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, Complexty

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→ Prediction on Crime
  - Risk factors on Environment→ Crime rate

The Social Layer in an Instrumented Interconnected World

12+ TBs of tweet data every day

? TBs of data every day

25+ TBs of log data every day

30 billion RFID tags today (1.3B in 2005)

4.6 billion camera phones world wide

100s of millions of GPS enabled devices sold annually

76 million smart meters in 2009… 200M by 2014

2+ billion people on the Web by end 2011

Definition of Cyber Bullying

- **Computers in Human Behavior**
  http://www.sciencedirect.com/science/article/pii/S0747563213003579

- 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

Psychological

Emotional

Physical

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

Cyber Bullying and the Effects

Psychological

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

Emotional

- Extreme sensitivity
- Aggression
- Helplessness

The Model

KIC 한국형사정책연구원 Korean Institute of Criminology

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

- 2011.1.1.~ 2012. 12. 31.
  - 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

Year	Searching Cyberbullying				Searching Suicide				Searching Stress			
	Mean ±S.D.	K[a)]	S[b)]	N[c)]	Mean ±S.D.	K[a)]	S[b)]	N[c)]	Mean ±S.D.	K[a)]	S[b)]	N[c)]
2011	270.13 ±302.27	21.87	3.87	98,082	7.36 ±33.49	41.44	6.29	8,033	4.90 ±4.04	13.07	3.09	1,788
2012	744.84 ±651.88	79.63	7.38	271,808	21.95 ±35.52	69.48	6.83	8,033	9.72 ±5.32	8.97	2.05	3,556

Year	Searching Alcohol				Searching Exercise				Searching Anger			
	Mean ±S.D.	K[a)]	S[b)]	N[c)]	Mean ±S.D.	K[a)]	S[b)]	N[c)]	Mean ±S.D.	K[a)]	S[b)]	N[c)]
2011	7.86 ±8.12	45.50	5.88	2,868	6.89 ±5.24	26.41	4.01	2,515	1.85 ±2.63	32.59	4.56	674
2012	12.22 ±7.16	9.41	2.34	4,473	11.07 ±6.45	24.42	3.37	4,050	2.65 ±2.43	18.55	3.00	971

* a) Kurtosis, b) Skewness, c) Number of case

<Table 2> The mediating effect between stress vulnerable factors (exercise, alcohol) and anger search

Route	Alcohol			Exercise		
	Total Effect	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect[1)]
Stress -> Cyberbullying	.488	.354	.135*	.488	.355	.133*
DP→MP[2)]	.488*→.354*			.488*→.355*		
Route	Anger			Cyberbullying		
	Total Effect	Direct Effect	Indirect Effect	Total Effect	Direct Effect	Indirect Effect[1)]
Stress -> Suicide	.595	.517	.078*	.595	.322	.273*
DP→MP[2)]	.595*→.517*			.595*→.322*		
Cyberbullying ->Suicide	.716	.654	.062*			
DP→MP[2)]	.716*→.654*					

* 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)

Parameter \ Model		Unconditional Model		Unconditional Slope Model		Conditional Model	
Fixed Effect		Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Level 1	Intercept (γ_{00})	2.58	0.07**	2.30	0.08**	2.31	0.05**
	Searching Suicide			0.09	0.02**	0.09	0.02**
	Searching Stress			0.02	0.02	0.02	0.02
	Searching Alcohol			0.05	0.02**	0.06	0.02**
	Searching Excercise			0.14	0.03**	0.14	0.03**
	Searching Anger			0.02	0.03	0.02	0.02
Level 2	Economically active population					-0.00	0.00*
	Rental Price Index					0.07	0.01**
	Suicide Rate					-0.01	0.01

a)**p<.01, * p<.05

<Table 5-2> Koreans' cyberbullying search determinants in the multi-level model (Random effect)

Model / Parameter	Unconditional Model		Unconditional Slope Model		Conditional Model	
Random Effect	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
level 2,	0.124	2961.63**	0.135	215.66**	0.045	100.28**
level 1,	0.030		0.018		0.018	
Searching Suicide			0.01	48.28**	0.01	52.71**
Searching Stress			0.00	30.39		
Searching Alcohol			0.00	13.39		
Searching Exercise			0.02	37.00*	0.02	47.33**
Searching Anger			0.01	38.50*	0.01	31.77
ICC	.805		.882		.714	

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