Analyzing Possible Human Trafficking Activity Using Large-Scale Information Extraction

Michael Cafarella

Joint Work with Jacob Shapiro, Chris Re, Ce Zhang, Jaeho Shin,

and many others

University of Michigan

Computer Science and Engineering
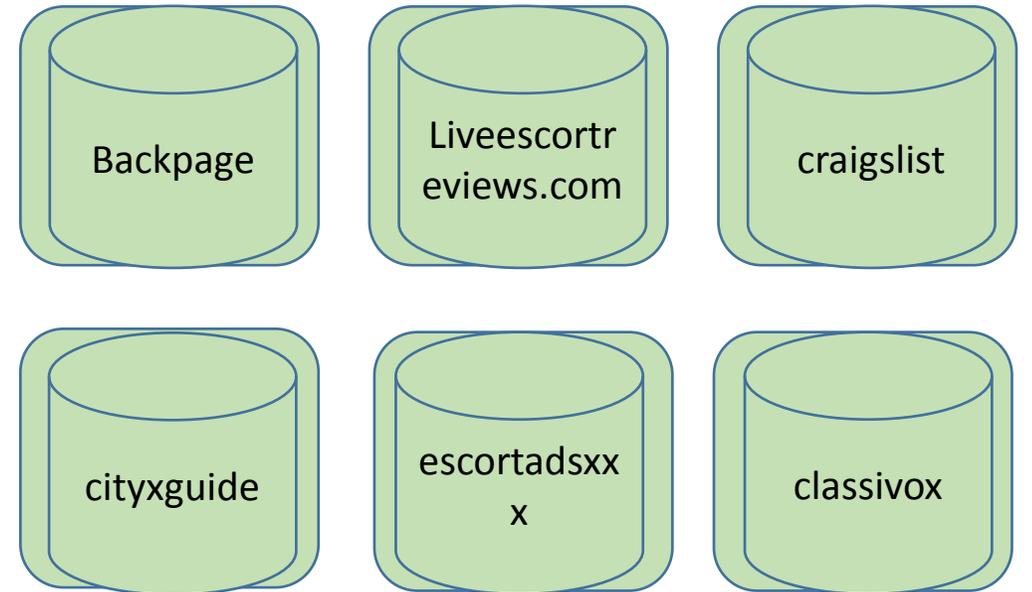
May 2, 2016

Human Trafficking

- Odious crime that obtains labor via physical, financial coercion
- Traffickers believed to channel many victims into sex work
- Trafficking phenomena very opaque
  - Unclear which sex workers are in fact trafficking victims
  - Unclear which sex work organizers engage in trafficking
  - Law enforcement lacks any good intelligence
  - Surprisingly old problem

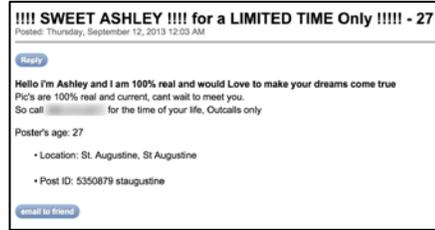Trafficking image, US Department of Justice

Human Trafficking and Web Data

- Sex work has gone online

- 80M advertisements posted on dozens of websites

- Advertisements contain:
  - Contact info
  - Personal details
  - Location
  - Price
  - Transaction details
  - Lots of other data

- Grim, but potential analytical gold mine

- Unfortunately, few or no H/T labels

Backpage

Liveescortreviews.com

craigslist

cityxguide

escortadsxxx

classivox

Our Goals

1. Transform 80M raw ads into high-quality database

!!!! SWEET ASHLEY !!!! for a LIMITED TIME Only !!!!! - 27
Posted: Thursday, September 12, 2013 12:03 AM

Reply

Hello i'm Ashley and I am 100% real and would Love to make your dreams come true
Pic's are 100% real and current, cant wait to meet you.
So call _____ for the time of your life, Outcalls only

Poster's age: 27

• Location: St. Augustine, St Augustine

• Post ID: 5350879 staugustine

email to friend

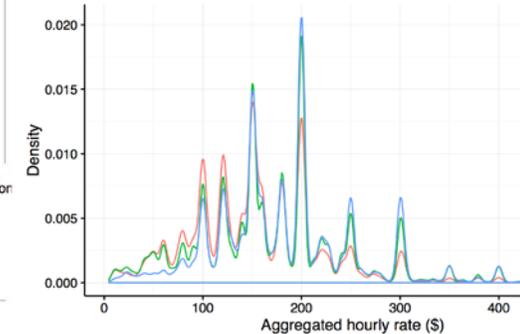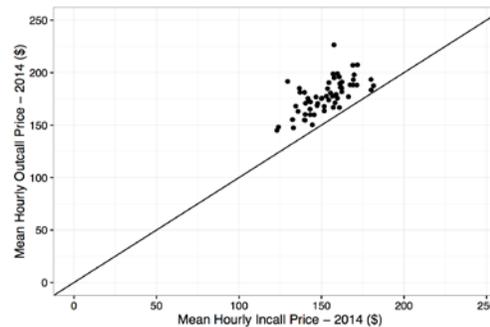| Name	Age	Location	MSA
Ashley	27	St Augustine	Jacksonville

2. Build database-bac[k]ed ~~ enforcement tools to ~~

Novel social science and tools made possible by computational methods
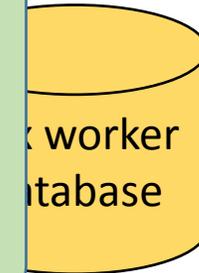
x worker
atabase

3. Long term plan: Use database for novel social science findings, further help law enforcement

Tools for Identifying Trafficking Victims with Data

- Is sex worker:
  - Offering surprisingly low price?
  - Willing to engage in risky services?
  - Moving unusually frequently?
  - Sharing contact info with suspected h/t victim?

- Does advertisement:
  - Use numerically-nearby phone numbers?
  - Reuse ad text from a suspected h/t victim?
  - Describe someone with possible warning signs (physical abuse, drug abuse, fear of law enforcement, etc)?

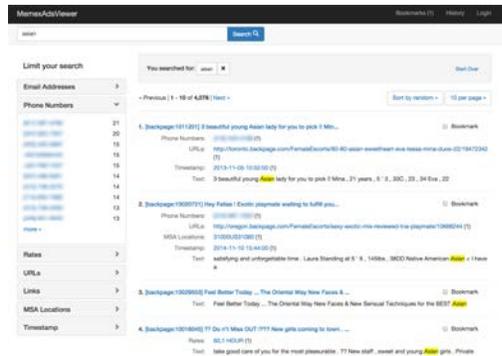60 Minutes Overtime, CBS.com

HT Advertising Table

Ad Identifier	Date Posted	(Craigslist) Region	...	Ad Text
10001	08/12/15	Wash		...
10002	08/15/15	NYC		...
10003	08/16/15	LA		...

Aspirational Table

Ad Identifier	Price/ Hour	Place(s)	Incall/ Outcall	Phone	Email
10001	100	Bethesda	Incall	555...	m@...
10002	225	Nyack	Outcall	555...	c@...
10003	...	...	...	...	...

Is the number of sex service ads increasing or decreasing?

What large region has the most ads?

What is the typical price for offered services?

Which providers are posting multiple times?
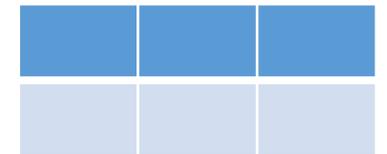
Which exact locality is growing in popularity?

Faceted Search

Data Analytics

Visualizations

Other Tools

Labor Market Interactions

- How do prices react to women's opportunities in broader market?
- If prices are insensitive to broader conditions, it might suggest a large amount of compelled participation

- Use Bartik wages [Bartik 1991] to measure labor market demand apart from local supply
- Build regression model to predict price in a given MSA and month
  - Control for posting venue and commute distance
  - Control for MSA-specific effects, seasonal time-of-year effects

Labor Market Findings

- After controlling for MSA-specific effects, robust *negative* correlation between female opportunity and hourlong price
  - $1 in overall wage increase predicts $0.43 **decrease** in hourly rate
  - Why??
- Divide market into low-, mid-, high-end prices (<25%, 25-75; >75%)
- Negative correlation between female opportunity and # mid-price ads
  - Predicts 7.1% drop in first-time advertisers at mid-price (signif at 0.01)
  - 1$ in overall wage increase predicts 3% drop in overall participation
- But, positive correlation between female opp'ty and # low-price ads
  - Almost identical rise in participation; 3.6% increase in first-time ads