An Analysis of the Factors that Influence Internet Searches on Cyber Bullying Using Social Big Data

Korean Institute of Criminology
International Center for Criminal Justice
Juyoung Song, Ph.D.
Introduction

• **Education**
  – Ph.D. in Criminal Justice, School of Criminal Justice, Michigan State University
  – M. A. in Criminal Law, College of Law, Hanyang University
  – B. A. in Law, College of Law, Hanyang University

• **Professional Background**
  – Associate Research Fellow, Korean Institute of Criminology
  – Assistant Professor, Department of Criminology, University of West Georgia

• **Research Interest**
  – Delinquency (Bullying, Cyber Bullying, Underage Prostitution)
“Data is the New OIL.”
What is Big Data?

“Every day, we create 2.5 quintillion bytes of data — so much that 90% of the data in the world today has been created in the last two years alone. This data comes from everywhere: sensors used to gather climate information, posts to social media sites, digital pictures and videos, purchase transaction records, and cell phone GPS signals to name a few.

This data is “big data.”
The 3 Vs: Volume, Velocity, Variety
+ The 2Vs, 1C: Veracity, Value, Complexity
Big Data Analysis Example

**US health care**
- $300 billion value per year
- ~0.7 percent annual productivity growth

**Europe public sector administration**
- €250 billion value per year
- ~0.5 percent annual productivity growth

**Global personal location data**
- $100 billion+ revenue for service providers
- Up to $700 billion value to end users

**US retail**
- 60+% increase in net margin possible
- 0.5–1.0 percent annual productivity growth

**Manufacturing**
- Up to 50 percent decrease in product development, assembly costs
- Up to 7 percent reduction in working capital

SOURCE: McKinsey Global Institute analysis
Whom does it matter?

- Research Community
  - Social Science
    - Criminology
      - Crime Prevention
- Business Community
  - New tools, new capabilities, new infrastructure, new business models etc.
- On sectors
The Research Project

• Risk Terrain Modeling
  – Big Data from police, criminal career
  – GIS software
  – Hot spot $\rightarrow$ Prediction on Crime
  – Risk factors on Environment $\rightarrow$ Crime rate
The Social Layer in an Instrumented Interconnected World

30 billion RFID tags today (1.3B in 2005)

4.6 billion camera phones worldwide

100s of millions of GPS enabled devices sold annually

12+ TBs of tweet data every day

25+ TBs of log data every day

200M by 2014

76 million smart meters in 2009...

billions of people on the Web by end 2011
Definition of Cyber Bullying

• **Computers in Human Behavior**

• Willful and repeated harm inflicted through the medium of electronic text (Hiduja and Patchin, 2006)

• Cyber bullying involves the intentional use of e-mail, networking sites, cell phones, chat rooms, websites and other forms of internal communications to harass, threaten, or victimize an individual or a group by an individual or group.
Cyber Bullying

- Involves the intentional use of e-mail, networking sites, cell phones, chat rooms, websites and other forms of internal communications to harass, threaten, or victimize an individual or a group by an individual or group.

- Two main forms:
  - Internet
  - Cell phones

The appeal of cyber bullying is that one may remain anonymous. Bullying a victim is easier when the effects are not seen by the bully.
Cyber Bullying and the Effects

- Physical: 
  - loss or gain
  - Bruises or cuts
  - Substance abuse
  - Death

- Psychological

- Emotional
Cyber Bullying and the Effects

- Anxiety
- Depression
- Loss of self worth
- Loss of self esteem

- Extreme sensitivity
- Aggression
- Helplessness
The Model

- Stress
- Drinking
- Exercise
- Anger
- Cyber-Bullying
- Suicide
Level 2
- Economically active population
- Rental Price Index
- Suicide Rate

Level 1
- Searching Suicide
- Searching Stress
- Searching Alcohol
- Searching Excercise
- Searching Anger

Cyberbullying Searching
Measurement

  - 371,209 Buzz
  - Suicide
  - Stress
  - Drinking
  - Exercise
  - Anger
- Multilevel Modeling
- Level 1: Suicide, Stress, Alcohol, Exercise, Anger
- Level 2: Economically active population, Rental price index, Suicide rate
### <Table 1> Multivariate normality of the main variables

<table>
<thead>
<tr>
<th>Year</th>
<th>Searching Cyberbullying</th>
<th>Searching Suicide</th>
<th>Searching Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± S.D.</td>
<td>K&lt;sup&gt;a)&lt;/sup&gt;</td>
<td>S&lt;sup&gt;b)&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±302.27</td>
<td>21.87</td>
</tr>
<tr>
<td>2011</td>
<td>270.13</td>
<td>7.36</td>
<td>41.44</td>
</tr>
<tr>
<td></td>
<td>±302.27</td>
<td>±33.49</td>
<td>±4.04</td>
</tr>
<tr>
<td></td>
<td>±302.27</td>
<td>±33.49</td>
<td>±4.04</td>
</tr>
<tr>
<td>2012</td>
<td>744.84</td>
<td>79.63</td>
<td>69.48</td>
</tr>
<tr>
<td></td>
<td>±651.88</td>
<td>±35.52</td>
<td>±5.32</td>
</tr>
<tr>
<td></td>
<td>±651.88</td>
<td>±35.52</td>
<td>±5.32</td>
</tr>
<tr>
<td>Year</td>
<td>Searching Alcohol</td>
<td>Searching Exercise</td>
<td>Searching Anger</td>
</tr>
<tr>
<td></td>
<td>Mean ± S.D.</td>
<td>K&lt;sup&gt;a)&lt;/sup&gt;</td>
<td>S&lt;sup&gt;b)&lt;/sup&gt;</td>
</tr>
<tr>
<td>2011</td>
<td>7.86</td>
<td>45.50</td>
<td>5.88</td>
</tr>
<tr>
<td></td>
<td>±8.12</td>
<td>±5.24</td>
<td>±5.24</td>
</tr>
<tr>
<td>2012</td>
<td>12.22</td>
<td>9.41</td>
<td>2.34</td>
</tr>
</tbody>
</table>

* a) Kurtosis, b) Skewness, c) Number of case
<Table 2> The mediating effect between stress vulnerable factors (exercise, alcohol) and anger search

<table>
<thead>
<tr>
<th>Route</th>
<th>Alcohol</th>
<th>Exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Effect</td>
<td>Direct Effect</td>
</tr>
<tr>
<td>Stress -&gt; Cyberbullying</td>
<td>.488</td>
<td>.354</td>
</tr>
<tr>
<td>DP→MP²</td>
<td>.488*→.354*</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Route</th>
<th>Anger</th>
<th>Cyberbullying</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Effect</td>
<td>Direct Effect</td>
</tr>
<tr>
<td>Stress -&gt; Suicide</td>
<td>.595</td>
<td>.517</td>
</tr>
<tr>
<td>DP→MP²</td>
<td>.595*→.517*</td>
<td></td>
</tr>
<tr>
<td>Cyberbullying -&gt;Suicide</td>
<td>.716</td>
<td>.654</td>
</tr>
<tr>
<td>DP→MP²</td>
<td>.716*→.654*</td>
<td></td>
</tr>
</tbody>
</table>

* a) Sobel Test: * p<0.01  
  b) Mediator Effect: DP(Direct Path coefficient), MP(Mediator Path coefficient)
### Table 5-1: Koreans' cyberbullying search determinants in the multi-level model (Fixed effect)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model</th>
<th>Unconditional Model</th>
<th>Unconditional Slope Model</th>
<th>Conditional Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Coef.</td>
<td>S.E.</td>
<td>Coef.</td>
</tr>
<tr>
<td>Fixed Effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept ( )</td>
<td>Level 1</td>
<td>2.58</td>
<td>0.07**</td>
<td>2.30</td>
</tr>
<tr>
<td>Searching Suicide</td>
<td>Level 1</td>
<td>0.09</td>
<td>0.02**</td>
<td>0.09</td>
</tr>
<tr>
<td>Searching Stress</td>
<td>Level 1</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Searching Alcohol</td>
<td>Level 1</td>
<td>0.05</td>
<td>0.02**</td>
<td>0.06</td>
</tr>
<tr>
<td>Searching Excercise</td>
<td>Level 1</td>
<td>0.14</td>
<td>0.03**</td>
<td>0.14</td>
</tr>
<tr>
<td>Searching Anger</td>
<td>Level 1</td>
<td>0.02</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>Economically active</td>
<td>Level 2</td>
<td></td>
<td></td>
<td>-0.00</td>
</tr>
<tr>
<td>population</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rental Price Index</td>
<td>Level 2</td>
<td>0.07</td>
<td>0.01**</td>
<td></td>
</tr>
<tr>
<td>Suicide Rate</td>
<td>Level 2</td>
<td></td>
<td></td>
<td>-0.01</td>
</tr>
</tbody>
</table>

a)**p<.01, * p<.05
<Table 5-2> Koreans' cyberbullying search determinants in the multi-level model (Random effect)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model</th>
<th>Unconditional Model</th>
<th>Unconditional Slope Model</th>
<th>Conditional Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>S.E.</td>
<td>S.E.</td>
<td>S.E.</td>
</tr>
<tr>
<td>Random Effect</td>
<td>Coef.</td>
<td>0.124</td>
<td>0.135</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>S.E.</td>
<td>2961.63*</td>
<td>215.66**</td>
<td>100.28**</td>
</tr>
<tr>
<td>level 2, Searching</td>
<td>Coef.</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
</tr>
<tr>
<td>level 1, Searching</td>
<td>Coef.</td>
<td>0.030</td>
<td>0.018</td>
<td>0.018</td>
</tr>
<tr>
<td>Searching Suicide</td>
<td>Coef.</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Searching Stress</td>
<td>Coef.</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Searching Alcohol</td>
<td>Coef.</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Searching Exercise</td>
<td>Coef.</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Searching Anger</td>
<td>Coef.</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>ICC</td>
<td></td>
<td>0.805</td>
<td>0.882</td>
<td>0.714</td>
</tr>
</tbody>
</table>

a)**p<.01, * p<.05
• Multivariate Analysis of Factors for cyber bullying using Social Big Data
The Closing Comment
Thank you

Juyoung Song

Contact Information
juyoung81@gmail.com