, **forced labour** or services, **slavery** or practices similar to slavery ...”

Common Misconceptions

- Occurs both across borders and within a country (not just cross-border).
- Is for a range of exploitative purposes (not just sexual exploitation).
-Victimizes children, women and men (not just women, or adults, but also men and children).
- Takes place with or without the involvement of organized crime groups.

Human Trafficking

"Human trafficking is modern-day slavery and involves the use of force, fraud, or coercion to obtain some type of labor or commercial sex act.

Every year, millions of men, women, and children are trafficked in countries around the world, including the United States. It is estimated that human trafficking generates many billions of dollars of profit per year, second only to drug trafficking as the most profitable form of transnational crime.

Human trafficking is a hidden crime as victims rarely come forward to seek help because of language barriers, fear of the traffickers, and/or fear of law enforcement."

https://www.dhs.gov/blue-campaign/what-human-trafficking
Human Trafficking (DHS)

https://www.dhs.gov/blue-campaign/what-human-trafficking

Code 8.7: Applying AI in the Fight Against Human Trafficking
Lopresti Slide 8
Human Trafficking (DHS)

https://www.dhs.gov/blue-campaign/what-human-trafficking

Code 8.7: Applying AI in the Fight Against Human Trafficking

Lopresti
Slide 9
Applying AI to Fight Trafficking

Our local collaboration:

- The WHY: Owens
- Lehigh CSE: Lopresti, Baumer, Heflin, Xie, et al.
- RIIC: Kocis, et al.
Jennifer Gentile Long (MGA, JD) currently serves as the Chief Executive Officer of AEquitas, which she co-founded in April 2009 and serves as an expert on issues related to the prosecution of sexual violence and gender-based violence and human trafficking. She began her career as an Assistant District Attorney in Philadelphia where she prosecuted cases involving adult and child physical and sexual abuse and served on a team in the Family Violence and Sexual Assault Unit. After her departure, served as an advocate for victims of domestic violence and child abuse in Bermuda and Philadelphia, respectively.

Jennifer graduated from Lehigh University with a Bachelor of Arts in English and East Asian Studies and the University of Pennsylvania Law School and Fels School of Government with a Juris Doctor degree and a Masters in Government Administration. She is a member of the Pennsylvania and New Jersey bars. She enjoys running, walking her dogs and making homemade pancakes for her family.
The WHY: Colleen Owens

Footwear designer, Fashion Institute of Technology Accessories Design Program and Bryn Mawr College alumnus, human trafficking researcher with 10+ years of expertise, career changer, risk taker, dreamer and doer, voracious reader, athlete, vegetarian, lover of all things Alexander McQueen and Philip Treacy, believer in fashion activism and the power of conscious consumerism to make a positive impact on the ethical treatment of people, animals and our planet.

https://www.linkedin.com/in/colleeneowens/

Combating Labor Trafficking: An Interview with Colleen Owens

When you hear the phrase "human trafficking," what immediately comes to your mind?

If you're like most people, you might think of sex workers or young girls in bondage. But this modern-day form of slavery goes far beyond the sex trade and has no place in our free society.

But how can law enforcement take a proactive approach to this seemingly nebulous, ever-expanding problem?

Be sure to check out this recorded webinar, as International Trafficking Expert Colleen Owens and Kaitlin Seale of DHS' Blue Campaign as they discuss:

- labor trafficking in the United States
- research results about US labor trafficking,
- information about DHS resources to combat the issue.

https://www.linkedin.com/in/colleeneowens/
The Regional Intelligence and Investigation Center (RIIC) in Allentown integrates police departments in the county and dozens of local, state and federal databases in one electronic crime-fighting resource. More than 650 investigators and other law enforcement personnel have been trained in using the information portal and search applications.

The RIIC opened in 2013 and already has revolutionized the way Lehigh County police departments analyze and share collected data to solve crimes. The primary purpose of the Center is to offer investigative support, strategic analysis and situational awareness to county law enforcement. Expert assistance by RIIC personnel and a consolidated technology platform deliver critical information quickly, conveniently and in a way not previously possible.

Criminal investigators can post and share information about alerts and suspects and can search more than 4 million police and prison records.

Director Julia Kocis, who led the development team for the RIIC, and crime analysts at the center have assisted investigators in solving six homicides at year end 2016.

https://www.lehighcounty.org/Departments/District-Attorney
The Four V’s of Big Data

Volume: Enterprises are awash with ever-growing data of all types, easily amassing terabytes – even petabytes – of information.

Velocity: Sometimes 2 minutes is too late. For time-sensitive processes such as catching fraud, big data must be used as it streams into the enterprise in order to maximize its value.

Variety: Big data is any type of data - structured and unstructured data such as text, sensor data, audio, video, click streams, log files and more. New insights are found when analyzing these data together.

Veracity: 1 in 3 business leaders don’t trust the information they use to make decisions. How can you act upon information if you don’t trust it? Establishing trust in big data presents a huge challenge as the variety and number of sources grows.

http://www-01.ibm.com/software/data/bigdata/
Table 1: Available datasets, either internal to RIIC or from public domains, that can lead to the identification and prosecution of various crimes, including human trafficking. The column Area indicates crime types covered by a dataset (S=sex trafficking, L=labor trafficking, D=drug, G=gang activity).

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Source</th>
<th>Area</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Police narratives</td>
<td>RIIC</td>
<td>S,L,D,G</td>
<td>Large</td>
<td>Natural language text containing people (criminals, victims, other involved parties), addresses, phone numbers, vehicles, drugs, etc.</td>
</tr>
<tr>
<td>Coded incidents</td>
<td>RIIC</td>
<td>S, L, D, G</td>
<td>Large</td>
<td>Coded perpetrator, victim(s), crime, date, time, location.</td>
</tr>
<tr>
<td>Gang Intelligence system</td>
<td>RIIC</td>
<td>S, L, D, G</td>
<td>Large</td>
<td>Integrate data from various law enforcement units in Lehigh County</td>
</tr>
<tr>
<td>records</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incarceration records</td>
<td>RIIC</td>
<td>S, L, D, G</td>
<td>Large</td>
<td>Coded jail phone calls, visitors and finance logs 4.</td>
</tr>
<tr>
<td>Human trafficking cases</td>
<td>Polaris</td>
<td>S, L, D, G</td>
<td>Small</td>
<td>Articles from the Polaris project</td>
</tr>
<tr>
<td>Social network</td>
<td>Facebook,</td>
<td>S, L, D, G</td>
<td>Large</td>
<td>Gang members’ and victims’ activities on social networks.</td>
</tr>
<tr>
<td></td>
<td>Twitter, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illicit websites</td>
<td>bedpage.com,</td>
<td>S, L, D, G</td>
<td>Large</td>
<td>Commercial sex advertisements in the format of short texts. Contain phone numbers, emails and address.</td>
</tr>
<tr>
<td></td>
<td>cityxguide.com,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>eroticmp, etc.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case law</td>
<td>DOL</td>
<td>L</td>
<td>Small</td>
<td></td>
</tr>
<tr>
<td>Wage and hour violation data</td>
<td>DOL</td>
<td>L</td>
<td>Small</td>
<td></td>
</tr>
</tbody>
</table>
Our Project

Considering All the Evidence: Human-in-the-Loop Artificial Intelligence to Combat Criminal Networks

Law enforcement data
- Police narratives
- Incarceration records
- Gang intelligence system record
- Social networking activity
- Case laws
- Reports from NGOs

AI models
- Entity extraction
- Relation extraction
- Graph mining
- Topic modeling
- Text classification

Human-interpretable results
- People
- Location
- Crime-type
- Topic 1: gun, shooting, ...
- Topic 2: sex trafficking, drug
- Topic 3: ...

Law enforcement actions
- Prosecution
- Investigating a particular case
- Closer monitoring
- Police patrol plan
- Bailout recommendation

Positive and negative feedbacks to models
Our Project

Human-in-the-loop training of a text classifier using dual supervision

RIIC narrative corpus

**Narrative 1:**
Natalie was being kept in a hotel against her will and forced into prostitution.

**Narrative 2:**
The workers were subjected to emotional and physical threats, and their identification was confiscated so they couldn’t leave.

**Narrative 3:**
A car accident is over in a matter of seconds, but the paperwork and phone calls with your insurance company can take weeks.

Classifier

**Recommended keywords:**
hotel, prostitution, kept, being, threats, confiscated

**Recommended narratives:**
Narrative 1, Narrative 3

Labeling features & narratives

**Selected keywords:**
hotel, prostitution, threats, confiscated

**Selected narratives:**
Narrative 1

Model re-training

**Relevant connections:**
- Recommendation with explanations
- Recommendation
- RIIC narrative corpus
- Classifier
- Labeling features & narratives

Code 8.7: Applying AI in the Fight Against Human Trafficking
8.7 Initiatives
“Sustainable Development Goal Target 8.7 calls on all to take immediate and effective measures to eradicate forced labour, end modern slavery and human trafficking and secure the prohibition and elimination of all forms of child labour as an essential step to achieving decent work for all, full and productive employment and inclusive and sustained economic growth.”

Delta 8.7 is a global knowledge platform exploring what works to eradicate forced labour, modern slavery, human trafficking and child labour, an aim set out in Target 8.7 of the UN Sustainable Development Goals.

Think of Delta 8.7 as the data supporting Target 8.7.
Story 1: Prevalence of Modern Slavery and Child Labour

https://delta87.org/earhtime/#anchor1
Story 2: Government Efforts to End Modern Slavery

https://delta87.org/earthtime/#anchor2
If Delta 8.7 is the data, Code 8.7 is the international research community in AI and computational science.
Optimize government data to measure prevalence and model vulnerability to modern slavery from law enforcement, prosecution, social welfare, unemployment and regional socioeconomic vulnerability factors:

- Structured and unstructured, temporal and spatial data and their coordination;
- Implementation of network analysis and graph theory;
- Communication to policymakers; and
- Analysis coordination.

Bias; fairness; it’s not just data - it’s people.

Using satellite imagery, drone footage and earth observation and mapping to identify human trafficking.

Brick kilns in India

https://delta87.org/2019/03/slavery-from-space/
https://www.sciencemag.org/news/2019/02/researchers-spy-signs-slavery-space
In most instances, the exploitation of an individual generates a financial transaction, and/or contributes to the financial gain of an individual or company:

- Structured data (commonly in the form of suspicious financial transactions);
- Unstructured data (information provided by civil society groups or service providers working with victims); and
- Publicity available information (social media or news).

Informal transactions; privacy; misinformation.

https://delta87.org/2019/03/fintech/
Anti-trafficking community must first establish some core building blocks that can serve as the foundation upon which new technologies can be developed and brought to bear:

- Shared strategy;
- A common infrastructure that allows for better and more sharing of data;
- A pipeline that shepherds ideas from basic research to applied research to operationalization and finally to demo and validation.

Data trusts; tech industry; AI is not a silver bullet.

https://delta87.org/2019/03/towards-pipeline-technology-techniques-training/
Challenge is to transform this emerging community of interest into a community of intent. Requires building a framework for collaboration that creates incentives for innovation while achieving collective impact:

- Must understand limits to the potential of AI;
- Constraining lack of standardization;
- Persistent barriers to information and data-sharing;
- Ongoing lack of strategic coordination among donors, funders and resource-allocators.

Planning for next steps now underway.

https://delta87.org/2019/03/how-we-can-advance-collaborative-problem-solving/
COMPUTING COMMUNITY CONSORTIUM

The Computing Research Association’s Computing Community Consortium (CCC) catalyzes the North American computing research community to enable the pursuit of innovative, high-impact research.

Who:
- Council – 20 members
- CCC/CRA Staff
- Chair, VC, & Director

What:
- Workshops & Conf. Blue Sky Tracks
- Whitepapers & Social Media
- Reports Out (esp. to government)
- Biannual Symposium

Professional Development:
- Early Career Workshops & Participation
- Council Membership
- Leadership w/ Gov’t (LISPI)
FOUNDATIONAL AI RESEARCH

From recent discussions within the AI research community:

- Open knowledge repositories
- Contextualized AI
- Human / AI collaboration
- Trust and responsibility
- Diversity of interaction channels
- Improving online interaction
- Robust and trustworthy machine learning
- Deeper learning for challenging tasks
- Integrating symbolic and numeric representations
It’s exciting to be able to play a role in such an important international effort. Check back with me as we progress!

Thank you!