Socialized Research
It's the End of Research as We Know it, but we Feel Fine

By Ipsos OTX, Michael Rodenburgh

We are now citizens of a socially networked mobile world where 'normal' changes faster than the seasons. Socialized Research is the promise on how to change market research forever by leveraging new data collection tools that are relevant in today's age. This paper focuses on one aspect of Socialized Research by examining how Facebook Social Graph data can add to the insights already being collected using traditional online surveys. It includes a case study focused on daytime television programs and discusses implications for future research as well as how to be prepared to integrate this data into other studies.

INTRODUCTION

The New Normal

We are now citizens of a socially networked mobile world where 'normal' changes faster than the seasons. The proliferation of digital technology and the pervasiveness of social media have changed the dynamic. Today's consumers are different. Everyone and everything is connected, allowing consumers to be both producers of content as well as distributors of other content.

The online world has created a new type of democracy where consumers have a voice like they have never had before. In fact, in the 60 seconds that you have spent reading this paper, there have been 175,000 Tweets, 700,000 Facebook messages sent, and 2 million YouTube videos viewed. Indeed, today's social, cultural, and commercial environments have no precedent in history, and they are dramatically changing the very fabric of society and commerce.

As with all businesses, the changes in the new digital society have opened tectonic gaps in our business models, and the market research industry is not alone in grappling with these challenges. However, even though our industry must embrace these challenges and move forward, our social science heritage holds us back and separates us from other industries in how we accept the new realities of the digital age. Instead of discussing how our industry can integrate new data sources into our surveys, we debate the merits of applying statistical margin of error calculations on online surveys.

Our clients are demanding our industry be more relevant in today's digital, socialized business environment. We are calling the next generation of market research, socialized research – a brave new world of immersion, passive data, geo-location, co-creation… the addition of a little "social" into everything we do.
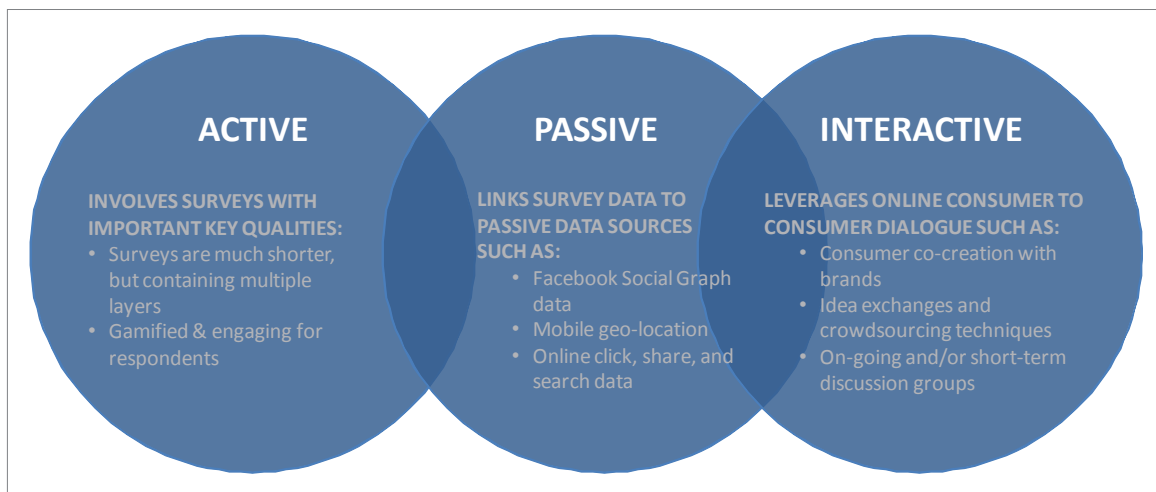
The Dawn of Socialized Research

Socialized Research isn't a methodology, it's a philosophy. It's designed to ensure that how we collect data in the 21st century is relevant to both respondents and clients alike. It requires researchers to retrain themselves and how they approach research problems. No longer can we approach our studies with the sterile gloves of the past, instead we must accept the realities of today's digital world and be prepared to get our hands dirty with our respondents, co-creating in immersive ways, sharing and examining passive data and combining it with our active surveys.

Socialized Research is the answer to many of the problems that are challenging modern marketing research:

- We rely too heavily on <u>survey</u> data, and ignore non-survey data such as what can be obtained from social media, passive web data, and other behavioural data

- Respondents are treated as scientific lab-rats, with survey length causing survey participation rates to spiral out of control

- Incentive structures for respondents are irrelevant and unmotivating, further hampering participation rates

- Surveys often ask the wrong questions and fail to illuminate on key client problems and issues

- Our heritage as social scientists force us to ignore immersive web experiences that are becoming the norm for both brands and consumers

Socialized Research answers these problems by blending active, passive, and interactive data in every study. All is not doom-and-gloom though. We are at the dawn of a new Socialized Research age. One where we can leverage our heritage as social scientists to provide important contextual insight around why people make the decisions they do. As much as our industry covets the latest digital industries and the data they can collect, we must remember that our training on methodological approaches to research can be combined with the new digital solutions for even better insights. And unlike today's digital industries, the marketing research profession is in a unique position to leverage these historical strengths by blending active, passive, and interactive experiences into all of our studies.

FIGURE 1

ACTIVE

INVOLVES SURVEYS WITH
IMPORTANT KEY QUALITIES:
- Surveys are much shorter, but containing multiple layers
- Gamified & engaging for respondents

PASSIVE

LINKS SURVEY DATA TO
PASSIVE DATA SOURCES
SUCH AS:
- Facebook Social Graph data
- Mobile geo-location
- Online click, share, and search data

INTERACTIVE

LEVERAGES ONLINE CONSUMER TO
CONSUMER DIALOGUE SUCH AS:
- Consumer co-creation with brands
- Idea exchanges and crowdsourcing techniques
- On-going and/or short-term discussion groups

Ipsos' experimentation into Socialized Research leverages the latest digital tools available, allowing us to deliver insights in a near real-time environment. Each study conducted with Socialized Research principles must attempt to:

- Include socialized sample – by capturing respondents in their native (social) environments online, we can offer better incentives and attach valuable secondary data to every survey we conduct. This means less reliance on panels and more on social sampling methodologies.

- Leverage socialized survey measures & principles – socialized research emphasizes both interactive and gamified question types, but it also focuses on asking only the absolutely necessary questions, eliminating all the wasted 'nice to have' questions to ensure a short and friendly respondent experience. It also means designing adaptive surveys that use intelligent heuristics to route respondents around questions when sufficient data has been collected on a given question.

- Leverage social principles for reporting – socialized reporting emphasizes the importance of dialogue between vendors and clients. Rarely are decisions made without such dialogue, and socialized research reporting principles ensure that all stakeholders are present in such dialogue as an outcome of a report or study. We believe that socialized reporting should extend the conversation with clients outside of the boardroom and inside a virtual community where all stakeholders can participate (such as ad agency, marketers, researchers, consultants, etc.). And these conversations occur beside the actual data and insights, allowing for instantaneous access should results need to be referenced in community conversations.

One area where Ipsos has been experimenting heavily as part of our focus on Socialized Research is the addition of 3rd party datasets to our standard survey data. Adding 3rd party data to survey research is not a new concept, and it has numerous benefits such as shortening survey length, bridging the opinion vs. behaviour gap, and enabling researchers to get insights that simply are unavailable in any other way.

In today's web-enabled world, respondent-level data is available through any number of sources. We are harnessing the ability to link our survey data to click, search, and share data collected through web properties such as search engines and social media sites. Obviously in the context of social media, our focus has been on the Social Graph and what it can tell us about a respondent. It is our work into understanding the modern Social Graph that is the focus of this paper.

The term "Social Graph" is widely used, although it is often interpreted and misinterpreted as it is repeated throughout business. Well known social media expert Jeremy Owyang defined the term "Social Graph" in 2007 as follows:

The Social Graph is the representation of our relationships. Today, these graphs define our personal, family, or business communities on social websites. [1]

However, since Owyang's 2007 definition, Facebook has expanded the meaning of the term Social Graph by including the concept of connections beyond just people to include brands, products, applications, or even games and hobbies. We believe that it is Facebook's representation of the Social Graph has the most utility for the marketing world since it is allows us to gather a broad view of all aspects of a user's life, not just their social connections between people. As illustrated in Figure 2, there are more connections available than simply being "friends" with someone. The act of "liking" a page or app is also a type of connection. By merging our survey data to these connections, we can start to harness the power of Facebook's social graph and add value by analyzing how aspects of the social graph including "Likes", and posts can provide insights in a way that is not possible with survey data alone. For example, do "likes" correlate with offline behavior as measured through traditional means and if so, what more do these tell us about current and future behaviour? In this way, we can extend the utility of our surveys beyond the standard demographic breaks and be more relevant for businesses today.

FIGURE 2
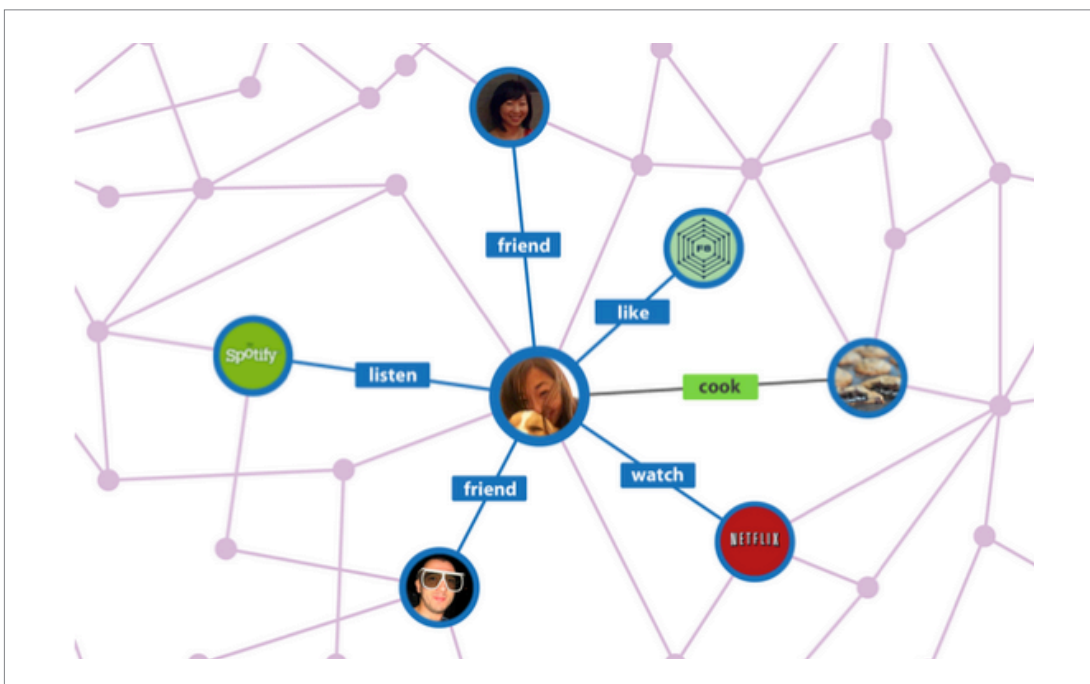
Above: Facebook's Social Graph [2]

The Social Graph is becoming an increasingly important element of the marketing mix for marketers to understand. For many years the standard in marketing was to buy media based on demographic segments. But in today's socialized world, this is an outdated notion and is irrelevant when web properties such as Facebook allow for significant targeting based on interests and observed behaviours. As such, Ipsos has been working on aligning our research processes with today's Social Graph to yield better and more relevant insights for our clients.

Innovating Socialized Research

Over the past 12 months Ipsos has been working on innovating Socialized Research by combining elements from social media into our surveys. Through a series of experiments we have attempted to merge social graph data to our surveys by leveraging Facebook's Open Graph (which refers to the ability to programmatically connect an app to a user's Facebook profile, allowing the Facebook profile data to be shared with the app). By authoring a small Facebook app and leveraging Facebook's "Login with Facebook" button as part of the survey experience, we are able to merge individual level data from Facebook with our survey data. The respondent experience is similar to the screenshots shown below.

FIGURE 3
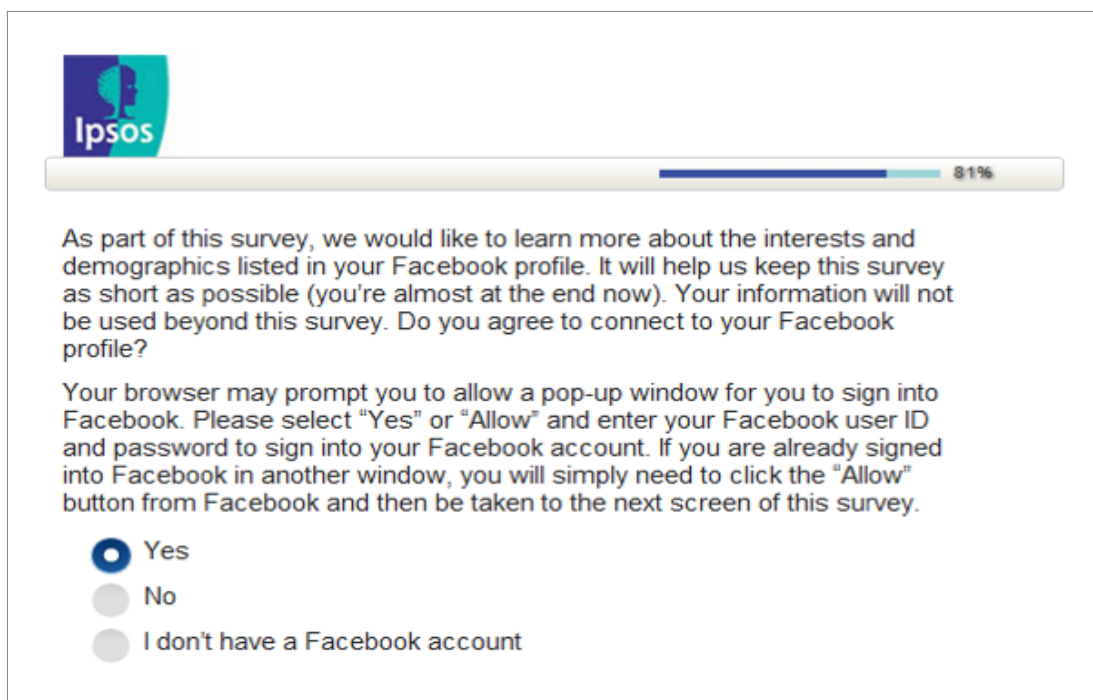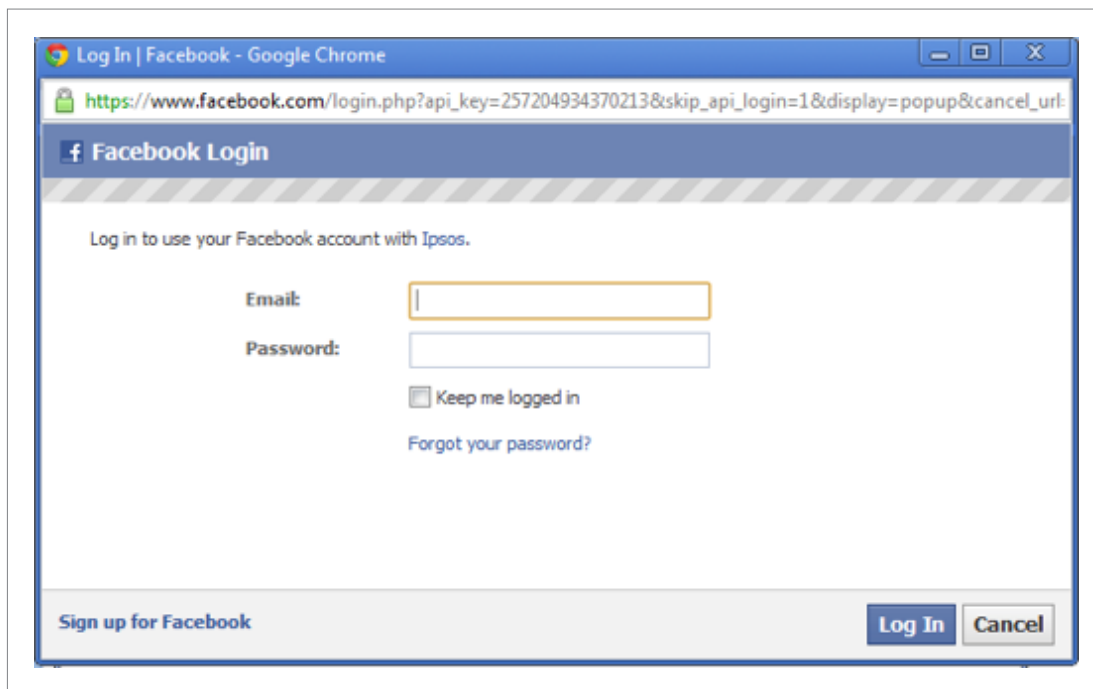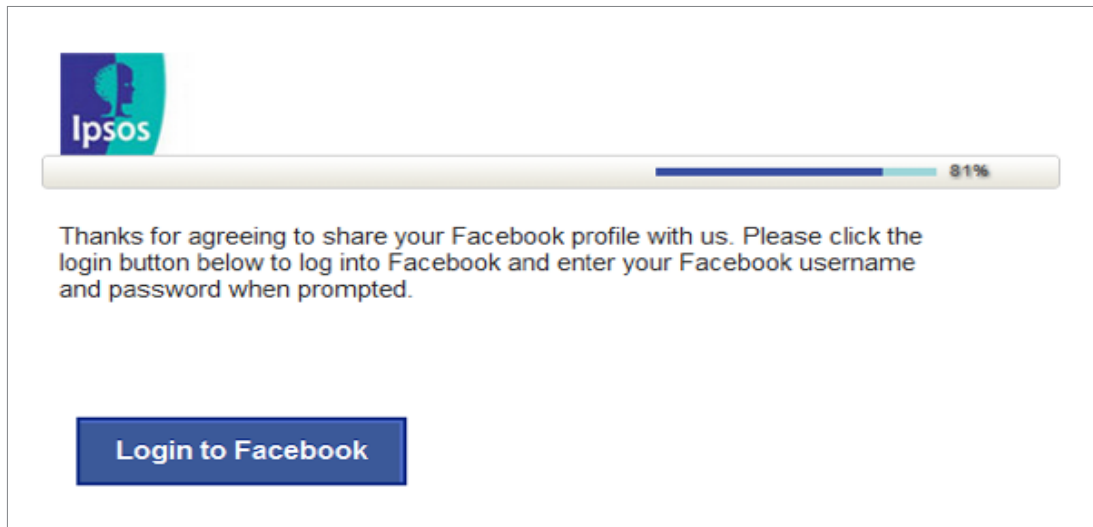
Our exploration of Facebook data and aligning our survey data with the Facebook Social Graph yields great promise, but takes us into new territory under the Socialized Research banner. It allows us to merge multiple data sets together for many benefits. However, merely being able to connect two pieces of data does not automatically mean that we will learn anything worthwhile for our clients. As such we conducted several studies to address two core research objectives:

- Does the addition of social graph data (such as "Like" data, number of friends, etc.) yield useful data that enlightens a research problem or adds a unit of analysis which extends the utility of a particular research study?

- To what extent are those who connect to Facebook different from those who do not connect to Facebook?

Throughout the remainder of this paper, we discuss the findings gathered from our experimentation to address these two questions. The results of which should provide a foundation of learning for the market research profession to build on over the next few years as social graph data and similar datasets become more widely used in the context of survey research.

What we Learned

In discussing our learning, we have addressed our findings in two main themes:

1) <u>Functional Learning</u>:
   Since Connecting to Facebook is a technical task normally built for apps, games, and websites, we quickly learned that normal market research processes, systems and IT infrastructures are not necessarily built in a way that allows us to achieve this task easily. These are important lessons to all researchers who wish to Socialized their Research by adding 3rd party web data such as Facebook.

2) <u>Research Learning</u>:
   Once we achieved a technical proficiency to implement the necessary scripts, merge the datasets and make the data connections, we focused our efforts on what additional research learning can be gained from merging social graph data to our survey data. While we conducted several studies where respondents Connected their Facebook profiles to our survey data, we have used one main case study in this paper to highlight why someone may choose to merge social graph data in their next survey.

Case Study: Daytime TV Tracking Study

Background

Each week, Ipsos MediaCT's TV Dailies survey tracks key marketing metrics for dozens of new and returning primetime programs. This data is often used by clients to fine tune and target their marketing plans if key metrics are not meeting expectations. During the summer months, many networks and studios are also preparing to launch new and returning programs in other day parts, specifically daytime. A majority of these shows are syndicated (air at different times on different stations across the country) and air multiple times per week. These shows often have a more targeted audience and are viewed and engaged with differently than their primetime counterparts. Because of this, TV Dailies is not always the best environment to track these programs. As a result, Ipsos MediaCT created a separate syndicated Daytime TV Tracking study designed to monitor the awareness, interest in viewing, frequency of viewing, and social engagement in daytime TV programs.

The Daytime TV Tracking Study tracks a wide variety of daytime TV programs, resulting in a relatively long survey (averaging roughly 20 minutes in length for most respondents to complete the survey). Given the length, there is limited ability to add more questions to the survey to increase the value for clients who choose to subscribe to this syndicated study. However, in experimenting with our Socialized Research initiatives, we felt that by adding Facebook social graph data to the respondents who complete the survey we could extend the value of the study without making a material impact on the length of the survey.

As part of our experiment, we had several questions regarding social graph data that attempted to add value to the study:

- To what extent do a viewer's Likes tell us more about viewers generally and specifically how to better reach them with messaging?
- Do certain shows attract viewers with varying sizes of social networks?
- Are certain shows better at engaging viewers with larger social networks?

Methodology

Normally Daytime Television studies focus on females aged 18 to 54 years of age. However, because we wanted to examine how gender might change the characteristics of our data, both male and female daytime television viewers aged 18 to 54 years of age were included in the sample. In total, 598 interviews were conducted from a mix of Ipsos iSay, Ampario (Ipsos' proprietary socially recruited sample), and external panel vendors. The study also included standard regional and ethnicity quotas to ensure a representative cross-section of America daytime television viewers was obtained.

The study was split into two design cells to help understand the best placement for the Facebook Connect opt-in message to respondents. Half of the respondents were prompted to Connect to Facebook at the beginning of the survey (after all screening questions were complete), with the remaining half being prompted at the end of the survey.

Execution Challenges

Unlike market research data, social graph data coming from Facebook is immensely complex and not specifically built for marketing research. When constructing the "Connect to Facebook" application, we had to modify our data collection systems to accept and handle JSON objects. JSON (which stands for JavaScript Object Notation) is a common lightweight data-interchange format in the web world. But what makes JSON unique for the market research world is that it is based entirely on text/string data. As such, to make JSON work within the context of most market research IT infrastructure, it requires storage into string variables, parsing tools to eliminate needless text, and recoding the strings into variables that can be easily 'counted' by standard market research software.

Our study utilized IBM SPSS Data Collection as an online data collection platform, and we leveraged custom-built scripts to connect our survey application to the Facebook API (Application Program Interface). For coding and data tabulation, we used Language Logic's Ascribe product to code and categorize the Facebook data and SPSS Data Collection's tabulation solution for the data tables.

It is noteworthy that these software packages are leading-edge research solutions, yet our teams still had challenges in working with the JSON objects from Facebook. For example, we had a challenge related to the character length of a JSON Like object, which averages around 100 characters. Knowing this, we scripted the study to account for Like data with a maximum of 100,000 characters, or roughly 1,000 different Facebook page Likes. Despite our planning and the allowance of strings up to 100,000 characters long, we found 9 truncated cases where the respondent had exceeded the 100,000 character limit. Our teams experienced similar challenges and delays in coding and tabulating the data due to the sheer size of the data file.

As such, one of the unexpected outcomes from our experiment actually was learning the level of effort required to connect Facebook data to survey data and to tabulate it with our existing infrastructure. As the market research industry moves towards adopting more Socialized Research principles in every study, it will be important to for our industry's IT vendors to accommodate these needs with more common digital data structures found online today and in the future.

Research Findings

The analysis of Facebook data was undertaken in two stages. First, we reviewed the demographic profile of the respondents who opted to share their Facebook data and compared them to those who did not opt to share their Facebook data. Next, we closely examined the Facebook Likes and Friend data to identify any patterns that revealed insight into the lives of various daytime TV viewers.

Overall, we had 46% of respondents agree to share their Facebook social graph as part of the survey experience. However, we discovered two variables which have a significant influence on a respondent's likelihood to agree to share their Facebook data: 1) Placement of the Opt-in Message; and 2) Source of sample.

Specifically, those respondents who were prompted at the beginning of the survey are much more likely to share their Facebook data (55%) compared to those who were prompted at the end of the survey (37%). Similarly, those respondents who were sourced using Ipsos' Ampario social sampling methodologies were substantially more likely to agree to share their Facebook data (61%) compared to Ipsos Isay panelists (43%). We expected

Ampario respondents to have a higher opt-in rate to share their Facebook data, since they are being recruited from social sites such as Facebook. Similarly, since our opt-in message suggested that the respondents would have a shorter survey experience if they connected with Facebook, we were similarly not surprised to observe a better rate of connection among those prompted at the beginning of the survey.

TABLE 4: Rate of Connecting to Facebook

Sample Source	Prompted at Beginning of Survey	Prompted at End of Survey	Total
Ampario	70%	53%	61%
Ipsos iSay	55%	30%	43%
External non-Ipsos Panels	49%	34%	42%
Total	55%	37%	46%

Demographic Analysis

The demographic profile analysis showed very few meaningful differences between those who did opt into sharing their Facebook data compared to those who did not share their data. Notably, the distribution of age, gender, income, and marital status did not differ significantly between the two groups of respondents. The only two significant differences observed were in relation to ethnicity and education, with African Americans and college graduates showing a slight propensity to opt out of sharing their Facebook data. Regardless of this slight bias, the magnitude of differences is reasonably small and could be addressed with commonly used weighting procedures.

TABLE 5: Demographic Analysis of Facebook Opt-in

		Opted into Facebook	Didn't Opt into Facebook[3]
Gender	Male	33%	33%
	Female	67%	67%
Age	18 to 24	18%	20%
	25 to 34	22%	22%
	35 to 44	27%	27%
	45 to 54	34%	31%
Ethnicity	Caucasian (not Hispanic)	68%	60%
	African American	19%	27%
	Asian/Pacific Islander	5%	4%
	Hispanic	7%	9%
	Other	1%	1%
Marital Status	Single, never married	35%	38%
	Living with partner	9%	12%
	Married	42%	36%
	Divorced/Separated	12%	13%
Income	Less than $25,00	31%	27%
	$25,000 to $49,999	32%	27%
	$50,000 to $74,99	21%	25%
	$75,000 to $99,999	10%	11%
	$100,000+	6%	13%
Education	Some high school	5%	3%
	High school graduate	29%	28%
	Associates/junior college	20%	14%
	Some college	27%	24%
	College graduate or higher	20%	31%

Prior to conducting the research, it was hypothesized that tech-savvy respondents might be more willing to share their Facebook data than those who are less tech-savvy. To that end, tech-savviness was evaluated by a common device ownership question. Contrary to what was expected, those who did not share their Facebook data showed the same levels of technology device ownership as those who did share their Facebook data.

TABLE 6: Technological Device Ownership of Facebook Opt-in

Devices Owned in Home	Opted into Facebook	Didn't Opt into Facebook
Electronic Reader (like Kindle, Nook or e-Reader)	23%	20%
DVD Player	76%	75%
3D TV	7%	7%
DVR (Digital Video Recorder)	46%	40%
HDTV (High Definition TV)	54%	58%
Basic mobile phone	51%	44%
Internet TV device (like Apple TV, Google TV, Roku)	11%	9%
Internet-enabled / smart TV	11%	14%
Smartphone (like Blackberry, HTC or iPhone)	50%	51%
MP3 player without video capability	38%	32%
MP3 player with video capability	34%	29%
Tablet (like iPad, Galaxy or PlayBook)	26%	22%
Handheld gaming device (like DSi, 3DS or PSP)	29%	23%
Gaming console (like PlayStation, Wii or Xbox)	58%	53%
Satellite radio (like SiriusXM)	11%	12%
VOIP (making voice calls via the internet)	8%	7%
Wireless internet connection	61%	56%
High speed / broadband internet connection	64%	60%
Windows PC laptop or desktop	74%	73%
Apple Mac laptop or desktop	15%	18%
Blu Ray Player	24%	28%

The conclusion therefore after reviewing the demographic profiles and tech device ownership rates is that those who opt into sharing their Facebook data are not meaningfully different than those who do not, and such data can be treated with confidence that significant biases exist among those who choose to share such data. Obviously as the use of this data increases among the marketing research industry, these biases should be monitored for stability, as it is still rare that companies are using Facebook as a tool purely for data.

Analysis of Friend Data

One of the reasons social media has been the focus of attention among marketers is their insatiable desire to understand how certain consumers spread content and personal recommendations throughout their social networks. At the center of this argument is the discourse around the value of a "Like". While there are many perspectives on the value of Facebook Liking behaviour, one way to measure the value is by examining Likes in terms of the number of impressions. Specifically, when a viewer Likes a Facebook page, all of their friends are potentially exposed to that Like in their Facebook newsfeed, which digital marketers track as reach, similar to how it is tracked by mainstream media.

Facebook page owners have the ability to track reach data for all content delivered to their fans from their page. Specifically, Facebook page owners have the ability to see the raw number of people who viewed a post or Like, similar to the example shown below. In the example, it shows that 422 people saw the post originating from the Digital Divas Facebook page. As one might expect, digital marketers are now using this information to optimize content for future posts & Likes.

FIGURE 7

As useful as this data is, it has limitations. For example, the aggregate reach statistics provided by Facebook are the only figures you get – you can't connect those reach statistics to survey data. So you aren't able to understand whether you're reaching brand loyalists or consumers who have never engaged with your brand before. But more importantly the data is only available to Facebook page administrators. This is a significant limitation if your objective is to understand the performance of your competitors, since you will obviously not be an administrator of your competitor's pages!

We decided to investigate whether these limitations could be addressed by connecting our survey data to the Facebook data. Specifically, we decided to attempt to calculate the theoretical reach of each daytime TV program, since it would likely have value to those in the industry. In calculating the reach figures, we first gathered the number of Liker from each TV show's Facebook fan page. Next, we used our survey data to quantify the number of friends each respondent had in our study to understand the average 'reach' for a Like across all Facebook newsfeeds. Next, we calculated the total number of theoretical Facebook wall feed posts by multiplying the total number of people who have Liked each show by the average number of friends among those who Liked the show. Finally, we weighting the total reach for each show by a factor of 17%, which represents the average performance of Facebook pages. [4] Table 6 illustrates the findings from this analysis.

One might expect that the number of Facebook page Likes would drive the overall number of impressions, and as such shows with a very large number of Facebook fans would also drive the largest reach. However, this is not always the case since some shows draw audiences who have more Facebook friends than other shows, causing the total reach to change significantly. Two shows which illustrate this relationship are the Dr. Oz show and Rachael Ray, both of which have benefited from roughly the same number of total Like impressions (87 million and 79 million respectively). However, the path to obtaining that reach differs. In the case of Dr. Oz, the show's reach is being driven by the total number of page Likes (2.3 million), whereas Rachael Ray's reach is being driven by the large number of friends that each viewer has (552 on average).

Table 8: Average Number of Facebook Friends (Among Those who Liked Each Show)

	# of Facebook Likes	Avg # of Friends Among those who have Liked the Show	Total Theoretical Reach
Ellen	8,059,371	260	356,772,235
Maury	1,662,850	393	111,208,082
Dr. Phil	1,258,692	482	103,158,620
Rachael Ray	936,527	552	87,947,378
Dr. Oz	2,328,067	200	79,312,587
TMZ	958,311	373	60,799,083
Jerry Springer	1,011,330	346	59,555,201
Good Morning America	846,166	353	50,735,267
The Wendy Williams Show	752,054	388	49,643,837
General Hospital	824,497	347	48,637,078
The Doctors	1,483,737	193	48,555,293
The Steve Wilkos Show	369,634	569	35,742,129
Family Feud	1,011,535	187	32,173,894
Judge Mathis	328,376	464	25,874,387
The Young and the Restless	779,255	154	20,374,401
Days of Our Lives	557,513	211	19,997,991
The Jeremy Kyle Show	661,086	135	15,171,924
Judge Judy	439,813	184	13,772,304
Today with Kathie Lee & Hoda	465,518	167	13,184,401
Judge Joe Brown	180,755	397	12,208,373
Divorce Court	170,461	364	10,548,127
The Price Is Right	238,459	227	9,194,025
The Chew	280,920	183	8,720,319
The View	335,039	144	8,224,537
Who Wants To Be A Millionaire	151,987	270	6,978,787
Anderson	155,591	248	6,546,491
Live with Kelly	265,590	139	6,289,437
The Talk	218,115	165	5,999,471
The Bold and the Beautiful	242,531	141	5,813,468
The Bill Cunningham Show	19,621	346	1,154,107
Let's Make A Deal	28,270	167	801,144
Judge Alex	28,475	108	522,801
People's Court	3,346	109	62,001

We obviously acknowledge that many factors will impact whether a Like appears in a person's newsfeed because it is subjected to Facebook's EdgeRank algorithm, so to a certain extent the figures may be somewhat misleading. But this is an important piece of competitive data that would otherwise not be available to digital & social media marketers.

The obvious conclusion from this analysis is that that one way to maximize your reach in Facebook is by engaging Facebook users with larger social networks. This advice would clearly be useful to the producers of Family Feud. Even though the Facebook page has over 1 million fans, the low average number of friends among Family Feud viewers limits the total reach through Facebook. At this point we don't have clear recommendations on how to accomplish this, but we strongly believe that there are gender & age skews (as well as other variables) that play into the variations by show. As such, that is part of the challenge that faces researchers in the future. The capture and integration of the social graph data is the first step in this process.

Analysis of Likes

Once we established that the Facebook opt-in process did not yield itself to a certain type of respondent, we pursued an analysis of the respondent's Likes. First, it is important to discuss the complexity of Facebook's Likes. In our survey of just 274 respondents who shared their Likes, we observed over 21,000 different page Likes – a tremendously rich and varied set of objects and user connections. The average respondent had 149 different Likes, again a number which was surprising and provides insight into just how deeply consumers are connecting with commercial interests (including music, TV shows, celebrities and brands) in this social media environment.

It is also important to note that even with the large number of Likes collected from Facebook, the proportion of respondents who had Liked any one given Facebook page was relatively low. The most frequently Liked Facebook page from the 274 respondents who opted to share their data was Swagbucks, with only 9% of the sample. Swagbucks is an extremely popular online rewards program in the United States, and was recently ranked by Alexa Internet as the 112th most visited website in the USA. [5] As such, it is not surprising the high rate of Liking this web property. The remainder of the top 10 Liked pages are among some of the largest brands in the country, which is again not surprising. However, there are some surprisingly small brands in the top 10, such as Wholly Guacamole, shopathome.com, and Ghirardelli Chocolate, suggesting that social media success is still attainable for smaller brands.

TABLE 9: Top Facebook Likes (Among All Respondents - Top 10 Mentions)

swagbucks	9%
target	8%
samsung mobile usa	8%
amazon.com	6%
tide	5%
walmart	5%
wholly guacamole	5%
eucerin us	5%
shopathome.com	5%
ghirardelli chocolate company	4%

Further analysis of the Likes revealed some interesting (and sometimes not surprising) findings when controlling for viewers of various daytime TV programs. In analyzing the patterns of Likes among regular and occasional viewers of each daytime TV show, an Affinity score was calculated to assist in identifying the extent to which viewers of a particular score were more likely to Like a particular page relative to all respondents who shared their Facebook data. This Affinity score is important because it helps digital marketers to create better targeting strategies to reach their core intended audience. Affinity scores of greater than 200% were used to identify patterns noteworthy in the data, and resulted in a number of interesting findings.

First, viewers of daytime TV programs which skew towards African American audiences are much more likely to have Liked a variety of classic African American icons. For example, regular and occasional viewers of the Wendy Williams show (a daytime talk show featuring an African American host) over-indexed on Liking the following Facebook pages of African American icons:

- Maya Angelou
- Fantasia
- Mary J. Blige
- Beyonce
- Fresh Prince
- Jamie Foxx
- 2pac
- Jada Pinkett Smith
- Oprah
- Taraji P. Henson
- Barack Obama

It is also noteworthy that the same pattern occurs for other shows that skew towards African American audiences, such as various courtroom drama shows such as Divorce Court, Judge Alex, etc.

Viewers of daytime dramas, which skew towards older Caucasian audiences, tend to have Facebook Likes that are consistent with that demographic. For example, viewers of Days of our Lives over-indexed on Liking pages that likely skew older female and Caucasian such as:

• Adam Sandler
• Official Grease Movie
• Dirty Dancing
• MircaleGro
• Dallas
• Nancy Grace
• Desperate Housewives
• Bon Jovi
• Johnny Cash

Not surprisingly other daytime dramas such as The Young and the Restless and General Hospital contained similar patterns of Liking pages that likely skew Caucasian. Interestingly, General Hospital viewers are also much more likely to have Liked the Facebook page of Rick Springfield, who was on the show during the 1980's and has recently made a come-back to the show and thrilled long-time audiences with his return.

Other show audiences drew consistent patterns that are more difficult to explain, but yield interesting insight into the audience even without a solid explanation for the pattern. For example, regular and occasional viewers of the Bill Cunningham show (a daytime talk show) show a particular affinity to a wide range of musicians/bands from a wide variety of genres and eras:

• Coldplay
• 2pac
• AC/DC
• Metalica
• Mary J. Blige
• Justin Timberlake
• Kiss
• ZZ Top
• Sade
• Alice Cooper
• Keyshia Cole
• Fleetwood Mac
• Black Sabbath

While the bands/artists listed as over-indexed among viewers of The Bill Cunningham Show range in genres, it is notable that most of the bands listed were popular in the 1970's and 1980's. It is therefore not surprising to observe that The Bill Cunningham Show performed very well in an older age demographic and won its timeslot among adults 25 to 54 years of age. [6]

Still other shows appear to draw audiences which show an affinity to Liking many different Facebook pages. For example, regular & occasional viewers of Anderson (a popular talk show) over-index on Liking 71 different Facebook pages. Reviewing the Likes for Anderson yields an extremely interesting perspective into the lives of Anderson viewers, and could potentially be used in decision making related to program content that will be of interest to its core audience. For example, there are a number of TV personalities listed who could be approached as potential guests, such as George Takei (popularly known as Sulu on Star Trek), Dog the Bounty hunter (a reality TV show star), Dolly Parton (a popular country music artist), just to name a few.

Similarly, Anderson viewers show an affinity to a number of popular consumer products that might provide some additional insight into their lives, such as:

- **They may be more likely to have pets:** there are three pet food companies listed – Wellness Natural, Beneful, and Hill's Pet Nutrition;
- **They care about their appearance:** there are a number of health/beauty brands listed – Alberto VO5, Degree, Secret, Clariol, and Burt's Bees;
- **They like baked goods:** there are three brands of well-known baked goods – Little Debbie's, Nonni's Biscotti, and Entenmann's.

It is not difficult to see how a media client such as the producers of Anderson could use this insight in identifying potential advertisers, sponsors, or potential show topics for future episodes. The full list of over-indexed Likes among Anderson regular and occasional viewers is in the table below.

TABLE 10: Over-Indexed Likes Among Regular/Occasional Viewers of Anderson

dog the bounty hunter on a&e	excedrin
express	secret
pinterest	butterfinger
george takei	all free clear
dolly parton	savealot food stores
inception	diet coke
maya angelou	fringe
fantasia	birds eye vegetables
lysol	clairol
lenovo	mary j. blige
cheezit	caesars casino
drugstore.com	tropicana
womanfreebies.com	usher
2pac	bingo blitz
fleetwood mac	slots farm
hill`s pet nutrition	sade
alberto vo5	fear and loathing in las vegas
white cloud	bingo bash
breyers	nonni`s biscotti
wellness natural pet food	amazon student
kirkland`s	doug
love and basketball	go red for women american heart association
entenmann`s	little debbie
abc7	stevie wonder
beneful	dentyne
maroon 5	tampax
hsn	jersey shore
law and order: special victims unit	seinfeld
gain	swiffer
modern family	mio
o`reilly auto parts	stouffer's
dave ramsey	shout
extra gum	burt`s bees
fox news	seattle`s best coffee
degree women	office
cost plus world market	

Inside Facebook's data structure, each "Like" is categorized into one of roughly 160 different categories (ranging from Musicians, TV Shows, Public Figure, Consumer Merchandise, etc.). And even these categories yield interesting patterns that inform the research and help the researcher understand a daytime show's viewers. For example, regular and occasional viewers of The Chew (a daytime cooking show) show much more affinity to Facebook pages that are categorized as "Food". Interestingly, they also have an affinity to pages categorized as Furniture, Home/Garden, and Museum/Art Galleries. This does not seem surprising considering the audience for The Chew draws a viewer who would be interested in these related interests. Other shows appeared to have viewers who show an affinity for Liking specific types of pages, as outlined in the table below.

TABLE 11: Top Liked Categories by Show

Today with Kathie Lee & Hoda	The Young & the Restless	The Bill Cunningham Show
• food	• furniture	• musical instrument
• amateur sports team	• appliances	• insurance company
• local/travel	• pet services	• studio
• arts/humanities		• tools/equipment
		• energy/utility
		• bank/financial institution
		• expertise

While it could be argued that this analysis of Facebook Like data contains largely intuitive findings, we believe that the findings are meaningful because they also contain insights that are not necessarily intuitive. For example, we might not have expected to see a relationship between viewing Anderson and Liking the brand Excedrin or the rap artist 2pac. However, these non-intuitive insights create a picture of viewers that is much more robust than we would have without linking the Social Graph data to our survey. And as mentioned earlier, this additional data could be easily be used to extend the utility of the research, for example by using it for identifying new advertisers/sponsors, identifying new program content, or new online engagement tactics. And this comes with a significant benefit, since adding the Facebook data to the survey experience has no material impact on the survey length. As a result, as the advertising industry debates the value of a Like, the marketing research community can clearly see value as a data point rather than a digital ad targeting strategy.

Analysis of Social Impact & Engagement

It is becoming commonplace among most major television programs to use social media to engage viewers outside of the core program. As such, it is important to understand the relevant relationships between this engagement activity and a more traditional performance metric in the media business such as reach. To explore these relationships, we explored the impact that audience size has on the likelihood to engage socially with the show. For each daytime TV show included in our study, we plotted reach against the net proportion of the audience who has engaged in some social activity with the show. [7]

As one would expect, there is a positive correlation between a show's reach and the size of audience who engage with the show in some social activity, with larger shows drawing more social activity than smaller shows. As an example, Figure 5 shows that 27% of daytime TV viewers are regular or occasional viewers of Judge Joe Brown and 13% of daytime TV viewers have engaged in some social activity related to that show. In contrast, Ellen draws a larger reach comprising 40% of all daytime TV viewers but also draws a larger share (31%) of social activity as well.

However, the relationships between reach and social engagement do not end there. One unique and unexpected finding from this analysis is that there appears to be a "double jeopardy" effect at play in the data. Double jeopardy is a statistical phenomenon identified by Andrew Ehrenberg's which finds that with all other things being equal, higher share brands in a category have proportionally higher brand loyalty than low share brands. Interestingly, the data from this study shows a similar pattern, with shows with larger reach obtaining proportionally greater social engagement than that of shows with smaller reach, as exemplified by the curve found in Figure 10.

The implications for content owners and producers is clear - larger shows are able to leverage social engagement tactics with less effort compared to shows with a smaller audience. Since social media engagement provides the show's producers with another means of reaching their audience and engaging them in immersive interactions, this is a meaningful step in retaining existing viewers and using the Facebook friend network to each potentially new viewers.

This does not suggest that larger shows don't have to have an effective social media presence. If the Facebook pages don't provide meaningful content, the viewers will disengage. Nor does is say that smaller shows don't get a benefit out of engaging their audiences with effective social activity, but it does suggest that they won't get the same benefit as shows with a larger audience. Moreover, it may be more effective to use traditional means of growing an audience until it reaches a critical threshold where the use of social media would become more cost effective.

FIGURE 12: Relationship Between Reach & Net Social Activity
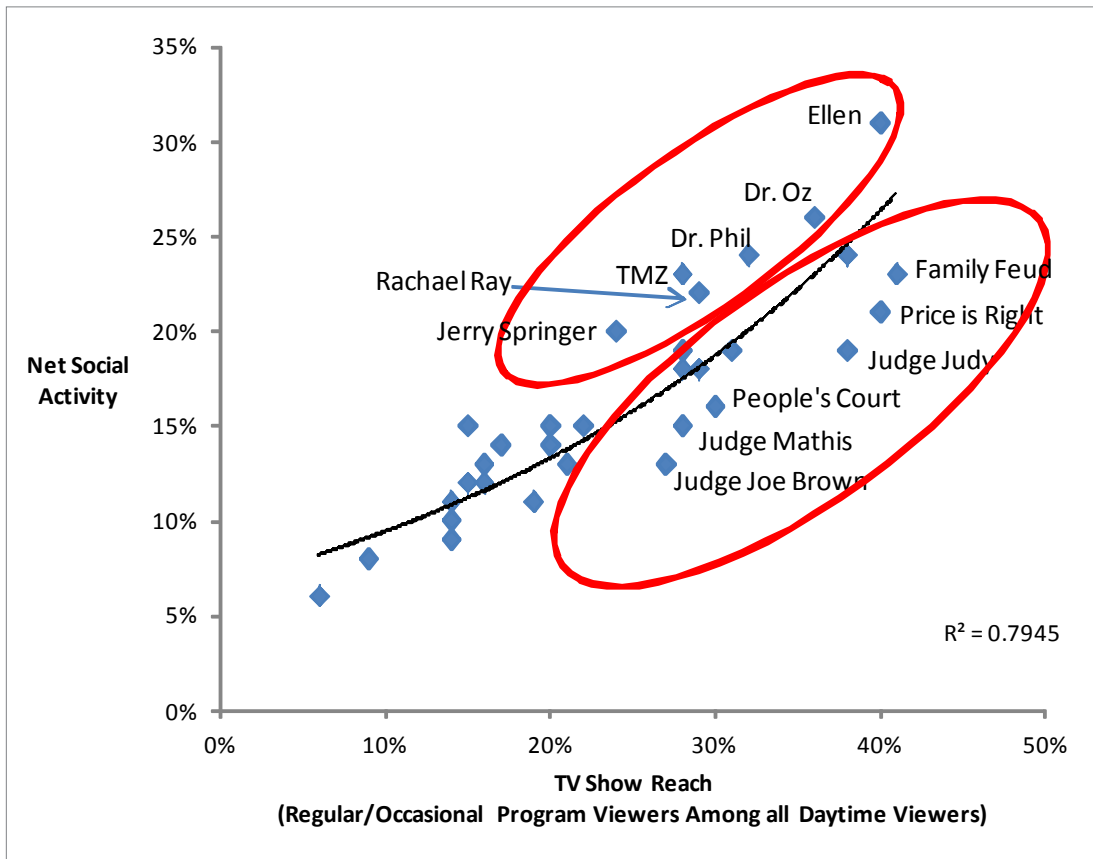
There is a third interesting relationship highlighted by the data in Figure 10 which shows that the emotional connection that a host has with his/her audience may dictate the extent to which the show can engage its audience in social activities. For example, shows which perform better than average in engaging their audience in social activities all have hosts with a strong emotional connection to their audience (such as Ellen, Dr. Oz, Dr. Phil, Rachael Ray, and Jerry Springer). Conversely, shows which perform worse than average all have hosts with lower emotional connections to their audiences, characterized by the various court shows or game shows. The one exception is TMZ, although one could argue that the content of TMZ (which is a tabloid format show focused on the lives of various Hollywood stars/personalities) is a reasonable surrogate for the host and thereby retains the same emotional connection with its audience.

This third finding is meaningful for content owners and producers because it provides them knowledge on how to improve the effectiveness of social media activities. Specifically, shows which under-perform in engaging their audience in social activities would benefit by first identifying ways in which they could improve the emotional connection the hosts have with their audiences. Presumably, this improvement in emotional engagement would likely yield dividends in social media receptivity and help improve the ROI for a show's social media activities.

Conclusions

As our experimentation has proven, it is not easy to apply Socialized Research principles to traditional studies. It requires thinking outside of the box, integrating non-traditional complex data structures, and applying operational processes that are not standard in the market research industry. Simply by focusing only on Facebook's social graph data, it can be seen that a wealth of information can be added to any given study that yields insight into the brand's social media followers and fans. While this is only a piece of the Socialized Research principles, it highlights the level of effort required to change the standard processes in our industry.

A consumer's social media presence, while not a perfect reflection of all aspects of their life, does identify meaningful patterns that can help brands build effective digital marketing strategies. It allows brands to move beyond the conventional audience measurement and media buying practices which focus solely on demographics, and moves the media buying practices closer to more meaningful and relevant behaviours and interests. And most importantly, adding this social graph data highlights patterns of data that would simply never be considered to be part of a traditional survey, leading to more unexpected findings by answering questions a brand would never ask. The 'double jeopardy' effect of show ratings and social media audience size is just one example of the findings that were unlikely, if not impossible, to be detected through standard survey approaches.

Market researchers need to pay attention to the level of effort required to integrate today's non-standard data structures into our survey data. Traditional research technologies cannot be utilized to efficiently combine these multiple data streams, and they will require specialized IT investments to achieve efficiency. It is also important to point out that Facebook and many of today's other social media sites have data structures that continually evolve. For example, Facebook is starting to push out updates to its site on a weekly basis, and changes in privacy rules can often change quickly due to consumer pressure or government regulation. The researchers of tomorrow will need to stretch their skill sets beyond the traditional research boundaries and understand these pressures so that they can quickly respond to changes in their environment.

Notwithstanding the challenges, however, the market research profession will be able to benefit from integrating social graph data into more datasets in the future. For example, we are able to leverage the demographic present in a respondent's Facebook profile to eliminate standardized profiling/demographic questions. Similarly, as marketers strive to understand how to reach and engage consumers in Facebook, we have the opportunity to understand a respondent's Likes more accurately through Facebook's API than relying on their ability to recall which brands (among their average of 149 Likes) they have liked in Facebook. Moreover, the Facebook data can be used as a quality check during the survey recruitment process, since the known panel information can be automatically compared to the respondent's Facebook profile data.

Finally, there are opportunities to use Facebook's social graph data to bridge the opinion vs. actual behaviour gap in typical market research surveys. For example, as consumers become more accustomed to 'checking-in' using social media sites, this data could become useful for those engaged in shopper research by linking a 'check-in' from Facebook to survey questions posed to respondents. Similarly, the ability to track which ads consumers have been exposed to in their Facebook newsfeed provides an opportunity for advertising research. This technology exists today; it just requires researchers to seize the opportunity by blending it together with traditional forms of data collection.

We believe that Facebook social graph data is merely the tip of the iceberg of 3rd party datasets. The way forward must also include other forms of passive measurement such as mobile geo-location data, passive media exposure, and online search and sharing behavior in other web properties. But even though the future of Socialized Research will inevitably include passive data collected from various digital properties, it is important to remember that there will always be an opportunity to merge what we observe with traditional survey data. Because like it or not, there still remains value in understanding what consumers think. And unlike various digital industries, the market research profession is in a unique position to be able to combine multiple datasets together to provide the most comprehensive outlook on consumer attitudes, opinions, and behaviours. We at Ipsos are preparing to seize the opportunity, and call on the rest of the industry to do the same.

Appendix: Socialized Research: A New Approach For A New Reality

Authors:
Ipsos OTX
Michael Rodenburgh is VP Research, Ipsos Open Thinking Exchange (OTX), Global

Footnotes:

[1] Source: http://www.web-strategist.com/blog/2007/11/10/what-is-social-graph-executives/

[2] Source: https://developers.facebook.com/docs/opengraph/

[3] Note: These respondents include those who do not have a Facebook account as well as those who chose not to use the Facebook Connect option in the survey.

[4] Source: All Facebook, the Unofficial Facebook Blog (http://allfacebook.com/facebook-page-17_b73948)

[5] Source: Wikipedia (http://en.wikipedia.org/wiki/Swagbucks)

[6] Source: Wikipedia (http://en.wikipedia.org/wiki/The_Bill_Cunningham_Show)

[7] For the purposes of this study, social activity includes all online social media activity, all online web activity related to the show, and any recommendation to others about the show.