Estimate Benefits of Crowdsourced Data from Social Media

Task 4 Project Report

www.its.dot.gov/index.htm
Final Report — February 2015
Publication Number: FHWA-JPO-14-165

U.S. Department of Transportation

Produced by Noblis, Inc.
U.S. Department of Transportation
ITS Joint Program Office

Notice

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-JPO-14-165		

4. Title and Subtitle	5. Report Date
Estimate Benefits of Crowdsourced Data from Social Media	December 18, 2014
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Jeffrey Adler, John Horner, Jeanette Dyer, Alan Toppen, Lisa Burgess, and Greg Hatcher	

9. Performing Organization Name And Address	10. Work Unit No. (TRAIS)
Noblis 600 Maryland Ave., SW, Suite 755 Washington, DC 20024	**11. Contract or Grant No.** DTFH61-11-D-00018

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
ITS-Joint Program Office 1200 New Jersey Avenue, S.E. Washington, DC 20590	Final Report
	14. Sponsoring Agency Code HOIT-1

15. Supplementary Notes

Jimmy Chu, FHWA, Government Task Manager

Marcia Pincus, ITS JPO, COR

16. Abstract

Traffic Management Centers (TMCs) acquire, process, and integrate data in a variety of ways to support real-time operations. Crowdsourcing has been identified as one of the top trends and technologies that traffic management agencies can adapt and take advantage of from outside the transportation community to improve data quality, coverage, and cost. Several traffic management agencies have active projects to integrate crowdsourced data into TMC operations. The purpose of this project is to identify, research, and synthesize the opportunity for agencies to work with crowdsourced data from available applications, and to develop measures that can articulate the value proposition to work with such resources and consider their impacts on TMC operations.

17. Key Words	18. Distribution Statement
Traffic Management, Social Media, Crowdsourced Data, SWOT Analysis	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	106	

Preface/Acknowledgements

The project team would like to thank Jimmy Chu, FHWA Office of Transportation Management, Government Task Manager, for his oversight and technical guidance during this project. We would like to thank James Pol and Michael Nesbitt of FHWA for their thoughtful comments on previous drafts of this final report. We would also like to acknowledge the following Transportation Management Centers Pooled Fund Study (TMC PFS) members for their contributions, support, and technical guidance during this project.

- Paul Arends, Michigan DOT
- Elizabeth Birriel, Florida DOT
- Vinh Dang, Washington DOT
- Tony Ernest, Idaho DOT
- Brian Hoeft, Regional Transportation Commission Southern Nevada/Fast
- Leslie McCoy, Pennsylvania DOT
- Lisa Miller, Utah DOT
- Jason Morgenroth, New Jersey DOT
- Cherice Ogg, Iowa DOT
- Sinclair Stolle, Iowa DOT
- Kelly Wells, North Carolina DOT
- John Hall, Tennessee DOT
- Victor Koo, City of Pasadena
- Vince Garcia, Wyoming DOT
- Silvana Croope, Delaware DOT
- Faisal Saleem, Maricopa DOT

Table of Contents

Preface/Acknowledgements i
Executive Summary 1
1 Introduction 4
1.1 Project Background 4
1.2 Terminology 6
1.3 Research Focus 7
1.4 Document Overview 8
2 Literature Review 9
2.1 Sources of Transportation-Related Crowdsourced Data 9
2.1.1 Active Sources 9
2.1.2 Passive Sources 11
2.2 How Crowdsourced Data is Obtained 12
2.2.1 In-House Applications 12
2.2.2 Private Sector Data 13
2.2.3 Mining Social Media 13
2.3 State of Practice at TMCs 14
2.3.1 Usage of Social Media 14
2.4 Performance Measurement 15
2.5 Crowdsourced Data Usage in Other State or Federal Agencies 16
2.6 Implementation Considerations 17
2.6.1 Gamification / Incentivizing the Crowd 17
2.6.2 Driver/Traveler Distraction 19
2.6.3 Data Fusion and Reliability 19
2.6.4 Data Privacy and Ownership 20
3 Assessing Current Practices and Perceptions 21
3.1 On-Line Survey 21
3.1.1 Survey Results 22
3.1.2 Key Findings from the survey 26
3.2 Phone Interviews with TMC Staff 27
3.2.1 Findings 27
3.3 Industry Expert Phone Interviews 29
3.4 Panel Meeting 31
3.4.1 Public Sector Round Table Discussion 31
3.4.2 Presentations by Industry Experts 32
3.4.3 Private Sector Crowdsourced Data Provider Discussion 33
3.5 Summary 33
4 Case Studies 35
4.1 Overview 35
4.2 Iowa DOT Social Media Program 35
4.3 Utah DOT's Citizens Reporting Program 40
4.4 Mining Social Media at DDOT 43
4.5 Florida DOT – Integration of Waze Data 45

4.6 SUMMARY AND KEY FINDINGS 46
5 SWOT Model 47
5.1 SWOT ANALYSIS 47
5.2 INTERNAL CAPABILITIES (STRENGTHS AND WEAKNESSES) 48
5.2.1 Operational Considerations 48
5.2.2 Institutional Considerations 50
5.2.3 Technical Considerations 51
5.3 OPPORTUNITIES 53
5.3.1 Achieving TSM&O Objectives 54
5.3.2 Overcome Data Deficiencies 55
5.3.3 Improve Institutional Cooperation 56
5.3.4 Social Media Monitoring Tools 56
5.3.5 Humanizing DOTs and Engaging Citizens 57
5.3.6 Incentivizing Citizens / Gamification 57
5.4 THREATS 58
5.4.1 Data Quality 59
5.4.2 Data Validity and Credibility 59
5.4.3 Data Ownership 60
5.4.4 Data Privacy 60
5.4.5 Distracted Driving 62
5.4.6 Emerging Innovations 62
5.5 IMPLEMENTATION CONSIDERATIONS 63
5.6 SUMMARY 64
6 Measures of Effectiveness 65
6.1 ASSESSING IMPACTS ON TMC OPERATIONS 66
6.1.1 Impacts on TMC/Agency Resources 66
6.1.2 Impacts on Agency Reputation and Public Engagement 66
6.1.3 Improve Ability to Meet TSM&O Operational Objectives 67
6.1.4 Improve Data Effectiveness 71
7 Conclusion 72
7.1 SUMMARY 72
7.2 FURTHER RESEARCH NEEDS 73
7.2.1 Dedicated Mining 73
7.2.2 Evaluate Accuracy of Citizen reports 74
7.2.3 Assessing Potential Biases in Social Media Data Sources 74
References 75
APPENDIX A List of Acronyms 79
APPENDIX B Glossary 81
APPENDIX C Online Survey Results 83
APPENDIX D Panel Meeting Attendees 95
APPENDIX E Project Team Participants 96
APPENDIX F Panel Meeting Agenda 97

List of Tables

Table 1-1: TMC Internal Capabilities 7
Table 5-1: Operational Considerations....50
Table 5-2: Institutional Considerations 51
Table 5-3: Technical Considerations 53
Table 5-4: Applicability of Crowdsourced Data by Type and Activity 54
Table 5-5: Overcoming Data Deficiencies....56
Table 5-6: SWOT Analysis Matrix....63
Table 6-1: Impacts of TMC/Agency Resources 66
Table 6-2: Impacts Agency Reputation and on Public Engagement 67
Table 6-3: Improve Ability to Meet TSM&O Operational Objectives....68
Table 6-4: Improve Data Effectiveness 71

List of Figures

Figure 1: SWOT Model Framework for Crowdsourced Data Initiatives 2
Figure 2: Incorporating Crowdsourced Data into TMC Operations 5
Figure 3: Metropia Mobile App 18
Figure 4: Priority of TSM&O Activities 22
Figure 5: Current Data Sources for TSM&O Activities....23
Figure 6: Perceived Issues with Current Data 24
Figure 7: Perception of Data from Crowdsourcing and Social Media Data....24
Figure 8: Perceived Threats to Using Crowdsourced Data 25
Figure 9: Agencies Motivations to Use Crowdsourced Data 25
Figure 10: Current Usage of Crowdsourced Data 26
Figure 11: Emerging Crowdsourcing Trends 28
Figure 12: Responsibility for Monitoring Social Media....29
Figure 13: Iowa DOT Twitter Post 36
Figure 14: Engaging a Traveler via Twitter....37
Figure 15: Social Media Dialogue 38
Figure 16: Traveler Engagement via Facebook....39
Figure 17: Using Radian6 to Monitor Social Media 40
Figure 18: UDOT Citizen Reports Mobile App 41
Figure 19: UDOT Citizen Reporter Training Screen 42
Figure 20: DDOT Operator View of Mining Results....44
Figure 21: SWOT Analysis Dimensions....47
Figure 22: SWOT Model Framework for Crowdsourced Data Initiatives 64

Executive Summary

Accurate and timely real-time data is critical for supporting traffic operations. Traffic Management Centers (TMCs) acquire, process, and integrate data from a variety of sources including: data manually collected by TMC personnel and recorded directly into TMC software; Intelligent Transportation Systems (ITS) field instrumentation; partner agencies or other stakeholder systems, and private sector providers. Public agencies that operate TMCs continually seek the latest advances in data acquisition and processing while striving to improve data quality, expand coverage, and minimize cost.

Crowdsourcing has been identified as one of the top trends and technologies that public agencies can adapt and take advantage of from outside the transportation community. Any equipped person (e.g., a person with a smartphone) or vehicle can be a potential data source to support TMC operations. The three most widely accepted approaches to generating crowdsourced data suitable for TMC operations are:

- Extracting data from social media platforms (e.g., Twitter or Facebook) by actively monitoring social media, engaging citizens in two-way dialogues over social media, or mining social media.
- Acquiring crowdsourced data from a private sector data provider/aggregator. Today, many traffic management agencies purchase from or barter with private sector companies (such as Google, INRIX, HERE, TomTom) to obtain data. These companies generate data using vehicles as probes leveraging smartphone applications and by relying on various forms of citizen reporting.
- Developing and deploying their own specialized mobile applications (Specialized Apps) and recruiting citizens to serve as data collectors.

Interest in crowdsourced data for TMC operations is growing. Several state and local agencies have active projects to integrate crowdsourced data into TMC operations. The purpose of this project is to identify, research, and synthesize the opportunity for an agency operating a TMC to work with crowdsourced data from available applications, and to develop measures that can articulate the value proposition to work with such resources and evaluate the impacts on TMC operations.

A thorough assessment of the current state of the practice as it relates to crowdsourcing and social media was accumulated by preparing a literature review. Through an on-line survey, a series of phone interviews, and a two-day workshop, TMC staff and other state agency personnel and leading industry experts were engaged to assess prevailing perceptions of crowdsourced data and social media to support TMC operations.

A SWOT (**S**trengths, **W**eaknesses, **O**pportunities, and **T**hreats) model framework is presented to provide guidance for agencies who want to advance crowdsourced data projects. The SWOT model enumerates the organization's internal capabilities to undertake the initiative by articulating its strengths and identifying its weaknesses. To assess the potential value of the proposed initiative, opportunities and threats are elaborated. Figure 1 summarizes this framework.

The Strengths and Weaknesses boxes within the SWOT Model are used to assess the internal readiness of a transportation agency to plan and execute crowdsourced data projects. This report

suggests a series of operational, institutional, and technical considerations that agencies can use to assess their strengths and weaknesses.

- ***Operational***: Attributes of the state agency within the TMC environment including, its staff, available resources, relationships with citizens and stakeholders, and core Transportation Systems Management and Operations (TSM&O) activities.
- ***Institutional***: Attributes and culture of the state agency external to the TMC environment.
- ***Technical***: Specific skills and capabilities within the TMC and its staff to perform the tasks necessary to acquire and integrate crowdsourced data.

Determining whether these implementation considerations are seen as strengths or weaknesses at a specific TMC is crucial to evaluating the capabilities and readiness for an agency to initiate a new project. To be considered a strength the agency must currently possess or willing and able to develop the capability and be ready to apply that capability in support of a new crowdsourced data initiative. A weakness would indicate that the agency lacks the necessary experience or resources.

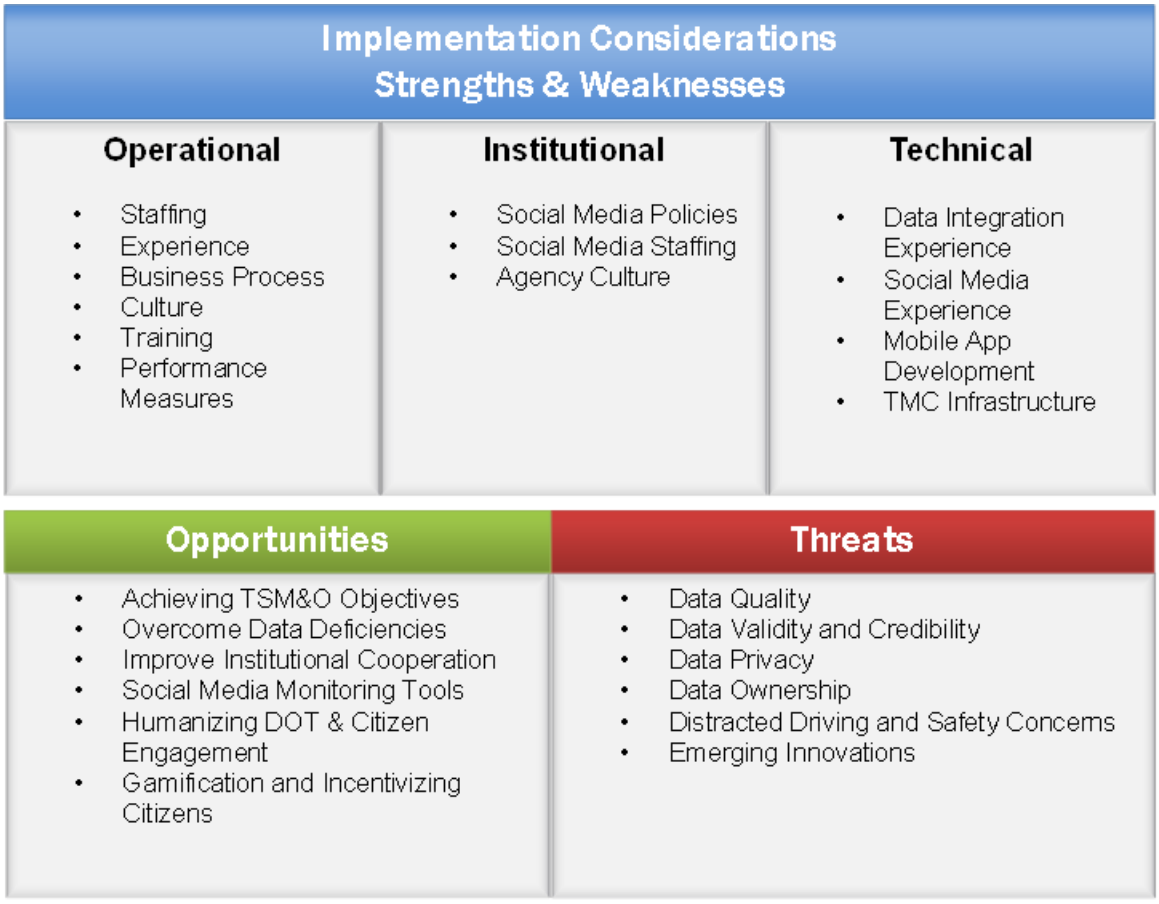


Figure 1: SWOT Model Framework for Crowdsourced Data Initiatives
(created by Open Roads)

The Opportunities and Threats quadrants of the SWOT Model focus on external considerations for integrating crowdsourced data into TMC operations. The opportunities quadrantseeks to quantify the potential benefits and advantages to be gained by undertaking a crowdsourced data project. Several

opportunities for integrating crowdsourced data into TMC operations emerged from the study. TMC staff recognized direct benefits for achieving TSM&O objectives and overcoming current data deficiencies. Additional opportunities were identified that would yield important institutional and societal benefits including improving cooperation across agencies, humanizing state agencies, and increasing citizen engagement.

The Threats quadrant enumerates issues that could present barriers to undertaking crowdsourced data projects. Data quality management, data overload, and data ownership considerations were cited as the most pressing issue for public agencies. Data privacy and safety concerns are the most critical issues as it relates to engaging citizens as data collectors.

It is recommended that systems engineering best practices be adopted for planning and implementing crowdsourced data projects. The value proposition for integrating crowdsourced data is linked to its cost efficiency to be acquired and integrated into TMC operations and its ability to enhance meeting TSM&O objectives:

- ***Cost Efficiency***: Evaluating the return on investment by considering the impacts on TMC operations and resourcing. Crowdsourced data is perceived by state agencies as being a cost-effective alternative to existing data sources with minimal integration needs. Agencies need to consider all of the direct and indirect costs associated with planning and implementing crowdsourced data projects.
- ***Improved ability to meet TSM&O Objectives***: Crowdsourced data is applicable to a wide range of TSM&O activities, most notably incident management and road condition reporting. The value for integrating crowdsourced data is correlated to the degree to which it will help agencies improve their ability to meet TSM&O objectives.

This field of study is still quite new and is rapidly evolving. It is premature to propose a specific roadmap for state agencies that wish to make use of crowdsourced data. This document will, however, provide guidance on how to identify the best opportunities for individual agencies, and build or acquire capabilities to exploit those opportunities.

1 Introduction

> *"The most transforming implications for TMC operations are based on the proliferation of wireless communication, the rise of social media, and the involvement of third parties. New classes of real-time holistic data become available to TMC operations, often through third parties."*
>
> *FHWA-HOP-13-008*
> *Impacts of Technology Advancements on Traffic Center Management Operations*

The rapid rise of social media and mobile applications (mobile apps) is transforming how information is collected and shared across society. Over the past few years, public transportation agencies have embraced social media as a platform for interacting with the traveling public and disseminating real-time traveler information. State Departments of Transportation (DOTs) have a strong social media presence to actively engage citizens in transportation-related issues and concerns. Many agencies that operate TMCs are actively using social media platforms, like Facebook and/or Twitter to inform travelers of adverse road conditions including heavy congestion, incidents, and weather impacts.

While social media has provided a valuable platform for pushing data and information from traffic management agencies to the traveling public, there is an untapped opportunity to embrace the crowdsourcing nature of social media to gather real-time information. In a recently published Federal Highway Administration (FHWA) report "Impacts of Technology Advancements on Traffic Management Center Operations" (FHWA-HOP-13-008), social media for traveler information and crowdsourcing is documented as a rising trend among TMCs (9). This report notes the opportunity to expand use of social media from merely a dissemination tool to a platform for receiving and distributing information among agencies, travelers, and third parties. Utilizing crowdsourcing for traffic information is cited as an emerging strategy to address the growing gap between mobile users and traffic management agencies as it relates to the limited ability for users to provide TMCs with real-time information.

With growing interest in crowdsourcing and social media, the FHWA Office of Operations in conjunction with the TMC PFS, commissioned a project to investigate crowdsourcing from social media and recommend best practices for TMC operations. The overall scope of this project is to: (1) develop a synthesis of how crowdsourced data from social media (i.e., content generated through end-user applications like Waze, Facebook, and Twitter) are applied at TMCs; and (2) determine measures of effectiveness that express the value for applying crowdsourced data.

This report summarizes the findings of the research effort. It is intended for the transportation community and to be used for continued technology transfer activities by the United States Department of Transportation (USDOT) related to applying crowdsourced data to TMC operations.

1.1 Project Background

TSM&O relies on the acquisition, processing, fusion, and analysis of data from multiple sources. TMCs traditionally acquire data from four sources:

- Data manually entered by TMC staff into software applications

- Data generated by TMC-managed field devices (e.g., sensors, camera, road weather information systems)
- Data acquired from partner agencies (e.g., state police, emergency management agencies, neighboring states, corridor coalition partners
- Data sourced from private sector or third-party providers

Agencies make decisions related to the sourcing and processing of data based on a number of factors including the cost of acquiring data, the cost of integrating and storing data, coverage area, the quality and accuracy of the data, and the impacts on data latency on real-time decision making. As the availability of new data sources has greatly expanded in recent years, traffic management agencies have steadily improved TSM&O performance. Opportunities remain to further improve TSM&O performance and agencies are eager to explore the availability of new data sources to further help the cause.

Crowdsourced data shows promise for supporting real-time operations. Any equipped person (e.g., a person with a smartphone) or vehicle can be a potential data source that an agency can tap directly or indirectly through a third-party aggregator. The three most widely accepted approaches to generating crowdsourced data suitable for TMC operations are:

- Extracting data from social media platforms (e.g., Twitter or Facebook) by actively monitoring social media, engaging citizens in two-way dialogues over social media, or mining social media.
- Acquiring crowdsourced data from a private sector data provider/aggregator. Today, many agencies purchase from or barter with private sector companies (such as Google, INRIX, HERE, and TomTom) to obtain data. These companies generate data using vehicles as probes and by relying on various forms of citizen reporting.
- Capturing crowdsourced data from citizens via the use of specialized apps.

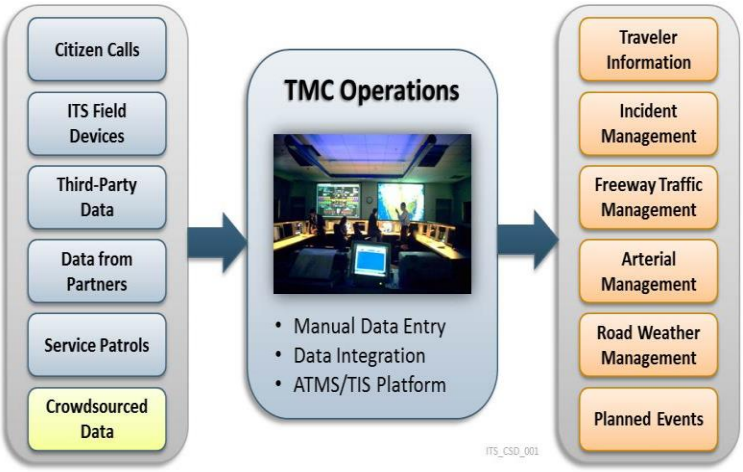


Figure 2: Incorporating Crowdsourced Data into TMC Operations (Source Noblis)

Each of these three approaches provides different opportunities to support TSM&O activities. In addition, planning and implementing projects around each approach presents and requires different institutional, operational, and technical considerations. This report analyzes the opportunity for an

agency operating a TMC to work with crowdsourced data from these available applications, and presents measures that can articulate the value proposition to work with such resources and consider their impacts on TMC operations.

1.2 Terminology

For the purposes of this report, the following key terms are defined.

- ***Crowdsourcing*** refers to the practice of obtaining services or content by soliciting contributions from a large group of people.
- ***Crowdsourced Data*** refers to data that are collected from a series of individuals. This term is used liberally in this report, making no distinction between raw data or a processed data set that has been sanitized and normalized.
- ***Active Data Collection*** refers to collecting data from people who voluntarily and manually contribute services or content. This can occur through agency-sponsored specialized apps where citizens provide real-time data to support TMC operations.
    - **Waze** is one of the most widely used community-based traffic and navigation app. Through the app, travelers can actively report and share real-time information on incidents, road hazards, and traffic jams with other members of the community. Waze also passively collects speed and location information from the device.
    - **FixMyStreet** is an app that enables citizens to report, view or discuss local problems such as graffiti, fly tipping, broken paving slabs, or street lighting.
- ***Passive Data Collection*** refers to the process of automatically acquiring data produced by system users whose mobile phones and vehicles act as probes. The app or system only needs to be "on" to be sending data. There is no extra intervention by the user required.
    - **Street Bump** is a mobile app that originated in Boston and passively collects data about the smoothness of the ride (e.g., potholes and speed bumps) and shares that information with city officials.
- ***Social Media*** is a platform for sharing user-generated content among a community of users. Within this report, social media is used to identify a group of internet-based applications enabling participation in social networking. Facebook and Twitter are two examples of social media platforms commonly used for interacting with travelers.
- ***Social Media Monitoring*** is the active monitoring of social media channels for what is being said on the internet about an organization or company. Tools exist, such as Radian6 and NetBase, to help organizations track what consumers are saying about their brands and actions.
- ***Social Media Dialogue*** is a two-way information exchange between two or more users on the social media network. For example, a citizen uses Twitter to send a tweet to a state DOT with information about a possible malfunctioning traffic signal. The state DOT responds to the tweet and includes the signal technicians at the TMC to relay that information. Other citizens may see the conversation, join the dialog, and contribute information.
- ***Social Media Mining*** is the process of discovering useful or actionable knowledge in social media sources by introducing specialized concepts and technical approaches for analyzing massive social media data. For example, tweets can be matched to traffic incidents and

weather road conditions by examining the content for key words and comparing the timestamps and locations of the tweets and incidents.

1.3 Research Focus

Integrating data into TMC operations is a complex undertaking. Many TMCs have staff with experience processing, fusing, integrating and archiving data from disparate sources. General approaches to data management have been well-documented in previous FHWA and TMC PFS reports. The intent of this report is to focus on the new and specific challenges that data from crowdsourcing and social media can present to TMC staff.

Crowdsourced data projects require agencies to have experience in social media and/or planning and developing specialized apps. The internal capabilities and along with institutional culture and experience will influence the readiness and abilities of an agency to undertake such efforts.

Table 1-1: TMC Internal Capabilities

Considerations	Examples of Capabilities
Operational Considerations (TMC and its staff)	• Staffing resources, experience, and skills • Business process to plan and execute technical projects • Experience and culture as it relates to social media usage
Institutional Considerations	• Agency-wide policies as it relates to social media and crowdsourcing • Agency policies as they relate to citizen engagement • Level of cooperation and coordination between TMC and agency
Technical/Tactical Considerations	• Capabilities to acquire and integrate new sources of data • Experience designing and/or developing mobile apps • Existing software and hardware platforms to support data integration, sharing, and warehousing

Once a decision has been made to move ahead with a crowdsourced data project, there are several implementation procedures, processes, and considerations including:

- Best practices for data fusion, processing, and warehousing
- Policy issues related to crowdsourced data including quality measurement, ownership rights and privacy
- Resource needs for deployment and maintenance including staffing, software, and equipment
- Training needs

This report seeks to provide guidance to agencies and their TMC staff on the current best practices for planning and executing crowdsourced data projects. This information is presented in four major parts:

- ***State of the Practice***: The report summarizes the current state of practice regarding integrating crowdsourced data into TMC operations.
- ***Case Studies***: The report presents a few selected case studies that demonstrate projects for each of the major approaches to generate or integrate crowdsourced data for TMC operations.
- ***Strategic Relevance and Value Proposition*****:** A SWOT analysis framework is proposed to help agencies assess their readiness and capabilities to plan and undertake projects to acquire and integrate crowdsourced data. This SWOT model is helpful for assessing the internal strengths and weaknesses of an agency and measuring the opportunities and threats posed by each approach to generate crowdsourced data. This model can serve as a template for agencies who want to create their own SWOT model and Concept of Operations to support new social media projects.
- ***Measures of Effectiveness (MOEs)*****:** Crowdsourced data from social media must provide agencies with a return on investment to justify the effort and cost required to undertake the effort. This report presents a series of MOEs directly related to crowdsourcing as well as the TSM&O activities to be supported.

1.4 Document Overview

This report includes documentation of the entire research project, the guidance concerning the MOEs that express the value of applying crowdsourced data as described in the SWOT analysis, and the implementation recommendations for integrating crowdsourced data within a TMC. The content of this report is organized as follows:

- Chapter 1 provides the project background, objectives and research approach, and the report organization.
- Chapter 2 presents the results of the comprehensive literature review.
- Chapter 3 presents an overview of the state of the practice gleaned from TMC staff and industry experts. The team undertook an on-line survey, conducted a series of phone interviews, and facilitated a two-day workshop to gather information on existing and planned efforts to make use of crowdsourced data in TMC operations.
- Chapter 4 presents a series of case studies that describe current projects by agencies across the country to acquire and integrate crowdsourced data into their operations.
- Chapter 5 presents the SWOT model. It discusses how agencies can assess their internal strengths and weaknesses by looking at operational, institutional, and technical considerations. Opportunities and threats are presented to help TMCs determine the value and assess risks associated with crowdsourced data projects.
- Chapter 6 discusses measures that can articulate the value proposition for an agency to work with such resources and consider their impacts on TMC operations.
- Chapter 7 provides a summary and conclusion for this work.
- References serve as a bibliography for the report and a listing of useful references.
- Appendices A-F provide supporting information such as a list of acronyms, glossary, survey results, and panel meeting materials.

2 Literature Review

The starting point for understanding the state of the practice in using crowdsourced data from social media is to review pertinent literature from traditional sources, such as scholarly journals, conference proceedings, and study reports, as well as websites and online blogs. This chapter presents a synthesis of the literature and summary of major findings. In general, the use of crowdsourced data from social media at TMCs is in the embryonic stage and there were limited resources on this specific subject. However, as social media and crowdsourcing applications and research topics are rapidly evolving, the study team uncovered a significant amount of literature on the topic as a whole which provided a rich background for this study. The References section contains the full set of references.

2.1 Sources of Transportation-Related Crowdsourced Data

Agencies operating TMCs seek to acquire a wide variety of traffic-related information, much of which has recently become available via crowdsourcing and is anticipated to become more widely available in the future. Crowdsourced traffic data can come from traditional social media sources such as Facebook or Twitter, specialized mobile applications, and connected vehicles. This data can be categorized into two types based on how it is produced: active sources and passive sources.

2.1.1 Active Sources

Active sources of crowdsourced transportation data are those in which a user voluntarily and manually provides information to a TMC or to a third-party, usually relevant to real-time conditions. This data is typically provided via a mobile device, whether through a social media application such as Facebook or Twitter or another application aimed at more specific information such as incidents or weather. Note that actively-crowdsourced information is not always provided in real-time. For example, a commuter could post on Twitter about their commute once at their desk at work. Traffic information from active participants can include incidents, work zones, detours, roadway conditions, weather data, and overall user sentiment of the transportation system.

2.1.1.1 Social Media

Traditional social media platforms such as Facebook and Twitter are currently the most useful and widely-used social media platforms for transportation agencies (18, 19, 23). The American Association of State Highway and Transportation Officials' (AASHTO's) 4th Annual State DOT Social Media Survey highlighted the use of social media by 43 state DOTs and the District of Columbia (23). The survey reported in 2013 that over 90 percent of states used both Twitter and Facebook, as opposed to less than half using Facebook and only 26 using Twitter in 2010. This increase in use has been "fueled by the proliferation of portable devices constantly connected to the internet" (23). Most agencies have personnel dedicated to maintaining their social media presence. Facebook and Twitter are used for traffic alerts, emergency notifications, posting photos and videos, and sharing information such as public meetings and public involvement opportunities, among others (16, 23). Facebook and Twitter also provide a direct mechanism for a transportation agency to interact directly with citizens and respond to questions and concerns. A 2013 FHWA report on the "Impacts of Technology Advancements on Transportation Management Center Operations" suggests that social media should encourage a *two-way* information exchange, and TMCs should utilize crowdsourcing for traffic and

incident information, infrastructure conditions including pavement roughness, and feedback on department performance (9).

Traffic and incident data may also be available on social media, especially Twitter, to parse and analyze. Incident data is largely reported manually and voluntarily. When this information is provided via crowdsourcing (as opposed to law enforcement / Closed Circuit Television recognition from a TMC operator), it typically begins from somewhere near the incident (32). Users can communicate directly with their local transportation agency and with other travelers to report an incident or service disruption via social media, typically using Twitter or perhaps Facebook. This helps get the word out in a timely manner without having to wait for the TMC. However, an ongoing study notes that "there is still no clear consensus among transportation managers on how social media can be used to collect and disseminate warnings to the public"; this study gathered information on best practices for using social media to support non-routine traffic operations, including how information from the public is best collected via social media (4). Incident data is also widely reported using mobile applications, as discussed in the next section.

Active contributors of transportation data also often post personal feelings on social media that may be relevant, such as displeasure with a service interruption. A passenger's user experience "can often be overlooked in transportation analysis, perhaps due to a lack of reliable information" (26). This trend is changing with the ability to mine social media for sentiment. Transportation agencies can also track information such as "likes" and "retweets" to see which information is connecting the most with users (9). Further discussion on how social media is mined for information such as customer sentiment can be found in Section 2.2.3.

2.1.1.2 Mobile Applications

A wide variety of actively-produced transportation data comes from user participation in mobile applications, some of which have been developed directly by a transportation agency and others which involve an outside party. For general traveler information, most transportation agencies have developed mobile-friendly web sites, and some have developed separate mobile applications for accessing their web content (23). Many states' 511 Traveler Information Systems now have mobile apps for accessing traveler information (23, 50). Some municipalities have developed applications to allow travelers to report on roadway conditions. For example, the Utah DOT has a Citizen Reporting App that allows trained volunteers to report on current road conditions throughout the state (1); several municipalities in Massachusetts including Boston have "Citizens Connect" apps to report potholes or infrastructure damage (48). The Wyoming DOT is in the process of developing a similar system (10). Beijing, China, has an online portal for cyclists and pedestrians to report on infrastructure conditions via their smartphone to help transportation planners identify problem areas and prioritize investments (36). Oregon DOT made a deal with Strava, a leading website and smartphone app used by cyclists to track their bike rides via Global Positioning System, to purchase a one-year license of a dataset to be use in a highway management context such as integration into travel demand models and forecasts, construction detour plans, and maintenance scheduling (40).

Most active-participant traffic data crowdsourcing involves third-party mobile applications. Waze, the "world's largest community-based traffic and navigation app" (43), allows users to share real-time traffic and road information, such as work zones, police locations, detours, etc. (2, 13, 21, 30, 31). In 2013, Google purchased Waze for over $1 billion and has been integrating Waze notifications into its mapping services (13). Apple is looking to integrate similar third-party supplied incident information with its mapping services as well (13). Trapster is a similar application which originally focused on alerting drivers to speed traps and speed/red light cameras but has grown to include incident data and speeds along major roads (35, 44). It has been integrated with Nokia's HERE mapping service and,

like Waze, allows road users to actively share information with others , providing users with information in return (35).

Crowdsourced weather data is also being collected from volunteers concerned about reporting and understanding what they are actually experiencing rather than “what the weatherman tells them they are experiencing”. Weather forecasts and reported conditions, such as precipitation radar, do not always reflect ground conditions (49). New active weather-reporting mobile applications, such as Weathermob (nicknamed the “Waze for Weather”), allows users to report and share what they are experiencing in their exact location (41). National Oceanic and Atmospheric Administration (NOAA) also has released a smartphone app called mPING that allows users to input weather conditions (49). Applications such as these create in essence a fleet of mobile weather stations. Because of the impacts of weather events on traffic, sharing of weather data gathered by crowdsourcing may have important benefits to transportation agencies and travelers.

2.1.2 Passive Sources

Transportation agencies and third parties can also obtain a variety of crowdsourced data via “passive” users whose mobile phones and vehicles act as probes for data collection. Available data includes speeds, travel times, pavement roughness, and some weather information. Technologies and applications are being developed to aggregate and infer further data, such as volumes, incident detection, and additional weather and infrastructure condition information.

It should be noted that personal information can harvested passively without user knowledge. Users routinely grant permissions to mobile apps without reading the fine print about what information they are providing. Furthermore, simply turning on location-based services allow the app to track the user. In some cases apps can collect data from the movement of the cell phone without the owner/user specifically entering data.

2.1.2.1 Mobile Devices

Speeds and travel times are perhaps the most established passive crowdsourced data source (9, 14). The private company INRIX uses commercial fleets, taxis, and individual users running their app as probes to obtain a wide coverage of traffic data (3, 12, 15). The private sector currently owns most of the market share in this area, due to privacy issues with the collection of individual vehicles’ movements and the much greater availability of data to achieve reliable results. INRIX obtains its information by agreeing to share its aggregated travel time information with the probes supplying the information; if a driver downloads the INRIX app, he agrees to let his mobile device act as probe (15). Applications such as Google Maps, Waze and Trapster obtain traffic information for roadways from application users’ phones as well, sharing the aggregated crowdsourced information in return (9, 22, 35).

Crowdsourced weather data that is passively collected via mobile probes is becoming more widespread as well. WeatherSignal is an Android-only mobile app that utilizes built-in sensors on phones to crowdsource weather data such as pressure, humidity, temperature, and light intensity, which can be integrated together to provide an accurate, localized picture of weather conditions (45). Passive data sources are also being developed for crowdsourcing information about infrastructure conditions. For example, Street Bump is a mobile phone app developed in Boston that passively collects road condition data as a user drives; however, this application must be enabled to “record a trip” at the beginning of a trip (46).

2.1.2.2 Connected Vehicles

In the future, connected vehicles may serve as an all-in-one probe for incident, infrastructure condition, safety, weather, and speed/travel time data (32). The precise geo-locating ability and anticipated broad coverage of connected vehicles could allow for a detailed data stream of lane-by-lane information. A 2013 report from the University of Virginia Center for Transportation Studies notes that as the types of data available "expands exponentially with the introduction of connected vehicle data," there will be a growing level of responsibility for obtaining, storing, managing, and integrating this data with other sources, such as existing probe data (32).

Connected Vehicle research is exploring the application of weather data collected from on-board vehicle systems, such as pressure, windshield wiper settings, anti-lock braking brake status, traction and stability control, and differential wheel speeds (8). The I-95 Corridor Coalition Vehicle Probe Project also is planning to include weather data sources from on-board vehicle systems, including braking and windshield wiper use (12). Ideally, this research will allow data to be "shared among other travelers or to the roadway agency operating the road" (8). Currently, INRIX already has agreements with some auto companies to offer connectivity and the sharing of travel information from specific vehicle models (15). Soon, vehicles themselves could be the probes that mobile devices currently function as, with the ability to provide even more information about roadway and weather conditions.

2.2 How Crowdsourced Data is Obtained

TMC operators can crowdsource data directly from in-house applications or indirectly from third parties; they can also actively parse social media to obtain further information. TMC managers must assess current and future data needs, including how this data will be acquired and stored. It is anticipated that the private sector, with the flexibility for much larger storage and more sophisticated data management, will be engaged to help (9). TMC staff will need to work collaboratively with third parties to optimize the collection and dissemination of information to the traveling public. As multiple sources noted, social media and mobile applications are constantly evolving; the most popular and useful tools currently may be replaced or overtaken in the near future, and agencies must adapt to keep up (9, 23).

Due to the raw nature of crowdsourced data, and the economies of scale needed to cost-effectively develop methods to extract meaningful information from it, it is likely third parties will continue to be intermediaries for the aggregation of this data. Probe data providers are current examples of this. INRIX reports that 46 states are currently using its free service for viewing real-time traffic (www.inrix.com). Many of those states are purchasing real-time probe data from INRIX and other companies.

2.2.1 In-House Applications

As mentioned, a few transportation agencies have developed in-house applications encouraging citizens to report traffic conditions directly to the agency. The Utah and Wyoming DOTs have citizen reporting programs allowing trained volunteers to report on road conditions (1, 10). The Utah DOT compiles incoming data "with carefully crafted logic" to determine road conditions and utilizes information differently depending on whether the reporter is a plow driver, law enforcement, or volunteer citizen (1). Training to become a citizen reporter includes a brief online or in-person course and quiz (1). Wyoming also has training for what information to report (pavement/weather conditions, debris/animals on roadway, incorrect traveler information on Dynamic Message Sign and Highway Advisory Radio) and how to report this information in a standardized fashion (10). The in-house

applications cited here are focused more on maintenance needs and less time-sensitive alerts than real-time operations data that would be used for TMC operations, however. Much of the traffic data used for TMC operations, at least of the crowdsourced variety, comes from the private sector.

In the future, connected vehicle data is likely to be obtained from both in-house sources and private sources. Using vehicle-to-infrastructure applications, TMC operators will have access directly to this data through their TMC software receiving data from the roadside (32). It is expected that the expansion of coverage and granularity of data will greatly enhance TMC operations while also requiring new software and personnel capabilities (32).

2.2.2 Private Sector Data

As mentioned, the private sector currently owns most of the market share for large-scale travel time data and other traffic-related data, including large vendors such as Google and INRIX (15, 35). Depending on the situation and need, agencies may need to purchase this information for TMC operations or they can share/exchange data with a third-party.

With third-party travel time information providers, transportation agencies can purchase this information but must specify standards and coverage area (14). This allows transportation agencies to avoid having to "put out for infrastructure" (15). Rather than having coverage limited to areas with Bluetooth readers or other data collection devices, agencies can pay a vendor such as INRIX for nationwide coverage of travel time data, including arterials (15, 20). Arterial data is rapidly improving from third-party data providers, although data accuracy is still a concern (32).

Agencies can also provide information from TMC operations back to the private sector. Many state DOTs, including Washington and Virginia, currently have third-party data agreements (9). For example, Caltrans pays Google for a traffic map incorporating data from its own loop detectors (which it shares with Google) with Google's vehicle probe data (34). Agencies must check with the conditions of their data agreements with third parties before publishing (9). As more information becomes available from third-party apps, agencies may need to focus on providing information that third parties do not have access to, such as construction information, so as to minimize redundant data feeds (9).

Connected vehicle data may also be funneled through the private sector as businesses with specialized expertise can convert the volumes of raw data into meaningful information to support TMC operations (32).

2.2.3 Mining Social Media

Social media can provide a large amount of transportation-related information, both directly and indirectly targeted toward the transportation agency. This data needs to be parsed and analyzed into "digestible" contexts and checked for accuracy and relevance. Social media can provide information about specific real-time incidents as well as overall traveler sentiment toward the transportation provider. Some agencies are correlating 511 usage (via web and phone) during peak periods or incidents with conditions of the transportation network; a future strategy for TMC operations should be to correlate social networking activity such as tweets with this data (9). As an example outside of TMCs, the Weather Channel mines tweets by location to infer weather conditions in localities; this information could potentially be obtained and utilized by TMC operations (9).

Mining social media can be especially powerful for understanding how users of the transportation system are experiencing real-time operations. Research has shown that basic algorithms can be successful in mining customer sentiment on Twitter for transit agencies and airlines (7). A study by the

New Cities Foundation concluded that sentiment analysis of user posts – in their study, from Waze and Roadify – is a practical and applicable tool for gauging commuters' emotions as well as learning about real-time infrastructure problems (2). The study recommends that transportation agencies obtain and aggregate commuters' sentiments for a variety of reasons, including the ability to target commuters with specific messages at specific times, the ability to gain real-time information on incidents or needed repairs, and the ability to evaluate the impact of investments (6). Ultimately, crowdsourcing can allow an agency to see how users are responding to information and the real-time impacts of their responses on traffic operations (9).

2.3 State of Practice at TMCs

Social media and, to a lesser extent, crowdsourced data usage have ramped up considerably for transportation agencies over the past few years. Many agencies are challenged to keep up with advancements in technology and demands for information (23). Usage of these technologies varies across agencies depending on staffing, data availability, and coordination with other parties, including the private sector.

Social media usage has been widely adopted by agencies operating TMCs, but generally in a one-way nature for disseminating information to the public more than for collecting information. Advancements in mining techniques and third-party assistance for analyzing social media usage can help TMC operators better understand customer sentiment and potentially the impact of their outreach on real-time operations.

2.3.1 Usage of Social Media

Traffic management agencies are widely using Facebook and Twitter to send out traffic alerts and emergency notifications. Due to its dynamic nature, Twitter is typically used "for real-time notifications such as crashes, closures or major weather impacts" (9). The AASHTO survey of state DOTs' social media usage reported that in 2013, nearly 60 percent of surveyed state DOTs used Twitter for emergency notifications and 93 percent used Twitter to post traffic updates; 70 percent used Facebook for emergency notifications and 63 percent used Facebook for traffic alerts (23). Transportation agencies have varying levels of support for these operations, but most transportation agencies are actively using social media and have personnel dedicated to maintaining their social media presence. For example, Michigan DOT has multiple staffers monitoring their Twitter feed in real-time to respond to real-time problems reported by users (17). A study of traveler information outreach for a major construction project in Texas showed that the Texas DOT plans on using Facebook and Twitter to disseminate information about incidents and delays (28).

Social media is especially used for a variety of information dissemination that is not on a real-time basis (23). The National Cooperative Highway Research Program (NCHRP's) 2013 State DOT Chief Executive Officer Leadership Forum released a whitepaper on "Technology and Business Practices that Work", highlighting Florida DOT's leadership and commitment to engaging the public via social media (16). The agency utilizes various Twitter, Facebook, and YouTube feeds differentiated by various sub-agencies (such as the state's districts) for more targeted messaging. Additionally, the state's transportation secretary engages the public via social media "and uses his tweets to elevate and stimulate public discussions, to shape the transportation debate, and encourage thought on critical topics in transportation in the state". The secretary also hosts webinars and podcasts intended to engage the agency employees but also shared publicly via YouTube. Washington State DOT has been successful at using social media to reach out to the public "to explain and promote Active Traffic

Management and other Active Transportation and Demand Management strategies" (9). The Texas DOT reached out to the public via social media to promote a survey of I-35 travelers, asking what types of information motorists would find most useful and what channels they would utilize to receive such information (28). This allowed the Texas DOT to get user feedback on what information they should disseminate in real-time once the project began.

Ultimately, Facebook and Twitter provide a direct mechanism for a transportation agency to interact directly with citizens and respond to questions and concerns. Ideally, this will be a two-way information exchange. Currently, however, there is still a gap between TMC operators and social media users in the information exchange from the user to the state DOT; there is much more "pushing" information from the state DOT than "pulling" information from the public (9).

2.4 Performance Measurement

As social media and crowdsourcing data become "the norm" for transportation agencies, it is important to quantify the effectiveness of these efforts and identify where messaging can be improved. The AASHTO 4th Annual State DOT Social Media survey noted an emerging trend of "increased emphasis on measurement and formally tracking social engagement" (23), although "the science of measuring social media use is still evolving" (21). In the past, there has been "little consistency" in this measurement outside of tracking the number of followers/friends (23), but:

> *"The pure number of Twitter followers or Facebook fans does not indicate impact. It is more important to understand who follows your Twitter or Facebook profile, what they do with the content, and who is in their network. Social networks have the ability to distribute information from friend to friend and to their friends reaching many more than those directly following your updates." (17)*

A New York University report on social media usage by transportation agencies in New York City (18) provides detailed recommendations for transportation providers' social media policies. Social media posts should be accessible, informative, engaging, and responsive. Part of performance measurement could include collecting and analyzing feedback:

> *"Data received via social media, whether a complaint about a bus driver or accolades for a new service, must be internalized to the agency. This information cannot stop with the public affairs officers who manage the social media account; it should be quantified, evaluated, and passed on to appropriate organization members. Frequent complaints can pinpoint a reckless train conductor, an especially friendly bus driver, or bring to light an incident that has not yet been reported by management. Using incoming feedback for improvement is a major step in social media success." (18)*

Basic free tools such as Facebook Insights and YouTube Insights can track the number of views for a post (21). Google Analytics also can help an agency track website visits, including which pages are the most searched-for and popular and the geographical location of website visitors (21). Klout is a well-known service used to generate a composite score of a user's social media presence and influence, integrating multiple accounts such as Twitter and Facebook if necessary, for the application site (21, 47).

More sophisticated (and often more expensive) analytic software is popular in the business community, but so far transportation agencies have been relatively slow to adopt (23). Radian6 is a

more expensive tool that “is particularly good at separating and categorizing sentiment”, allowing a company – or transportation agency – to more easily analyze their strengths and weaknesses in the eyes of the customer (25). Sprout Social is a less expensive tool that provides a social interaction score for each platform that a company or agency uses as well as useful demographic information to help “understand the effect of individual engagements so that you can more effectively plan future campaigns” (27). Other popular tools include HootSuite and Rignite, as well as many others, some of which can target a specific aspect of social media monitoring (25).

Finally, for a transportation agency to quantify the effectiveness of its social media usage, it must be able to quantify the cost. More than a third of state DOTs had exclusively-dedicated staff for maintaining and managing their social media activities in 2013, although shrinking resources including workforce cuts can be a challenge (23). Many public transportation agencies can generally estimate the number of employee-hours dedicated to social media, but “most did not put a price tag on this effort” (21). Staffing and social media monitoring tools are two important costs that a transportation agency must incur to expand and implement a social media presence.

2.5 Crowdsourced Data Usage in Other State or Federal Agencies

Outside of the transportation sector, federal, state, and local government agencies are beginning to use the wealth of crowdsourced data to assist with their operations. For example, the District of Columbia launched a program called Grade.DC.gov. The program is gathering data from social media as well as any comments posted to any of the D.C. government websites. The data is collected and run through text-analytics to extract significance. Certain key words are assigned various levels to determine the sentiment of each message collected. The text analytics is backed by a human team to help find and locate slang used in social media (42). In this program each D.C. government agency receives a monthly grade based on the data, as well as daily reports that are sent to agency managers. The monthly grades have seen a positive trend since the beginning. At the start, D.C. saw 4 of 5 agencies with C- or C+ grades. Now the program has expanded to 15 agencies with nine A’s, two C’s, and one D that only had three reviews (42).

In another example, the CitySourced platform allows citizens to report non-emergency issues such as public works issues back to a central system. Citizens can send in audio, video, pictures, or text of any issues they find. The city can collect these reports, send them to the correct departments and respond to the citizens and keep them updated on their progress to fix the issue.

In Emergency Management, crowdsourcing is becoming a very effective tool. Several recent large natural disasters have caused major destruction and caused large emergency responses. The Haiti earthquake and Superstorm Sandy disaster responses both used crowdsourced data to more efficiency respond to the needs of the people. After the earthquake in Haiti, “640 volunteer mappers from OpenStreetMap traced high-resolution satellite imagery released by Digital Globe and GeoEye. They made 1.2 million map edits.” The volunteers “built a free and open atlas of roads and critical infrastructure that is among the most detailed in the world.” The United Nations used this map data to build their maps used for the missions in Haiti. Also in Haiti, 90,000 text messages were collected and geocoded and made available to rescue teams, to provide them with information straight from the citizens of Haiti (33).

2.6 Implementation Considerations

Several issues have been identified that would impact how to plan and implement projects that enable agencies to fully utilize crowdsourced data. There is debate over how to best incentivize the crowd to contribute; at the same time, if a crowd of roadway users is actively contributing information, there can be issues with driver distraction and safety. Additionally, if and when crowdsourced data becomes more widely available, how can transportation agencies efficiently parse data for what is most relevant and merge together data from multiple sources? Finally, privacy concerns over crowdsourced data, especially locational data and connected vehicle data, remain a popular debate topic.

2.6.1 Gamification / Incentivizing the Crowd

One way to incentivize participation in crowdsourcing is via a concept called gamification. Gamification is applying the principles of game design to non-game related applications, primarily to make them more attractive and engaging to users. Applying gaming concepts to non-related applications such as transportation is a way to engage users and motivate them to stay actively involved. Keeping users involved is a key aspect of gamification: if the game is well designed, it will keep more users active for a longer amount of time, allowing the game to satisfy the users' needs as well as other goals the application may have, such as increased transit use, or greener travel. Gamification has already been applied successfully in the transportation industry as well as many others, borrowing elements from gaming such as scoring, peer interaction, and immediate feedback, without appearing to be an actual game (e.g., Waze ranking system).

Waze is the perhaps the most active participant utilizing gamification in the transportation industry. Waze relies on crowdsourcing data from all of its users and encourages users to report congestion, incidents, disabled vehicles, closed roads, and other conditions or situations drivers may find on the roadway. Each time a user submits data, or completes a challenge, they gain points and move up in the Waze rankings. Collecting more points allows you to unlock different features to reward members for continued use of the application. Active participation allows a more detailed level of data collection; for the Waze application this may be road debris, disabled vehicles, traffic incidents, and more. This data can be very valuable for TMC operations. This real-time data from the roadway can greatly increase an agency's ability to collect and disseminate more data. TMC operators have already begun to see this as an important source of data [GA NaviGAtor 511].

Gamification can help provide effective incentives to change behavior, which is a difficult task for a transportation agency. For example, the City of Enschede in the Netherlands has implemented an application to encourage and incentivize participants, with real world rewards, to take alternative transportation methods such as transit, biking, or walking. The application platform is called Sustainable Social Networking Services for Transport (SUNSET) (37).The application is actively seeking to change travelers' behavior to solve congestion and environmental issues in the city (24). The more a user chooses to take public transportation or ride a bike to work, the more points they earn. A user can get bonus points for activities such as biking to work when it's raining or cold; the points add up to various prizes such as having a transit route named after the user for a period of time (24). Not only is the application helping users find the most efficient and effective way to work, it is helping to change long term behavior by getting more people using alternative transportation choices and solving the larger issues in the city of congestion and emissions. Transportation agencies can use these types of applications to target specific audiences and customers, find the appropriate incentive, and help to change their travel behavior in both the short term and long term.

Metropia applies various active gamification strategies in engaging users. The gamification includes providing a user's total points earned, and total travel time saved. Further, Metropia also partners with American Forest (54, 55) to graphically show the CO2 savings and Metropia notifies the user that Metropia has planted a tree for the user when the CO2 savings reaches every multiple of 100lb. Leader boards of these categories are posted online and the user can share such info on social media.

When Metropia is fully integrated in a TMC environment, additional gamification and user engagement capabilities are enabled. A partnering agency can send out the "agency alert" by various geo-fenced criteria to Metropia users to alert them of various upcoming events such as work zones, weather, incidents, etc. The geofenced filtering capabilities allow the TMC operator to deliver alerts only to relevant users at the right time. Metropia also allows a TMC operator to trigger routing for users based on certain criteria during an extraordinary situation like an incident. For example, Metropia presently triggers an "enroute alert" for a user to re-route if the user's experienced travel time becomes longer than the pre-planned travel time by a certain pre-defined threshold due to unexpected events. This re-routing can be triggered by a TMC operator if there is a need for diverting users away from a hazardous area.

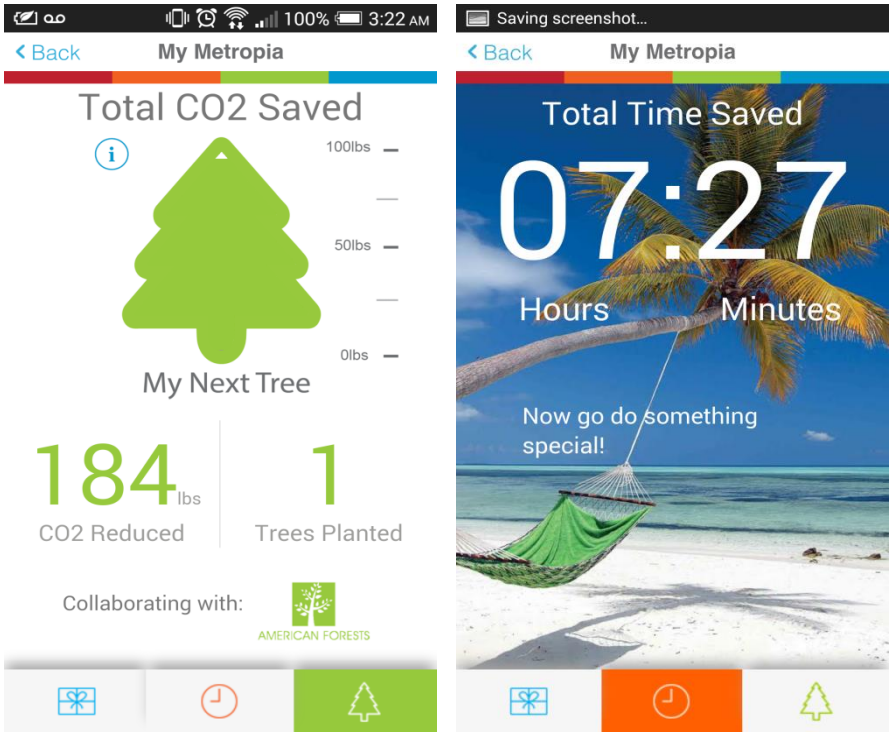


Figure 3: Metropia Mobile App
(*Source: Metropia*)

TMC operators can only do so much with roadside sensors, cameras, service patrols, and other technology. Gamification takes this non-gaming context and allows people to become involved in a meaningful way – whether contributing data to support operations or providing agencies with effective mechanisms for encouraging individuals to change their traveler behavior in the short and long term in order to reach public policy objectives.

2.6.2 Driver/Traveler Distraction

Incentivizing a crowd of active participants to contribute data about roadway conditions necessarily invites the issue of traveler safety. Many agencies may be hesitant to engage in activities that encourage mobile phone usage while driving, such as real-time incident dissemination (23). One agency interviewed by AASHTO stated, "Because people are so reliant on smart phones, we feel somewhat conflicted in providing our information via that platform. On one hand, we are communicating in the manner that increasing numbers of people prefer. But we're doing it in a way that can be a cause in distracted driving crashes" (23). The growing number of in-vehicle mobile applications can be distracting to drivers; there are new applications applying gamification to safe driving to combat this. One specific example involves an application called Road Wars which allows teenage drivers to "take over" neighborhoods where they are the "safest" driver; activities such as picking up the phone (detected from the phone's accelerometer), speeding, opening a text message, or taking a call can all lead to negative points in the game (39). The FHWA's report on technology advancements for TMCs advises that information provided through social media or mobile apps should currently be focused on pre-trip planning to minimize driver distraction for the near-term; however, in the long-term, as voice activation becomes more common, en-route social media usage could be less distracting (9). It should be noted that recent studies have cast doubt on whether voice activation really reduces driver distraction. A recent study sponsored by AAA on the mental workload of common voice-based vehicle interactions concluded that "common voice tasks are generally more demanding than natural conversations, listening to the radio, or listening to a book on tape." On a positive note, the additional cognitive load was reduced for well-designed systems (38).

2.6.3 Data Fusion and Reliability

Data fusion refers to the ability of state agencies to integrate and manage crowdsourced data alongside other data that is used to support TMC operations. Crowdsourced data that is collected by agencies or acquired from third-party sources needs to be incorporated into the operational framework. There are two primary issues of note: the tactical approach to process and aggregate the data, and the concern over data quality, resulting in the need to verify and validate crowdsourced data.

Currently, most state DOTs do not have an established methodology to technically combine the contributions from crowdsourced data. While some states have developed tools to capture the crowdsourced data, the majority of states use visual inspection to associate crowdsourced entries published by users around the same location or area. This is challenging primarily because the data is unstructured. Contributors do not provide precise location data, spelling may be incorrect and creative abbreviations and grammar are used when confined to 140 characters (5). Various solutions have been proposed for this problem, but these methods fall short when the task of collecting information has to be performed continuously and in real-time, by an always changing crowd (11). In terms of assessing the quality of citizen reporting, as mentioned, both the Utah DOT Citizen Reporting program and the new Wyoming DOT Enhanced Citizen-Assisted Reporting system require trained volunteers (1, 10); Utah's system consists of an online training session and quiz. The FHWA recommends that state DOTs develop pre-qualifications or standards for data providers, including items such as data format, coverage, and latency (9).

For data purchased from a vendor, data verification and validation is an important challenge. Best practices dictate that TMC operators should visually validate any event reported by a third-party. Law enforcement and other business partners are often accepted as "trusted partners" that agencies rely on to help validate field reports (9). Agencies need the ability to develop the same trusted relationships with their vendors to obviate verifying or validating events from crowdsourced data. With smartphone applications, coverage issues can arise, especially when a small subset of the driving population is

providing data and roadway conditions can change suddenly (3, 15); INRIX claims that their reported traffic speeds are within 5 mph of the actual speed 98 percent of the time (15). Another study on passive traffic applications accuracy reported that Google "accurately reported" 52% of traffic jams, while INRIX achieved this 38% of the time (29). Accuracy of travel time data should continue to improve as more data becomes available and more complex aggregation algorithms are developed, but the issue of comparing data across sources remains, especially if a TMC operator or software application must make the "final call".

2.6.4 Data Privacy and Ownership

Privacy is a major concern among many users of social media and crowdsourced mobile applications, especially as the amount of information being exchanged rapidly increases. Protecting the privacy associated with data captured through crowdsourcing is an important consideration for state agencies that intend to mine social media and/or deploy applications to directly collect crowdsourced data. From the agency perspective, policies that clearly articulate how data will be handled and protected to prevent the misuse of personal data need to be put into place. It is recommended that TMC managers understand privacy issues and be able to both manage data properly and develop contracts that protect the privacy of collected data, regardless of whether that data is collected through private vendors or by mobile communications (9). It is further advised that all TMC personnel be restricted from accessing personal data records unless there is a compelling reason. Established privacy protocols would require a legal agreement to be drafted between the agency and necessary personnel. Researchers at Rensselaer Polytechnic Institute suggest that at present, social media is not used universally as an official channel of communication, especially by the government (4). There is still no clear consensus among transportation managers on how social media should be utilized to collect and disseminate warnings to the public. Guidelines are warranted to dictate the collection and use of information. Agencies will need to transparently demonstrate how information is acquired and used for TMC operations, allow for an opt-in or opt-out, define a privacy policy, and emphasize how the real-time information they are gathering is utilized to improve system operations. (9)

3 Assessing Current Practices and Perceptions

To assess the potential for using social media and crowdsourced data for TMC operations, TMC staff and industry experts were engaged through an on-line survey, phone interviews, and an in-person workshop. This chapter discusses the various data acquisition methods and summarizes the key findings.

To understand the potential for using crowdsourced data for TMC operations, the data acquisition process focused on exploring answers to the following questions:

- How do state agencies currently use real-time data to support mission critical TSM&O activities and what are the perceived gaps in the current state of available data that would create suitable opportunities for integrating crowdsourced data?
- What is the perception within state agencies as it relates to social media and crowdsourcing?
- Are state agencies currently using social media and/or crowdsourcing (both inside and outside of the TMC)?
- Do state agencies have experience planning and executing technology projects such as deploying apps, disseminating information via social media, or integrating data from external systems?
- What are the perceived challenges and threats to acquiring and integrating crowdsourced data into a TMC?

3.1 On-Line Survey

Based on the limited findings of the literature review, a short on-line survey was prepared and distributed to 88 members selected from the TMC PFS and the national ITS deployment survey located in large metro, small urban and rural areas. A copy of the online survey is presented in Appendix C.

The survey was designed to investigate a small number of critical issues as it relates to crowdsourced data and TMC operations including:

- Assessment of the most relevant TSM&O activities within a TMC and the most common data sources to support these activities
- Stated challenges with current data sources
- Perception of crowdsourced data and motivation to support TSM&O activities
- Status of crowdsourced data use by agencies operating TMCs
- Presence on social media
- Responses were received from 32 participants representing 21 states.

3.1.1 Survey Results

The survey results confirm a strong interest in leveraging crowdsourced data for TMC operations even though there is currently minimal use of this data. There are strong indications that TMC staff perceive limitations in current data sources used to support TSM&O activities and believe that crowdsourced data can supplement existing data and support real-time operations.

3.1.1.1 Current TMC Activities and Data Sources

Figure 4 and Figure 5 illustrate the criticality of TSM&O activities along with the current status of data sources used to support these activities. Figure 4 shows both the incremental and total number of survey responses at different levels of priority. For example, 26 respondents indicated that emergency management was a high priority for their agency, five respondents indicated emergency management to be medium priority, and one participant said not applicable, for a total of 32 responses. The top three ranked TSM&O activities that have the most prevalence or have the highest importance are Freeway Operations and Management, Incident and Emergency Management, and Traveler Information. Work Zone Management, Road Weather Management, and Maintenance Management also were considered to be important functions, with many respondents rating them as having either high priority or medium priority within their operational setting.

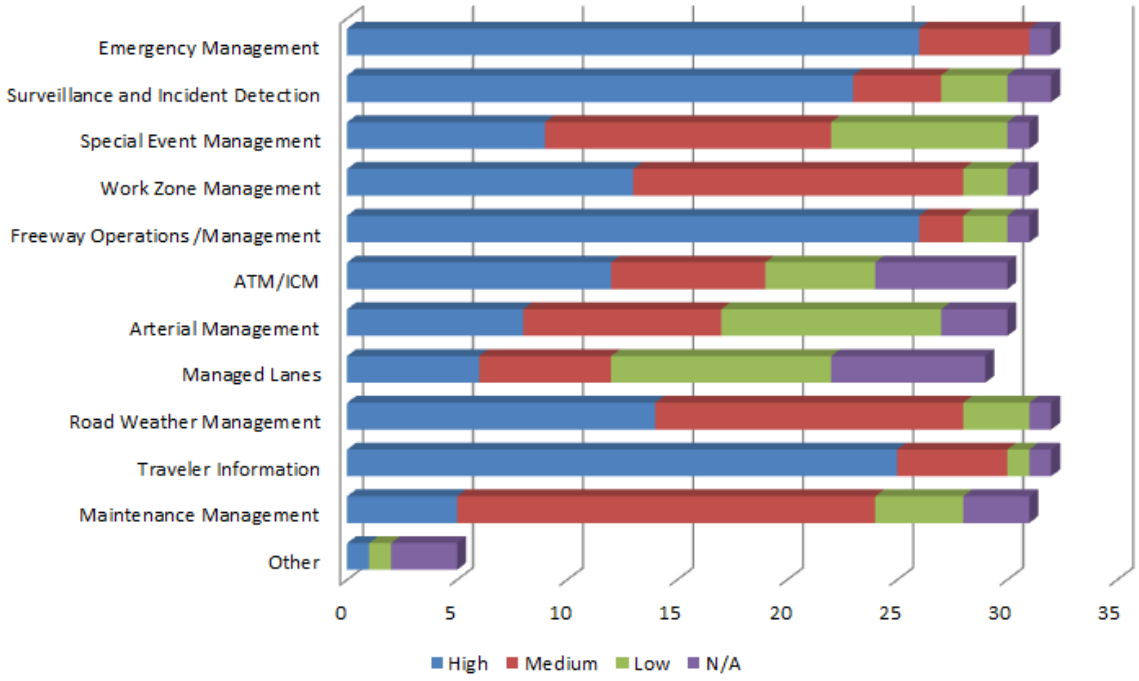


Figure 4: Priority of TSM&O Activities (32 respondents)
(created by Open Roads)

Further analysis revealed that for surveyed agencies indicating that they use crowdsourced data, the priority shifted where Incident Management was placed as the area with the highest priority, followed by Traveler Information, and then Freeway Operations and Management. This could be attributed to the fact that both Incident Management and Traveler Information impact a wider range of travelers in a more direct way than with Freeway Operations and Management.

TSM&O activities are supported by data acquired from multiple sources. Data is acquired by manually data entry by agency personnel (Manual), deployed ITS field devices (Self-Generated), partner agencies (Shared), directly or indirectly from the general public (Crowdsourced), and private sector or third-party vendor (Acquired). Data acquired manually and/or self-generated is most often used to support TSM&O activities. The use of crowdsourced data was reported across most of the TSM&O activities, with anywhere between 2 and 8 respondents indicating this to be the case.

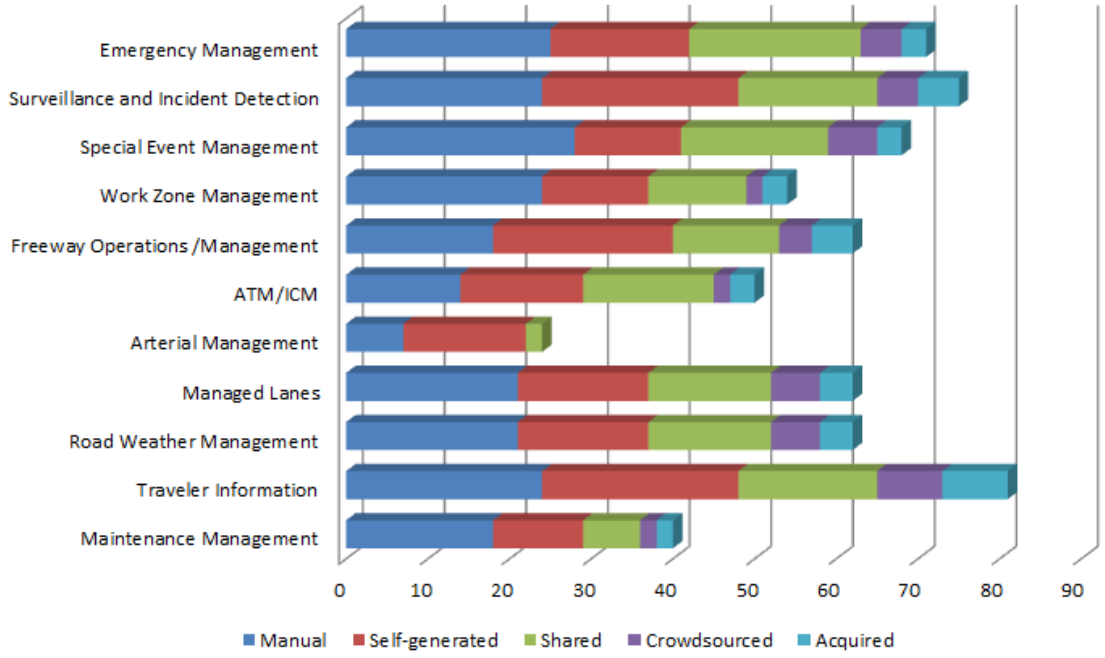
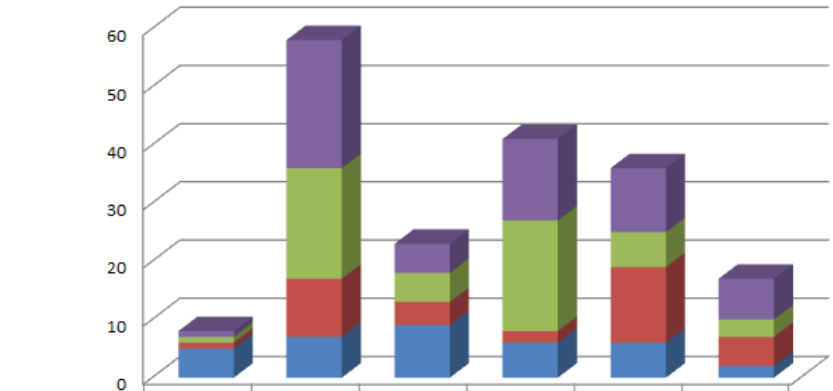


Figure 5: Current Data Sources for TSM&O Activities (32 respondents)
(created by Open Roads)

3.1.1.2 Challenges with Current Data Sources

Figure 6 illustrates the perceived challenges with current data sources. The survey indicated that coverage area, cost, and effort to integrate current data sources are the most challenging for state agencies. State roadway networks are expansive and it is very costly for state agencies to provide sufficient deployment of ITS devices to cover the entire network. It is also difficult to provide extensive staffing to manually monitor the entire network. Today, most agencies are focused on the most critical freeway and arterial segments.

The survey also indicated that when state agencies attempt to fuse and integrate data from other parties, either by acquiring data from partner agencies (shared data) or purchasing data from third-party providers, the level of effort to integrate the data is challenging. Some states continue to shy away from acquiring third-party data due to concerns over data quality.



	No Barriers	Limited Coverage	Poor Quality	Too Costly	Tough Integration	Other
Manual Data	1	22	5	14	11	7
Self-Generated	1	19	5	19	6	3
Shared Data	1	10	4	2	13	5
Acquired Data	5	7	9	6	6	2

Figure 6: Perceived Issues with Current Data
(created by Open Roads)

3.1.1.3 Perception of Crowdsourced Data

There was a strong correlation between the stated challenges of current data sources and the perception of crowdsourced data to compensate for these challenges. Figure 7 shows that the respondents perceive that crowdsourced data would provide significant benefits to overcome the cost and coverage issues with current data sources that are being experienced by state agencies.

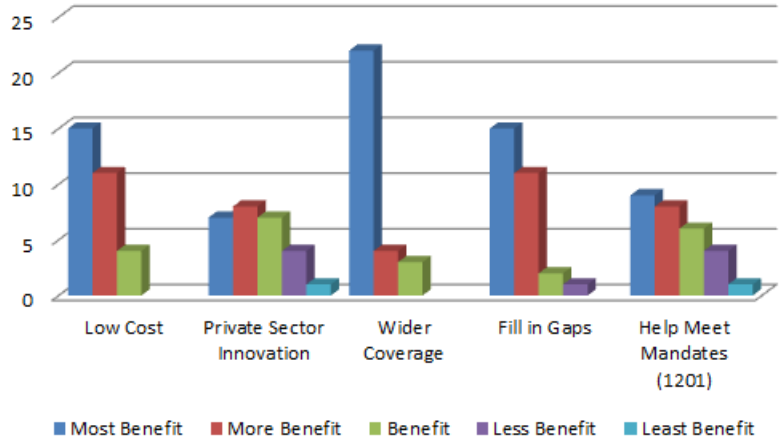


Figure 7: Perception of Data from Crowdsourcing and Social Media Data
(created by Open Roads)

Figure 8 illustrates the perceived threats to using crowdsourced data. TMC operations are increasingly data driven. Many agencies rigorously validate information prior to making it available to the public. Agencies are most concerned with the accuracy and credibility of crowdsourced data. The survey also indicated that agencies are most concerned with the challenges to offer in-house support to build and maintain mobile applications. In addition, state agencies are concerned with distracted driving with

over half the respondents stating this was of most or more concern to them. Interestingly, data privacy did not seem to be a primary concern to most respondents.

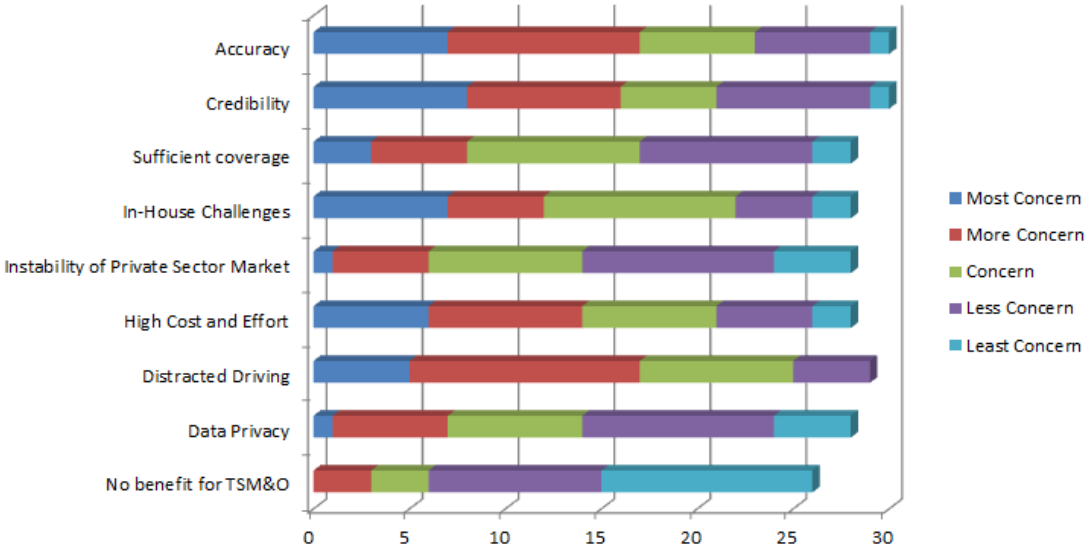


Figure 8: Perceived Threats to Using Crowdsourced Data
(created by Open Roads)

3.1.1.4 Status of Agencies Currently Using Crowdsourced Data

Over half of the respondents do not currently use or collect crowdsourced data (54.8%). State agencies motivations to use crowdsourced data is shown in Figure 9. Of the state agencies collecting crowdsourced data, the primary motivations are to explore new technologies, to help with insufficiencies with existing data, and to investigate cost-effective alternatives to acquiring and disseminating information to the traveling public. Figure 10 illustrates the current usage of crowdsourced data. The majority of these agencies collect crowdsourced data by the use of their website and by purchasing from the private sector.

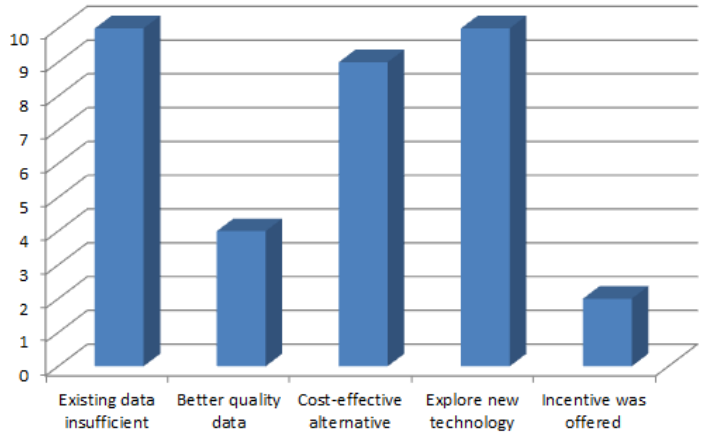


Figure 9: Agencies Motivations to Use Crowdsourced Data
(created by Open Roads)

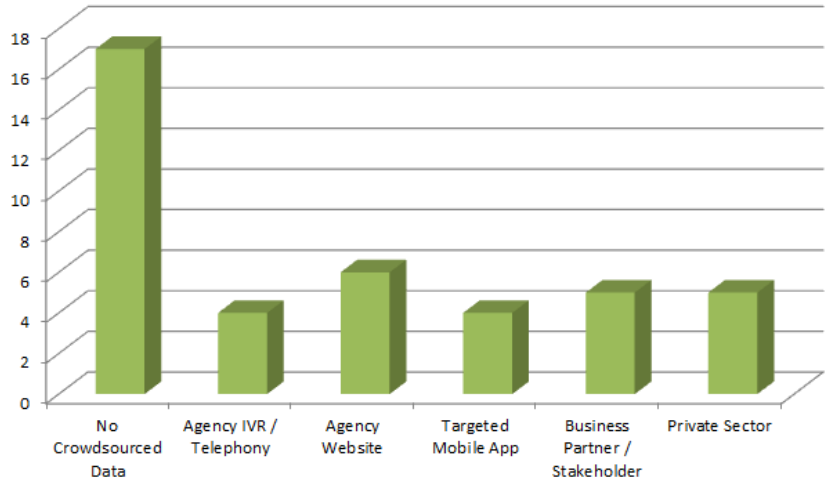


Figure 10: Current Usage of Crowdsourced Data
(created by Open Roads)

3.1.1.5 Presence on Social Media Platforms

Today, most state agencies have a presence on social media platforms. Most state agencies are currently using social media platforms to disseminate real-time traveler information (80%) and general information on projects, events, and initiatives (83%). Only 30% of the state agencies monitor social media content at all times. The remaining 70% is either monitoring during business hours, occasionally, or on a routine schedule.

Three quarters of the state agencies monitoring social media indicated that it is purely a function of public information or communications staff. Only 6% indicated that it is a function of the operations staff only. The remaining agencies indicated that the responsibilities are shared between operations and communications staff.

3.1.2 Key Findings from the survey

The survey provided a strong baseline for understanding TSM&O priorities, effectiveness of current data sources, and perceptions toward crowdsourced data. The key findings of the online survey include the following:

- The priority of TSM&O activities for surveyed agencies indicating that they use crowdsourced data is Incident Management, Traveler Information, and Freeway Operations and Management.
- Crowdsourcing is perceived as a cost-efficient approach to collecting and sharing data and has potential for broader geographic coverage than traditional methods.
- The most cited impediments to the use of crowdsourced data to support TMC operations are: (1) validity and reliability of data; (2) the challenges to offer in-house support to build and maintain mobile applications to collect data; and (3) the effort to fuse and integrate collected data with current data sources.
- Many agencies use existing infrastructure (e.g., 511 phone system and websites) and third-party involvement to collect crowdsourced data.

- There is currently a strong use of social media for disseminating traveler information.

3.2 Phone Interviews with TMC Staff

Following the online survey, phone interviews were conducted to verify the results from the online survey and to provide a more detailed view of the respondents experience with crowdsourced data and their plans to apply crowdsourced data to their TMC operations. Eight survey respondents were selected for follow-up phone interviews, using the following criteria: (1) known history with crowdsourced data; (2) interesting responses to online survey; (3) willingness to participate in a brief phone interview. Informal interviews were conducted by telephone with two project team members and a respondent from eight agencies. These included Washington State DOT, Wyoming DOT, Idaho Transportation Department, Delaware DOT, Georgia DOT, Utah DOT, North Carolina DOT, and Maricopa County DOT.

3.2.1 Findings

The findings of the phone interviews identified emerging crowdsourcing trends that are currently being implemented by the state agencies. It also identified the evolving role of social media monitoring and the current obstacles in the use of crowdsourced data in TMC operations.

3.2.1.1 Emerging Crowdsourcing Trends

Dissemination of traveler information through social media platforms such as Twitter and Facebook is widely used among the interviewed agencies. All participants primarily use Twitter as a means to acquire and disseminate real-time traveler information, and Facebook to disseminate general information on projects, events, and initiatives. This coincides with the survey results of where Incident Management was placed as the area with the highest priority, followed by Traveler Information among agencies using crowdsourced data. Incorporating automated tools to disperse information out to the public while relieving the load on the TMC staff was one trend identified during the phone interviews. Half of the surveyed participants stated that they have implemented or are moving towards automating distribution of information through the use of social media platforms.

Other crowdsourcing trends within the state agencies identified during the phone interviews are illustrated in Figure 11. One trend identified was the experimentation with existing infrastructure to collect traveler information by leveraging their 511 phone systems and website to acquire crowdsourced data. Another trend identified was passively collecting crowdsourced data from travelers. There is a large involvement of private sector and third-party data collection used for traveler information strategies within the TMCs (e.g., probe data and detectors).

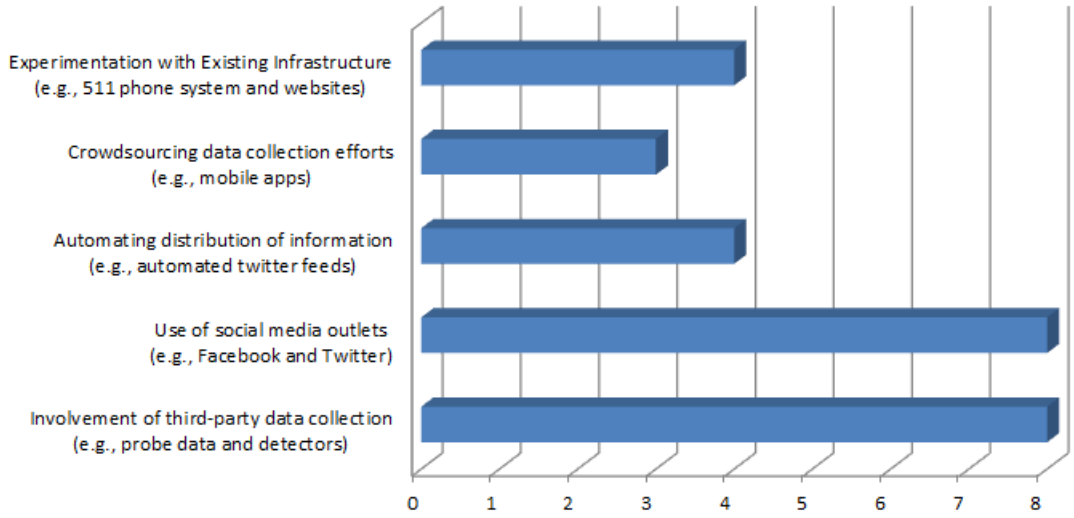


Figure 11: Emerging Crowdsourcing Trends (number of agencies identifying these trends out of the 8 interviewed by phone) ***(created by Open Roads)***

Fifty percent of online surveyed respondents identified crowdsourced data as being useful for road condition reporting. Data collection efforts by some state agencies where road conditions and weather reporting are of high importance were identified during the phone interview process. One state agency has added real-time incident and weather conditions reporting to their agency 511 website. A mobile citizen reporting app that allows trained volunteers to report on weather and road conditions was developed by another state agency. A third state agency created an enhanced citizen-assisted reporting program that enables trained volunteers to call in road condition reports due to inclement weather and other hazardous conditions, such as roadway debris and large potholes.

3.2.1.2 Evolving Role of Social Media Monitoring

Social media monitoring and publishing responsibilities are typically handled by the Public Information or Communications staff. Figure 12 shows the percentage representation of the staff responsible for monitoring social media within the respondent's agency. Results from the phone interviews indicate that this role is evolving due to the real-time publishing to social media. This role that generally begins as a communications staff task is morphing to either a communications staff residing within the TMC or a fully operationalized publishing role. Some state agencies have a Communications representative physically in the TMC working together with the operators while others are transitioning the communications responsibilities over to operations. One state agency stated that they were in the process of trying to fill that role with someone with social media expertise rather than media and communications experience.

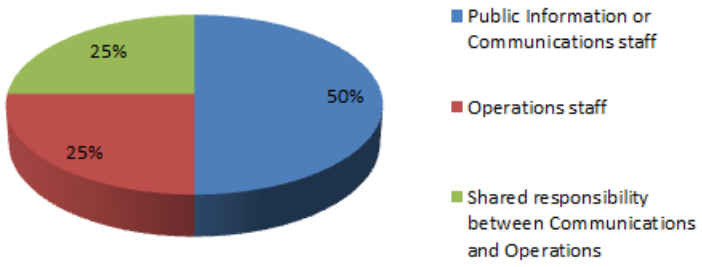


Figure 12: Responsibility for Monitoring Social Media
(created by Open Roads)

3.2.1.3 Crowdsourcing Obstacles

The online survey identified that over fifty percent of the state agencies thought it was most useful for incorporating crowdsourced data to support Surveillance and Incident Detection operations. However, most phone interview participants whose state agencies were using some form of crowdsourced data stated that incorporating new incident detection methods involving crowdsourced data depends on the validity and reliability of the data. Most state agencies do not have a motivation to change how incidents are detected because they believe their current process and acquired data (e.g., Computer-Aided Dispatch Resources) is sufficient.

Other obstacles in the use of crowdsourced data identified during the phone interviews included the fusion and processing of data into the current infrastructure, the meeting of state agencies safety and privacy objectives, identifying market strategies to encourage more citizen involvement, and keeping pace with the rapidly moving private sector.

3.3 Industry Expert Phone Interviews

In parallel to the data collection from TMC staff, a series of phone interviews with selected industry experts from academia, non-TMC public sector staff, and the private sector was conducted. The purpose of these phone interviews was twofold:

1. Understand the perception of industry experts as to the current state of the practice for using social media and crowdsourced data to support real-time operations.

2. Explore best practices outside of TMC operations for using social media and crowdsourced data in a command and control environment to support real-time decision making.

The list of participants included:

- *Sarah Kaufman*, Adjunct Assistant Professor of Planning and Digital Manager at the NYU Rudin Center for Transportation. Ms. Kaufman previously worked at Metropolitan Transportation Authority where she led the open data program and assisted in the development of their social media program.
- *Lea Shanley*, Presidential Innovation Fellow at NASA working on Crowdsourcing to Improve Government. Ms. Shanley is chair of the Federal Community of Practice on Crowdsourcing

and Citizen Science and former Director of the Commons Lab of the Science and Technology Innovation Program at the Wilson Center. The Commons Lab (www.wilsoncommonslab.com) has advanced research and policy analysis on the use of social media and crowdsourcing.

- *Di-Ann Eisnor*, VP Platform & Partnership at Waze (Google). Ms. Eisnor is currently spearheading Google's Connected Citizen program, a program to exchange data with public sector agencies.
- *Mike Haas*, Director of Business Development, Castle Rock Consultants. Castle Rock partners with several state DOTs to deliver 511/traveler information services and is currently deploying apps that collect crowdsourced data.
- *Dr. William Wallace*, Rensselaer Polytechnic Institute. Dr. Wallace is a Principal Investigator on a U.S.DOT Research and Innovative Technology Administration (RITA) project to explore the role of Social Media to improve the efficiency of traffic operations during non-routine events.
- *Dr. Qing He*, University of Buffalo Transportation Informatics University Transportation Center. Dr. He has been working on mining transportation information from social media for planned and unplanned events (http://www.buffalo.edu/transinfo.html).

Discussions with these experts confirmed many of the findings of the survey and the literature review. The main points from these discussions are as follows:

- Today there is limited use of social media or data mined from social media for tactical operations today – as social media usage expands and new tools are generated this could increase.
- There is still no clear consensus among transportation managers on how social media can be utilized to collect and disseminate warnings to the public.
- Context and content are important considerations to understanding the behavior of people using social media.
- The responsibility for monitoring social media at government agencies tend to fall under an "information officer" who is housed in a group outside of the operations group. This results in a disconnect between the information staff and the operations staff.
- There is limited evidence of successful use of gamification by state or federal agencies.
- An agency's fear of public humiliation is often the driving motivation for social media use.
- Many agencies really do not know what percentages of their customers are engaged in social media.
- Growing use of geotagging is making it more possible to extract usable data from social media.
- The private sector is better equipped to acquire and process crowdsourced data on a large scale. There should be growing opportunities for the private sector to support agency activities.
- Difficult to measure and assess the impact of social media and crowdsourcing on operations.

3.4 Panel Meeting

A two-day panel meeting was convened to assess the current state and future potential for use of crowdsourced data in a transportation operations environment. The meeting was held on July 23 and 24, 2014 in Reston, Virginia. Attendance was by invitation only with invitees compiled based on input from the FHWA and the project team. Invitees were highly qualified and experienced professionals drawn from one of the following categories:

- Public sector operational leaders with experience using cutting edge technology
- Private sector vendors with a proven track record of quality products
- Academics with an established research record related to crowdsourced data for transportation operations

The list of participants in the panel meeting and the complete agenda are provided in the report Appendices.

The panel meeting allowed for detailed discussion and further investigation of points and trends identified in previous stages of the project. The goals of the panel meeting included:

- Provide a forum for peer exchange and review amongst experts with various backgrounds relevant to the project objectives
- Analyze and validate the SWOT model
- Elaborate on operational, institutional, and technical considerations that would impact a state agency's ability to successfully undertake crowdsourcing projects
- Formulate Measures of Effectiveness to help state agencies evaluate the value of undertaking crowdsourced projects

The following sections provide key finding from the information sharing sessions.

3.4.1 Public Sector Round Table Discussion

During the first session of the workshop, each public sector representative provided an overview of current activities and future plans for using social media and crowdsourced data. The following trends were observed:

- Citizen reporting on weather road conditions is being collected and integrated with 511 and operations by two agencies by use of a mobile app and a state agency's 511 website.
- Some state agencies expressed interest in creating apps to report incidents and work zones.
- There is a strong use of Facebook and Twitter for disseminating traveler information to the public. Most state agencies have multiple social media accounts (e.g., statewide, regional, interstates). Information is primarily being "pushed" out on social media than "pulling" information from the public. This is due to limited resources available to monitor all accounts.
- Social media platforms are primarily monitored and maintained by Communications staff during business hours. Some state agencies have social media monitoring by TMC staff outside of business hours.
- State agencies are leveraging 511 phone systems to collect traveler information.

- Automated feeds have been implemented from some state agencies 511 systems to Twitter.
- There is an interest in some state agencies to standard hashtags to identify tweets associated with particular regions or corridors. Once established the public can use them to direct specific information both to the state DOT and other interested citizens.
- Florida DOT signed an agreement with Waze to share information to enhance their offerings to the public. They will have access to the Waze data for use in managing traffic and assisting in the identification of events on non-covered roads.
- Radian6, a program that brings all their social media feeds into a centralized platform, provides an easy way for Iowa DOT staff to track, monitor, and respond over social media, identify trends, and cross-tab different public responses.
- Dlvr.It, a complete content sharing solution, is used by Iowa DOT to automatically link their Twitter and Facebook feeds.
- YouTube is used by most state agencies for educating on state DOT activities and general information sharing.

3.4.2 Presentations by Industry Experts

Three representatives from the private sector and academia presented their perspective on using social media applications to influence traffic operations:

- *Susan Grant-Muller (University of Leeds)*: Dr. Muller, a leading researcher in crowdsourcing traffic data, discussed the state of the practice.
- *David Zaharchuk (IBM)*: Mr. Zaharchuk discussed opportunities and challenges that government agencies face to leverage big data and the role of crowdsourcing.
- *Dr. Yi-Chang Chiu (University of Arizona)*: Dr. Chiu discussed Metropia (www.metropia.com) and his vision for revolutionizing how drivers and transportation agencies rethink traffic congestion and work together to solve the rising traffic congestion problem.

Key findings from this session include:

- Incentivizing engagement influences traffic operations. Based on the research conducted in the SUNSET project, applications that use gamification methods of recognition and rewards influence participant travel choices. Early results indicate that a significant segment of the public is willing to share profile information and are responsive to reward systems.
- Crowdsourced mobile apps can be used to achieve Active Demand Management strategies. Users receive rewards (personal and altruistic) if they choose routes and departure times which help to optimize system performance. An analysis of a real world case study in Colorado indicated that even modest response rates to these kinds of incentivizes produces flow improvements comparable to those associated with major infrastructure capacity expansion projects.
- Social media data have little or no inherent context. Without context, the potential value of the collected social media data may not be fully realized. Data mining methods to dynamically construct contextual information can be used to interpret social media data.
- There are four key activities to detecting complex patterns in data: (1) Collect all relevant data from a variety of sources; (2) Turn information into knowledge by drawing correlations

between the data; (3) Analyze data in context to help form decisions and actions to be performed on the data; and (4) Use context to deliver insights for supporting intelligent systems or human operations.

3.4.3 Private Sector Crowdsourced Data Provider Discussion

Representatives from three of the leading private sector companies (INRIX, HERE, TomTom) who specialize in aggregating crowdsourced data that support TMC operations presented their perspectives on the current state of the marketplace and their thoughts on future innovations. Key findings from this session included:

- There are several advantages to relying on third-party data providers for processing and aggregating crowdsourced data. These companies are experts in processing and fusing large amounts of data. They also sanitize the data, removing any privacy issues that agencies might be concerned about.
- Private companies have dedicated resources to acquire data from multiple sources (e.g., probe vehicles, crowdsourced apps, state data feeds). This economy of scale enables them to perform complex data analytics and offer a wide range of data services to the TMC marketplace.
- With a world-wide footprint, private sector data providers have better capacity and incentive to create a variety of data products and services for the marketplace.
    - All of the data providers are expanding into coverage for secondary roadways and arterials.
    - Each company is developing proprietary tools for traffic monitoring, queue detection, and incident detection.
- Competition across companies fosters innovation and the need to maximize data quality.
- The industry would benefit from the development of and adherence to data standards.
- Private sector is better equipped to acquire and fuse new data sources, like connected vehicles.
- Private sector data is better suited for selected TSM&O activities related to traffic management. These companies currently do not offer data to support asset management or road condition reporting.

3.5 Summary

This assessment suggests that traffic managers perceive real value in collecting crowdsourced data and applying it to support TSM&O activities. Agencies operating TMCs have started experimenting with their existing infrastructure to collect traveler information and developing new data collection efforts (e.g., citizen reporting apps to report on road and weather conditions). Public agencies should be aware that third-party data collection for speed and travel time reporting is now widely available as an option to consider. Facebook and Twitter are widely used by transportation agencies to disseminate traveler information for planned and unplanned events or incidents.

The expanding usage of the traveling public in social media creates opportunities for agencies to monitor social media and extract data to support TSM&O activities such as Incident Detection, Active

Traffic Management, and Road and Weather Condition Reporting. However, mining social media for real-time operations is difficult and requires additional resources to help add context to the collected data. Private sector firms can be leveraged to perform data aggregation and fusion of this context data in state agencies lacking the resources and expertise in data analytics and big data that may be required to execute these new data integration projects.

Traffic management agencies also see value in collecting crowdsourced data using websites and specialized apps. The use of specialized apps is considered a cost-efficient approach to collecting and sharing information, influencing public behavior, and broadening geographic coverage. Most agencies lack the resources and the knowledge required to develop an effective mobile app. The agencies that have data collection efforts being developed have partnered with private sectors and third-party providers that have the experience for crowdsourcing app development. Partnering with private sectors and third-party providers also eliminates agencies issues relating to the privacy and the legal challenges involved with deploying crowdsourcing apps and actively collecting data. It is important to note that incentivizing the crowd using gamification methods within the specialized apps can encourage public engagement.

Agencies continue to express concerns with the reliability and quality of crowdsourced data. Agencies operating TMCs traditionally employ strict data verification measures, and that level of trust with crowdsourced data has not yet been achieved. Agencies operating TMCs would like to encourage and empower citizen reporting. More effort needs to be placed on how citizens are recruited, trained, and incentivized to improve confidence in the resulting data.

4 Case Studies

4.1 Overview

This section contains four case studies that represent recent and salient work in each of the major categories of crowdsourced data applications for transportation operations. The categories and case studies are:

- Extracting data from social media platforms
  - Iowa DOT has an advanced social media monitoring and engagement program (section 4.2).
  - District of Columbia DOT uses sophisticated data mining techniques to extract real-time incident information from social media platforms (section 4.4).
- Developing and deploying specialized mobile applications
  - Utah DOT has developed and deployed a mobile application for reporting road conditions (section 4.3).
- Acquiring crowdsourced data from a private sector data provider/aggregator
  - Florida DOT has integrated Waze incident data directly into their operational systems (section 4.5).

4.2 Iowa DOT Social Media Program

Iowa DOT has a strong social media engagement program. They monitor and respond to social media outlets around the clock directly from the Transportation Operations Center (TOC). The program grew out of their early experience with automatically populated Twitter and Facebook feeds. Their traveler information system automatically posted traffic information in real-time, but the feeds were not monitored by DOT so there was no dialogue.

A severe winter storm struck Iowa in December 2012, and it changed the way Iowa DOT managed social media. In the after-action reviews of the storm, some DOT staff members noted that there was a great deal of useful dialogue about the storm in the social media universe as the storm happened, and that DOT's automated feeds really failed to tap into that effectively. They began to actively monitor and respond to social media during extended business hours from the communications office. Based on the program's success, DOT decided to extend the hours and move the program into the TOC to get even more immediate interaction with other operational processes. Social media training for TOC operators is in the process of being developed as of this writing.

In addition to the purely operational benefits of the program, Iowa DOT sees value in the way that the program humanizes the DOT to the general public. Several examples follow which show the kind of positive interaction between citizens and the DOT which is common under this program. Figure 13 shows an example of information that the DOT posts to Twitter. This case includes an image and map to provide context. A member of the general public has replied to the tweet expressing his appreciation for the information. Note that in slightly more than ten minutes this message has been retweeted

(forwarded to additional information streams) fourteen times. Retweets are a strong indication that consumers see value in the information. If the consumer did not see value, they would not dilute their own data stream with the information.

Figure 13: Iowa DOT Twitter Post
(Source: Iowa DOT)

In Figure 14 a Twitter user has directed a tweet at the DOT by using the DOT identifier @iowadot. The DOT presence in the social media world made her aware of that outlet and she has made use of it. The interaction is positive and satisfying to the member of the public.

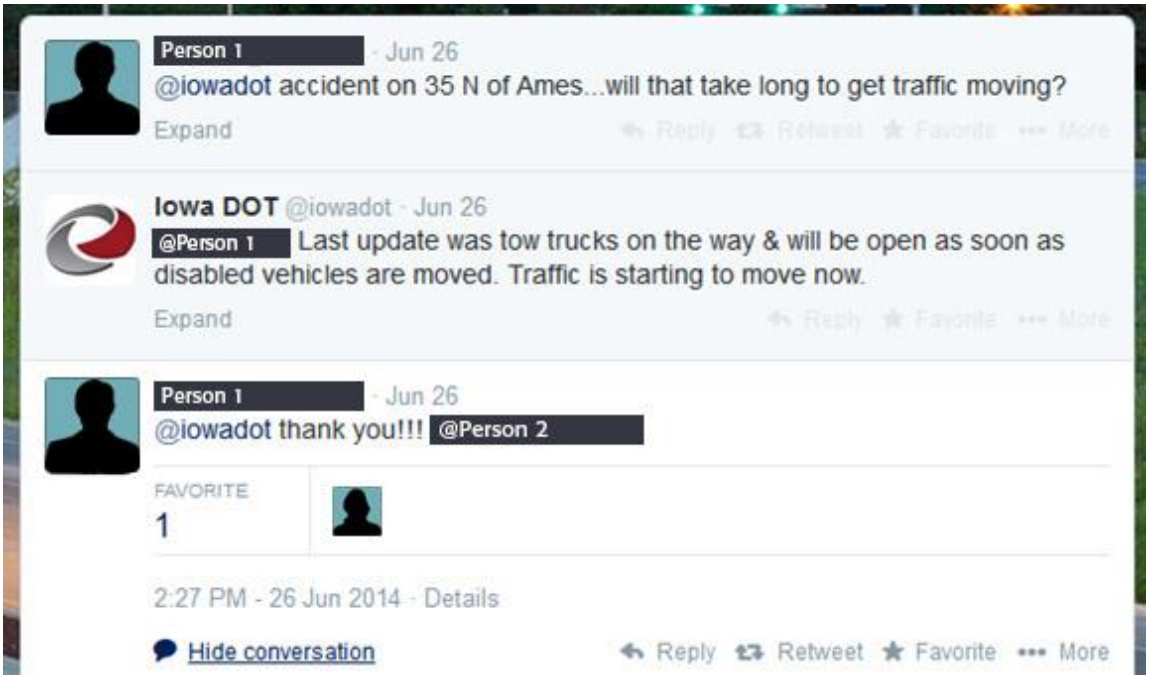


Figure 14: Engaging a Traveler via Twitter
(Source: adapted from Iowa DOT)

Figure 15 shows a situation where a member of the public is volunteering information. Note that this post is not directed to DOT; it does not use a DOT identifier. The user is tossing the information into the universe hoping that it will be useful to someone. This post has come to the attention of DOT because their use of 'I-80' matched a monitoring filter in use at DOT. In this case the information from Twitter got to the TMC before the crash was detected through other means. Note that a second member of the public joins in the dialogue.

Figure 15: Social Media Dialogue
(Source: adapted from Iowa DOT)

In Figure 16 a member of the public has requested information through Facebook. Several other people join in the conversation.

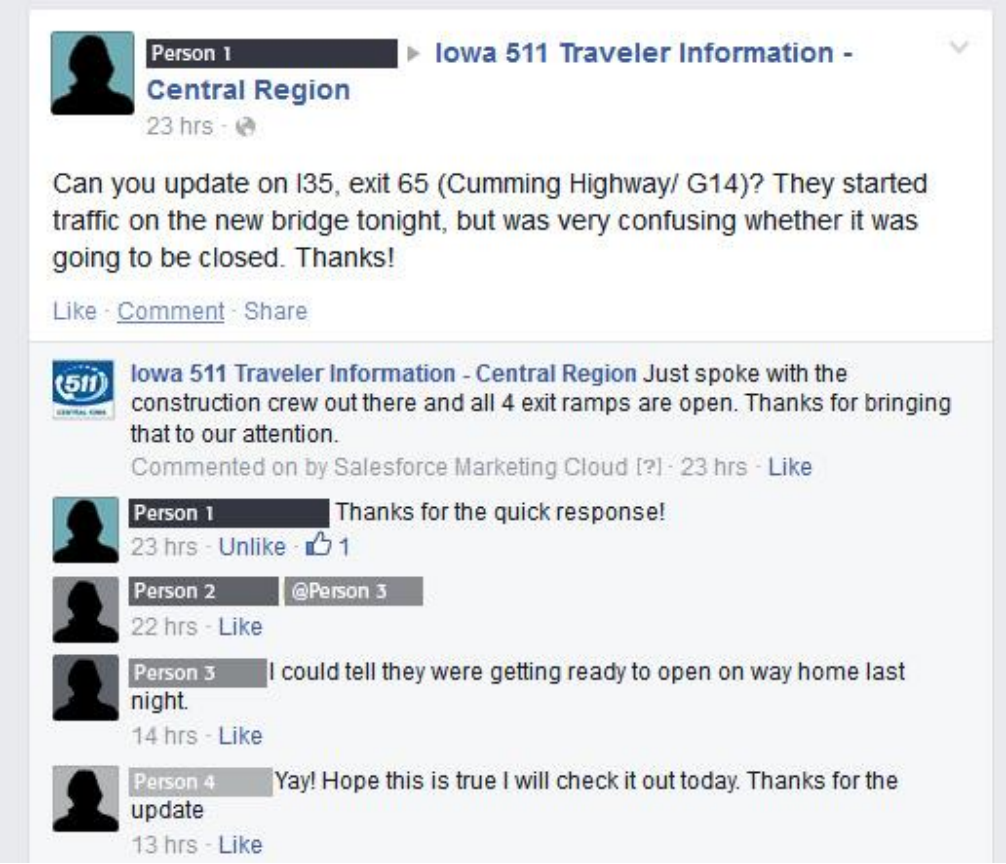


Figure 16: Traveler Engagement via Facebook
(Source: adapted from Iowa DOT)

Iowa uses the tool Radian6 to manage their engagement. The DOT recognized that acquiring the correct support tools was a critical part of making their social media policy work. They evaluated multiple vendors in a competitive process before choosing Radian6. This tool allows DOT to select keywords related to their operations, and the user display automatically updates with current information from multiple sources relevant to the keywords. Figure 17 shows how they use it to simultaneously monitor many data streams. DOT has plans to periodically evaluate the actual performance of their toolset and make any necessary adjustments.

Figure 17: Using Radian6 to Monitor Social Media
(Source: adapted from Iowa DOT)

4.3 Utah DOT's Citizens Reporting Program

The best example of a state DOT sponsored mobile application for crowdsourced information is the Utah Citizen Reporting Program. Under this program, citizens can use a mobile application to report road conditions. Citizen reports are sent directly to traveler information outlets without need for any additional human intervention. This app is the first of its kind and it was recognized with a Best of ITS award from ITS America in 2014. Fully operational, public participation began during the winter of

2013/14. As of this writing there are approximately 550 registered citizen reporters who have submitted over 1800 individual reports.

Utah has over 5000 miles of rural, state-maintained highway, and is subject to extreme winter weather. The total cost of extending and maintaining their Road Weather Information System (RWIS) system throughout that entire network would have been extremely high. Utah Department of Transportation (UDOT) developed its Citizen Reporting Program as a way to extend road condition data coverage without incurring the full cost of RWIS systems.

The UDOT Citizen Reports App is available for download at general purpose outlets such as Google Play. The app provides a simple, intuitive interface which allows users to quickly create and submit road condition reports. Citizen reports are automatically posted to traveler information outlets as soon as they are submitted. UDOT had dedicated meteorological staff which performs spot checks on randomly selected reports.

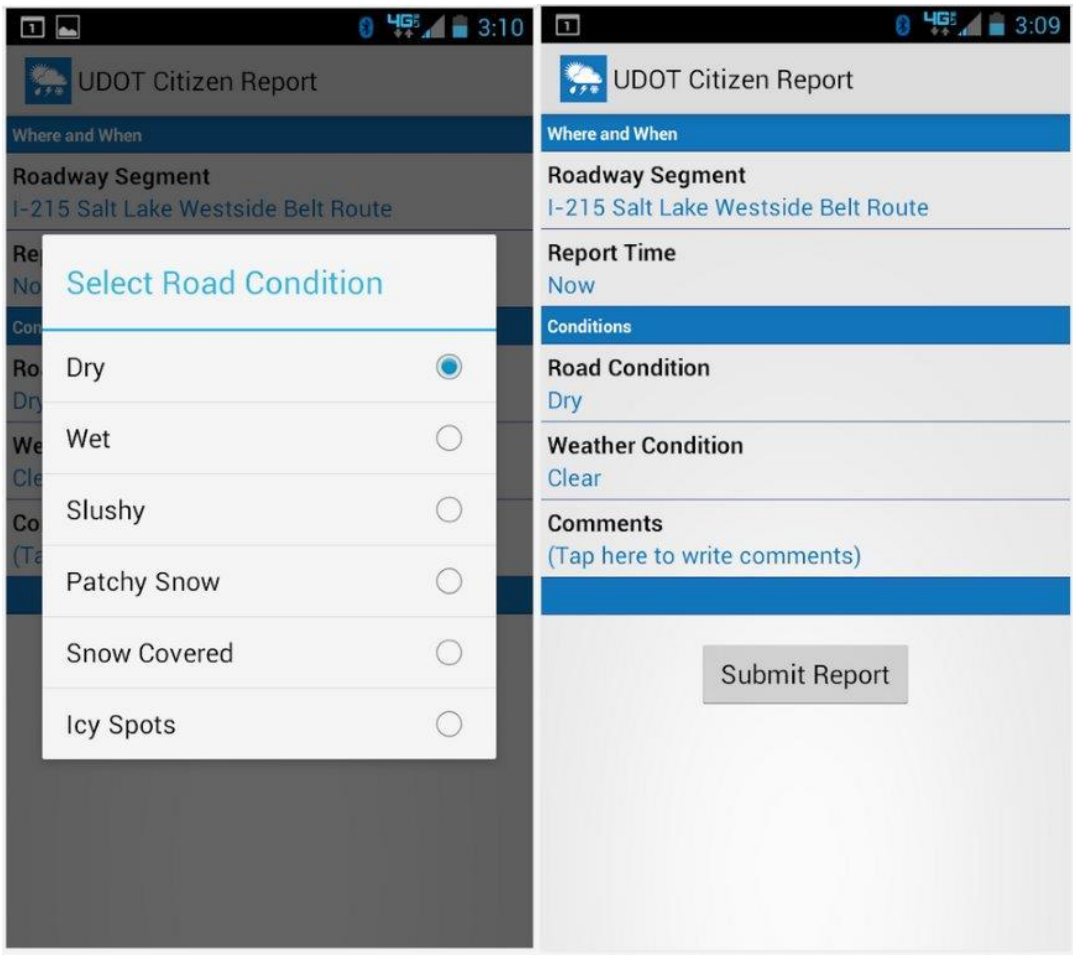


Figure 18: UDOT Citizen Reports Mobile App
(Source: Utah DOT)

User management processes are a critical part of the Utah Citizen Reporting system. Users must go through a registration and training process before they are allowed to submit reports. The registration

process weeds out potential users with a frivolous interest, and gives UDOT a mechanism for monitoring individual user input. If spot checks indicate that a particular user repeatedly submits inaccurate reports, then administrators can eliminate that user's ability to post reports. The training process helps to ensure that observations and reports use consistent definitions from user to user. Utah has strict laws against using mobile devices while driving, and the training emphasizes the importance of not using the app while driving. The user management processes have proven to be very successful. Reports from the public have been 99.7% accurate. To date, not a single user has had their posting privileges revoked due to inaccurate reports.

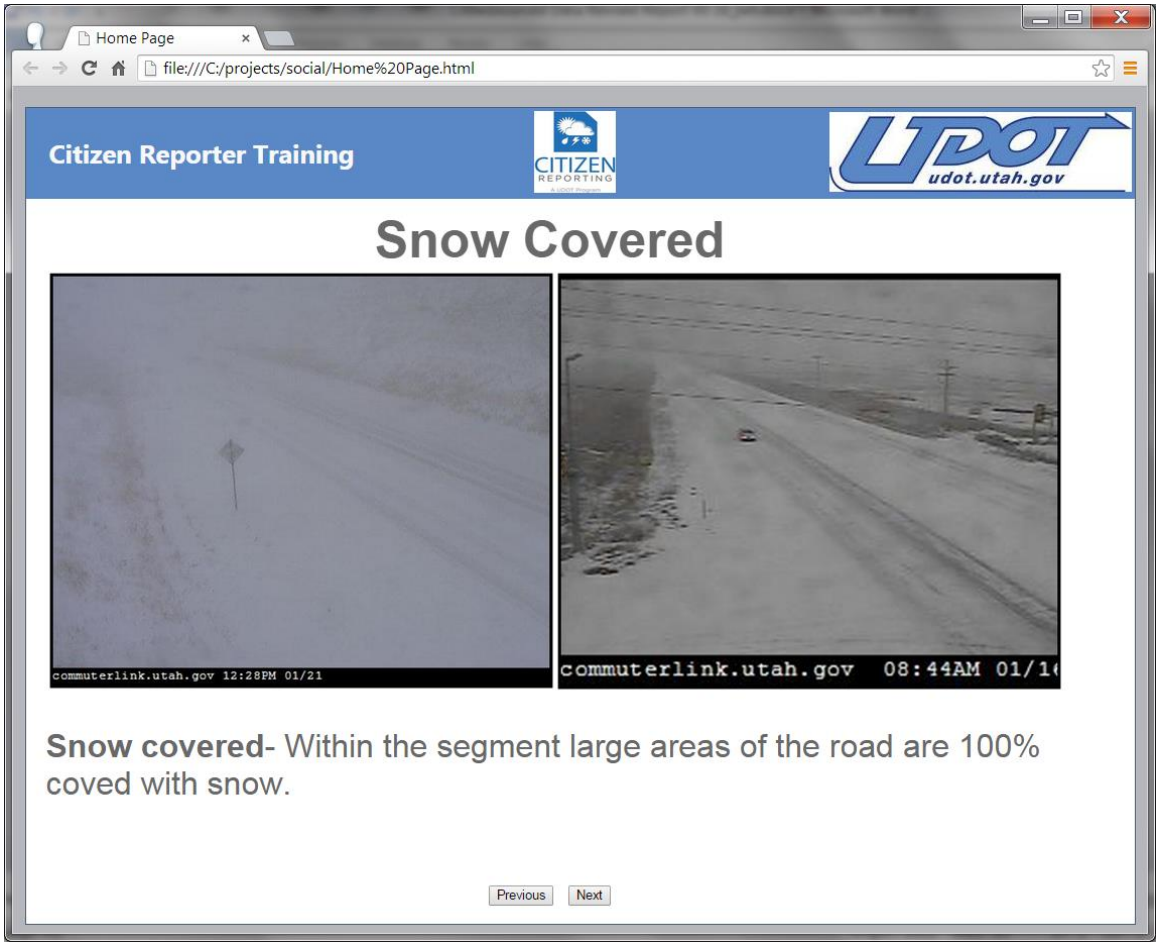


Figure 19: UDOT Citizen Reporter Training Screen
(Source: Utah DOT)

Building institutional confidence and support for the system was a multi-stage process. The system designers first analyzed relevant pre-existing systems. Wyoming has had a citizen's reporting system based on voice phone calls for several years. The Wyoming experience with managing reporters from the general public contributed to the design of Utah's user management processes. The Utah system designers circulated information on the Wyoming system within UDOT to build confidence in a citizen based system. The testing and validation of the system prior to official launch was extensive. UDOT did a soft launch of the system using only reporters internal to the DOT. The soft launch lasted for a full winter season, and included over 100 different reporters. During this period, UDOT validated the technical and operational aspects of the system without publishing reports to public outlets. Testing

included users who deliberately posted false reports to confirm that verification processes were sufficient to detect errors. Upon completion of this extensive validation period the system enjoyed broad support within UDOT and moved ahead with a public launch.

The goal of the program is to extend data coverage in rural areas, and UDOT has made an effort to reach citizens best positioned to do that. They have actively invited participation from professional communities such as:

- Trucking Associations
- Highway Patrol Officers
- Snow Plow Operators
- Ham Radio Operators
- Adopt a Highway Coordinators

In addition, they have used digital marketing campaigns targeted at rural areas. The majority of current users are in urban areas, but the marketing efforts are ongoing.

UDOT developed the app with a combination of its own funds and some early support from FHWA. Consultants were used for some of the technical work, but UDOT maintains ownership of the source code. UDOT intends to continue developing the system. Plans already exist for two-way data sharing with other systems including the National Weather Service and Waze. UDOT is considering extending the system's geographic coverage and making it a multi-state project.

4.4 Mining Social Media at DDOT

The District of Columbia Department of Transportation (DDOT) has recently deployed a pilot project which performs real-time data mining on Twitter feeds to extract traffic incident information. This process has some similarities with social media monitoring projects discussed in section 4.2, but differs in that DDOT has implemented the algorithm for identifying relevant tweets themselves. The algorithm assigns tweets a score based on content, source, geolocation, timing and relationships to other tweets. Off the shelf tools such as Radian6 implement general purpose algorithms which can be configured based on local needs. Direct implementation of the algorithm gives DDOT much more control and power over the scoring process than they would get from an off the shelf tool.

The operator interface automatically displays the high scoring tweets as they are posted. As the tweets age they move down the display and eventually drop off. Operators can choose to retain tweets on the display if they wish. Figure 20 shows an example of the operator interface containing relevant tweets from the Washington DC Metropolitan Area. As of this writing the initial pass of operator training and feedback is in progress. Training processes will be refined and standardized based on the results of the initial operator experience.

The initial results of the pilot have been very positive. Only about 7% of the tweets which score high enough to be displayed on the operator interface are not relevant to traffic incident management. The data feed has proven useful in incident detection, management and after-action reviews. The greatest benefit though has been in incident management. It provides another data stream with information about ongoing incidents, and sometimes includes photos. The timeliness of the tweets has proven to be very good. Other researchers have not always found this to be the case, so it is important to

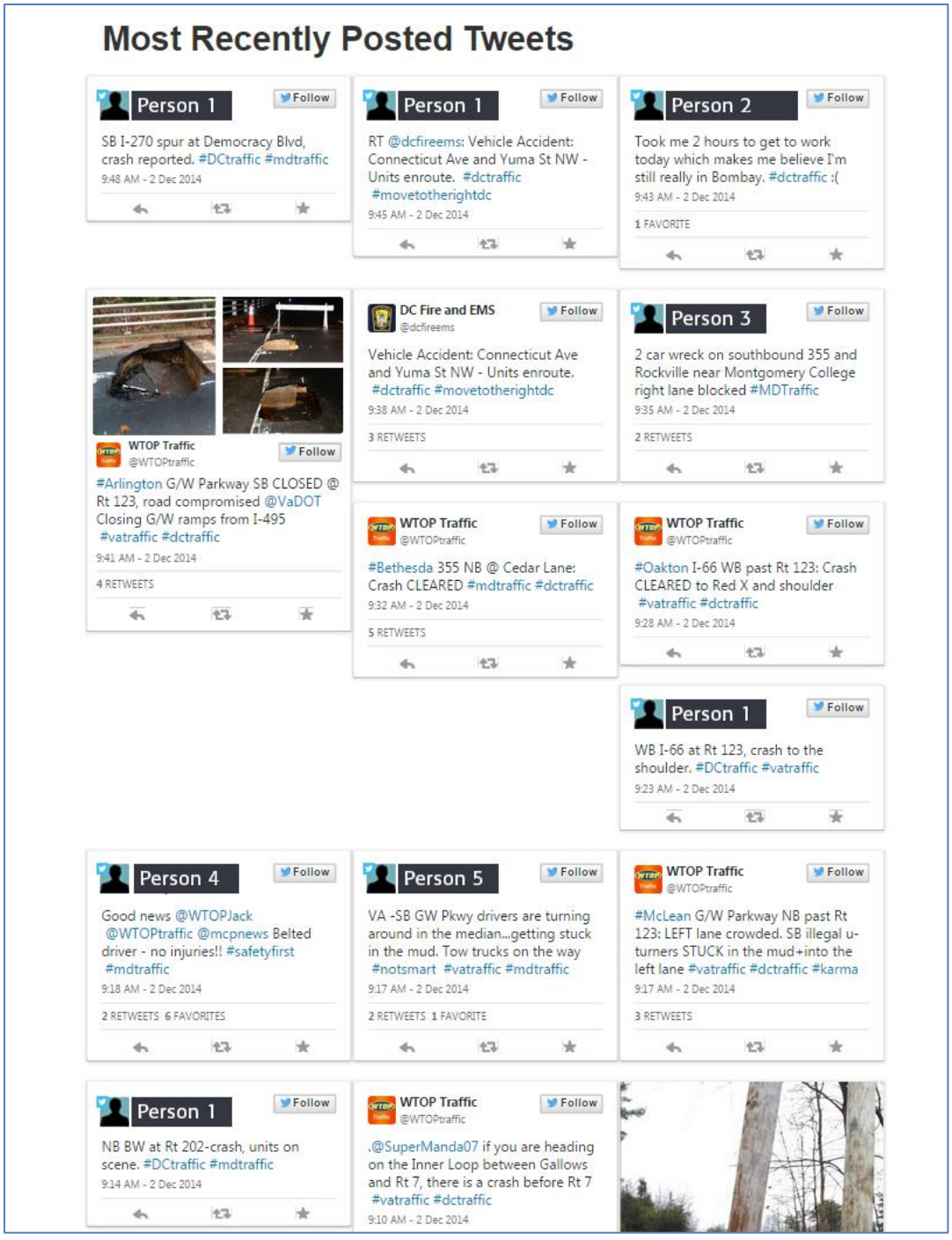


Figure 20: DDOT Operator View of Mining Results
(Source: adapted from District of Columbia DOT)

understand why it works well in DC. It seems that urban density and the presence of significant numbers of pedestrians and transit riders is an important factor determining tweet immediacy. When an incident occurs on a busy DC street, there are always many potential tweeters, including pedestrians or bus riders who do not have to cope with the distraction issues that restrict auto drivers.

The development effort for this project was significant. DDOT spent about six months on the algorithm. Developers estimate that about half of that effort was devoted to identifying the DC-specific sources, terms and locations which typically indicate DC incidents. That portion of the effort is not transferable to other locations and would have to be done again for any other implementations. Development work was performed by DDOT staff, but was funded by a federal grant. DDOT has put the source code in the public domain and it is freely available to anyone. The project was presented at the 2015 meeting of the Transportation Research Board (53).

4.5 Florida DOT – Integration of Waze Data

Many states consume speed data from private vendors. Florida has gone a step further and is consuming incident data entered by the public in the Waze mobile application. They are the first state DOT to integrate Waze data directly into their operational systems. The primary motivation for Florida Department of Transportation (FDOT) was to improve incident detection and management on arterials. Florida has extensive instrumentation along its interstates, but far fewer devices along arterials. Waze data gives them visibility into arterials at less cost than additional sensors.

The agreement between Waze and FDOT is for mutual data sharing with no fees charged by either party. Both sides are required to attribute data to the other party, and this was a point FDOT was careful to include in the agreement. Negotiations on the agreement began in the spring of 2014, and phase one of the project was operational by the end of 2014.

While there is no charge for the data, FDOT has expended significant effort to complete the technical and operational aspects of the integration. FDOT receives raw, unfiltered data from the Waze data feed. FDOT has developed systems which filter the incident feed based on incident type and geography, and then inject relevant incidents directly into the Advanced Traffic Management System (ATMS). The Waze interface is new, and is still evolving. FDOT has had to make adjustments to their integration systems as the Waze interface has changed. Mapping has presented some challenges as well. Road names are not necessarily identical across different sets of map data, and FDOT has had to make adjustments to bring the road names into agreement.

The new data integration has caused some changes for the operators. When a Waze incident comes into the ATMS it appears on the operator interface as a new icon. The operator can inspect the new incident report and determine if they want to forward the report to traveler information outlets or take other action. The volume of the new data presented to the operators is an important issue to consider. The Waze incident feed contains a confidence level for each incident. FDOT uses this value in their filters and has tuned the acceptable confidence threshold as a way to prevent the operators from being overwhelmed. Operators statewide currently receive approximately 450 alerts per day from the Waze feed. In phase one of the integration the Waze report appears automatically in the ATMS, but the operator must manually transfer that data to the traveler information system if they wish to do so. Phase two of the integration is underway, and that will give the operator a way to trigger an automatic transfer of data from the Waze incident to the traveler information system. While the system will transfer all the data, the decision to execute the transfer will still be at the operator's discretion.

4.6 Summary and Key Findings

These crowdsourcing examples cover widely varied applications and circumstances, yet there are common elements. Chief among them is the lesson that just because the data providers are unpaid does not mean that the data are free. There may be integration costs, development costs and operational costs associated with new data, including training operators on the new procedures. Some of these may be ongoing costs which will have to be allocated from operational budgets. Software and hardware must be maintained, and any increased operational workload must be sustained indefinitely.

None of the projects described here are universally applicable. These projects have succeeded because they were designed to suit the particular conditions at hand. The DDOT project depends upon the critical mass of social media providers present in a dense, urban environment, and the Utah project is cost effective because it is located in large, rural areas where traditional instrumentation is cost prohibitive. When evaluating a project executed by another agency, it is important to identify any conditions which were critical to the project's success. Before emulating a project, the agency should be sure that the same conditions apply in their local situation.

The other common theme is that all of these projects sought to fill in a gap in the state agency's data portfolio. In no case did the crowdsourced data replace an existing data source. This is likely to be a good rule of thumb for state agencies considering crowdsourced projects. With the possible exception of commercially available speed data from passive collectors, we have not seen a case where crowdsourced data is a viable replacement data source.

5 SWOT Model

This chapter presents a SWOT Model to help agencies frame the discussion of the value proposition for acquiring and integrating crowdsourced data into real-time operations. The model structure presented in this chapter is designed as a framework to guide agency decision making. The model presents implementation considerations that assess the strengths and weaknesses of the agency to plan and implement crowdsourced projects. The model also presents the opportunities and threats associated with each approach to acquiring crowdsourced data and supporting TMC operations.

The SWOT model framework is generic and applicable to all of the predominant approaches to acquiring crowdsourced data as described in earlier chapters.

- ***Social Media*** - Extracting data from social media platforms (e.g., Twitter or Facebook).
- ***Third-Party Data*** - Acquiring crowdsourced data from a private sector data provider/aggregator.
- ***Specialized App*** - Capturing crowdsourced data from citizens via the use of specialized apps.

5.1 SWOT Analysis

SWOT (**S**trengths, **W**eaknesses, **O**pportunities, and **T**hreats) analysis is a strategic planning tool for analyzing the capabilities of a business or organization to undertake a new initiative. The SWOT model enumerates the organization's internal capabilities to undertake the initiative by articulating its strengths and identifying its weaknesses. To assess the potential value of the proposed initiative, opportunities and threats are elaborated. Figure 21 portrays these dimensions.

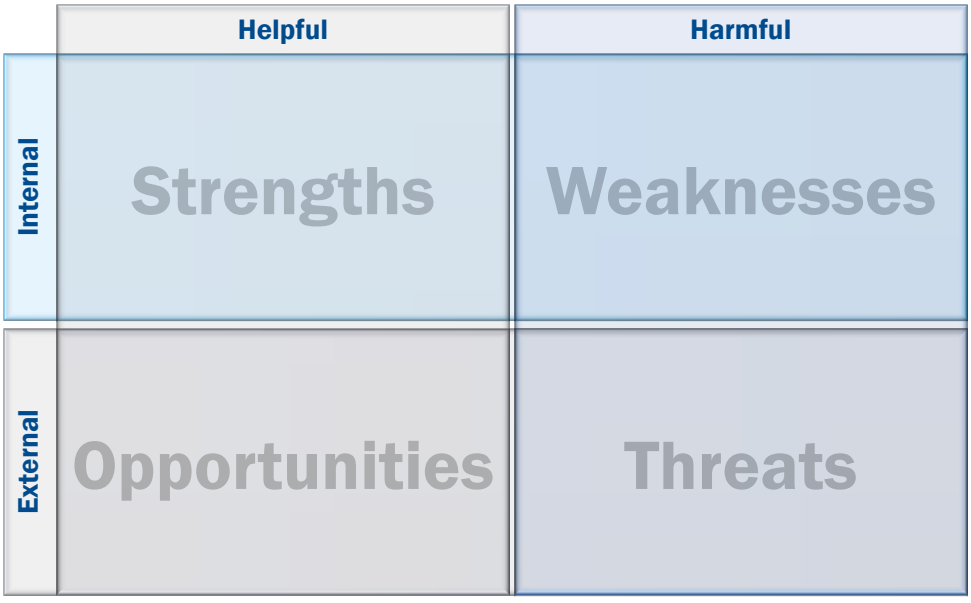


Figure 21: SWOT Analysis Dimensions

To apply the SWOT modeling framework to crowdsourced data projects, the following structure is recommended:

- ***Internal Capabilities (Strengths and Weaknesses)***: The internal capabilities for agencies to undertake crowdsourced data projects focuses on the operational, institutional, and technical implementation considerations to be assessed. This chapter will present a discussion on the implementation considerations that should be considered. Individual agencies, when performing a SWOT analysis, should perform a self-assessment to determine whether these considerations should be considered strengths or weaknesses.
- ***External Considerations (Opportunities and Threats)***: This chapter will offer a set of core external considerations to be assessed when evaluating the potential value of implementing a crowdsourced data project.

The SWOT analysis should be used by agencies as part of the systems engineering process during the early planning stages of a project, prior to the development of a Concept of Operations document.

5.2 Internal Capabilities (Strengths and Weaknesses)

There are operational, institutional, and technical considerations that may be used to assess the capabilities and readiness of an agency to initiate a project to acquire and integrate crowdsourced data. For purposes of the SWOT model the following definitions will be adopted:

- ***Operational***: Attributes of the agency within the TMC environment including, its staff, available resources, relationships with citizens and stakeholders, and core TSM&O activities.
- ***Institutional***: Attributes and culture of the state agency external to the TMC environment.
- ***Technical***: Specific skills and capabilities within the TMC and its staff to perform the tasks necessary to acquire and integrate crowdsourced data.

Determining whether these implementation considerations are seen as strengths or weaknesses at a specific TMC is crucial to evaluating the capabilities and readiness for the state agency to initiate a new project.

- ***Strength***: To be considered a strength the agency must currently possess or willing and able to develop the capability and be ready to apply that capability in support of a new crowdsourced data initiative.
- ***Weakness***: Lacking the necessary experience or resources or the willingness to apply existing resources would be considered a weakness.

5.2.1 Operational Considerations

Organization willingness and readiness are the central operational considerations that would measure the strengths and weaknesses of an agency to plan crowdsourced data projects. The following sections present the operational considerations that agencies should evaluate.

5.2.1.1 Organization and Staffing

All new projects require the availability of appropriate personnel to plan and execute the project. Participation from primary and secondary TMC stakeholders may be needed depending on the scope of the effort. Personnel needed to successfully plan and execute crowdsourced data projects may include:

- Project Champion – Crowdsourced projects, like all critical projects undertaken by traffic management agencies, require a project champion who will provide the energy and leadership to drive the project from planning through execution. The project champion will serve as the primary stakeholder and will manage relationships with all participants and stakeholders.
- Operations Staff – It is important for the users of the data to be part of the planning process making sure it is understood how the data will be acquired and used to support TSM&O activities.
- Maintenance Staff – Maintenance staff should be involved if the crowdsourced data project will be related to weather or maintenance activities.
- Information Technology (IT) – Many agencies operating TMCs rely on an IT group to manage systems and software. IT representation on the project team is strongly recommended.

Organization and staffing would be considered a strength if the appropriate personnel can be mobilized and motivated to support the project.

5.2.1.2 Experience

There are several skill sets that are required to achieve project success. This may include experience using and monitoring social media, integrating and processing disparate data, developing apps, and demonstrating successful execution of similar types of projects.

5.2.1.3 Business Processes

This area of capabilities is associated with the internal agency processes to plan and execute technology projects. Included is experience with the systems engineering process, capital planning for initial deployment and longer-term operations and maintenance. Within the systems engineering process, the ability to generate requirements and successfully verify and validate that the requirements have been met are cornerstone activities that must be demonstrated in order to have a successful project.

5.2.1.4 Culture

This refers to the intrinsic values, focus, and values of the agency that would enable and support new technology and data projects. Does the agency have a commitment to be a leader in exploring new technologies? Is there a project champion who will motivate staff and see the project through to completion? Does the agency have the ability and motivation to adopt/modify policies and procedures as it relates to social media usage and the ownership and sharing of data?

5.2.1.5 Performance Measures Program

Having a strong performance measures program will enable agencies to initiate new projects with crowdsourced data. Capabilities and policies for data acquisition, processing, and management are a benefit. The existence of performance measures to assess TSM&O activities will help measure the benefits and value of crowdsourced data to improve operations. Specific examples of performance measures are discussed in Chapter 6.

5.2.1.6 Training Program

As discussed in the case studies, agencies considering integrating crowdsourced data into their operations must keep in mind that operations staff will require training to best make use of this new data source. Those agencies seeking to implement their own specialized app must keep in mind that

individual participants will need to be trained on how to report traffic or weather/road conditions information depending on the particular application. Even in the case of mining social media for crowdsourced data, agencies should consider providing guidelines on their website to aid in the standardization of geolocation information, types of information desired, terminology, etc.

Table 5-1: Operational Considerations

Capability	Examples of Strengths
Organization and Staffing	• Availability of TMC primary and secondary stakeholders to plan and execute crowdsourced data project • Availability of private sector consultants to support project planning and execution if agency experience is limited • Presence of project champion who is strongly invested to making the project a success • Capabilities to facilitate staff training and professional development to enable TMC staff to learn about new technologies and data sources
Experience	• Existing experience within TMC staff or consultants with crowdsourced data and/or using social media • Dedicated Staff member who monitors social media feeds or who is actively involved in promoting operations via social media • Existing presence on social media (e.g., using Twitter for 511) demonstrates some experience with social media platforms • Demonstrated past success in executing similar projects
Business Processes	• Strong understanding of best practices, including systems engineering process • Established program management for scoping, planning, implementing, and verifying • Capital budgeting for initial deployment and longer-term Operations and Maintenance
Culture	• Demonstrated leadership (e.g., project champion) and support for the new project • Funding mechanism to support new initiatives • Commitment to adapt TMC policies associated with social media usage and data ownership
Performance Measurement	• Established program and policies for data acquisition, data analysis, and how data is used • Use of measures of effectiveness to evaluate TSM&O activities
Training Program	• Demonstrated experience and commitment to training operations staff on new systems • Existence of formal training programs within the agency

5.2.2 Institutional Considerations

TMCs do not operate in a vacuum. The culture and policies of their associated state agency can positively or negatively impact a TMC operator's ability to work with new data, embrace new technologies, and interface with citizens. This series of considerations focus on institutional issues that can positively or negatively impact the planning of crowdsourced data projects.

5.2.2.1 Social Media Policies

TMC operations staff is often bound by the social media policies of their sponsoring state agency. The ability for TMC staff to access social media during business hours or the guidelines for communicating via social media may be dictated by institutional policy. These policies will include rules of engagement for:

- Who is responsible for monitoring and posting to social media
- Restrictions on the content that may be shared via social media platforms

5.2.2.2 Social Media Staffing

State agencies are actively increasing their staff and staff capabilities to manage social media presence. According to the 2014 AASHTO State DOT Social Media survey, 44% of state DOTs have dedicated staff for social media outreach. In other states, the responsibility is shared between staff members in the communications office. The relationship between these institutional staff positions and the TMC operations staff strongly impacts the ability of an agency to undertake projects that involve social media.

5.2.2.3 State Agency Culture

The state agency culture as it relates to social media is rapidly evolving across the country. While most state agencies embrace social media for to communicate with citizens, there may be restrictions on who can use social media because it is perceived by many to be a distraction and a security risk. Institutions that are more progressive on embracing social media and who have programs in place to provide access and training is an advantage to agencies looking to initiate crowdsourced data projects.

Table 5-2: Institutional Considerations

Capability	Examples of Strengths
Social Media Policies	• Policies that enable TMC operations staff to have access to social media and support two-way engagement with citizens • Defined policies for data use, data ownership, and privacy
Social Media Staffing	• State agency commitment to allocated dedicated staff for social media-related activities • TMC and state agency share resources who specialize in social media
State Agency Culture	• State agencies that embrace social media and who have programs in place to educate and train their staff • State agencies that are looking to engage the public through social media.

5.2.3 Technical Considerations

Technical considerations focus on the skills and capabilities within a TMC to implement crowdsourced data projects including technical skills to process, integrate, and archive data, existing hardware and software platforms that can be used, and capabilities to design and implement specialize apps.

5.2.3.1 Data Integration and Processing Experience

Having the requisite skills and experience, either in-house or through contracted private sector consultants, to fuse and integrate new data with current data sources will impact decision making as it relates to crowdsourced data projects. Among the capabilities that would support moving forward with new crowdsourced data initiatives include:

- Existing data fusion hub or system that handles data from disparate sources
- Traffic data management plans
- Framework for data quality management including defined data quality targets
- Existing data center or warehouse with easy retrieval and reporting capabilities

Data overload is another consideration. The increasing number of available data sources and amount of data entering a TMC is a concern both in terms of processing and storing the raw data and extracting critical information from and interpreting the data. TMCs with strong capabilities to process and interpret data may have fewer concerns or issues with adding a new data source.

5.2.3.2 Experience Monitoring and Mining Social Media

Integrating crowdsourced data increases the coverage of the transportation network. Data mining social media can help state agencies gather crucial information for real-time operations. Analyzing data within the constructed context can provide stronger support to intelligent systems and human operations. However, crowdsourced data is difficult to mine because of the lack of resources required to perform the added responsibilities and the level of effort to acquire the data and make sense of it.

The acquisition, processing, and analysis of social media data require a host of highly specialized technical skills. Social media systems may have volumes of data which dwarf even large scale traffic data. Data mining (as opposed to social media monitoring performed by operators) and data cleansing depend on sophisticated statistical methods. State agencies may not have the skill set to take on and maintain complex social media projects that may require highly specialized technical skills. Technical leadership will have to examine their department's skill set to determine if they can undertake and maintain complex social media projects. Additional cost may be associated with bringing in additional resources to support integrating crowdsourced data into the existing infrastructure.

5.2.3.3 Mobile App Planning and Development

Specialized apps are an emerging technology to collect crowdsourced data. Successfully designing and implementing a specialized app program requires skills in software development, marketing, and training. Agencies will have to decide whether they can take on the effort in-house or engage external partners (e.g., universities, consultants, and marketing firms) to assist with various aspects of the project depending on the level of experience and capabilities within the TMC staff. Regardless, there are a core set of skills and capabilities that TMC staff will need to have in order for the project to be successful:

- Technical Project Management – Capabilities to plan and manage a software development and data collection project.
- Software design –Development of the specifications and design of the app.
- System Validation – Testing the app and verifying that it works as designed.

- Marketing – Involvement in marketing the app and recruiting citizens.
- Training – Developing a training program for citizen reporters who will be engaged in active data collection efforts using the app.

5.2.3.4 TMC Infrastructure (Hardware/Software Platforms)

Crowdsourced data projects will support mission critical traffic operations activities. Data will need to be integrated into existing operations platforms, such as an ATMS or Condition Reporting System (CRS). The data may be used to support traveler information or shared with business partners and stakeholders. It is also likely that the agency will want to archive and report on the data. Existing infrastructures for processing, fusing, applying, and warehousing the data is a core strength.

Table 5-3: Technical Considerations

Capability	Examples of Strengths
Data Integration and Processing Experience	• Dedicated staff or consultants who currently manage data program
Experience Monitoring and Mining Social media	• Dedicated staff who monitor social media • Technologies in place for interfacing with social media
Mobile App Planning and Development	• In-house staff who can design, build, and maintain apps • Experience working with partners who develop apps for the TMC
TMC Infrastructure	• Existing ATMS or CRS platform that easily integrated new data sources • Data hub or data sharing platform • Data warehouse and reporting capabilities

5.3 Opportunities

The opportunities quadrant of the SWOT model focuses on the external elements related to crowdsourced data and social media that agencies could exploit to further their goals and objectives. Included in this discussion are the following considerations:

- The ability to improve real-time operations through crowdsourced data projects
- Using social media and crowdsourcing to improve communication with citizens and promote the humanizing of the state agency
- Undertaking activities to boost the potential value and effectiveness of using crowdsourced data
- Implementing new tools or techniques to enable better integration of crowdsourced data

- Identify new funding sources or programs (public-private partnerships) to support crowdsourced data initiatives

5.3.1 Achieving TSM&O Objectives

The primary opportunity for crowdsourced data is the potential to support TMC operations. A state agency's TSM&O priorities will help to determine what type of crowdsourced project is most suited to their needs. Table 5-4 presents a list of the most common TSM&O activities, their perceived criticality for TMC operations, and the perceived applicability of crowdsourced data. These measures are based on the survey responses and discussions with TMC staff in the panel meeting. The criticality column is a qualitative measure of how useful crowdsourced data would be to support TSM&O activities.

Table 5-4: Applicability of Crowdsourced Data by Type and Activity

TSM&O Activity	Criticality	Social Media	Third-Party Data	Specialized Apps
Incident Management	Very High	Moderate	Very High	High
Traveler Information	Very High	Moderate	High	High
Work Zone Management	High	Moderate	High	Low
Special Event Management	High	Moderate	Medium	Low
Freeway Operations	High	Low	High	Low
Arterial Management	High	Moderate	High	Low
Road Weather Management	High	Moderate	Low	High
Managed Lanes	Low	Low	Low	Low
Maintenance Mgmt.	Low	Moderate	Low	Moderate

The applicability of crowdsourced data columns measure how useful the three predominant crowdsourced data sources would be to support TSM&O activities:

- ***Social Media*** – Most relevant for event-based reporting associated with TSM&O activities such as incident detection, work zone status monitoring, and road condition reporting.
- ***Third-Party Data*** – Probe data for travel times and speeds is most useful to support freeway and arterial monitoring and a primary source of real-time information for travelers

- ***Specialized App*** – This has the most limited applicability and is best suited for acquiring data on roadway conditions and disseminating traveler information.

5.3.2 Overcome Data Deficiencies

The survey and subsequent panel discussion revealed some dissatisfaction with current data sources. Figure 6 presented the perceived challenges with current data sources with coverage, cost, latency, and integration cited being most critical. Figure 7 suggested that TMC staff perceived crowdsourced data as providing an opportunity to offer wider network coverage or fill-in data gaps. In addition, crowdsourced data is seen as being a lower cost option compared to conventional data sources (e.g., ITS field devices).

5.3.2.1 Expand Coverage Areas and Fill in Gaps

The lack of coverage and data accuracy was identified as a challenge with the traditional data sources used to implement TSM&O strategies. Crowdsourcing provides good spatial and temporal coverage, and results in real-time content for road condition and weather reporting and incident detection. Crowdsourced data has the potential to expand coverage areas and fill in gaps in the transportation network on roadways where data is limited or not available (e.g., secondary routes and arterial). In some cases it is challenging for state agencies to operate in an environment where multiple state agencies share responsibility due to jurisdictional or functional boundaries. This may be because of institutional history or the difficulty of enforcing technical or operational uniformity in a heterogeneous environment. Crowdsourced systems cross boundaries very naturally. They may provide a powerful tool for streamlining operations in regions with these kinds of challenges.

5.3.2.2 Lower Cost

There is a perception that crowdsourced data is a lower cost source compared with conventional data sources. This may be attributed to seeing lower direct costs associated with mining social media or acquiring data from third-parties through a data exchange program (e.g., Waze's Connected Citizens Program). The reality is that each of the three predominant crowdsourced data sources will require some direct and indirect costs. Agencies should perform a detailed cost assessment as part of the planning process.

5.3.2.3 Reduce Latency

Latency, measured as the time between when an event occurs and the time that the TMC becomes aware of the event, is one of the primary weaknesses of TMC operations. Incident Management and Road Condition Reporting are two TSM&O activities which suffer from latency issues. Crowdsourced data may provide an opportunity for reducing latency by hastening the transmission of data from the field to the TMC.

Table 5-5: Overcoming Data Deficiencies

Data Sources	Social Media/Specialized Apps	Third-Party Data
Expanded Coverage	• Use citizens to provide wider coverage of network	• Data offerings cover freeway networks with emerging arterial coverage
Lower Cost	• Limited direct costs for monitoring social media • Apps require up-front investment and on-going maintenance	• May be more cost efficient than installing and maintaining field devices
Reduced Latency	• Citizen reporting may be very fast but depends on TMC ability to extract and integrate data	• Small latency (1-5 minutes) with private sector data

5.3.3 Improve Institutional Cooperation

Social media engagement often starts off as a task within the communications department of the transportation agency. As the engagement becomes more and more real-time and oriented toward specific incidents, this role tends to become operationalized. This movement has taken different forms in different agencies. In some cases, the communications staff has established a presence in the TMC, and in others they have handed off part of their role to operations staff.

No clear formula has yet emerged for how to balance the communications and operations aspects of social media engagement. There will probably always be a mixture of both functions within the social media engagement role. Quality social media involvement has a marketing and promotional aspect to it that should be guided by communications professionals. However, the content of state agency social media dialogue is increasingly focused on real-time operations. Both communications and operations will probably have to modify training and staffing policies to facilitate sharing of skills and expertise.

5.3.4 Social Media Monitoring Tools

Monitoring social media without advanced tools is like looking for a needle in a haystack. Manually examining social media feeds is far too labor intensive to be productive. Several different tools exist for making this task manageable. Depending on the product and feature set selected, the capabilities of these tools may range from assistance with managing simple text searches across multiple feeds, to sophisticated real-time statistical analysis of large volumes of social media data. Some of these tools blur the line between monitoring and data mining.

- Radian6 – Provides an easy way to review information from multiple feeds in one tool.
- Dlvr.it – Auto posts to Twitter and Facebook
- Hootsuite – Manages social networks, schedules messages, engages audiences and measures return on investment. Able to manage Twitter, Facebook, Google+, and LinkedIn profiles in one platform.
- Tweetdeck – Provides the ability to manage multiple accounts from one platform, organize and build custom timelines, and create searches to track topics, events, and hashtags.

- Trendsmap – Shows latest trends from Twitter.

Selecting the correct toolset and then training staff to use them effectively are critical steps in building an effective social media engagement program.

Social media tools are also able to measure the impact of the data a state agency posts to social media. Many tools can report the exact number of users who have access to posts either directly or indirectly. This is called the 'reach'. The change in reach over time is a particularly good measure of the quality of information a state agency posts. If readership is increasing either by growth in followers or from republishing, then indications are that readers find value in the information. If reach is dropping over time, then it follows that users are not finding value.

Evaluation tools can also measure the average emotional content for posts associated with an organization. The most common use case for this feature is a corporation measuring the response to a marketing campaign. 'Sentiment analysis' quantitatively measures how positive or negative public posts related to an organization are. The same tool can be used to determine how the public feels about state DOTs in general and even about particular projects.

5.3.5 Humanizing DOTs and Engaging Citizens

The internet, websites, apps, Social media, and open data programs have provided exciting ways for greater engagement with citizens. State government websites and traveler information websites enabled traffic management agencies to share information with its citizens. Social media and specialized apps support two-way engagement which is transforming the relationship between government and its citizens. Government is leveraging these platforms to put a human face on government activities and enable greater citizen participation in shaping community life.

A secondary benefit of undertaking crowdsourced data projects is the ability to provide additional opportunities for governments to expand the interaction with its citizens. Humanizing a state DOT using social media increases both the quantity and quality of relationships with other organizations and the traveling public. It gives state agencies the chance to connect with the community by engaging in social media dialogue. Online social media users like knowing that there are actual people behind the accounts they are interacting with. Communicating with the community not only breeds loyalty and trustworthiness, but also enhances the agency's reputation.

- **Monitoring Social Media**: As shown in the Iowa DOT case study, having state agency or TMC staff that monitor social media and actively engage in social media discussions with citizens helps to increase trust.
- **Specialized Apps**: Engaging citizens to report data to the TMC through specialized apps, like Utah DOT's condition reporting app, empowers and incentivizes citizens to contribute to the public good.

5.3.6 Incentivizing Citizens / Gamification

Gamification refers to the application of game design principles to non-game activities, incentivizing user engagement by appealing to a sense of fun and competition. The approach often involves awarding points, badges, or comparable status symbols for good performance. Competition is a central element of gamification and is often encouraged through leader boards or some other means of performance comparison. Gamification can be used to engage, persuade, and empower participants to solve large-scale, real-world challenges. The general principle behind gamification is to introduce persuasive mechanisms that appeal to natural instincts and behaviors of human beings.

Waze is one of the most recognized gamified transportation applications. The application takes the mundane experience of driving and turns it into an immersive, rewarding experience for its users. Where most navigation apps simply provide directions, Waze takes the experience further by enabling social components to create localized communities of drivers. The application is driven by crowdsourced data as drivers constantly contribute updates such as road conditions, debris, traffic, and other very specific details that might be helpful to other drivers using the app. Users also use the app to verify inputs from other users using thumbs up or thumps down buttons. This helps to validate the information, improving data quality. Additionally, when users contribute information, they are awarded points for their good deeds, further motivating their participation in the Waze experience. As a result, users stay engaged in the long-term. Sustaining long-term, high quality data from crowdsourced apps is important and can be very valuable for TMC operations.

Understanding the potential of gamification, state agencies planning to implement gamified methods may want to consider the quality of experience the state agency seeks to offer its citizens. For public state agency apps, gamification provides an opportunity to encourage greater citizen participation and sustain their involvement. Some citizens may naturally contribute their time and efforts towards participating in crowdsourcing data collection efforts with gamification methods. With others, it may be beneficial to integrating key game mechanics to make the experience more immersive and compelling.

Gamification can also be a successful demand management strategy. Individual travelers are empowered with the ability to affect a transportation system, based on their unique decisions. As a result, the ability to incentivize certain aspects of human behaviour as they relate to transportation systems relates provides great potential for an operating state agency and its ITS practitioners to improve the mobility and efficiency of the system, as well as provide improved experiences for each individual driver.

Many industry leaders believe that if transportation goals can be aligned with human stimulants commonly utilized in traditional gaming mechanics such as community building, rewards systems, public praise, and residual user benefits, gamification can be utilized to address certain transportation issues. In contrast to congestion charging schemes that penalize bad behavior, incentives-based applications seek to reward good behavior. Gamification methods can provide effective incentives to influence citizens' short-term and long-term behavior. Real world rewards can deliver demand management across wide areas with far reaching impacts. Gamified applications like Metropia, Chromaroma (www.chromaroma.com), and From5To4 (www.f5t4.com) can support demand management strategies based on game mechanics. Leveraging smartphones and other mobile devices these applications allow users to earn rewards when they are willing to make travel decisions the help to mitigate congestion or reduce emissions. Incentives can be used to persuade citizens to use alternative modes and reduce driving. They can also be used to change modify departure time to help reduce traffic congestion. For example, gas credit from a local municipality, can incentivize citizens to leave early or to assign citizens into car-pooling groups to help keep the roads clearer.

5.4 Threats

This quadrant of the SWOT Model focuses on issues that would prevent an agency from considering to undertake a crowdsourced data project or negatively impact the value of crowdsourced data to support TSM&O activities. The issues listed in this section can threaten any crowdsourced data project, however, in most cases, agencies can mitigate the threat by understanding the posed threat and applying organizational strengths.

5.4.1 Data Quality

Ensuring the quality of traffic data is critical for supporting real-time operations. When planning a crowdsourced data program, agencies need to establish guidelines for assessing data quality. Each TMC applies different approaches to manage data quality. Ahn, Rakha, and Hill (52) recommend six characteristics that can be used to assess data quality: accuracy, completeness, validity, timeliness, coverage, and accessibility. These are good starting points for aiding TMC staff in determining crowdsourced data quality.

- **Accuracy**: TMC staff must have confidence that crowdsourced data matches actual conditions.
- **Completeness**: The data should contain all required attributes in all readings. For example, extracting an incident report from Twitter that is lacking a location or details will not be of value for TMC operations.
- **Validity**: The percentage of instances that the data is valid. Ideally one would like to strive for 99% validity.
- **Timeliness**: The percentage of times that the necessary data is available when it is required. Third-party probe data has a high degree of timeliness since it is updated every few minutes. Data from social media or specialize apps is likely to be sporadic and may not meet the needs for specific TSM&O operations.
- **Coverage**: The data needs to have sufficient geographic coverage to meet the operational needs of the TMC. One of the perceived advantages of crowdsourced data is to expand network coverage compared to current TMC data sources.
- **Usability**: The relative ease with which the data can be processed and manipulated. Obtaining crowdsourced data in standardized formats would be the ideal state for TMC use.

5.4.2 Data Validity and Credibility

The survey revealed that data validity and credibility is the most significant concern regarding the use of crowdsourced data for TMC operations. Data validity is not a specific issue with crowdsourced data as TMCs rely on data accuracy and timeliness for traffic management and operations. Crowdsourced data can evoke several concerns as it relates to validity and credibility:

- **Limited data points**: When non-recurring events or abnormal conditions are acquired through specialized apps or the monitoring of social media, it may be difficult for TMC staff to respond unless there are multiple reports that corroborate the event. When there are few data points, TMC staff may need to rely on secondary data sources (e.g., ITS field devices, manual reports from staff) to validate the event. For example, a citizen reporting a freeway event on Twitter is similar to a single call to 911. TMC staff will generally require visual corroboration from a trusted source.
- **Inconsistent Terminology**: Data used in TMC operations needs to be fairly precise. To accurately define an event requires specific attributes on the location, timing, and description. While transportation agencies have standard terminology to describe the roadway network and report events, the traveling public is less likely to use conventional terminology. For example, when multiple citizens “tweet” about the same event, it is very possible that they may use different terms which will have to be reconciled by TMC staff.

- **Data Spoofing / Malicious Reporting**: Crowdsourced systems are always subject to the effects of malicious or incompetent users who report inaccurate data. Inaccuracies are not, of course, unique to crowdsourced systems. All data systems have some potential for errors. Anecdotal evidence on the accuracy of crowdsourced traffic data systems seems to suggest that the accuracy rates may vary dramatically from system to system. Some user communities are more conscientious than others, and some systems have a strong self-policing dynamic. A better understanding of what conditions foster accurate crowd based reporting might provide valuable guidance. Each state agency will have to decide what their tolerance is for reporting information which may not be 100% accurate.

Data acquired from the private sector raises different issues as it relates to data accuracy and validity:

- **Travel Time / Speed Data**: There are several companies that sell travel time and speed data to support TMC operations. Market conditions and competition between private sector providers have improved the level of quality control. When travel time data was first available commercially, several states undertook detailed studies to assess the accuracy of these data sets. The reliance on probe data led to some noticeable data reliability issues, most notably during late night hours when there were fewer probes and companies relied on historical data to fill in the gaps. Today, as the industry has advanced, there are fewer gaps. Still, TMC staff, especially in more rural areas, should be prepared to investigate the accuracy of third-party data.
- **Event Data**: As there arise more opportunities in the marketplace to acquire third-party event data, agencies need to consider the implications of data accuracy. As was discussed in Chapter 4, Florida DOT shares data with Waze but makes a distinction on their 511 website between events reported by Waze that have and have not been validated by TMC staff.

5.4.3 Data Ownership

Data ownership for crowdsourced data is largely dependent on the method of acquisition. It is not uncommon today for agencies operating TMCs to use data provided from external sources and to have limited or shared rights for data ownership. Data that is acquired from private sector providers (e.g., vehicle probe data, event data) is subject to data use agreements that stipulate the rights of the agency/TMC staff to access, use, and distribute the data. Before procuring data from third parties, agencies need to understand the licensing agreements and be prepared to negotiate terms. Pack and Ivanov (51) suggest that agencies should be prepared to seek the right to use the data for any and all internal purposes in perpetuity; the right to share the data with operational partners, and the right to summarize the data for operational uses. Agencies should not expect to obtain rights to sell the raw data or share raw data with developers not working for hire.

When crowdsourced data is acquired directly from specialized apps or from monitoring social media, the agency that owns the app would gain ownership of the data extracted. As it relates to social media sites like Twitter, most standard tweets are not owned by anyone and the data from tweets about traffic conditions or events are in the public domain.

5.4.4 Data Privacy

Data privacy is a significant concern for ITS applications. Anytime that data is collected which has the potential to identify a specific traveler or vehicle there is a concern about protecting data privacy. For crowdsourced data projects each application presents a different perspective on data privacy:

- **Extracting data from social media**: When TMC staff monitor social media platforms, like Twitter, the information is likely to contain the name/handle of the person who posted the data and possibly some information about their location. Using tools that automatically mine social media may help to minimize the collection of private data by extracting the content without identifying the source.
- **Data acquired from Private Sector Companies**: Data privacy issues are typically minimized when acquiring data from the private sector. The private sector assumes responsibility for cleansing the data of any personal attributes and by the time the data reaches the TMC it contains no personally identifiable information.
- **Data acquired through specialized apps**: Data that is acquired by TMCs directly from specialized apps may present the largest challenge. In this case the agency will need to take the appropriate steps in the design of the apps to minimize the risk:
    - Have a proactive plan to embed personal data protection into the design of the app
    - Enact an end-to-end plan to protect data from collection to deletion
    - Ensure that the app includes the ability to inform users of the collection and use of data
    - Consider a model whereby the app contains a permission-based access approach for users to "opt-in"
    - If app development is outsourced, the procuring agency needs to include privacy provisions in the agreement with the app developer and incorporate the appropriate validation procedures to ensure that the agreed upon privacy considerations were correctly implemented.

A full discussion on ITS data privacy is available through the ITS e-primer at http://www.pcb.its.dot.gov/eprimer/module12.aspx#capabilities. The e-primer suggests several strategies for mitigating privacy issues. As it relates to crowdsourced data projects, the following strategies should be considered:

- **Preparing formal data sharing agreements:** Having formal agreements in place will provide the necessary structure for protecting personal data. There are two types of data sharing agreements that may come into play for crowdsourced data projects:
    - ***Agreements between private Sector and State Agencies/TMCs***: Agencies operating TMCs pay for acquired data from the private sector and/or share data generated by the TMC with the private sector. Specific agreements should be put into place to safeguard the data and rights of each party.
    - ***Agreements between TMCs and other Stakeholders***: Separate agreements are needed between the agency operating the TMC and partner stakeholders with whom crowdsourced data will be shared.
- **Create a formal written data policy that includes procedures for disclosure of information:** Agencies should enact policies that cover all aspects of data management including collection, processing, sharing, and disclosing.
- **Limit the amount and type of personal information needed:** Strive to collect data anonymously and, if necessary, strive to minimize the amount of personal, sensitive data that is used.

- **Consider an opt-in solution where possible to provide users choices:** Provide the capabilities for citizens to understand how data is being collected and used and offer the opportunity for them to opt-in.
- **Limit unnecessary retention of personal data:** Be wary of storing any personal information. If, for some reason, personal data is needed, it should be discarded as soon as it is no longer relevant.

5.4.5 Distracted Driving

Distracted driving is extremely dangerous and the use of smartphones and social media applications is a rising cause of accidents across the world. Agencies need to think about how to encourage citizen reporting without compromising safety.

- **Monitoring Social Media**: TMC operations staff initiating programs to monitor and mine social media for data and/or engage citizens should work with their parent agencies to establish clear guidelines for social media interactions. Staff should never engage any citizen on social media who is driving. Any campaign aimed at encouraging participation should warn of the dangers of distracted driving.
- **Specialized Apps**: Agencies that deploy specialized apps for acquiring crowdsourced data should consider ways to design the app and manage citizen reports to prevent distracted driving. For example, when citizens are recruited as reporters their training should include a section on the hazards of distracted driving and they should be required to sign an agreement stating that they will not use the app while driving and, if caught, should be removed from the program. TMCs should consider designing safeguards into the app to forbid interactions by the driver while the car is moving.

5.4.6 Emerging Innovations

The rapidly evolving social media marketplace and the user interest and activity are challenges that make it difficult for agencies to integrate crowdsourced data or justify funding short-term projects. Social media tools, like many technologies used by agencies, also are evolving. There may be a limited shelf life for crowdsourced data from social media which results in added risks to system and application integration.

Data system investments may quickly be made obsolete. It is imperative that agencies time their decisions to ensure that they are not too soon or too late. The current glut of social media choices may dilute the value of any single integration forcing a state agency to build and maintain multiple integrations. Effective use of data standards may mitigate this risk, but there is currently no evidence that social media applications are converging toward any data standard which would be of particular use to a TMC.

Innovations are driving the marketplace and changing business models. Emerging data sources, e.g., connected and autonomous vehicles, and next generation social media may erode the value of investment in mining social media. The impacts of emerging data sources are unknown. Most likely, agencies will need to adapt to the changes required to obtain, store, manage, and integrate the emerging data sources with their existing sources.

5.5 Implementation Considerations

The results from conducting a SWOT analysis will provide strategic direction for agencies in the planning of crowdsourced data projects. Table 5-6 presents an assessment on how agencies can use the SWOT analysis to better plan their projects:

- Select crowdsourced data projects that play to the agency's operational, institutional, and technical strengths. This will help steer agencies to consider whether it is best to monitor social media, purchase data from third-party providers, or undertake projects to design and develop specialized apps.
- Avoid projects that will expose the TMC weaknesses and present threats that are not easily mitigated.
- Challenge staff to embrace new opportunities and dedicate resources to expand the skills and capabilities of the TMC staff and turn weaknesses into strengths.

Table 5-6: SWOT Analysis Matrix

	Strength	Weakness
Opportunity	Embrace projects that are best suited to TMC strengths and provide the best opportunities for success	Undertake challenging projects that present market opportunities and will build TMC skills and experience
Threat	Leverage internal strengths to mitigate and overcome barriers	Avoid high-risk projects that expose TMC and agency weaknesses

For example, engaging in the monitoring of social media to extract crowdsourced data is best suited for TMC staff who have strengths in using social media, agency policies that support social media usage, and a culture that embraces the challenge of working with new data sources and technologies. These TMCs are best positioned to use social media to both improve TSM&O performance and humanize their agency by engaging citizens.

Agencies operating TMCs who do not have current strengths in using social media should consider this as an occasion to enhance their capabilities and turn weaknesses into strengths. This may include:

- Working to modify TMC/agency policies as it relates to social media.
- Working with their state agency to co-locate information staff in the TMC or work more closely with TMC staff.
- Identify staff members who would like to work with social media and provide appropriate training and opportunities.
- Seek funding to support new social media initiatives.
- Invest in new technologies that enable social media interaction and data mining.

5.6 Summary

This chapter presented a SWOT analysis framework for helping agencies assess the opportunity to work with crowdsourced data. This framework is summarized below in **Error! Reference source not found.**. The strengths and weaknesses are based on operational, institutional, and technical considerations within the agency.

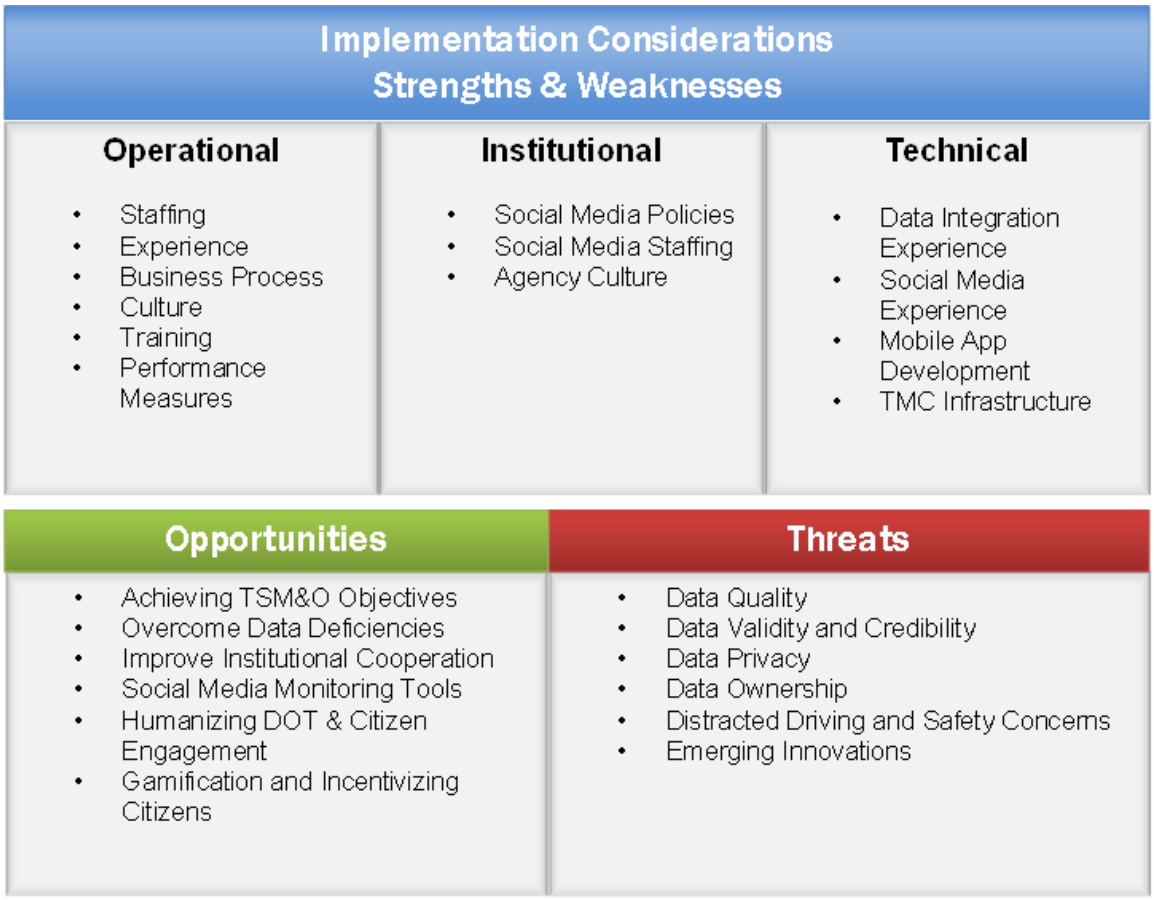


Figure 22: SWOT Model Framework for Crowdsourced Data Initiatives
(created by Open Roads)

Each of the three crowdsourced data application areas: monitoring social media, acquiring data from private sector, and developing specialized apps present different opportunities and challenges for agencies to consider. State agencies need to assess their specific strengths and weakness when planning crowdsourced data projects.

6 Measures of Effectiveness

Crowdsourced data from social media has the potential to greatly expand the available data for state agencies to use for real-time management, performance monitoring and reporting, as well as to support customer-focused activities like traveler information. Given the current state of practice and future plans of agencies, crowdsourced data will likely be one of many data types available to support operations and management, as well as performance monitoring. Integrating this data into TMC operations will need to yield benefits to the agency, in terms of broadening coverage of real-time network status and ease of use (among others), in order for it to be viewed as a viable data source.

Measures of Effectiveness, in the Systems Engineering context, refers to how well a system or a data source or an operational process performs in relation to achieving broader system or agency goals. A measure of effectiveness is different than a performance measure. A good description of the differences between these two concepts is found in "Decision Making in Systems Engineering and Management," (second edition, 2011):

- A *measure of performance* is a quantitative expression of how well the operation of a system meets its design specification
- A *measure of effectiveness* is a quantitative expression of how well the operation of a system contributes to the success of the greater system

To evaluate the impact of integrating crowdsourced data from social media into TMC operations, it is important to consider how this relatively new data source can contribute to specific TMC operating needs, as well as how it can be integrated into broader performance management strategies (such as for a corridor, traveler information system, regional congestion management program, among others).

This section examines the following:

- Overall impacts and benefits on TMC/Agency resources, processes and operational capabilities to acquire and integrate crowdsourced data into system operations
- Perceived benefits to the amount and quality of data available to TMC staff
- Improved agency reputation and enhanced public perception through more proactive public engagement
- Improved efficiency in achieving operational objectives, including incident management/response, work zone management and system maintenance

While there are many MOEs that can be identified and evaluated, there are other benefits that will take time to fully capture. In the interim, qualitative assessments such as surveys, after-action reviews and qualitative reviews of the data can give clues to effectiveness. In terms of the amount of data needed for evaluation there are parallels with the Connected Vehicle program. Only with time can the amount of data needed to fully evaluate new systems be understood and accepted by the industry. That is an area for future study.

6.1 Assessing Impacts on TMC Operations

This section presents potential measures of effectiveness for assessing the impacts of crowdsourced data on TMC operations. The value proposition is framed by identifying three key components:

- ***Objective***: Defining the objective to be achieved
- ***MOE***: Presenting a series of MOEs that can measure the success of meeting the objective
- ***Data Needs***: The data necessary to measure the "before and after" impacts of integrating crowdsourced data into TMC operations

6.1.1 Impacts on TMC/Agency Resources

State agencies have limited resources for traffic operations. New initiatives are successful when they can provide a significant return on investment and enable the TMCs to operate more cost effectively and efficiently. This MOE focuses on whether the integration of crowdsourced data is a cost-effective option for a TMC.

Table 6-1: Impacts of TMC/Agency Resources

Objective	Potential Measures of Effectiveness	Data Needs
Perform TMC operations in a cost-effective manner.	• Monetary costs associated with social media activities, including staffing, software, etc.	• Capital and operating program costs

6.1.2 Impacts on Agency Reputation and Public Engagement

Increased public engagement is a primary motivation for public agencies to embrace and expand their presence in social media. This series of MOEs proposes various ways to measure the extent to which public engagement is increased.

Table 6-2: Impacts Agency Reputation and on Public Engagement

Objective	Potential Measures of Effectiveness	Data Needs
Increase level of engagement with traveling public by using social media	• Measure the reach of social media outlets. Changes in reach are direct indication of information value -- users are seeking the information • Measure use of agency supported hashtags, which indicates value of information focus and structure. • Increased user confidence in real-time information obtained from agency with social media strategies • Demonstrate ability to show how agency responds to information it receives from users through social media through improved operations at specific locations, on specific corridors, or addressing specific maintenance needs	• Number of followers • Number of times agency posts are shared • Number of unique user sharing agency posts with others

6.1.3 Improve Ability to Meet TSM&O Operational Objectives

State agency personnel believe that that crowdsourced and social media data can support important TSM&O activities. At the heart of its value proposition is the potential for crowdsourced data to increase data coverage in a cost effective manner. Infrastructure based detection systems are costly to deploy and maintain. If they can be replaced in whole or part by crowdsourced systems, then there is potential for significant savings. Even in situations where infrastructure based systems will not be replaced there is potential value in adding crowdsourced data to the overall data portfolio. New data sources can validate and refine existing data sources, and the presence of multiple inputs makes for a more robust and resilient system as a whole. Table 6-3 presents typical TSM&O activities with a sample operational objective against which potential MOEs and data needs may be defined.

Table 6-3: Improve Ability to Meet TSM&O Operational Objectives

TSM&O Activity	Operational Objective	Potential Measures of Effectiveness	Data Needs
Incident Detection	Detect incidents as quickly as possible.	• Incident detection time. Social media outlets may allow agencies to detect incidents more quickly. • Incident detection time on corridors where traditional detection/monitoring devices are not available	• Number of incidents detected first via social media compared with other traditional means • Differential in detection time via social media vs other means • Need to track incident detection source for all incidents.
Incident Management	Improve safety for responders and other travelers	• Identify how using crowdsourced data can improve the TMCs situational awareness near the incident scene, thereby improving the TMC's decision support for response strategies	• Surveys • Qualitative reviews of archived social media data and records of actions taken
Incident Management	Verify reported incidents	• The ability for an agency to verify reported incidents is important for credibility. While more incidents may be detected via social media, they may take longer to verify on camera or via courtesy patrols.	• Time to verify incidents detected via social media vs. those detected by other means • Need to track verification times for all incidents
Workzone Management	Improving traffic flow in and around the workzone	• Crowdsourced data can be utilized to monitor how traffic is responding to work zone traffic management strategies, and identify any impacts of queuing near the work zone • Crowdsourced data can be an important source of information on the surrounding roadway network to identify if there are other nearby impacts that should be considered or that might be resulting from the work zone	• Surveys • Outcomes of after-action reviews • Qualitative reviews of archived social media data and records of actions taken

TSM&O Activity	Operational Objective	Potential Measures of Effectiveness	Data Needs
Freeway Operations	Hasten detection of and response to congestion choke points (e.g., bottlenecks and queues)	• Utilize crowdsource data to expand available travel times • Improve situational awareness of freeway operations, speeds, queues and bottlenecks • Improve response to freeway conditions through specific information from crowdsourced data	• Surveys • Outcomes of after-action reviews • Qualitative reviews of archived social media data and records of actions taken
Freeway Operations	Improve Performance Reporting	• Examine how crowdsourced data can improve freeway performance reporting and freeway strategy development over existing traditional data sources	• Comparison of performance reports across different data sources
Road Weather Management	Increase coverage and frequency of road condition segment reports	• Improve accuracy of road conditions information available to the TMC through using crowdsourced data • Evaluate conditions information received from crowdsourced data vs. forecast information received through other sources (such as NOAA) • Evaluate effectiveness of crowdsourced data compared to traditional RWIS data • Improve bi-directional reporting from the TMC to other divisions/agencies by sharing information on weather and road conditions received through social media	• Surveys • Outcomes of after-action reviews • Qualitative reviews of archived social media data and records of actions taken

TSM&O Activity	Operational Objective	Potential Measures of Effectiveness	Data Needs
Traveler Information	Improve the timeliness and accuracy of information dissemination to the traveling public	• Reduce time lag from detection to dissemination by using crowdsourced data • Improve quality and precision of traveler information by integrating crowdsourced data • Increase the number of corridors where real-time information is available for traveler information updates/reports as a result of using crowdsourced data • Assess how effective crowdsourced data is to support traveler information on corridors where there is/was not another real-time data source • Increase user satisfaction and perception of quality of information received as a result of integrating crowdsourced data from social media	• Surveys • Outcomes of after-action reviews • Qualitative reviews of archived social media data and records of actions taken
System Maintenance	Reduce detection and response times to maintenance issues	• Improve maintenance needs detection through social media/crowdsourced data • Reduce verification and response time by maintenance crews by having more precise information available through crowdsourced data • Measure relationship of improved system or device up time as a result of using new social media sources	• Surveys • Qualitative reviews of archived social media data and records of actions taken

6.1.4 Improve Data Effectiveness

Acquiring and integrating crowdsourced data into TMC operations is only beneficial if it enhances an agency's data program. Table 6-4 presents a series of MOEs that focus on improving data effectiveness.

Table 6-4: Improve Data Effectiveness

Operational Objective	Potential Measures of Effectiveness	Data Needs
Increase quality of data available to TMC Staff	• Evaluate quality of crowdsourced data compared with other data sources, and ability of crowdsourced data to improve the quality of performance reporting • Social media data can provide additional details about an incident that assist with the response.	• Details about incidents collected from social media vs other means. • Need to track source of incident attributes to determine contribution of social media to the overall picture of the incident.
Data Reliability	• The reliability of social media reports. This can be measured by how many false positive incident reports are received. This may be exacerbated by the "echo chamber" effect, where incorrect information is unknowingly redistributed. • The accuracy of information incidents obtained from social media. This affects the ability to rely on social media reports for responders.	• Number of false positive incident reports. • Percentage of new information about incidents that is correct vs. incorrect.
Reduce Data Latency	• Reduce the time between when an event occurs in the transportation network and when the TMC operational staff is aware of the event	• Before and after analyses • Comparing detection time across data sources

7 Conclusion

7.1 Summary

Crowdsourcing is developing rapidly and state agencies are recognizing the potential of seeking the collective intelligence of the public to efficiently perform TSM&O activities and deliver transportation services faster and better. Extracting data from social media, acquiring crowdsourced data from a private sector data provider/aggregator, and capturing crowdsourced data from citizens via the use of specialized apps are three most widely accepted approaches suitable for TMC operations.

The capabilities and readiness of a traffic management agency and its TMC operations staff should be assessed before integrating new crowdsourced data sources into their current data environment to support TSM&O activities. This can be accomplished by conducting a SWOT analysis to determine whether their operational, institutional, and technical implementation considerations are considered strengths or weaknesses and to assess the potential value of the proposed initiative by identifying opportunities and threats.

Agencies operating TMCs should embrace projects that are best suited to their strengths and provide the best opportunities for success. The following are some implementation procedures and considerations for agencies planning and executing crowdsourced data projects:

Build Social Media Capabilities

- **Develop a social media policy for staff members.** The policy should include guidelines for professional and personal use of social media by agency staff. The social media environment is very fluid and fast paced. It is challenging to maintain a professional approach which is still appropriate to a largely informal environment. Other type of agencies (e.g. public safety) have faced this issue and have experience and examples with may be of use to transportation agencies.
- **Plan the division of responsibilities between communications staff and operations staff.** Typically the communications staff has valuable experience with social media. Also, communications-oriented social media functions are not going to go away even as the operational aspect of social media engagement increases. Agencies should determine what functions will remain with communications and which will migrate to operations.
- **Research options for monitoring tools.** Many third-party tools are available to assist in monitoring and posting to multiple social media feeds and sources at the same time. Any contemporary social media monitoring program should take advantage of such a tool. Technical staff should work together with operational staff to determine which set of tools is most appropriate for the agency.
- **Allocate resources for operations staff.** At a minimum these resources will include training. Cross-training with communications staff may reduce the overall cost of new training. Some agencies have approached staffing for social media by adding new functions to existing operational staff, while others have placed communications staff in the TMC and expanded their role to include operational duties. One challenging aspect to staffing for social media is that the workload varies dramatically according to conditions. During a major event it can consume all of a person's time, but there may be nothing at all to do during slow times.

- **Establish operational procedures including social media.** The agency should determine what verification and response procedures are necessary for events detected through social media. This should be revisited periodically because the nature of incoming information may change as the overall social media program evolves.
- **Develop a marketing plan for social media initiatives.** Create market strategies and branding plans to encourage more citizen involvement, and enhance the agency's image and reputation. Establish clear guidelines for social media interactions to eliminate distracted driving.
- **Set goals for performance metrics.** At minimum the following metrics should be monitored:
    - Reach for social media outputs
    - Sentiment analysis for agency projects
    - Incidents detected through social media
    - Incidents where social media contributed to incident management effort
    - False positives

Build Technical Capabilities

- Integrate vendor probe data for travel times and speeds to support freeway and arterial monitoring and provide real-time information for travelers.
- Develop and deploy specialized apps (e.g., gamified transportation applications) and recruit citizens to serve as data collectors. Incentive citizens to engage, persuade, and empower participants.
- Establish programs and policies for data acquisition, data analysis, and how data is used.
- Expand ATMS capabilities to integrate/process multisource social media data sources.
- Develop and implement data standards (e.g., standard hashtags) to improve the quality of fused data outputs and data mining methods.

7.2 Further Research Needs

While there is great interest across state agencies to integrate crowdsourced data into TMC operations, there are several questions that warrant further research to advance the state of the practice and help state agencies better plan new crowdsourced data projects.

7.2.1 Dedicated Mining

Data mining is a set of sophisticated statistical analysis techniques suited for very large volumes of data. It is not simply scanning a data feed for particular values. The social media monitoring tools use some aspects of data mining techniques, but generally fall short of true data mining systems. Monitoring tools can identify valuable information within a social media data feed, but cannot identify the complex correlations and associations which can be revealed with data mining. Mining represents another layer of complexity beyond current monitoring tools. Most data mining systems are used for post hoc analysis, and some research has indicated that mining social media data is useful for transportation planning. The DC DOT project is an example of research along these lines.

7.2.2 Evaluate Accuracy of Citizen reports

The accuracy of citizen reported events and conditions seems to be very uneven. For some systems, the results are extremely accurate, and for others, much less so. It would be useful to get a better understanding of what factors and conditions influence accuracy from one system to another. This would help agencies to better estimate the value of a project before committing to it.

7.2.3 Assessing Potential Biases in Social Media Data Sources

Even if crowdsourced data is accurate, it still may not provide a complete picture. Smartphone usage is increasing in all sectors of society, but is still not uniformly distributed across age and socio-economic groups. Research is needed to determine if these data sources will provide the same level of service in all regions, and to identify mitigating strategies if not.

References

1. L. Miller, “UDOT Launches a new Citizen Reporter Program,” Utah Department of Transportation, November 2013, http://blog.udot.utah.gov/2013/11/udot-launches-a-new-citizen-reporter-program/.
2. “Connected Commuting: Research and Analysis on the New Cities Foundation Task Force in San Jose,” New Cities Foundation, 2012, http://www.newcitiesfoundation.org/wp-content/uploads/New-Cities-Foundation-Connected-Commuting-Full-Report.pdf , referenced in C. Schuetze, “Crowdsourcing Your Commute,” The New York Times, December 2012, http://rendezvous.blogs.nytimes.com/2012/12/10/crowdsourcing-your-commute/?_php=true&_type=blogs&_r=0.
3. J. Turner, “Crowd-sourced traffic data: driven by innovation,” roadtraffic-technology.com, February 2013, http://www.roadtraffic-technology.com/features/featurecrowd-sourced-traffic-data-android-smartphone/.
4. W. Wallace and X. Ban, “The Role of Social Media in Improving the Safety and Efficiency of Traffic Operations during Non-Routine Events such as Incidents and Planned Special Events, “ USDOT Research and Innovative Technology Administration (RITA)/Rensselaer Polytechnic Institute, November 2014, http://www.utrc2.org/research/projects/social-media-role-improving-safety-and-efficiency-traffic.
5. I. Leontiadis et al., “On the Effectiveness of an Opportunistic Traffic Management System for Vehicular Networks,” IEEE Intelligent Transportation Systems Society, December 2011, http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=5970119.
6. M. Moss and S. Kaufman, “Use of Web-Based Rider Input for Transit Management in the New York City Region,” RITA /New York University, January 2014, http://www.utrc2.org/research/projects/use-web-based-rider-input-transit-management.
7. L. Schweitzer, “How Are We Doing? Opinion Mining Customer Sentiment in US Transit Agencies and Airlines via Twitter,” TRB, 2012, http://trid.trb.org/view/2012/C/1129878.
8. R. Young, “Connected Vehicle Weather Data for Operation of Rural Variable Speed Limit Corridors,” RITA, August 2013, http://trid.trb.org/view/2013/P/1258532.
9. A. Mizuta et al., “Impacts of Technology Advancements on Transportation Management Center Operations,” FHWA/Parsons Brinckerhoff, FHWA-HOP-13-008, January 2013, http://www.ops.fhwa.dot.gov/publications/fhwahop13008/fhwahop13008.pdf.
10. V. Garcia, “ECAR: Enhanced Citizen-Assisted Reporting,” Wyoming Department of Transportation, Elevate ITS 2013, http://nritsconference.org/downloads/Presentations13/Garcia_A1.pdf.
11. A. Mashhadi and L. Capra, “Quality Control for Real-time Ubiquitous Crowdsourcing,” University College London Department of Computer Science, 2011, http://www.crowdsourcing.org/document/quality-control-for-real-time-ubiquitous-crowdsourcing/5440.
12. S. Young, “The Next Generation of Probe Based Outsourced Traffic Monitoring Systems,” ITS America, 2012, http://trid.trb.org/view/2012/C/1215867.

13. J. Koetsier, “Apple files patent for crowdsourced real-time mapping, event, and traffic data,” VentureBeat, July 2013, http://venturebeat.com/2013/07/04/apple-files-patent-for-crowdsourced-real-time-mapping-event-and-traffic-data/.

14. J. Singer et al., “Travel Time on Arterials and Rural Highways: State-of-the-Practice Synthesis on Rural Data Collection Technology” FHWA /SAIC/Westat, FHWA-HOP-13-028, April 2013, http://www.ops.fhwa.dot.gov/publications/fhwahop13028/fhwahop13028.pdf.

15. P. Marshall, “How to build a crowdsourced traffic safety network,” GCN, February 2014, http://gcn.com/articles/2014/02/11/crowdsourced-traffic.aspx.

16. T. Warne, “Technology and Business Practices that Work,” AASHTO/National Cooperative Highway Research Program (NCHRP) Transportation Research Board, NCHRP Project 20-24(84), April 2013, http://downloads.transportation.org/LeadershipForum/Technology%20and%20Business%20Practices%20That%20Work%20White%20Paper.pdf.

17. D. Miller, “The Evolving DOT Enterprise: Today Toward Tomorrow,” AASHTO/NCHRP/Cambridge Systematics, NCHRP 20-24(84), April 2013, http://downloads.transportation.org/LeadershipForum/The%20Evolving%20DOT%20Enterprise%20White%20Paper.pdf.

18. S. Kaufman, “How Social Media Moves New York—Part 2: Recommended Social Media Policy for Transportation Providers,” Rudin Center for Transportation, NYU Wagner School of Public Service/RITA, December 2012, http://wagner.nyu.edu/files/rudincenter/how_social_media_moves_new_york_part2.pdf.

19. S. Kaufman, “How Social Media Moves New York: Twitter Use by Transportation Providers in the New York Region,” Rudin Center for Transportation, NYU Wagner School of Public Service/RITA, October 2012, http://wagner.nyu.edu/files/rudincenter/how_social_media_moves_new_york.pdf.

20. P. Marshall, “Can transportation agencies call on smartphones for traffic data?” GCN, February 2014, http://gcn.com/Articles/2014/02/10/connected-vehicles.aspx?Page=3.

21. S. Bregman, “TCRP Synthesis 99: Uses of Social Media in Public Transportation, A Synthesis of Transit Practice,” Transportation Research Board (TRB)/Federal Transit Administration (FTA), 2012, http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_syn_99.pdf.

22. D. Barth, “The bright side of sitting in traffic: Crowdsourcing road congestion data,” Google Official Blog, August 2009, http://googleblog.blogspot.com/2009/08/bright-side-of-sitting-in-traffic.html.

23. “Fourth Annual State DOT Social Media Survey,” American Association of State Highway and Transportation Officials (AASHTO) Subcommittee on Transportation Communications, September 2013, http://communications.transportation.org/Documents/Social%20Media%20Survey_final_Oct2013.pdf.

24. M. Meeuwissen and R. Harris, “Challenges and competition to change mobility behavior,” 9th ITS European Congress, June 2013, http://gallery.mailchimp.com/f15b6ba26fe7837737124a612/files/5._Marcel_Meeuwissen_ERTICO_Gamificatio_webinar_2_October_2013.pdf.

25. J. Janney, “HootSuite vs. Rignite vs. Sprout Social vs. Radian6 – Social Media Management Systems Compared,” JohnJanney.com, December 2013, http://johnjanney.com/hootsuite-vs-rignite-vs-sprout-social-vs-radian6-social-media-management-systems-compared/.

26. E. Mai and R. Hranac, “Twitter Interactions as a Data Source for Transportation Networks,” Berkeley Transportation Systems, Inc./Transportation Research Board (TRB) 92nd Annual Meeting, January 2013, http://docs.trb.org/prp/13-1636.pdf.

27. K. Townsend, “Social media analytics tools - tested and rated,” USM - Useful Social Media, December 2011, http://usefulsocialmedia.com/taxonomy/term/211/social-media-analytics-tools-tested-and-rated.

28. R. Brydia et al., “Construction Traveler Information Systems for I-35 Widening in Central Texas,” TRB/Texas Transportation Institute, 2013, http://docs.trb.org/prp/13-3108.pdf.

29. R. Edelstein, “Envisioning the TMC of the Future,” ITE Journal, January 2014, http://www.ite.org/membersonly/itejournal/pdf/2014/JB14AA42.pdf.

30. K. Fitchard, “The Advantages of Crowdsourced Maps,” Bloomberg Businessweek, September 2013, http://www.businessweek.com/articles/2013-09-16/the-advantages-of-crowdsourced-maps.

31. New Scientist, “Google Street View and crowdsourcing could soon fix potholes “, Dvice, February 2012, http://www.dvice.com/2013-2-10/google-street-view-and-crowdsourcing-could-soon-fix-potholes.

32. “Traffic Management Centers in a Connected Vehicle Environment: Task 3. Future of TMCs in a Connected Vehicle Environment,” University of Virginia Center for Transportation Studies/Noblis & Kimley-Horn and Associates, Inc., December 2013, http://www.cts.virginia.edu/wp-content/uploads/2014/05/Task3._Future_TMC_12232013_-_FINAL.pdf.

33. J.Crowley, “Connecting Grassroots and Government for Disaster Response”, Woodrow Wilson International Center for Scholars, October 2013, http://jcrowley.net/assets/themes/read/docs/g2gv0.91.pdf.

34. J. Pritchard, “How Dull Smart Freeways Really Are”, Associated Press, November 2013, http://bigstory.ap.org/article/dim-traffic-sensors-dull-how-smart-freeways-are.

35. A. Martonik, “Trapster 4.0: navigating through the latest update,” AndroidCentral, July 2013, http://www.androidcentral.com/trapster-40-review.

36. N. Berg, “China Crowdmaps Pedestrian and Cyclist Problem Areas”, The Atlantic Citylab, July 2012, http://www.citylab.com/commute/2012/07/china-crowdmaps-pedestrian-and-cyclist-problem-areas/2699/.

37. “Sustainable Social Networking Services for Transport,” sunset, http://sunset-project.eu/.

38. J. M. Cooper, H. Ingebretsen, and D. L. Strayer, “Mental Workload of Common Voice-Based Vehicle Interactions across Six Different Vehicle Systems,” AAA Foundation for Traffic Safety, October 2014, https://www.aaafoundation.org/sites/default/files/Cog%20Distraction%20Phase%20IIA%20FINAL%20FTS%20FORMAT.pdf.

39. B. Berman, “Driving And Gaming Still Don't Mix—Though App Makers Keep Trying,” ReadWrite, November 2013, http://readwrite.com/2013/11/12/driving-gaming-road-wars#feed=/series/drive&awesm=~ouyhRcZi3J9dUW.

40. J. Maus, “ODOT embarks on “big data” project with purchase of Strava dataset”, May 2014, http://bikereportland.org/2014/05/01/odot- embarks-on-big-data-project-with-purchase-of-strava-dataset-105375.

41. J. Holleman, “Social Media Changes Emergency Operations, Adds Immediacy,” Government Technology/McClatchy News Service, February 2014, http://www.govtech.com/internet/Social-Media-Changes-Emergency-0perations-Adds-Immediacy.html.

42. P. Marshall, “DC Transforms Social Media Site with Citizen-Service Analytics,” GCN, January 2014, http://gcn.com/Articles/2014/01/13/DC-social-media.aspx.

43. Waze Mobile, https://www.waze.com/.

44. Trapster, https://www.trapster.com/.

45. WeatherSignal, http://www.weathersignal.com/about/.

46. Street Bump, City of Boston (Massachusetts) Mayor’s Office of New Urban Mechanics, http://www.streetbump.org/about.

47. Klout, http://klout.com/home.

48. Citizens Connect (Boston), http://www.cityofboston.gov/doit/apps/citizensconnect.asp.

49. NOAA mPING Weather App, http://www.noaanews.noaa.gov/stories2013/20130206_mping.html.

50. VDOT 511 App, http://www.virginiadot.org/travel/511.asp.

51. M. Pack and N. Ivanov, “Sharing Operations data Among Agencies”, NCHRP Synthesis 460, February 2014, http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_460.pdf.

52. K. Ahn, H. Rakha, and D. Hill, “Data Quality White Paper”, Report NO. FHWA-HOP-08-038, June 2008, http://www.ops.fhwa.dot.gov/publications/fhwahop08038/pdf/dataqual_whitepaper.pdf.

53. K. Fu, R. Nune and J. Tao, “Social Media Data Analysis for Traffic Incident Detection and Management”, TRB 2015, to be published

54. American Forests, http://www.americanforests.org/corporate-partners-page/

55. Metropia, http://www.metropia.com/blog/plant-tree-save-earth-drive-smarter

APPENDIX A List of Acronyms

Acronym	Meaning
AASHTO	American Association of State Highway and Transportation Officials
ATMS	Advanced Traffic Management System
CRS	Condition Reporting System
DDOT	District of Columbia Department of Transportation
DOT	Department of Transportation
FDOT	Florida Department of Transportation
FHWA	Federal Highway Administration
FTA	Federal Transit Administration
IT	Information Technology
ITS	Intelligent Transportation Systems
MOE	Measure of Effectiveness
NCHRP	National Cooperative Highway Research Program
NOAA	National Oceanic and Atmospheric Administration
RITA	Research and Innovative Technology Administration
RWIS	Road Weather Information System
SUNSET	Sustainable Social Networking Services for Transport
SWOT	Strengths, Weaknesses, Opportunities, Threats
TMC	Traffic Management Center
TMC PFS	Traffic Management Center Pooled Fund Study
TOC	Traffic Operations Center
TRB	Transportation Research Board

Acronym	Meaning
TSM&O	Transportation Systems Management & Operations
UDOT	Utah Department of Transportation
USDOT	United States Department of Transportation

APPENDIX B Glossary

Item	Definition
Active Data Collection	Collecting data from people who voluntarily and manually contribute services or content. This can occur through agency-sponsored specialized apps where citizens provide real-time data to support TMC operations. Examples of active data collection apps are Waze and FixMyStreet.
Crowdsourcing	The practice of obtaining services or content by soliciting contributions from a large group of people.
Crowdsourced Data	Data that is collected from a series of individuals. This term is used liberally in this report making no distinction between a raw data or a processed data set that has been sanitized and normalized.
Institutional Capabilities	Attributes and culture of the state agency external to the TMC environment.
Measure of Effectiveness	A quantitative expression of how well the operation of a system contributes to the success of the greater system.
Measure of Performance	A quantitative expression of how well the operation of a system meets its design specification.
Operational Capabilities	Attributes of the state agency within the TMC environment including, its staff, available resources, relationships with citizens and stakeholders, and core TSM&O activities.
Passive Data Collection	The process of automatically acquiring data produced by system users whose mobile phones and vehicles act as probes. The app or system only needs to be “on” to be sending data. There is no extra intervention by the user required.
Radian6	A program that brings social media feeds into a centralized platform providing an easy way for TMC staff to track, monitor, and respond over social media, identify trends, and cross-tab different public responses.
Social Media	A platform for sharing user-generated content among a community of users. Within this report, social media is used to identify a group of internet-based applications enabling participation in social networking. Facebook and Twitter are two examples of social media platforms commonly used for interacting with travelers.

Item	Definition
Social Media Dialogue	A two-way information exchange between two or more users on the social media network. For example, a citizen uses Twitter to send a tweet to a state DOT with information about a possible malfunctioning traffic signal. The state DOT responds to the tweet and includes the signal technicians at the TMC to relay that information. Other citizens may see the conversation, join the dialog, and contribute information.
Social Media Mining	The process of discovering useful or actionable knowledge in social media sources by introducing specialized concepts and technical approaches for analyzing massive social media data. For example, tweets can be matched to traffic incidents and weather road conditions by examining the content for key words and comparing the timestamps and locations of the tweets and incidents.
Social Media Monitoring	The active monitoring of social media channels for what is being said on the internet about an organization or company. Tools exist, such as Radian6 and NetBase, to help organizations track what consumers are saying about their brands and actions.
Specialized App	A specialized application software designed to capture real-time data from citizens to support TMC operations.
Technical Capabilities	Specific skills and capabilities within the TMC and its staff to perform the tasks necessary to acquire and integrate crowdsourced data.
Third-Party Data	Acquiring crowdsourced data from a private sector data provider/aggregator.
Waze	One of the most widely used community-based traffic and navigation app. Through the app, travelers can actively report and share real-time information on incidents, road hazards, and traffic jams with other members of the community. Waze also passively collects speed and location information from the device.

APPENDIX C Online Survey Results

Q1 What Transportation System Management and Operations (TSM&O) strategies are your agency most involved with?

Answered: 32 Skipped: 0

	High Priority	Medium Priority	Low Priority	Not Applicable	Total
Incident/Emergency Management	**81.25%** 26	**15.63%** 5	**0.00%** 0	**3.13%** 1	32
Surveillance and Incident Detection	**71.88%** 23	**12.50%** 4	**9.38%** 3	**6.25%** 2	32
Work Zone Management	**41.94%** 13	**48.39%** 15	**6.45%** 2	**3.23%** 1	31
Special Event Management	**29.03%** 9	**41.94%** 13	**25.81%** 8	**3.23%** 1	31
Freeway Operations/Management	**83.87%** 26	**6.45%** 2	**6.45%** 2	**3.23%** 1	31
Active Traffic Management or Integrated Corridor Management	**40.00%** 12	**23.33%** 7	**16.67%** 5	**20.00%** 6	30
Arterial Management	**26.67%** 8	**30.00%** 9	**33.33%** 10	**10.00%** 3	30
Managed Lanes	**20.69%** 6	**20.69%** 6	**34.48%** 10	**24.14%** 7	29
Road Weather Management/Operations	**43.75%** 14	**43.75%** 14	**9.38%** 3	**3.13%** 1	32
Traveler Information	**78.13%** 25	**15.63%** 5	**3.13%** 1	**3.13%** 1	32
Maintenance Management	**16.13%** 5	**61.29%** 19	**12.90%** 4	**9.68%** 3	31
Other (use text box to specify)	**20.00%** 1	**0.00%** 0	**20.00%** 1	**60.00%** 3	5

Q2 *Please identify how data is acquired to support your TSM&O strategies.*
Manual: Entered by agency personnel
Self-generated: Generated by agency deployed Field Devices (e.g., sensors, cameras, RWIS)
Shared: Acquired from partner agencies (e.g., state police, emergency management, neighboring states, corridor coalition partners, etc)
Crowdsourced: Data acquired directly or indirectly from the general public
Acquired: Data sourced from a private sector or third-party vendor

Answered: 32 Skipped: 0

	Manual	Self-Generated	Shared	Crowd sourced	Acquired	Total Respondents
Incident/Emergency Management	78.13% 25	53.13% 17	65.63% 21	15.63% 5	9.38% 3	32
Surveillance and Incident Detection	64.52% 20	77.42% 24	54.84% 17	16.13% 5	16.13% 5	31
Work Zone Management	80.00% 24	43.33% 13	40.00% 12	6.67% 2	10.00% 3	30
Special Event Management	87.50% 28	40.63% 13	56.25% 18	18.75% 6	9.38% 3	32
Freeway Operations and Active Traffic Management	62.07% 18	75.86% 22	44.83% 13	13.79% 4	17.24% 5	29
Arterial or Corridor Management	51.85% 14	55.56% 15	59.26% 16	7.41% 2	11.11% 3	27
Managed Lanes	41.18% 7	88.24% 15	11.76% 2	0.00% 0	0.00% 0	17
Road Condition Reporting	70.00% 21	53.33% 16	50.00% 15	20.00% 6	13.33% 4	30
Traveler Information	75.00% 24	75.00% 24	53.13% 17	25.00% 8	25.00% 8	32
Maintenance Management	69.23% 18	42.31% 11	26.92% 7	7.69% 2	7.69% 2	26
Other (use text box to specify)	0.00% 0	33.33% 1	66.67% 2	33.33% 1	0.00% 0	3

Q3 What challenges with Manual data have you experienced in trying to successfully implement TSM&O Strategies?

Answered: 32 Skipped: 0

Answer Choices	Responses	
No barriers experienced	**3.13%**	1
Not enough c overage of the transportation network	**68.75%**	22
Poor data quality	**15.63%**	5
Costly to deploy and maintain infrastructure	**43.75%**	14
Difficult or costly to fuse and integrate	**34.38%**	11
Other (use text box to specify)	**21.88%**	7
Total Respondents: 32		

Q4 What challenges with Self-Generated data have you experienced in trying to successfully implement TSM&O Strategies?

Answered: 29 Skipped: 3

Answer Choices	Responses	
No barriers experienced	**3.45%**	1
Not enough coverage of the transportation network	**65.52%**	19
Poor data quality	**17.24%**	5
Costly to deploy and maintain infrastructure	**65.52%**	19
Difficult or costly to fuse and integrate	**20.69%**	6
Other (use text box to specify)	**10.34%**	3
Total Respondents: 29		

Q5 What challenges with Shared data have you experienced in trying to successfully implement TSM&O Strategies?

Answered: 27 Skipped: 5

Answer Choices	Responses	
No barriers experienced	**3.70%**	1
Not enough c overage of the transportation network	**37.04%**	10
Poor data quality	**14.81%**	4
Costly to deploy and maintain infrastructure	**7.41%**	2
Difficult or costly to fuse and integrate	**48.15%**	13
Other (use text box to specify)	**18.52%**	5
Total Respondents: 27		

Q6 What challenges with Acquired data have you experienced in trying to successfully implement TSM&O Strategies?

Answered: 23 Skipped: 9

Answer Choices	Responses	
No barriers experienced	**21.74%**	5
Not enough c overage of the transportation network	**30.43%**	7
Poor data quality	**39.13%**	9
Costly to deploy and maintain infrastructure	**26.09%**	6
Difficult or costly to fuse and integrate	**26.09%**	6
Other (use text box to specify)	**8.70%**	2
Total Respondents: 23		

Q7 Does your organization currently use crowdsourced data to support TSM&O functions at the Traffic Management Center (TMC)?

Answered: 31 Skipped: 1

Answer Choices	Responses	
Purchase crowdsourced data from private sector	**16.13%**	5
Obtain crowdsourced data from another agency or agency business partner	**16.13%**	5
Deployed specific crowdsourcing mobile applications to collect data	**12.90%**	4
Use agency website to collect data	**19.35%**	6
Use agency telephony service to collect data	**12.90%**	4
We do not use/collect Crowdsourced data	**54.84%**	17
Total Respondents: 31		

Q8 What was the motivation for using crowdsourced data (check all that apply)?

Answered: 14 Skipped: 18

Answer Choices	Responses	
Incentive was offered	**14.29%**	2
Explore new technology	**71.43%**	10
Cost-effective alternative	**64.29%**	9
Better quality data	**28.57%**	4
Existing data insufficient	**71.43%**	10
Total Respondents: 14		

Q9 For what purposes are you currently enabling the public to directly provide information which supports operations (check all that apply)?

Answered: 13 Skipped: 19

Answer Choices	Responses	
Incident/Emergency Management	**61.54%**	8
Surveillance and Incident Detection	**46.15%**	6
Work Zone Management	**15.38%**	2
Special Event Management	**23.08%**	3
Freeway Operations and Active Traffic Management	**30.77%**	4
Arterial or Corridor Management	**15.38%**	2
Managed Lanes	**0.00%**	0
Road Condition Reporting	**76.92%**	10
Traveler Information	**76.92%**	10
Maintenance Management	**15.38%**	2
Other (use text box to specify)	**7.69%**	1
Total Respondents: 13		

Q10 Is your organization considering the possibility of using crowdsourcing to support TMC operations in the future? (select the answer that best describes your organization's status)

Answered: 29 Skipped: 3

Answer Choices	Responses	
Actively planning or developing a related project or initiative to implement	**24.14%**	7
Actively evaluating this as an option	**24.14%**	7
Interested in learning more	**44.83%**	13
Previously assessed and decided not to pursue	**0.00%**	0
Not interested at this time	**0.00%**	0
Unsure	**3.45%**	1
Other (please describe)	**3.45%**	1
Total Respondents: 29		

Q11 How useful do you believe it would be to incorporate crowdsourced data to support TMC operations?

Answered: 30 Skipped: 2

	More Useful	Useful	Less Useful	Uncertain	Total
Incident/Emergency Management	**40.00%** 12	**43.33%** 13	**13.33%** 4	**3.33%** 1	30
Surveillance and Incident Detection	**51.72%** 15	**34.48%** 10	**6.90%** 2	**6.90%** 2	29
Work Zone Management	**26.67%** 8	**53.33%** 16	**6.67%** 2	**13.33%** 4	30
Special Event Management	**27.59%** 8	**51.72%** 15	**10.34%** 3	**10.34%** 3	29
Freeway Operations and Active Traffic Management	**39.29%** 11	**39.29%** 11	**17.86%** 5	**3.57%** 1	28
Arterial or Corridor Management	**33.33%** 9	**40.74%** 11	**7.41%** 2	**18.52%** 5	27
Managed Lanes	**12.50%** 3	**33.33%** 8	**37.50%** 9	**16.67%** 4	24
Road Condition Reporting	**51.72%** 15	**41.38%** 12	**0.00%** 0	**6.90%** 2	29
Traveler Information	**57.14%** 16	**35.71%** 10	**7.14%** 2	**0.00%** 0	28
Maintenance Management	**8.00%** 2	**44.00%** 11	**36.00%** 9	**12.00%** 3	25
Other (use text box to specify)	**33.33%** 1	**0.00%** 0	**0.00%** 0	**66.67%** 2	3

Q12 What do you consider to be potential benefits or opportunities for using crowdsourced data to support TMC operations?

Answered: 30 Skipped: 2

	5 Most benefit)	4	3	2	1 (Least benefit)	Total
Low cost data source	**50.00%** 15	**36.67%** 11	**13.33%** 4	**0.00%** 0	**0.00%** 0	30
leverage private sector innovations	**25.93%** 7	**29.63%** 8	**25.93%** 7	**14.81%** 4	**3.70%** 1	27
Provide wide coverage area	**75.86%** 22	**13.79%** 4	**10.34%** 3	**0.00%** 0	**0.00%** 0	29
Growing market can fill the gaps in data	**51.72%** 15	**37.93%** 11	**6.90%** 2	**3.45%** 1	**0.00%** 0	29
Assist with meeting mandates such as performance-based management, reporting requirements, section 1201 rule, etc.	**32.14%** 9	**28.57%** 8	**21.43%** 6	**14.29%** 4	**3.57%** 1	28
Other (use text box to specify)	**100.00%** 2	**0.00%** 0	**0.00%** 0	**0.00%** 0	**0.00%** 0	2

Q13 What do you consider to be impediments to the use of crowdsourced data to support TMC operations?

Answered: 30 Skipped: 2

	5 (Most concern)	4	3	2	1 (Least concern)	Total
Data is not accurate	**23.33%** 7	**33.33%** 10	**20.00%** 6	**20.00%** 6	**3.33%** 1	30
Data is not credible	**26.67%** 8	**26.67%** 8	**16.67%** 5	**26.67%** 8	**3.33%** 1	30
Data does not provide sufficient Coverage	**10.71%** 3	**17.86%** 5	**32.14%** 9	**32.14%** 9	**7.14%** 2	28
Challenging to offer In-house support to build and maintain mobile applications	**25.00%** 7	**17.86%** 5	**35.71%** 10	**14.29%** 4	**7.14%** 2	28
Instability of the market for private sector applications	**3.57%** 1	**17.86%** 5	**28.57%** 8	**35.71%** 10	**14.29%** 4	28
High Cost and effort to fuse and integrate with other data sources	**21.43%** 6	**28.57%** 8	**25.00%** 7	**17.86%** 5	**7.14%** 2	28
Concerns over distracted driving	**17.24%** 5	**41.38%** 12	**27.59%** 8	**13.79%** 4	**0.00%** 0	29
Concerns over data privacy	**3.57%** 1	**21.43%** 6	**25.00%** 7	**35.71%** 10	**14.29%** 4	28
Benefits to T SM &O not apparent	**0.00%** 0	**11.54%** 3	**11.54%** 3	**34.62%** 9	**42.31%** 11	26
Other (use text box to specify)	**100.00%** 1	**0.00%** 0	**0.00%** 0	**0.00%** 0	**0.00%** 0	1

Q14 Does your organization maintain a presence on social media platforms (Check all that apply)?

Answered: 30 Skipped: 2

Answer Choices	Responses	
Disseminate real-time Traveler Information via social media (e.g., twitter, facebook)	**80.00%**	24
Disseminate general information on projects, events and initiatives via social media (e.g., twitter, facebook)	**83.33%**	25
Agency monitors social media content at all times	**30.00%**	9
Agency monitors social media during business hours	**43.33%**	13
Agency monitors social media occasionally or on a routine schedule	**30.00%**	9
Other (use text box to specify)	**6.67%**	2
Total Respondents: 30		

Q15 Within your agency who is responsible for monitoring social media?

Answered: 30 Skipped: 2

Answer Choices	Responses	
Public Information or Communications staff	**76.67%**	23
Operations Staff	**6.67%**	2
Shared responsibility between Operations and Communications	**33.33%**	10
Not applicable	**3.33%**	1
Other (please specify)	**3.33%**	1
Total Respondents: 30		

Q16 Does your organization believe that there is value in monitoring social media as a source for data that could be used in support of TSM&O activities?

Answered: 30 Skipped: 2

Answer Choices	Responses	
Yes	**93.33%**	28
No	**6.67%**	2
Total Respondents: 30		

Q17 What is the name of you organization?

Answered: 30 Skipped: 2

Q18 What is your role at your organization?

Answered: 30 Skipped: 2

Answer Choices	Responses	
Operational Staff based at T M C	**43.33%**	13
Operational Staff based outside the T M C	**23.33%**	7
Communications Staff	**0.00%**	0
Executive Leadership	**13.33%**	4
Other (please specify)	**20.00%**	6
Total Respondents: 30		

Q19 Would you be willing to participate in a brief telephone interview on this topic?

Answered: 29 Skipped: 2

Answer Choices	Responses	
Yes	**82.76%**	24
No	**17.24%**	5
Total Respondents: 29		

Q20 Please provide your contact information:

Answered: 29 Skipped: 2

Answer Choices	Responses	
Name	**100.00%**	29
Email	**100.00%**	29
Phone	**100.00%**	29

APPENDIX D Panel Meeting Attendees

Name	Affiliation	Name	Affiliation
John Hall	Tennessee DOT	Pete Costello	INRIX
Elizabeth Birriel	Florida DOT	Monali Shah	HERE
Paul Arends	Michigan DOT	Andrew Locke	TOM TOM
Vinh Dang	Washington DOT	Dr. Susan Grant-Muller	University of Leeds
Tony Ernest	Idaho Transportation Department	David Zaharchuk	IBM
Victor Koo	City of Pasadena	Dr. Yi-Chang Chiu	University of Arizona
Cherice Ogg	Iowa Department of Transportation	Matthew Burt	Volpe Center
Sinclair Stolle	Iowa DOT	Jimmy Chu	FHWA
Brian Hoeft	RTC Southern Nevada		
Jason Morgenroth	New Jersey DOT		

APPENDIX E Project Team Participants

Name	Affiliation
Jeff Adler	Open Roads Consulting
John Horner	Open Roads Consulting
Jeanette Dyer	Open Roads Consulting
Drennan Hicks	Noblis
JD Schneeberger	Noblis
Lisa Burgess	Kimley-Horn and Associates
Alan Toppen	Kimley-Horn and Associates

APPENDIX F Panel Meeting Agenda

Day	Time	Session
1	8:00 – 8:30 am	Registration
1	8:30 – 9:00 am	Session 1 – Welcome 1A Workshop Overview and Introductions (Jimmy Chu & Greg Hatcher) 1B Public Agency Roundtable
1	9:00 – 9:45 am	Session 2 – Project Status Update Highlights from Lit Review (Alan Toppen) Insights from the Survey and Phone Interviews (John Horner)) Program Overview (Jeff Adler)
1	9:45 – 10:00 am	Break
1	10:00 – 11:00 am	Session 3 –Social Media and ITS– Innovations and Opportunities Susan Grant-Muller – U. of Leeds David Zaharchuk – IBM Dr. Yi-Chang Chiu – University of Arizona
1	11:00 am – 12:00 noon	Session 4 – Private Sector Roundtable – Crowdsourced Data for TMC Applications –Emerging Trends Pete Costello – INRIX Monali Shah – HERE Andrew Locke – TomTom
1	12 noon – 1:00 pm	Lunch
1	1:00 – 3:00 pm	Session 5a: Opportunities and Challenges – SWOT Analysis: Breakout Session
1	3:00 – 3:30 pm	Break
1	3:30 – 4:15 pm	Session 5b – Review of Breakout Session
1	4:15 – 4:30	Day 1 Wrap Up/Day 2 Preview
1	4:30	Adjourn

Day	Time	Session
2	8:00 – 10:00 am	Session 6 – Integrating Crowdsourced Data - Implementation Procedures and Considerations : Breakout Session
2	10:00 – 10:15 am	Break
2	10:15 – 11:30 am	Session 7 - Measures of Effectiveness
2	11:30 - 12 noon	Session 8 - Wrap up and Next Steps
2	12 noon – 1:00 pm	Lunch – (Optional)
2	1:00 pm	Adjourn

U.S. Department of Transportation
ITS Joint Program Office-HOIT
1200 New Jersey Avenue, SE
Washington, DC 20590

Toll-Free “Help Line” 866-367-7487
www.its.dot.gov

FHWA-JPO-14-165

U.S. Department of Transportation