



HMCollab Approach to Social Data:

How digital can inform your campaign and what social media can tell us about tobacco behavior

Glen Szczypka, Deputy Director, HMCollab
@glenszczypka



NCTOH 2017



Road Map



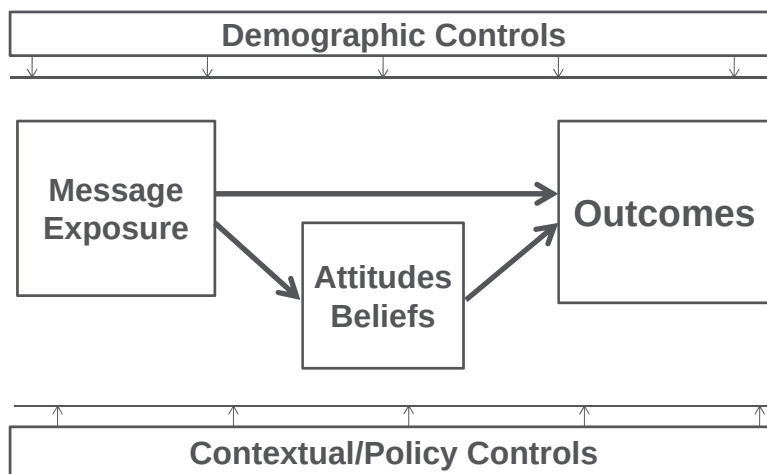
- Background: Brief history of HMCollab
- Social Data 101
- Examples/Applications



A Brief History of the HMCollab

3

Conceptual Framework Pre-2011



4

Social Media is the Real World for Youth and Young Adults

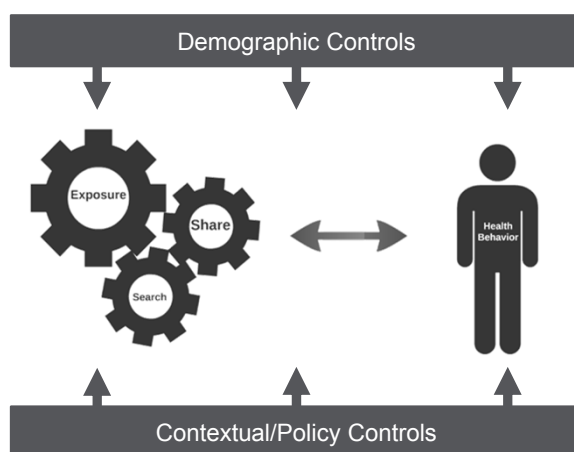
65% of ALL American Adults

90% of 18-29 Year Olds

77% of 30-49 Year Olds



New Paradigm → New Conceptual Framework



The HMC Data Archives

08/2014

 20,000 Pages
 3 Million
 Comments

12/2011

 250 Million
 Tweets &
 Metadata



2001-Present

 TV Ratings

07/2014

 5.5 Million


 May 13: 1M Videos
 Aug 14: 9M Comments

2013-Present

 Print, Outdoor
 & Internet Ads


 at the UNIVERSITY of CHICAGO

7

Q: How do you make sense of so much data?

A: Combination of Human and Computer Brains

Machine Classification

- ▶ Data cleaning
- ▶ Content/source analysis

Topic Modeling

- ▶ Content discovery
- ▶ Confirmatory analysis of content classification

Iterative process of human
 labeling and computer
 modeling


 at the UNIVERSITY of CHICAGO

8

Social Media 101

The Anatomy of a Tweet – Not just 140 Characters

The screenshot shows a Twitter interface. On the left, a vertical list of tweets is partially visible. The main focus is a tweet from Joy McCarthy (@joyoushealth) posted on May 31, 2014, at 9:00 PM. The tweet text is: "Looking to quit smoking for World No Tobacco Day? Consider herbal remedies like New Roots Stop Smoking: ow.ly/xs96s". The tweet has 2 retweets and 1 like. There are two orange arrows pointing to the retweet and like counts. Another orange arrow points to the tweet's text. A fourth orange arrow points to the bottom of the tweet, near the interaction icons. The bottom of the image features a large orange text overlay: "Tweets are 140 c plus 232 data points known as m".

Dimensions to Consider on Social Media

- Amount
- Content
 - Product type
 - Valence: pro/anti
 - Credibility
- Source
 - Commercial
 - Social
 - Vested Interest/Advocate
 - Earned Media
- Diffusion/Network

The Importance and Challenge of Identifying of Message Source

Organic Users

Influential Users

Commercial Users



Mining the Mountain of....Information

Why “HOW” Matters

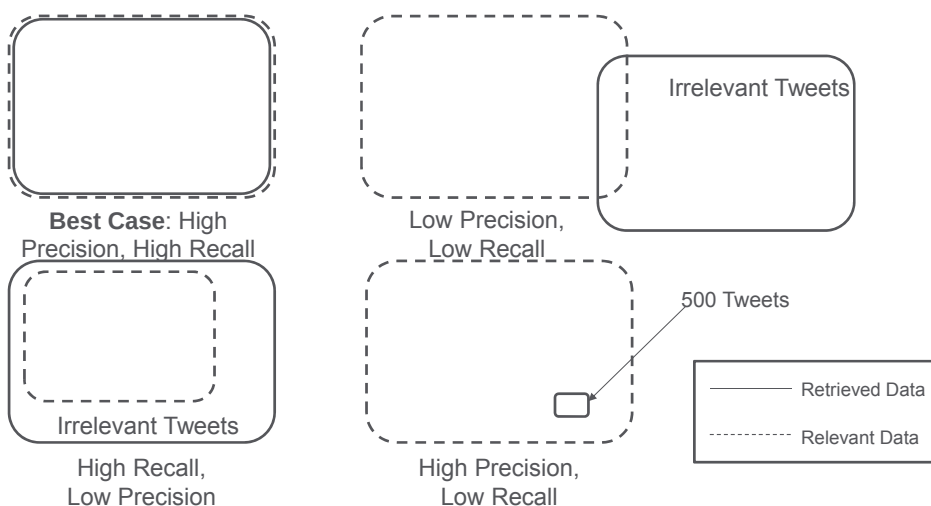
13

Garbage In, Garbage Out



- Keywords are how we get all social data
- KWs depend on expert knowledge and are rarely knowable prospectively
- Measuring data quality is critically important but rarely done (so far)
- Human judgment is a key component of the 'computational' process

Precision and Recall of Social Data



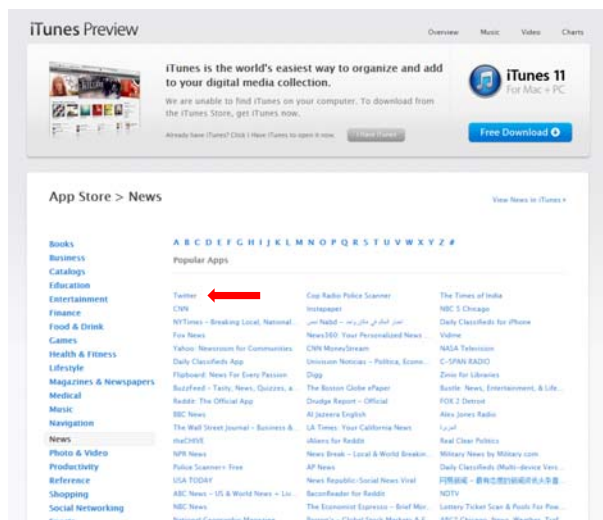
15

Lessons Learned So Far: Rules for Social Media Data

- Analysis is “easy”
 - Data collection and management represent 90% of the work
- Qualitative data at a quantitative scale
 - Much human judgement required
- Social media data rarely stand alone
 - Natural partnership with survey data

16

Social media platforms evolve



17

Examples/Applications

18

ORIGINAL ARTICLE

Are You Scared Yet? Evaluating Fear Appeal Messages in Tweets About the Tips Campaign

Sherry L. Emery¹, Glen Szczypka¹, Eulalia P. Abril¹, Yoonsang Kim¹, & Lisa Vera²

¹ Health Media Collaboratory, Institute for Health Research and Policy, University of Illinois at Chicago, Chicago, IL 60612, USA
² Moores Cancer Center, University of California, San Diego, CA 92093, USA

In March 2012, the Centers for Disease Control and Prevention (CDC) launched "Tips from Former Smokers," a \$54 million national campaign featuring individuals experiencing long-term health consequences of smoking. The campaign approach was based on strong evidence that antitobacco ads portraying fear, graphic images, and personal testimonials are associated with attitudinal and behavior change. Yet it was also controversial, critics cited the danger that viewers might reject such intensely graphic messages. Tinkled with informing this debate, our study analyzes the corpus of Tips campaign-related tweets obtained via the Twitter Firehose. We provide a novel and rigorous method for media campaign evaluation within the framework of the Extended Parallel Process Model. Among the relevant tweets, 87% showed evidence of message acceptance, whereas 7% exhibited message rejection.

doi:10.1111/jcom.12083

Nearly 5 decades after the first U.S. Surgeon General's report on smoking and health (U.S. Department of Health and Human Services, 2014; U.S. Public Health Service Office of the Surgeon General, 1964), an estimated 443,000 Americans will die each year from smoking-related disease. Thus, tobacco use remains the leading preventable cause of death and disease in the United States (Centers for Disease Control and Prevention, 2011) and in the world (World Health Organization, 2011). For this reason, on March 15, 2012, the Centers for Disease Control and Prevention (CDC) launched "Tips from Former Smokers" (Tips), a \$54 million national media campaign that used real-life stories from smokers suffering from the long-term health consequences of smoking, including amputation, tracheotomy, paralysis, and heart surgery (Block, 2013). The combination of graphic images, fear appeals, and personal testimony in the ads was expected to elicit a negative reaction to smoking or to the

Corresponding author: Sherry L. Emery; e-mail: slemery@uic.edu

Journal of Communication 84 (2014) 278–295
 Published 2014. This article is a U.S. Government work and, as such, is in the public domain in the USA.

Using Social Data for Evaluation

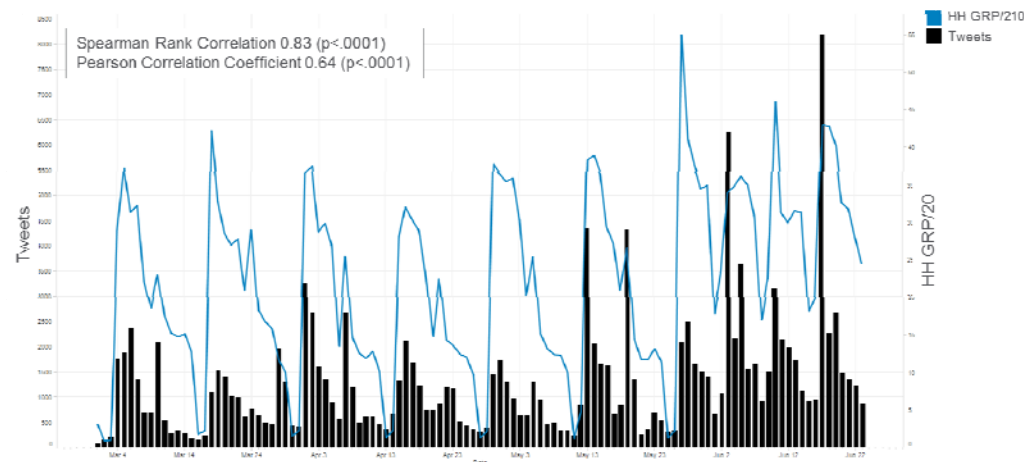
CDC Tips 2012:

Messages both extend reach of ad exposure and inform us about how audience reacted to ads

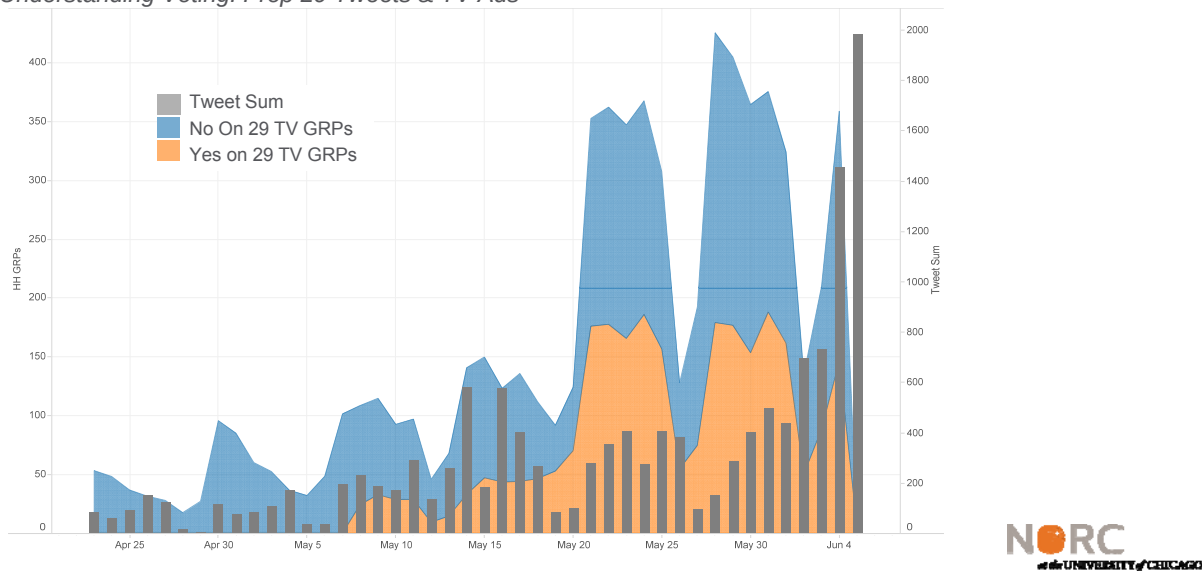
37 keywords retrieved ~ 250,000 potentially relevant messages

- 79% precision; 94% recall
- 90% 'organic' users
- 87% message acceptance

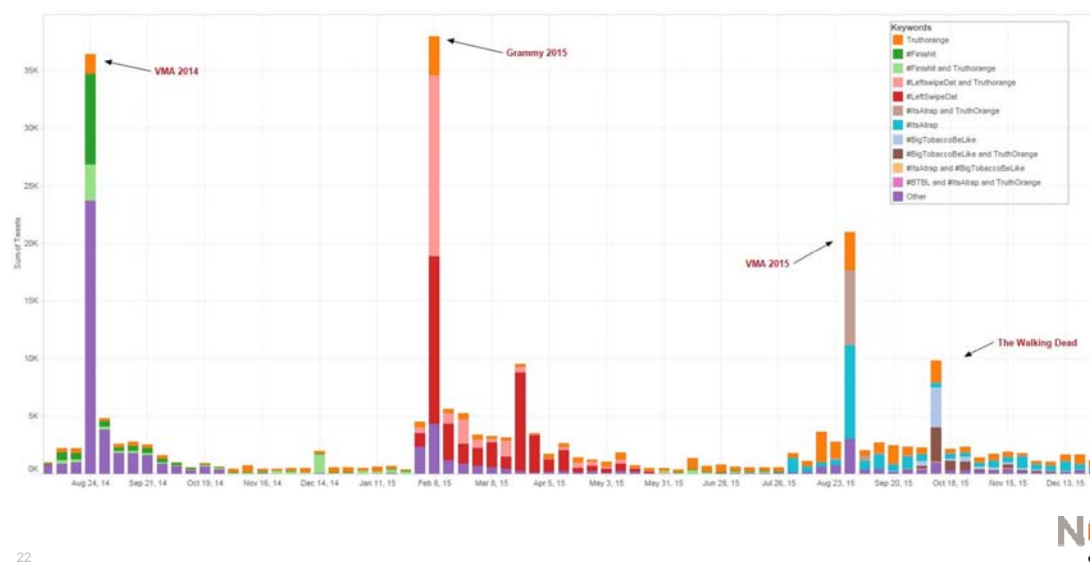
CDC Tips 2012: Relationship Between TV Ad Exposure & Tweets



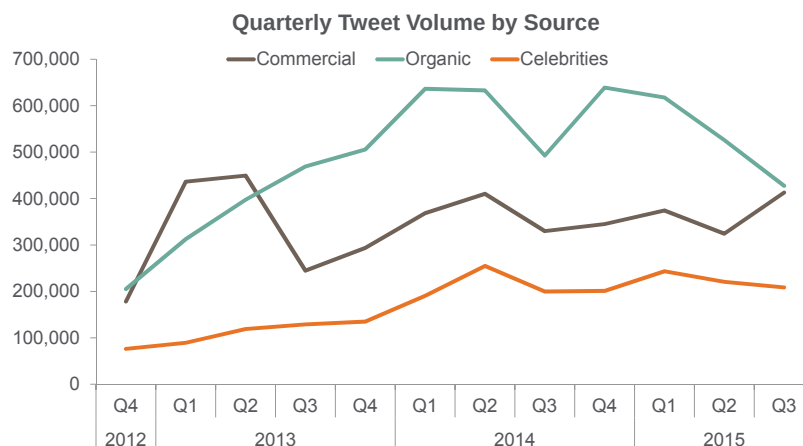
Understanding Voting: Prop 29 Tweets & TV Ads



Truth Tweets Over Time (Aug 2014 – Dec 2015)



3 Years of E-cigs on Twitter



23

Social Media as Discovery Tool



81 likes

splitarillos_usa (@DJSHOCKY) Presents the official @GUTTABLLOS USA poster "PUFF PUFF PASS" Featuring all of your favorite hip hop bangers and hottest artists produced by the high end Folioz Entertainment, Inc. and the producers to stay up to date with our full line of quality tobacco products! #PuffPuffPass #PuffPuffPass



30 likes

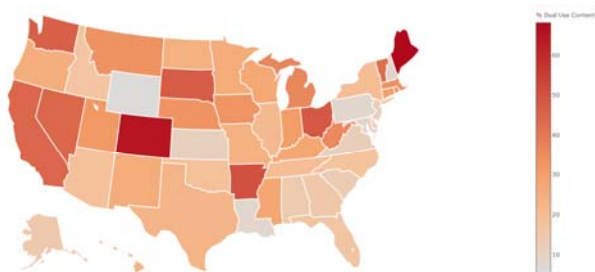
goodtimesusa Brand ambassador contest! We are seeking submissions! @goodtimesusa Chicago #goodtimesusa #goodtimesusa2015

Post a video or image tagging @goodtimesusa showing why you should be a good times girl. #goodtimesusa Let's work together!

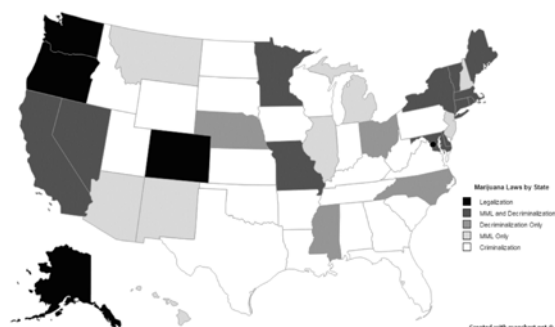
onlyforeshore I missed it. Next go round!!

24

Percentage of LCC Instagram Posts with Marijuana References



Legal Status of Marijuana 2016



Thank You!

