The role of real- time crowdsourced information and technology in supporting traveller information and network efficiency June 2016

Dale Harris, Dave Smith and Carl O'Neil
Abley Transportation Consultants Limited, Christchurch

Jeremy Severinsen
Southern Spatial Solutions, Christchurch

NZ Transport Agency research report 593
Contracted research organisation – Abley Transportation Consultants Limited, Christchurch

ISBN 978-0-478-44560-2 (electronic)
ISSN 1173-3764 (electronic)

NZ Transport Agency
Private Bag 6995, Wellington 6141, New Zealand
Telephone 64 4 894 5400; facsimile 64 4 894 6100
research@nzta.govt.nz
www.nzta.govt.nz

D Harris, D Smith, C O'Neil and J Severinsen (2016) The role of real-time crowdsourced information and technology in supporting traveller information and network efficiency. *NZ Transport Agency research report 593*. 116pp.

Abley Transportation Consultants Limited was contracted by the NZ Transport Agency in 2014 to carry out this research.

Keywords: crowdsourcing, intelligent transportation systems, network efficiency, real-time data, road reports, social media, travel time, traveller information, web application

An important note for the reader

The NZ Transport Agency is a Crown entity established under the Land Transport Management Act 2003. The objective of the Agency is to undertake its functions in a way that contributes to an efficient, effective and safe land transport system in the public interest. Each year, the NZ Transport Agency funds innovative and relevant research that contributes to this objective.

The views expressed in research reports are the outcomes of the independent research, and should not be regarded as being the opinion or responsibility of the NZ Transport Agency. The material contained in the reports should not be construed in any way as policy adopted by the NZ Transport Agency or indeed any agency of the NZ Government. The reports may, however, be used by NZ Government agencies as a reference in the development of policy.

While research reports are believed to be correct at the time of their preparation, the NZ Transport Agency and agents involved in their preparation and publication do not accept any liability for use of the research. People using the research, whether directly or indirectly, should apply and rely on their own skill and judgement. They should not rely on the contents or the research reports in isolation from other sources of advice and information. If necessary, they should seek appropriate legal or expert advice.

Acknowledgements

The project team would like to thank the New Zealand Transport Agency research owner, Sonia Pollard, and members of the steering group for their time and helpful feedback including Russell Pinchen and Ben White (NZ Transport Agency), Tim Hatherley-Green (LINZ), Catherine Kim (Ministry of Transport), Steffan Thomas (Christchurch City Council) and Peter King (AA).

The project team would also like to thank our peer reviewers, Dr Femke Reitsma (University of Canterbury) and Ken Lee-Jones (Auckland Transport) for their guidance and advice, and to Professor Susan Grant-Muller (University of Leeds) for sharing her knowledge and research with our team.

We would also like to acknowledge the support of Michele Poole and Andy Tebay at the Queenstown-Lakes District Council in the rollout of the real-world trial. In addition, we also recognise Tresca Forrester (Christchurch Transport Operations Centre) for her support in developing the trial.

Finally, we would like to thank all the stakeholders who kindly provided their time to share their thoughts during the stakeholder consultation phase of the research.

Contents

Executive summary

This report explores the potential for crowdsourced information and technologies to support traveller information and network efficiency outcomes across all transport networks in New Zealand. The objectives of the research were to:

- determine the extent to which New Zealand's legal and government policy framework supports crowdsourcing as a means of collecting data for use by government
- determine the value of real-time or near real-time crowdsourced data within a transport context, and to what extent crowdsourced data can contribute to better service delivery and outcomes for the travelling public and the NZ Transport Agency
- explore the role of the NZ Transport Agency in creating the technical framework and strategy required to support the practical implementation of a crowdsourced ecosystem.

An international literature review is presented to explore the role of crowdsourcing in traveller information and intelligent transportation systems (ITS). This review includes a summary of the relevant literature and legislation regarding privacy management, ownership and licensing of crowdsourced data, safety, data quality validation and practical implementation factors. A range of crowdsourcing applications or projects across all transport modes (private vehicle, public transport and active modes) have also been reviewed to demonstrate these matters in practice.

The strategic transport direction of New Zealand government agencies and the traveller information needs of the New Zealand public in relation to real-time crowdsourcing applications are presented, focusing on the extent to which crowdsourcing supports real-time network efficiency outcomes and the provision of traveller information. An initial list of high-level information needs that crowdsourcing could potentially address is then identified.

The literature review and strategic assessment of needs is supplemented through engagement with key stakeholders. During February and March 2015, 24 stakeholders from a range of agencies and user groups were interviewed to

- understand transport information needs and consider whether crowdsourced information could meet these needs
- identify the data quality needs of stakeholders and end users
- explore concerns regarding privacy, safety, data and technology, practicality and institutional barriers to crowdsourcing.

The results of the stakeholder engagement are summarised by topic area and build on the findings of the literature review and strategic assessment.

The identified information needs are evaluated against a selection of potential crowdsourcing engagement approaches, including social media monitoring, social media mining, developing a custom application, procuring third-party data or supporting private sector-led development. The evaluation was undertaken through a SWOT (strengths, weaknesses, opportunities, threats) analysis. Among other findings, the evaluation demonstrated that both social media and customised applications are most likely to meet the traveller information and network efficiency information needs, but on balance all approaches show potential value with respective strengths outweighing weaknesses and external positive factors (opportunities) generally being greater than those negative factors (threats) identified.

A real-world crowdsourcing trial was piloted in the Queenstown-Lakes District. The purpose of the trial was to test a custom web application designed to collect and report on incidents such as crashes, road damage, snow and ice. At the same time, the research team tested the effectiveness of social media monitoring (using Facebook) and social media mining (Twitter data) approaches.

The success of the Queenstown-Lakes District trial was measured through engagement analytics, follow-up surveys and interviews. Given the small-scale of the trial, the results from the trial are considered to be indicative rather than conclusive. The findings from the trial highlighted the value of combining traveller information and crowdsourcing in a single platform and found that both the application and social media (Facebook) monitoring provided crowdsourced information that was not picked up from any other information channel.

The social media (Twitter) mining component of the trial yielded no relevant information when specific transport-related keywords were used. In addition, very few Tweets included geotagged location details. Future applications of this approach may be improved by targeting higher-density urban networks, purchasing access to a larger proportion of a social media feed, and applying more sophisticated filtering techniques.

The research concludes that the Ministry of Transport and NZ Transport Agency provide a clear mandate for the use of technology in transport through their respective strategic documents and the broader data management goals of the New Zealand government also support the use and sharing of data to deliver social and economic value. There is, however, no clear direction regarding the role of crowdsourcing information in New Zealand's transport system but in commissioning this research the NZ Transport Agency is demonstrating leadership in this area.

The New Zealand government's legal framework does not restrict public agencies from crowdsourcing information, with the Privacy Act 1993 being the applicable legislation in this regard. Legislative restrictions on the use of mobile devices in vehicles will affect the use and type of crowdsourcing applications that can be promoted or used in New Zealand and at the time of writing the Ministry of Transport is working towards legislation to address this matter. Government policy and direction in areas such as social media use, web application development, open data principles, licensing and copyright all provide some direction on matters that need to be considered when undertaking crowdsourcing.

A broad range of applications of real-time or near real-time crowdsourced information was specified in the research with public transport service occupancy/capacity, congestion reporting, planned event monitoring and incident and hazard reporting identified as high priority information needs. Other real-time needs that could be met through crowdsourcing included reporting dangerous drivers, parking availability and weather/road information reporting, with a number of non-real-time information needs also identified that would support traveller information for all modes of transport.

Crowdsourcing applications are likely to deliver greatest value when coupled with a traveller information service. This is effective as users who submit information receive gratification by seeing their report publicised. The crowdsourcing collection method is promoted directly or indirectly through the service interface, and the service creates awareness of the value of crowdsourced reports for other transport users.

A number of approaches to crowdsourcing have been reviewed in this research including passive and active approaches, social media, existing applications and third-party data sources, including commercial and community-based applications, and custom-developed applications.

As determined through the literature review, SWOT analysis and real-world trial, each approach has strengths and weaknesses depending on the audience, application and location context. Given the range

of information gaps crowdsourcing can potentially fill, there is no one-size-fits-all engagement approach that can be recommended from this research but instead an open-minded approach is recommended to embrace a range of initiatives where there is a valuable level of engagement.

The NZ Transport Agency can best support transport crowdsourcing initiatives by:

- monitoring international best practice
- building and maintaining relationships with leading transport agencies and academia
- fostering relationships with TOCs and RCAs to understand information needs, gaps and local crowdsourcing projects.
- supporting open data practices
- providing leadership on how to establish trust and the measurement of crowdsourced information quality.

A key role for the NZ Transport Agency is to provide leadership in terms of establishing 'trust' or quality ratings for crowdsourced information. Given the range of crowdsourcing approaches and applications, there is no one-size-fits all model for validating data and understanding trust. Once an agency makes a series of important policy decisions about the way it collects and delivers data, the mediums for this, and the threshold for each type of event that requires human verification, then a large body of crowdsourced data can be assessed for trustworthiness through automated and algorithmic models. Trust and validation matters should be considered early in any crowdsourcing project, although the exact specifications or policy for this may need to be developed over time due to uncertainty.

Specific recommendations for the NZ Transport Agency towards the implementation of this research and future study are as follows:

1 Establish a formal process for reviewing privacy issues before a crowdsourced application is developed, or where crowdsourced data will be sourced from a third-party supplier. This process can be equally relevant for other ITS applications where personal data is collected, including the movements of individual transport users. This process shall include:

   a privacy statements and compliance with the Privacy Act 1993

   b security testing to ensure there will be no loss or misuse of personal information

   c a legal opinion regarding the re-release of data to the wider public to address licensing, data ownership and the requirements of the OIA and Privacy Act 1993.

2 Consult/collaborate with the Ministry of Transport regarding potential safety issues of any proposed crowdsourcing application or service being undertaken by the NZ Transport Agency or RCA, including:

   a whether the proposed safety measures are appropriate for the context and user type

   b whether the application will put users in breach of existing or proposed legislation.

3 Prioritise identified high priority crowdsourcing information gaps for further scoping or development of pilot projects

4 Re-run the QLDC crowdsourcing trial over next winter, taking into account the recommended enhancements discussed in section 7.5 of this report.

5 Consider potential crowdsourcing options when developing traveller information systems (and vice versa) to integrate crowdsourcing data collection and traveller information provision into a single easy-to-use platform

6 Investigate potential for crowdsourcing in meeting non-real-time needs, especially regarding public transport, walking, cycling, ridesharing and other emerging applications.

7 Ensure crowdsourced data is made public and reusable through a creative commons licence, taking into account the privacy of individual contributors (as required under the Privacy Act).

8 Continue to follow social media mining trials undertaken elsewhere and consider piloting this technique on a dense urban network

9 Monitor the usage of third-party applications (and any new applications that arise) to determine future potential.

10 Build upon the findings of this research by assembling a guidance document for the industry to promote 'best practice' for crowdsourcing data to support transport objectives. A guidance document should provide direction on how to derive value from existing social media channels and address the validation of crowdsourced information.

Abstract

Crowdsourcing is the practice of engaging the services of a large and undefined group of people (the 'crowd') to provide information or input into a particular task or activity. In transport applications, crowdsourcing has the potential to meet information needs across all modes for a range of purposes, including the provision of traveller information, and supporting transport network operations, road safety and asset management.

This research investigates the role of crowdsourcing in New Zealand as an input to inform traveller information systems and in doing so supporting road network efficiency outcomes. The primary purpose of this research is to identify the strategic, legal and policy considerations necessary to enable road controlling authorities and government agencies to lead or support crowdsourced data initiatives. These considerations include the management of privacy, safety, data collection, storage and retrieval, the use of incentives, data quality assurance and organisational barriers to new technologies and data collection methods.

The outputs of the research include practical recommendations regarding potential applications of crowdsourcing in the transport sector, the role of NZ Transport Agency and other road controlling authorities in supporting a crowdsourcing data ecosystem, and the technical requirements for crowdsourcing data to support traveller information systems.

1 Introduction

The NZ Transport Agency ('the Transport Agency') contracted Abley Transportation Consultants to explore the potential for crowdsourced information and technologies to support traveller information and network efficiency outcomes across transport networks.

The objectives of this research were to determine:

- the extent to which New Zealand's legal and government policy framework supports crowdsourced data as a means of collecting data for use by government
- the value of real-time or near real-time crowdsourced data within a transport context, and to what extent can crowdsourced data can contribute to better service delivery and outcomes for the travelling public and the Transport Agency
- the role of the Transport Agency in creating the technical framework and strategy required to support the practical implementation of a crowdsourced ecosystem.

The research underpinning this report was undertaken in late 2014 and 2015 and involved eight stages of work, starting from an initial literature and crowdsourcing application review, and finishing with the testing of a crowdsourced application in the Queenstown–Lakes District in the South Island of New Zealand. Each phase of the project was separately reported to the Project Steering Group. This report therefore represents both an overview of the research process as it was developed, and the outputs (including recommendations) that were generated.

1.1 Report structure

The report is organised as follows:

- Chapter 2 provides background for the research by summarising the findings of international and New Zealand literature on crowdsourcing applications.
- Chapter 3 describes the needs of the New Zealand public and the Transport Agency in relation to real-time crowdsourcing applications.
- Chapter 4 describes stakeholder's views on a range of issues relating to the collection and dissemination of crowdsourced information.
- Chapter 5 provides direction to the Transport Agency and other stakeholders as to how the information needs can be met, the preferred engagement strategies and a quantitative assessment of a broad range of internal and external influential factors.
- Chapter 6 describes a real-world trial which was developed to incorporate the preferred engagement options and address the research objectives and test the research findings up to that point.
- Chapter 7 analyses the success of the trial in chapter 6 and describes the lessons learnt during the process.
- Chapter 8 brings together the research findings to address the research objectives.
- Chapter 9 presents the conclusions and recommendations.

2 Literature review

This chapter explores the literature on crowdsourcing, with a particular focus on the role of crowdsourcing in transport applications and projects. In addition to a literature search, a selection of international and New Zealand-based crowdsourcing applications and projects have been reviewed to demonstrate the type of applications available.

This chapter begins by exploring the definition of crowdsourcing and describing and contrasting two different approaches to data collection: passive and active. The application of crowdsourcing within the context of this study – to support traveller information and network efficiency – is also discussed, including a conceptual diagram that demonstrates how these activities are integrated.

A review of crowdsourcing applications comes next and is followed by a review of the literature on:

- privacy, ownership and licensing
- safety
- data quality validation
- practical implementation factors.

2.1 Crowdsourcing in transport

Crowdsourcing, in the widest sense, is the practice of enlisting the services of a large and undefined group of people, usually via web technologies, to provide information or input into a particular task or activity that would otherwise be too difficult or impossible to solve using traditional outsourcing or data collection methods (Howe 2009). In return for their input, the 'crowd' receives some type of reward which may include monetary recompense, access to data or services, recognition or personal satisfaction.

In transport applications, crowdsourcing for data collection relies on the use of mobile devices (eg smartphones) acting as sensors to 'fill a gap' where traditional traffic monitoring sensors do not exist, or cannot report in real time. For some applications, crowdsourcing is a relatively low-cost method for collecting transport data, using human inputs rather than, or alongside, specialised sensory equipment (Ali et al 2012; Misra et al 2014; Marshall 2014). In addition to filling a gap, crowdsourced data can sit alongside traditional transport data sources to help with verification or to provide additional context. Crowdsourcing also brings together a large group of people on the same platform to address common issues that affect its members – particularly for stakeholders who have a small, but engaged user base (for example cyclists and public transit users) (ibid).

While the collection of real-time data is an obvious application of crowdsourcing, the crowd can also assist in the development of transportation software applications and in the planning and design of transport systems (Sorenson 2010). As discussed in section 2.2.4, events such as HACKAKL and OpenHack Christchurch demonstrate how a pool of interested volunteers (the 'crowd') can be mobilised to develop vehicle routing and public transport tracking applications from open data initiatives.

2.1.1 Passive versus active approach to crowdsourcing

In transport applications, information garnered from the crowd involves:

- 'passive' data collection using mobile device technologies such as Global Navigation Satellite System (GNSS) sensors and Bluetooth logging, and/or

- 'active' data collection whereby a user sporadically enters information about activity on the transport network.

Crowdsourcing applications are not limited to either passive or active data collection, but can use a mix of both techniques. Table 2.1 summarises the key differences between active and passive data collection.

Table 2.1 Passive versus active crowdsourcing data collection methodologies

	Passive data collection	Active data collection
Data collection method	Continuous collection using inbuilt sensor technologies	Active and sporadic user input
Data quality	Data quality checked using automated quality algorithms. The standardised collection method ensure a generally high quality of data	Quality can vary and is dependent on collection technique (eg selecting a predefined event type versus free-text forms). Data validation requires quality assurance algorithms and/or moderation
Safety issues	Background data collection requires no active user input	Data collection requires user input and may be a source of distraction for drivers in particular instances.
Data completeness	High	Variable depending on level of engagement
User requirements	Opt-in through terms of use	Active input and engagement required, with or without incentives.

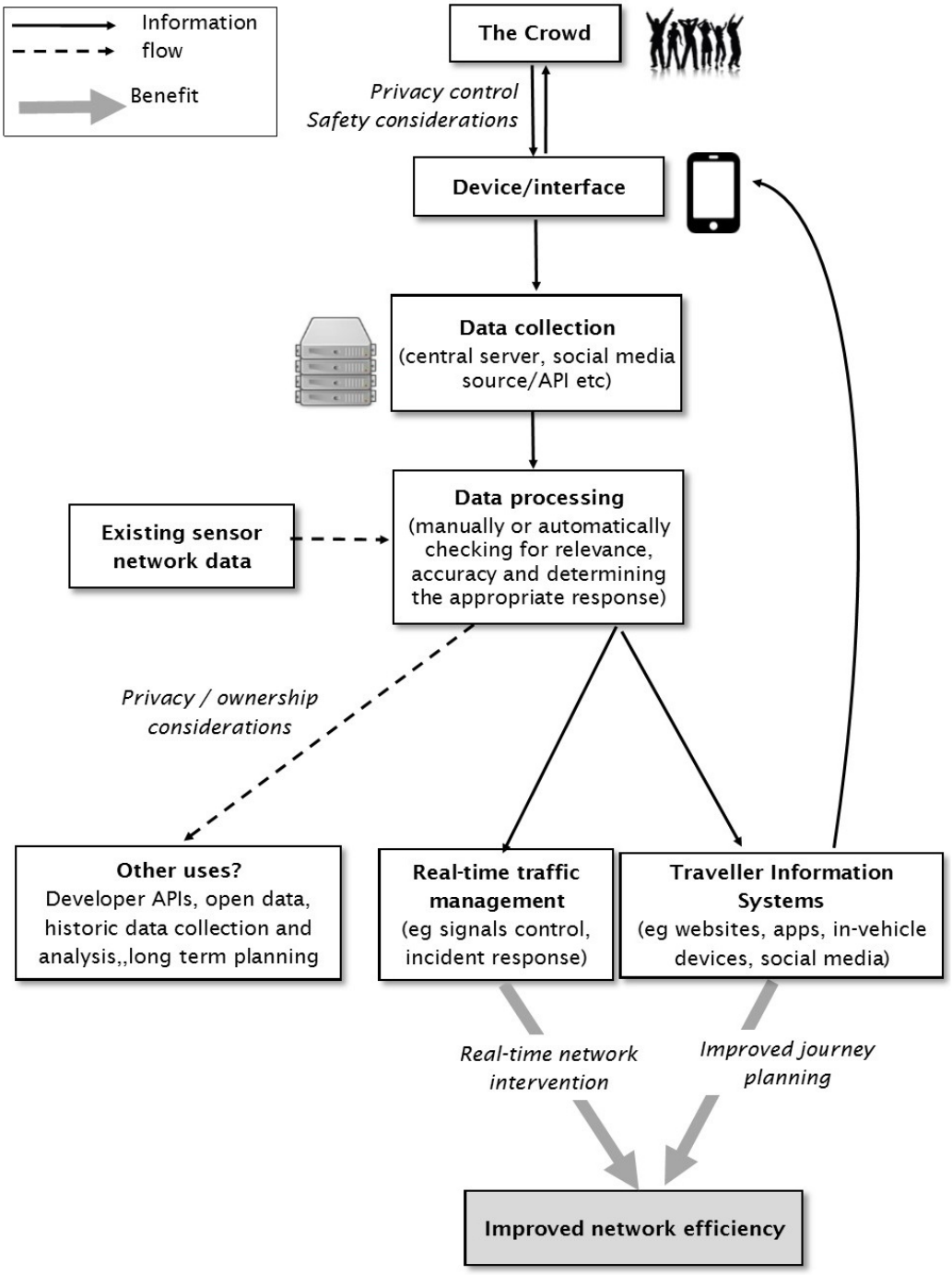
The passive approach uses in-built technologies to continuously collect and transmit data about a user's location, speed and direction. Other data from device sensors, including accelerometers, may also be collected. This approach usually requires the user to permit this collection of data by a third party by 'opting-in' and in return, the user can then receive localised crowdsourced and other information that helps them understand the transport network, for example through improved routing, access to mapping applications or improved access to public transport services.

The active approach requires users to actively input data on events, incidents or the quality of service that are affecting the transport network. This could include selecting from a pre-defined list of incidents or quality statements, or providing a qualitative commentary on transport network activities. Compared with the passive approach, this requires an active and engaged user base, and additional data validation issues arise (refer section 2.5). Safety factors regarding the interaction with mobile devices must also be considered (section 2.4).

2.1.2 Crowdsourcing framework

The purpose of this research is to identify how real-time crowdsourced information and technology support traveller information and network efficiency. Figure 2.1 demonstrates the relationship between crowdsourcing, traveller information and network efficiency, including how crowdsourced data may relate to existing sensor networks and traveller information systems (TIS).

Figure 2.1 A framework of crowdsourcing for traveller information systems and network efficiency



The crowd, through an appropriate interface, provides information that is stored and processed in a central location. The crowd may also be employed to check and validate other user submitted information. This may be supplemented by data from the existing sensor network (eg loops, Bluetooth sensors and CCTV) and processed for use in TIS. The information can also be used to support real-time interventions in the transportation system (eg responding to weather events or crashes, influencing signals control). The combination of TIS, real-time interventions and data collection all support improved efficiency of the transport network.

2.1.3 Traveller information systems (TIS)

TIS are the mechanism by which information about the road network and transport systems is distributed to travellers. Successful TIS deliver network efficiency benefits by providing information to road users that enables them to be informed about the modes of transport available, travel time and route choice (Raine et al 2014). Systems for delivering traveller information include traditional sources such as free-phone services, websites, radio and VMS, as well as emerging systems such as social media and mobile applications, vehicle information communication systems, vehicle-to-vehicle communication and travel time signage (ibid).

The customer requirements of TIS in New Zealand were explored by Chang et al (2013) and Chang et al (2015). This included identifying the type of information that users would like to receive from crowdsourced origins, including information on public transport, travel times, network events, road conditions, parking availability and information that assists with routing. These information types and requirements are explored in more detail in chapter 3.

2.2 Review of crowdsourcing applications and projects

This section provides an overview of crowdsourcing applications with relevance to this research. There are a number of crowdsourcing applications available, including transport applications, therefore this review has focused on those that represent a broad spread of transport modes, user types and providers. This is not intended to be an exhaustive review of all the applications available, and there may be other relevant (or similar) applications that exist.

A detailed review of each application, including references, is provided in appendix A.

Due to the ever-evolving nature of many of these applications and the speed at which technology improvements occur, it is noted that information on them has primarily been gleaned from web-based sources in November and December 2014. Where possible, specific details about each application has been sourced directly from the provider's website; however, in some instances this information has also been sourced from third-party sources, including technology reviews and news websites. The research team has not approached any of the application owners directly to verify that this information is correct and current, therefore the team is relying on the accuracy and currency of the online sources referenced in this research.

2.2.1 Private transport

Private transport applications use crowdsourcing to assist with real-time routing and congestion reporting, identifying and sharing hazards, obstacles and speed traps, to collect data on road quality and for parking assistance. These applications can be community or commercially based, although the distinction between the two types can sometimes be blurred.

2.2.1.1 Traffic congestion and incident reporting

Waze is a popular global community-based application for traffic congestion, navigation and incident reporting. Users can directly report events such as crashes, speed cameras, congestion and flooding through the application interface, while vehicle speed and location data is also collected. The Florida Department of Transportation (2014) has explored the potential for integrating Waze data into traffic operations management. Under a two-way data sharing agreement, each party provides information through a customer-to-customer plug-in connection. This approach improved coverage along interstate

highways; however, integration of data (interoperability) with existing systems was complex and costly (ibid).

The **Utah Department of Transport** (UDOT) **Citizen Reporter Program** has successfully developed a smart phone application to collect citizen reports on weather and road conditions. This information supplements existing sources (including meteorological data) to improve the frequency and accuracy of weather and road condition reporting for both traffic operations centres and the public.

Applications developed by commercial companies include **HERE Drive** (owned by Nokia), a routing and navigation application that uses GNSS traces to improve travel time estimation and routing. TomTom, a leading supplier of location and navigation products, has a **TomTom Map Share** platform that enables its community of users to report and update road changes, including speed limits, new streets, blocked roads and altered turn directions. These updates are then used for improved routing through TomTom in-car navigation systems. Waze also supports a mapping 'crowd' community for correcting road network data.

Commercial providers such as HERE and TomTom also provide data services combining crowdsourced vehicle-based probe data with other sources of data, including public sensor data (Center for Automotive Research 2015). The potential use and applications of these sources are discussed in detail in *NZ Transport Agency research report* 559 (Smith et al 2014).

2.2.1.2 Parking and other applications

A number of parking applications have also been developed to help users find parking vacancies as close as possible to their destination. Few of these rely on crowdsourced data and many are city specific. The experimental **Google Open Spot** (Zibreg 2010) was unsuccessful as it required users to actively identify free parking spots; however, applications that rely on passive detection (eg GNSS) techniques, such as **Parko**, are proving more successful (Blum 2012; Center for Automotive Research 2015).

An innovative application is **Street Bump**, which collects data on 'bumps' in the road using in-built accelerometer and GNSS sensors. Road authorities use this information to identify and prioritise short and long-term road repair programmes. The main difficulty in this approach to crowdsourcing is recruiting enough drivers to use it (Center for Automotive Research 2015).

2.2.2 Public transport

Much of the literature on crowdsourcing for transport applications has focused on public transport (PT) applications, primarily the use of crowdsourcing for:

- *Passenger information systems* – users with mobile devices report and receive information on network conditions, including delays and the quality/capacity of the service. This provides a source of real-time travel data (Edwards et al 2011; Corsar 2013, Nunes et al 2014, Souliotis et al 2014).
- *Tracking PT services in rural areas* – providing information on PT services in areas where traditional sensor networks are unavailable (Edwards et al 2011; Corsar 2013).
- *Understanding service quality* – crowdsourcing as a means of reporting service quality issues to bus operators and other users on the network (Hopkin et al 2014, Nunes et al 2014).
- *Empowering and fostering community among PT users* (Nunes et al 2014; Filippi et al 2013).

Two specific PT applications have been identified in appendix A: **Tiramisu**, which was developed as part of a research study at Carnegie Mellon University, and **Moovit**, a third-party application that provides and shares transit information for more than 400 cities worldwide, including some major centres in New Zealand. Both applications use real-time crowdsourced GNSS traces to improve arrival and travel time predictions. Tiramisu provides a mechanism for users to actively provide and share feedback on the

quality of service, including a “fullness” rating. Similarly, Moovit provides a mechanism for users to rate the quality of the service they are using, in transit.

2.2.3 Active transport

Active transport crowdsourcing applications gather data both passively by recording the movements of cyclists and runners, and proactively where users report problems related to the pedestrian and cycle networks.

Many community or commercial applications, for example **Strava**, are used by individuals to record and share cycling and running trips and times, both for personal use or for sharing with the user communities. Information from individual Strava users is combined and can be viewed on a public heat map, and is available for purchase to undertake more detailed analysis.

Cycle Atlanta was developed by the City of Atlanta to record cycle trips and enable cyclists to report problems such as potholes and obstructed bike lanes. Transport planners at the City of Atlanta use this data to plan strategic improvements to bicycle infrastructure.

Other applications that we have not reviewed but may be relevant to this project include cycling applications that enable users to report crashes and safety hazards. For example, **Collideoscope**, an interface for cyclists to report crashes and near-miss events in the United Kingdom.

2.2.4 Crowdsourcing community events

Crowdsourcing is not limited to the on-going contribution of passive or active information inputs, but can also include engaging an interested community (or ‘crowd’) to contribute to one-off events or projects. Two examples of this are **HackAkl:Transport** and **OpenHack Christchurch**. Both were collaborative events that brought mobile developers and interested parties together to develop innovative transport applications using open data sources. Outputs from these events included vehicle routing and real-time bus tracking applications.

2.2.5 Ridesourcing and ridesharing applications

A number of crowd-based transportation services are emerging that offer ridesharing options. This type of service is usually offered in one of two ways:

- ridesourcing – app-based, on-demand ride vehicle services where the drivers do not share the destination with the passenger (Rayle et al 2014), or
- ridesharing – applications that arrange shared rides on demand based on the source and destinations of the users (ie carpooling).

An example of a ridesourcing application is **Uber**[1], which offers on-demand rides using a mobile application interface for requesting and tracking a driver, and fare payment. Much in the same way as taxi companies, drivers are recruited and vetted before being allowed to pick up passengers. Uber is currently only available in Auckland, although the company has shown an active interest in expanding to other parts of the country, including Christchurch, Wellington and Queenstown (Stewart 2015).

An example of ridesharing applications include **SnowPool**[2], a website that allows users to coordinate travel to popular skifields. **Chariot**[3] is a New Zealand-only ridesharing application currently under

[1] www.uber.com

[2] www.snowpool.org

[3] www.getchariot.com

development. This application also allows for payment of drivers, but only to cover the genuine costs of carpooling (Chariot 2015).

Very little research has been undertaken to understand the customer information and network efficiency benefits of these types of applications. One report suggests that while ridesourcing offers better choice and cheaper travel options for users, the wider impacts of these services are not yet well understood, including the impacts on overall vehicle kilometres travelled and traffic volumes (Rayle et al 2014). While genuine ridesharing applications can reduce vehicle trips by increasing the number of passengers per vehicle, ridesourcing applications such as Uber are more like unlicensed taxi services and do not necessarily reduce the number of vehicle trips on road networks. Ridesourcing applications may also have detrimental impacts as users choose a low cost ridesourcing option instead of walking, cycling or taking public transport (Hawkins 2015).

The University of California-Berkley and the Natural Resources Defence Council is currently undertaking research to study the environmental impacts of application-based ridesourcing and ridesharing (Hawkins 2015). This research will involve analysing activity data from ridesourcing providers Uber and Lyft, and undertaking user surveys to understand how the services have changed the way people travel (ibid).

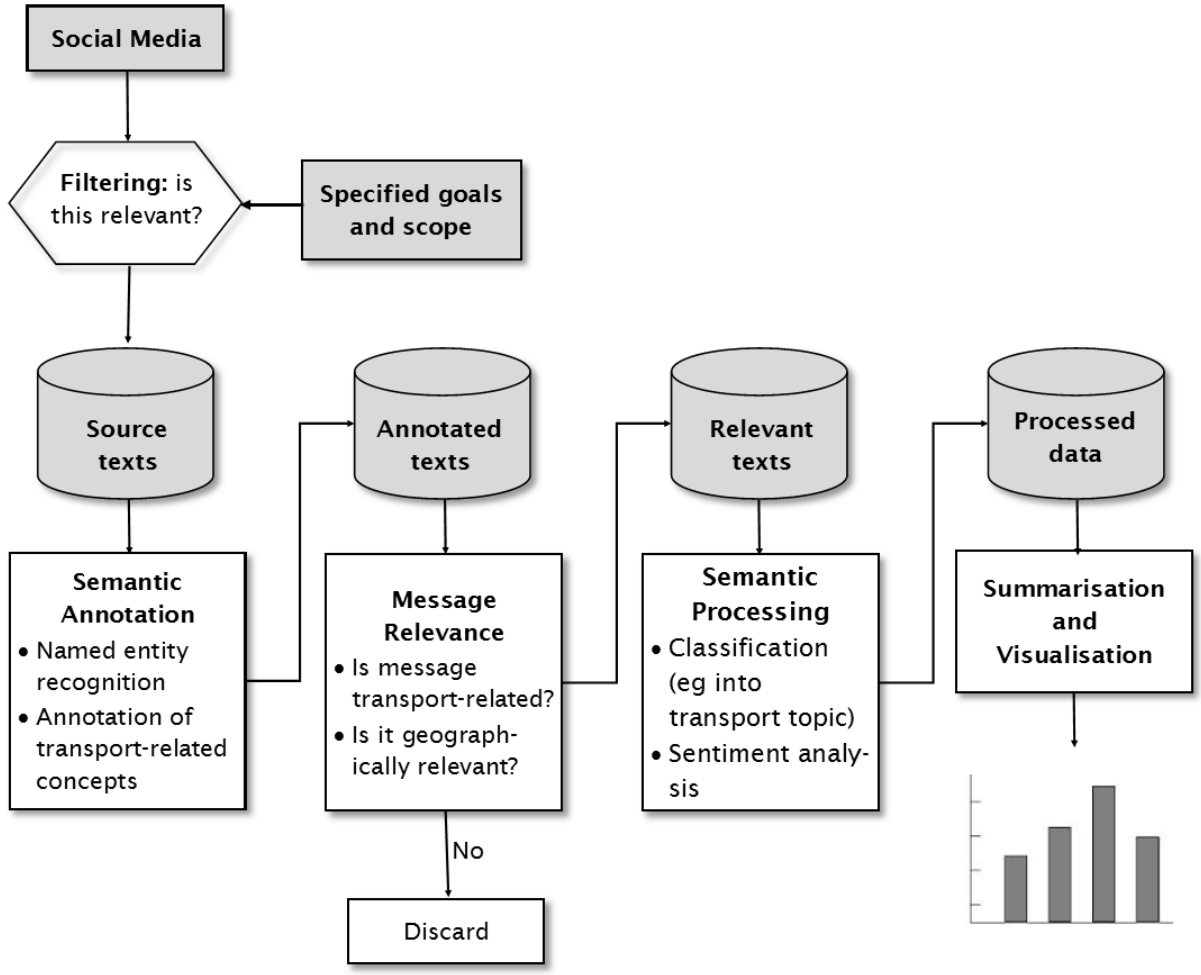
2.2.6 Social media monitoring and mining

Many New Zealand transport agencies use social media (eg Facebook and Twitter) to communicate events or incidents affecting the road or public transport network. Social media platforms can also generate vast amounts of crowdsourced user content. This content can be actively monitored, or potentially mined to yield useful information, including geotagged or descriptions that provide location information. Most of this content is unstructured and therefore difficult to interpret using automated tools. Data mining using semantic annotation, transport-based ontologies and geotagging can be used to interpret and convert this information into meaningful geospatial data[4]. This technique is used, for example, in emergency management to provide real-time, geo-referenced disaster reports (Yin et al 2012).

There is little research and few examples of social media mining specifically to support transport systems. Such methods are generally in the early stages of development and require complex methods of data mining and sentiment analysis (Center for Automotive Research 2015). One relevant article explored whether social media sources such as these could produce sufficient information of the necessary quality for transport system operators and TIS (Grant-Muller et al 2014). The authors have established a methodology for extracting transport data from social media sources, shown diagrammatically in figure 2.2. The conclusion from this research is that while social media data sources have potential within the transport space, there are particular challenges regarding data quality.

[4] Semantics is the study of meaning. In the web context, semantics refers to the interface between machines and humans, and the ability of machines to interpret and process unstructured human-generated content. This is supported through the use of ontologies – methodological frameworks for representing contextual information as a networked structure of objects or concepts (Grant-Muller et al 2014)

Figure 2.2 Social media text mining process for generating transport data (adapted from Grant- Muller et al 2014)



A social media analysis trial for traffic management purposes using Twitter was undertaken at the District of Columbia Department of Transportation (DDOT) in late 2013 (Fu et al 2015). Initially, DDOT extracted a set of keywords from historical traffic-related tweets by influential users. These were then used to extract tweets in real-time from the Twitter application programming interface (API)[5]. A keyword association algorithm was applied to construct Twitter queries using 'wordsets' rather than single keywords, as it was found that single keywords resulted in 'noise' in the extracted tweets. The trial included a user interface that traffic management centre operations staff used to obtain supplemental information that may be missing from existing incident profiles. Validation of the trial found that tweets were received much earlier than events recorded from other sources, and captured more detailed contextual information.

2.2.7 Other applications

A number of other applications and projects not specifically transport related were reviewed, exemplifying other types of crowdsourcing projects, applications and platforms employed in New Zealand and abroad. This is not an exhaustive review of other applications, but provides context as to the breadth of crowdsourcing undertaken in New Zealand.

[5] https://dev.twitter.com/overview/api

The Transport Agency has developed a zero harm strategy (NZ Transport Agency 2014c) to create a zero harm culture both within the agency and across the wider roading sector. To assist the agency in reporting on safety incidents a **Zero Harm** reporting tool has been made available free to Transport Agency contractors and industry partners. The tool is a cloud-based application that can be accessed through a range of devices. Whist the Zero Harm application is primarily focused on reporting workplace safety incidents, it is also being used more widely by the public including to report crashes, potholes and damage to roadside infrastructure, essentially providing crowdsourced information which may be of value to road controlling authorities. As the app becomes used more widely within the transportation sector and by members of the general public it has the potential to unlock further benefits.

Building our Footprints was a collaborative crowdsourcing project involving Land Information New Zealand, Environment Canterbury and the University of Canterbury engaging Canterbury high school students to map building footprints across the region. Nearly 20,000 footprints were mapped through a web application, which will be used to improve existing data sources.

Share an Idea was a community engagement campaign run by the Christchurch City Council following the February 2011 Christchurch earthquake. Ideas for the rebuild of the central city were gathered through a range of methods, including through the Share an Idea website, written submissions and sticky notes posted at a community expo. Over 100,000 ideas were generated and processed using specialist software and teams of data analysts into a format that could be used in the development of the Central City Recovery Plan.

Ushahidi is a free and open source software platform that supports crowdsourcing. Primarily designed as a flexible platform for empowering community-based applications, it includes tools for information collection, visualisation and mapping. Following the February 2011 Christchurch earthquake Ushahidi was used to develop the Christchurch Recovery Map which displayed both official information about essential services, as well as crowdsourced information regarding, for example, the location of free barbeques and water sources (McNamara 2011). The only transport application found in this review was a pilot project implemented by the Beijing Transport Research Center. This included an online urban transport public feedback portal for users to report issues regarding cycling and walking infrastructure via the web, smartphone applications, SMS (short message service) and social media (The World Bank 2012).

2.3 Licensing, ownership and privacy considerations

This section discusses the licensing, ownership and privacy matters that need to be considered regarding use of crowdsourced information. While this section covers a number of legal matters, is not intended to be legal advice and should not be relied on for this purpose.

Crowdsourcing applications can be developed and administered through public agencies, or developed by a third party using government open data sources or private data collections, including crowdsourced data. Third-party applications can be country specific, or have a global application. The choice of application and the types of data that underpin this application can therefore have different standards for licensing, ownership and privacy.

2.3.1 Relevant legislation and guidance

2.3.1.1 Privacy and official information

In New Zealand, information collected from individuals by public and private agencies is subject to the provisions of the Privacy Act 1993. In this Act, ‘personal information’ is ‘information about an identifiable

individual' and the privacy of this information is governed by 12 privacy principles that cover the following matters:

- the collection of personal information, including why it may be collected, where and how it is collected (Principles 1 to 4)
- how personal information is stored (Principle 5)
- the rights of individuals to access and correct information about themselves (Principles 6 and 7)
- how people and organisations can use or disclose personal information (Principles 8 to 11)
- how 'unique identifiers' can be used (Principle 12).

There are two laws that govern access to official information in New Zealand, the Official Information Act 1982 (OIA), which governs all official information held by the government, and the Local Government Official Information and Meetings Act 1987 (LGOIMA), which applies to local government bodies. Both Acts have a similar purpose in that they:

- make official information available to enable more effective public participation
- promote the accountability of the Ministers of the Crown, local authority members and officials
- protect sensitive information where necessary in the public interest or in the preservation of personal privacy.

If a government agency or local authority holds information about an individual, that individual has rights under the Privacy Act 1993 to access and correct this information. Requests for official information about an individual who is not the person requesting that information may be refused on the grounds that it breaches that person's privacy. In addition to privacy, requests for information can be denied for a number of reasons, including information collected by individuals for the purpose of public consultation is considered official information under the LGOIMA and the OIA.

For example, a request for the raw data collected during the Share an Idea project was considered by the Christchurch City Council under the LGOIMA in November 2012 (Gilbert 2012). This request was rejected on the basis of protecting personal privacy (section 7(2)(a)), confidentiality (section 7(2)(c)) and the substantial effort required to collate the information (section 17(f)). The informal nature of how the information was collected (for example – sticky notes, voicemail messages) means that personal information including name and contact details were scattered throughout the database. A process of stripping personal information from this database is currently underway so that the information can be publically released (Kath Jamieson, pers comm 10 December 2014).

2.3.1.2 Copyright and data ownership

Copyright is the property right that exists in some original works, as regulated by the Copyright Act 1994. Under this law, copyright automatically applies to original works, including written works (eg letters and emails) and artistic works (eg photographs, diagrams and maps). When public users add their own copyright content to a social media or other web forum, they retain ownership of the copyright unless expressly agreed otherwise. This content can be reused by the site or application owner only if the owner, either expressly or by implication, confers a licence to use that content, for example through the Terms of Use (DIA 2011b). For example, the terms of use for the 'Building our Footprints' project acknowledge the copyright of the individual who submits information, and the user agrees to license this work to the public domain under the Creative Commons Licence[6].

[6] https://creativecommons.org/about/cc0/

The New Zealand Government Open Access and Licensing framework (NZGOAL) fosters the sharing of government material for re-use by others, and provides guidance for government agencies to follow when releasing copyright and non-copyright works (State Services Commission 2010). Regarding copyright works that do not belong to government agencies, for example social media content, NZGOAL directs agencies to obtain an assignment of copyright from the relevant copyright owner(s), or obtain the right to sub-license this work on the terms of the copyright owners preferred terms (such as the Creative Commons Licence or as mandated through the Terms of Use).

In general, crowdsourced data should be made public and reusable through a creative commons licence, in terms with the principles of open data. An agency considering crowdsourcing should establish policy about how the data was to be held following collection, and how it was to be re-released without impacting on the privacy of individual contributors.

2.3.2 Responsibilities of public agencies

2.3.2.1 New Zealand Government Web Toolkit

The New Zealand Government's Web Toolkit (New Zealand Government 2014a) identifies minimum standards and guidance for the development of public agency websites, including interactive websites and web applications. A web application developed by the Transport Agency for the purpose of collecting crowdsourced information would need to meet these standards. This includes the Web Usability Standard 1.2 (New Zealand Government 2014b), which sets out the minimum standards, among other things, for:

- *Copyright* – websites must include a copyright statement, including rights to third-party copyright material.
- *Privacy* – websites must include a privacy statement that complies with the Privacy Act 1993 and details out how personal information is collected and used, and the rights of users to access this information.

Additional guidance is provided in the Web Toolkit for developing Terms of User for interactive websites, where users can contribute their own content (New Zealand Government 2014c). According to this guidance, Terms of Use may need to include:

- warranties that site users contributing third party copyright content have the right to use such material
- ownership and licensing of users' contributions.

In addition, government agencies need to consider how to identify and restrict access to private information that may be collected so they can meet their obligations under the LGOIMA and OIA.

2.3.3 Third-party applications

Usually third-party crowdsourcing applications set their own terms of use, offering privacy and copyright terms on a 'take it or leave it' basis. There may be little (if any) scope for a government body such as the Transport Agency to negotiate alternative terms of agreement, or to apply additional terms within an existing application. Other jurisdictions may also be exempt from New Zealand privacy and copyright laws. These limitations of third-party applications do not necessarily exclude them from being a source of crowdsourced information; however, there are specific matters that must be considered before engaging with them.

In 2011 the Department of Internal Affairs (DIA) provided guidance on the use of social media by government agencies. This also included detailed advice on the use of third-party social media sites where users can register and share information. Social media sites are similar to crowdsourcing applications in

that the user submits information that may be subject to privacy and copyright law. These guidelines are also considered relevant when a government agency is considering third-party crowdsourcing applications. Key points from this guidance are presented in table 2.2.

A clear message from the DIA guidelines is that some third-party applications can be suitable for use by government agencies, provided key matters such as those listed in table 2.2 are assessed and addressed.

Table 2.2 Matters to consider when using a third- party crowdsourcing application

Matter	Consideration
Terms of use	The agency should consider the extent to which it is able to include its own terms of use or a separate set of rules covering the provisions of Terms of Use Standard [(a)]
Copyright	If possible, agencies should include terms requiring site users to license the agency's publication of their content, to the extent it is protected by copyright, for the agency's intended purposes.
Privacy	Agencies should include a privacy policy, if possible, or link to a privacy policy, which is fit for purpose, on another site the agency controls. This privacy policy must be consistent with any privacy policy on the third party site. The agency must also be confident that there will be no loss or misuse of personal information that may be collected from citizens through its use of the service.
Records and requests	The third-party applications should not prevent agencies complying with record-keeping obligations under the Public Records Act 2005 or their ability to respond to requests for information under the Official Information Act 1982.
Communication	The agency must consider whether the service provider reserves the right to communicate direction with users (eg commercial messages) and whether this is appropriate given the context.
Credibility	The agency must be confident that the service provider is credible, trustworthy and likely to represent stable propositions for the duration of the project. The agency should also consider the jurisdiction in which the service provider is based, particularly in regard to the treatment and protection of personal information and other data, and the enforcement of contractual rights.

[(a)] Now superseded by the New Zealand Government Web Toolkit (New Zealand Government 2014a)

2.3.3.1 Business models for procurement of third- party data

The procurement of third-party crowdsourced data may require new thinking about the development of business cases for sourcing data, particularly when there are potential concerns around privacy and data ownership. There is currently little guidance in this area within the New Zealand context.

In *NZ Transport Agency research report 559*, Smith et al (2014) investigated risk management and business case development for the procurement of emerging digital data sources from commercial GNSS and Bluetooth data providers. The authors found that the Transport Agency's existing forms of contract for the procurement of professional services are generally appropriate as a basis for digital data supply agreements. Key actions identified to reduce risks associated with third-party data procurement for emerging data sources included:

- early engagement with suppliers
- a flexible approach to contractual dealings, particularly during initial phases
- short term contracts that allow flexibility to fine-tune the specifications of the data stream initially
- ensuring the Transport Agency is not overly constrained in the use of this data.

2.4 Safety considerations

Crowdsourcing applications and TIS provided through a mobile interface can be a source of distraction for motorists. Between 2011 and 2013, 163 drivers in fatal or injury crashes contributed to the crash by being distracted by a cell phone (Ministry of Transport 2014a). Despite the risks attributed to diverted attention, public understanding of distraction is low (Ministry of Transport 2010).

A comprehensive literature review on distracted driving and in-vehicle devices was recently prepared by Rickard-Simms (2014) as an internal report for the Ministry of Transport. This section relies heavily on this review and this resource is acknowledged accordingly.

2.4.1 Distraction

Drivers can be distracted in a number of ways, affecting the ability of the driver to control a vehicle and respond to conditions in the road environment. Ideally, crowdsourced data collection technologies and applications should not increase the workload of the driver or cause an unacceptable level of distraction. Types of driver distraction can be loosely classified into one of three groups (Rickard-Simms 2014):

Physical distractions – distractions that require a driver to physically engage with and use, such as food, devices or objects. These distractions are particularly problematic due to 'glance time' – where a driver's attention is diverted away from the road and onto the device or object.

Cognitive distractions – distractions which may take the driver's mind away from the task of driving, for example, having a conversation via a hands-free device, or listening to music.

Emotive distractions – any distraction that engages with the driver and is capable of affecting a driver's emotions, disrupting the driver's focus at very high levels (Chan and Singhal 2013; Astleitner 2004). Emotive distractions are the most problematic and risky for drivers, and inexperienced and young drivers are particularly at risk (Astleitner 2004).

The type of distraction involved in active crowdsourced information collection will vary according to the device and application used. Rickard-Simms (2014) provides a number of relevant conclusions from the literature that will need to be considered in more detail when developing a crowdsourcing application:

- The driver is more at risk if the distraction involves more than two seconds of glance time, or if the driver glances at the device repeatedly over a short-period (Zhang et al 2013).
- Audio output from navigation and music systems does not negatively affect crash rates (Ünal et al 2013; Sodnik et al 2008).
- Voice recognition inputs are less risky than physical input methods, but can be less accurate (Tsimhoni et al 2004). Cognitive distraction through common voice tasks is generally more demanding than natural conversations, listening to the radio, or listening to a book on tape (Cooper et al 2014).
- The screens of mobile devices can be small, and therefore require longer periods of glance time (Owens et al 2011; Ishiko et al 2013).

2.4.2 Legal issues and implementation

In 2009, the Land Transport (Road User) Rule 2004 was amended to prohibit drivers from using mobile phones for phone calls, text messaging and for other types of communications and purposes. Drivers may, however, use these devices if:

- the phone is secured in a mounting fixed to the vehicle
- if the driver manipulates or looks at the phone, he or she does so infrequently and briefly.

Studies following the implementation of this legislation found there has only been a slight reduction in usage while driving (Wilson 2013). In a review of legislation other jurisdictions, Rickard-Simms (2014) found similar outcomes that could be attributed to a number of reasons, primarily:

- the perception that the risk associated with driving distraction is small
- drivers do not consider the penalty high enough to change behaviours
- the legislation is difficult to police
- the ubiquitous nature of mobile phones and a high level of desire to use them in-vehicle.

As the range of technological devices available for use in-car continues to grow and evolve, the legislation must also adapt and respond. Because of this, the Ministry of Transport intends to review New Zealand legislation relating to driver distraction from in-vehicle devices by the end of 2016 (MoT 2014b).

2.4.3 Safety mitigation measures

There are many suggestions in the literature for making in-vehicle devices safer for drivers to use, including forcing drivers to have these devices mounted, restricting touch screen input while the vehicle is moving, and using audio/voice functionality rather than the visual display (Jamson 2012; Tsimhoni et al 2004; van Nes et al 2012).

Of the applications reviewed in appendix A, those most likely to distract drivers are those designed for private vehicle use and require active and ongoing user inputs. Depending on how they are used, these applications display many potentially distracting items on a single screen, including road hazards, traffic conditions and navigational instruction. In some applications, this hazard is mitigated, to some degree, by restricting text-input while the car is in motion, and allowing drivers to enter alerts using voice commands.

2.5 Validating crowdsourced information

Unstructured (and potentially malicious) submitted content requires administration and checking for data quality. The type of evaluation employed depends on the type of data collected and the standard of quality required for the output information. Depending on the application and type of information, user contributions can be checked using either manual screening or moderation of user inputs, and/or automated screening using predefined criteria or algorithms. Malicious activity can be detected using both manual and automated processes. In some applications, the users themselves verify other users' content through reporting or verification processes.

Most studies on this topic focus on quality assessments for crowdsourced information, or addressing the question of how an organisation can trust these data for incorporation into their own datasets, or release to the public alongside their own data (Johnson and Seiber 2012, Severinsen 2015). When considering the integration of crowdsourced data, government agencies must therefore make a policy decision, addressing the risk of relying on certain pieces of data, while also enabling the realisation of its value.

Severinsen (2015) proposed a trust model to address this balance, which was deployed during the Building Our Footprints case study. The trust model used a number of case specific parameters to algorithmically generate a trust 'rating', and an arbitrary threshold was set for re-release to other members of the public. Data which met this threshold was automatically considered reliable, while the agency's quality assurance

resource could then be more efficiently deployed to assess data which fell below this threshold. A different quality threshold could be set or each agency or data type, depending on the need of its end consumer.

In the case of real-time traveller information, a logical distinction could be made between events or reports classed as severe, such as a crash, or those classed as more common or benign, such as wet weather. It is logical that a road controlling agency could consider the latter example using algorithmic trust models, whereas the former situation should receive human intervention to verify its accuracy. The controlling agency must therefore make a considered policy decision about where to impose such a threshold.

In developing a crowdsourced application, a critical decision must be made about where to set a threshold for trust in data. This question could be phrased as, 'which types of report data can we trust through automation, and which require human verification?' Once this question has been answered, there are a number of ways to algorithmically determine trust in these data.

2.5.1 Trust modelling

Assessing trust in, and inferring quality of, crowdsourced data is a topic that has been heavily researched in recent years. Research has focused on a number of different trust indicators, including trust in a data author, assessing the spatial precision of data, modelling its temporal quality, and inferring quality through semantic attributes. The first three methods of quality assessment have been thoroughly investigated, while the latter – semantic considerations – is still a relatively new area of study. Severinsen (2015) amalgamated data author, spatial, and temporal elements of trust into a single model, which could be further enhanced through the assessment of semantic trust and crowdsourced ontologies.

All of these previous studies have relied on the post processing of crowdsourced data to assess its quality, or trustworthiness, or in the case of Severinsen (2015), to critically examine the veracity of a trust model. This research focuses on real-time traveller information applications that could include a trust model to run on contributors' volunteered data. The research in this area is sufficiently mature to inform the development of a trust model for real-time traveller information, and could be applied most appropriately where an agency has developed a bespoke application for collecting information.

A base trust model in this domain could consider some or all of the following parameters:

1 Information about the author of data, including:
    a familiarity with the local area (this could be inferred from their address).
2 An assessment of spatial precision, including the:
    a GNSS metadata of a particular device, and/or
    b proximity of the contributor to the actual event, and
    c proximity of the event to the transportation network.
3 Temporal quality indicators, particularly:
    a number of like or similar contributions, and
    b change over time.

2.5.2 Assessing the contributor

The need to collect this information is linked to the engagement model used for crowd participation. For example, an effective way to collect information about a data author is to operate a log-on/user profile system through a web or mobile application. When creating a profile/account for the application, a user would be required to enter certain details about themselves, such as a home address. In reality, the provision of personal information can often be seen as a barrier to participation in crowdsourcing applications, where users often desire anonymity. This can be contrasted with open applications where users can contribute and consume data without a sign-in requirement or the need to release personal information. A decision would need to be made as to whether the benefit of collecting this information outweighed the possible impact on widespread participation in the activity.

Goodchild (2007) discussed the concept of a person's 'activity space', which suggests that a person has valuable and accurate knowledge about the area in which they spend the majority of their time, and any contributions that person makes within the area are likely to be of good quality. In this case, a person's home and work addresses could be considered activity spaces for the purposes of modelling trust, and value could be added to contributions from those people in the areas.

An assessment of a data author is considered valuable in terms of estimating overall trust; however, an agency deploying a trust model must weigh this benefit against any perceived barriers to participation, and also account for additional policy and infrastructure required to secure personal information.

2.5.3 Assessment of spatial precision

Spatial precision or spatial trust is an important factor when assessing the reliability of crowdsourced geographic information (CGI). CGI inherently contains data or metadata of a uniquely spatial nature, and as such the accuracy of its location or extent is important, particularly when considering reuse of that information. In addition to other factors influencing trust for general crowdsourced information, spatial precision or trust must be taken into account for CGI.

Spatial precision can be inferred by a number of methods. These methods differ depending on whether a mapping and reporting application is used via a mobile device or on a desktop computer.

For data collected from a mobile device, it is a straightforward task to collect metadata about that device's spatial precision. This information could include the manufacturer's margin of error for GNSS devices, or whether the device is collecting location from direct GNSS measurement, the triangulation of cellular network signals, or a mixture of both.

If a piece of data is digitised on a screen using a computer, the inferred spatial precision of that feature is directly correlated to the zoom scale at which that feature was created, ie the more zoomed-in the screen was at the time of submission, the more likely it is that the user correctly placed the feature spatially. This is often based upon the enhanced quality of imagery at smaller scales.

If an agency wanted to only collect data related to a transportation network, it would be useful to ensure all crowdsourced features 'snapped' to the existing transportation network data, rather than allowing the user to create features at places spatially removed from the transportation networks.

A further key measure is to assess the location of the contributor in relation to the report that they have created. Contributions are to be more highly trusted if the contributor is directly observing the event in question, rather than reporting from a distance and at a potentially later and less relevant time. This information could easily be gathered from a device's location services, or from a computer's internet protocol (address) at the time of contribution. In essence, spatial trust in a feature is inversely proportional to the distance between that feature and the location of its creator at the time of creation.

2.5.4 Temporal trust

The change in status or location over time is an important concept to consider when investigating the precision and accuracy of a crowdsourced feature. Rapid change to a feature can signal that it is dynamic in the real world, confirming crowdsourcing as a valuable source of up-to-date data. It is, however, important to consider that not all changes over time are the result of poor quality, or that a lack of change means that the original data is any less accurate.

The lineage of data becomes particularly important when it is considered that 'real-time' information is the most valuable to travellers and RCAs. A number of studies make use of OpenStreetMap (OSM) data to explore this lineage, and to therefore make inferences about the quality of those features (Mooney et al 2010, Haklay 2010, Du et al 2012). Crowdsourced information in its most basic form is time-stamped at its creation (Aragó et al 2009), but of particular use in the case of many applications is that all edits are logged and can be tracked with associated timestamps. There have been two approaches to considering the changes that can affect a feature at any stage after its inception. The first, as considered by Haklay et al (2010) is that each amendment can be considered a refinement of the original feature, and that the associated improvement in overall quality is a reflection of the applicability of Linus' Law. The second approach, as defined by Mooney et al (2010) and refined by Trame and Keßler (2011), interprets each time-stamped amendment to a feature as a reflection of a real world change to the object represented, and as such, captures and allows modelling of change over time.

Aragó et al (2009) propose two formulae for assessing another theory on temporal quality – that no change over time may represent no change in the real world, or in fact an endorsement of the quality of the original feature. The first they label 'change ratio', which assesses the number of changes made to a feature in the time between its creation and final edit. The second they describe as 'contribution ratio', which seeks to measure the total number of both changes and endorsements to a feature within that same time period. Both ratios account for times a feature has been viewed but not changed, and therefore endorsed. These representations of process can serve to identify areas of rapid change, or conversely show areas were little change is evident.

Ye et al (2011) suggest using timestamps to classify features and explore semantic inferences using user check-ins at points of interest. The authors argue that it is not simply physical space that defines any given feature, as that space may change in its use and purpose at particular points in its lifecycle. For example, during the day or business hours, a particular establishment may be classified as a cafe or restaurant; however, as time progresses into evening, this could more appropriately be considered as a bar or nightclub.

2.5.5 Semantic quality

Of all the measures of trust for crowdsourced data, semantic trust has been the least studied, but perhaps holds the greatest potential. Free text contributions often hold a wealth of information which is difficult to extract, but may not explicitly define a location, event, issue or timestamp.

Adams and McKenzie (2013) emphasise that people's descriptions of events capture subjective experiences, and therefore can be used to develop a large number of analytics about any given crowd of contributors. Of most relevance to real-time traveller information is that a semantic analysis can also reveal spatio-temporal themes, ie an assessment of how widespread a phenomenon might be, such as road ice. Equally, semantic analysis can predict how localised a particular event is, such as a vehicle collision or a slip/washout.

Crowdsourced text content also reveals significant regional variation, where people's description of various phenomena is often affected by their own context and paradigms. For example, a person's definition of a mountain versus a hill will differ depending on the size and nature of the terrain with which they are most familiar. The same is true for traveller information.

Kalbasi et al (2014) propose and demonstrate a tool that can be used to analyse content semantically, through the development of crowd ontologies[7]. Kalbasi et al (2014) recognise that while an agency can initialise the development of an ontology, these basic building blocks could never fully describe the level of complexity required to build an effective ontology for widespread crowdsourcing applications. Instead, they propose a modular framework known as 'collaborative ontology modelling' which they demonstrate as effective through the use of specific software packages. The collaborative ontology modelling framework would not only allow for the crowdsourcing of geographic data, in this case traveller information, but also for the crowd to build and develop its own ontologies for the translation of subjective meaning held within textual threads. This system could differentiate a text thread about road ice from a text thread about having an icy drink on a road trip, for example.

The use of semantic understanding would add real value in the case of traveller information, as this would enable the use of existing social media such as Facebook or Twitter. To date it has been difficult to extract meaningful data from these services due to blurring of context (the icy road vs icy drink example above), and the inability to infer location from text. Semantic and ontological trust modelling would enable this value to be extracted from text data and social media more widely, and would augment any other trust model or bespoke application developed.

2.5.6 Data integrity

In addition to determining the quality of, or trust in, crowdsourced information, an underlying risk that must be addressed is the potential for individuals to use crowdsourcing platforms for personal gain or malicious intent (Goodchild 2007; Coleman et al 2009).

As social media and collaborative web projects come of age, an increasing number of businesses rely on crowdsourced information for promotion and advertising. Tripadvisor, for example, is an extremely influential crowdsourcing application in the hospitality and tourism industry, where hotel or restaurant customers are able to rate, or provide feedback about, their experience in a particular establishment. Other users of the site, who are prospective customers of a business, often use this publicly available feedback to influence their choice of accommodation or restaurant in the future – a feedback loop that can have significant financial implications for business owners. Harris (2011) asserts that people who stand to benefit from crowdsourced information can and do submit false information for personal gain. A business owner in the hospitality industry would benefit from an increase in positive reviews, while the users of this information would have difficulty distinguishing genuine from false content.

A similar concept could therefore be applied for transportation information. A business owner, for example, on a particular route could supply misinformation signalling congestion on other roads, thereby encouraging increased traffic flow past their own business, as travellers attempt to rely on crowdsourced information in order to avoid delays in their journey. Other malicious uses of data also exist – it would not be difficult for a disgruntled member of the public to use crowdsourcing sites to publicise grievances against a particular agency or individual, or, at the extreme end of the spectrum, to plant a software virus within datasets (Kershi and Clark 2012). Such misinformation – both positive or malicious – needs to be identified and addressed if any authoritative agency or roading authority is to rely on crowdsourced

[7] Ontologies are methodological frameworks for representing contextual information as a networked structure of objects or concepts (Grant-Muller et al 2014)

information for integration into their own datasets (Wihbey 2014; Starbird et al 2014; Elwood et al 2012; Harris 2011; Kanhere 2011; Mashhadi and Capra 2011; Nunes et al 2014; Doan et al 2011).

Heipke (2010) observed that the field of crowdsourcing was in its infancy, and as such there had been no known attacks on any crowdsourcing platforms or projects, with members of the crowd demonstrating overwhelmingly benevolent behaviours. Conversely however, and perhaps not intentionally, there can be a clear pattern of misinformation present on social networks or in any crowd generated content (Starbird et al 2014). Ideas can spread virally with relative ease throughout a crowd or network when emotive topics are involved (Starbird et al 2014). Similarly, when the information is spread by a 'viral super user' of a network – one who is very well connected and highly visible within the network, misinformation can proliferate (Doan et al 2011),

Starbird et al (2014) conducted a series of experiments on Twitter data that was, at the time of analysis, known to be false. Their studies identified that for the majority of misinformation spread through the public domain, there existed clear corrections to these data by other users, although number of corrections to false data is startlingly low when compared to the viral rate at which misinformation is spread.

Most studies do however present many unified themes about how misinformation (positive or malicious) can be identified and addressed in crowdsourcing systems. Threats to data integrity could arise from fake or 'bot' profiles, or simply from information that is inherently wrong. Most studies identify one or all of the following methods for addressing the threat of misinformation, which are addressed further below:

1 Use a predefined set of business rules to verify the validity or logic of data

2 Establish a reputation based assessment of data contributors, such as trust in a person

3 Implement a stratified system of peer review for data, as effected by OSM.

2.5.6.1 Use of a predefined set of business rules to verify the validity or logic of data

Elwood et al (2012) suggest that crowdsourced information should be compared for integrity against other readily available related datasets, and based on some form of logic. In the transportation domain, a report of heavy snow blocking a road, or severe icy road conditions, could be verified against weather reports from the time period of capture. Similarly, a report of a car crash could be assessed for its proximity to a road or highway, where it is most likely that such an incident would occur, or reports of congestion could be correlated against mobile phone density data for relevant cellular towers.

Kanhere (2011), Mashhadi and Capra (2011), Elwood et al (2012) and Severinsen (2015) focus on the proximity of a contributor to their contributed data as a means to determine its quality. Linked to the idea of Goodchild's (2007) 'Activity Space', a simple proximity calculation could be derived in order to determine if a contributor is in fact personally viewing the event in question. Modern mobile devices all make use of location services or geotagging, and a computer's location can be ascertained through IP address. This method could also be used to identify and address positive misinformation, such as a large number of contributions originating from a single device or location.

A further trigger for investigation is the presence of corrected data – deliberate or accidental misinformation could be identified when members of the crowd are actively altering the contributions of others (Starbird et al (2014). By using this attribute as a trigger for investigation, a possible solution for addressing viral data surges is proposed. Starbird et al (2014) identify that large surges of misinformation are almost always countered by corrections from other members of the crowd, even if the number of corrections is relatively small. The presence of data corrections can serve as an indicator for possible misinformation, triggering further investigation.

Such business rules are an effective means of identifying suspicious data but should not be used to reject crowdsourced contributions arbitrarily. Rather, the use of business rules is an effective way of identifying where to best deploy quality assurance resources, particularly for government agencies working with crowdsourced information.

2.5.6.2 Establish a reputation based assessment of data contributors

Doan et al (2012), Mashhadi and Capra (2011) and Harris (2011) suggest that a reputation system based on user profile is the most effective means of handling misinformation. There are several factors contributing to online reputation, which are summarised by Severinsen (2015) and addressed elsewhere in this report.

An important policy decision for any agency implementing such a system is whether to make use of 'identity intensive' or 'identity extensive' security systems (Harris 2011). By collecting more information for a user profile, including verifiable but necessarily private information, an inference can be made about that user's reputation and therefore the quality of data that they contribute. Such a system also gives assurance that a user profile is attached to a real person, rather than a fabricated identity or a computer 'bot'; however, this must be considered in light of any barriers to participation that these layers of security may present. The extent to which an identity intensive profiling system is implemented is a policy decision that must be made on a case-by-case basis by any agency wishing to leverage the value of crowdsourced information, and is most useful when applied to facilitated crowdsourcing applications (Severinsen 2015).

2.5.6.3 Implement a stratified system of peer review for data

Elwood et al (2012) and Doan et al. (2011) suggest that the simplest and most effective means of reducing malicious or accidental misinformation in a crowdsourcing system is through user stratification. This system relies on a layer of 'super-users' who are, by virtue of qualification or experience, trusted as the 'gatekeepers', or approvers of information in a crowdsourced information database. Effectively implemented by OSM and Google, this system is probably the easiest solution for regulating data quality, but comes at the cost of speed and efficiency. In times of high data flow, a stratified system could result in bottlenecks for data processing and delays in the completion of feedback loops through the delivery of data.

Issues of data integrity and misinformation therefore present difficult-to-quantify, but very real risks for crowdsourcing applications, and particularly agencies wishing to rely on crowdsourced information. Previous studies reveal a strong tension between methods to mitigate these risks – either through regulation, such as strict profiling or detailed business rules, or by establishing trust in person or data. It is clear that to gain the most value from the crowd and promote freedom of participation, crowdsourcing agencies must move away from a regulatory environment and towards a verification system based on trust. It is also evident, however, that an agency must continue to effectively manage its own risk when using these data, and that a realistic crowdsourcing system should strike a balance between regulation and trust as means to examine data integrity, which will vary depending on the needs of each situation or case study.

2.6 Practical implementation factors

Successful crowdsourcing involves more than releasing an application or project and expecting an engaged crowd to adopt and use it in sufficient numbers to generate useful outputs. Effective crowdsourcing requires developing a product that provides an incentive for users to be involved in and collects information in a format and of a quality that is useful to the agency. Doan et al (2011) identify four major areas that crowdsourcing systems need to address to be successful: a) recruiting and retaining

the participant base; b) assessing user capabilities; c) aggregating the information provided by users; and d) evaluating the contribution of users.

2.6.1 Encouraging participation

There must be a compelling reason for people to want to participate. Methods to recruit and support ongoing involvement include:

- Marketing promotions to make people aware of the application or project, including the benefits of being involved and the effort required. Recurring campaigns and marketing strategies will also ensure that people remain curious and involved (Misra et al 2014).
- Incentives are a key driver for attracting ongoing participation and may include financial rewards as well as less tangible incentives such as instant gratification, user rating systems and competition.
- Building a social community around the application and engaging with social media works for many applications.
- The application or project should be fun and entertaining to engage with and designed with ease-of-use in mind.

There are numerous examples of how these methods have been put to use in crowdsourcing applications. Building our Footprints, for example, used competition, prizes and rankings to encourage users. Applications such as Waze, Moovit and Street Bump use point or ranking systems.

An effective method of building engagement is to combine both the 'push' (customer information) and 'pull' (crowdsourced data) in a single platform. Severinsen and Shah (pers comm) developed an application 'Rate-it' in post-earthquake Christchurch that demonstrated this. Rate-it was a mobile mapping application built for Android devices that targeted tertiary students in post-earthquake Christchurch. In the post-disaster city, as all commerce and occupancy moved from the central city to satellite 'hubs' in various suburbs, social venues such as cafes and bars also moved to disparate clusters. It was identified that tertiary students do not usually possess sufficient financial means to regularly travel between these hubs, therefore this group desired real-time intelligence about the atmosphere of venues before making a decision to travel there.

Rate-it captured location points for existing venues, and plotted these on a map. Users of the application could then rate each venue according to how busy it was, on a simple three-point scale – quiet, moderate, or busy – recognising that no one category is more desirable than the others. For example, a student may be wanting a quiet café in which to work, or a student may be looking for a busy nightclub in which to dance. Different contributions were assessed for trust using an assessment of similar contributions, temporal relevance, or currency, and the proximity of the observation to the subject location.

Critically, the application provided a single map interface. Each venue was symbolised according to its business rating, as either a small circle (for quiet), a medium diameter circle (for moderate), or a large circle (for busy). Users could contribute by simply clicking on these symbols and volunteering data. The effectiveness of this application was built on the simplicity of its interface – simultaneously collecting and delivering real-time data.

2.6.2 Assessing user capabilities

In developing or considering crowdsourcing options, it is important to consider the types of users and their abilities, including how demanding the activity and how users can contribute (Doan et al 2011). The information collection activity (or activities) should match the abilities of the users. One method of

supporting different users is to provide multiple levels of involvement, with users either choosing their level of involvement, or a ranking or point system that unlocks opportunities for higher levels of activity.

Some applications use a point system to track edits to their base map. The more points a map editor accumulates, more areas of the map are unlocked for them to edit. Another method of reducing the cognitive load on the user is to use predefined attribute-value pairs to limit what the user can report. For example, only allows users to select predefined types of map corrections, for example '(un)block street', 'reverse traffic direction', 'edit street name', 'change speed limit' and 'missing street'.

Doan et al (2011) also distinguish between the divisions of tasks for machines versus humans. For example, measuring congestion using travel time data collected by GNSS is simpler and more objective than relying on the users to report this on the road. Human inputs are better at identifying quality or identifying events, for example, reporting a hazard on the road or rating the quality of a bus service.

2.6.3 Aggregating information

Aggregating data from crowdsourcing participants can be a complicated task given the volume of responses received from a diverse pool of crowd participants (Misra et al 2014). Integrating these inputs with other external data sources such as sensor network present an additional challenge. Aggregating data can be labour intensive or cost intensive as complex data management systems and processes are developed to reduce sources of human error (ibid)

The Transport Agency will also need to consider how the crowdsourced information is referenced against existing transport datasets, for example traffic count sensor, CCTV inputs and GNSS/Bluetooth sensor networks. Issues with the integration of GNSS and Bluetooth data sources, which are provided with varying levels of detail and granularity, with the Transport Agency's existing road centreline dataset(s) were highlighted in Smith et al (2014).

Issues with integrating information sources, including crowdsourced information, for the purposes of customer information delivery were explored by Chang et al (2015). A number of tasks were identified to improve information provision, including the establishment of a transport industry working group to share learnings about information projects, the normalisation of data sources (including crowdsourced information) and agreement on data exchange standards.

Protocols for information sharing such as APIs and open standards can be used to support the integration or 'mash-up' of multiple sources of data. The HackAkl and OpenHack Christchurch events used standardised data feeds such as XML and APIs for real-time bus and congestion data to develop innovative transport applications.

2.6.4 Organisational factors

Sorenson (2010) explores the types of challenges organisations, particularly transport agencies, may face when implementing crowdsourcing applications or technologies, particularly:

- the ability of the organisation to embrace and implement new technologies
- the extent to which internal guidelines, procedures and organisational culture may preclude crowdsourcing
- risk management processes
- accepting that there is no guarantee that a suitable and capable crowd will participate in the project/application.

Another challenge for organisations involved in crowdsourcing is a workforce that has both the skills and capacity to implement and administer new data collection applications and technologies. This is recognised by the Ministry of Transport's ITS Action Plan (2014b) as a challenge for the wider ITS industry.

Government agencies in New Zealand have generally embraced social media applications (Controller and Auditor-General 2013) and subscribe to open data principles that generally support crowdsourcing projects. Existing government guidance regarding the use of social media within government organisations addresses these matters, including management, the implementation of social media projects and the culture around its use (DIA 2011a; DIA 2011b; DIA 2012 and Controller and Auditor-General 2013). For example, the DIA (2012) published a guide on how to handle mishaps such as hacking and abuse in social media channels. Controller and Auditor-General (2013) reports learnings from public entities' use of social media, discussing matters such as risk management, leadership and implementation in some detail.

3 Strategic assessment

This chapter considers the strategic transport direction and needs of the New Zealand public and the Transport Agency in relation to real-time crowdsourcing applications, focusing on the extent to which crowdsourcing supports real-time network efficiency outcomes and the provision of traveller information.

3.1 New Zealand strategic context

In New Zealand, the Ministry of Transport and the Transport Agency set the national strategic direction for the road network. The primary drivers for both organisations are to improve the performance of the transport network and deliver value for money from transport investment by the government.

Within the broader strategic transport context in New Zealand, there are synergies between the strategic approaches taken by the Ministry of Transport and the Transport Agency in the use of new and emerging technologies to provide and improve travel time reliability, traveller information and more efficient provision of transport. The general consensus is that as the cost of information technology and information systems decreases over time, this offers 'new opportunities to lift the performance of the transport system' (Ministry of Transport 2014d) and influence the choices made over the coming years to improve the efficiency of the transport system.

3.1.1 The role of technology in transport

The Government Policy Statement (GPS) on Land Transport 2015/16 – 2024/25 (MoT 2014c) acknowledges that responding to developments in technology is a challenge for RCAs and the Transport Agency. Similarly, the Connecting New Zealand document (MoT 2011) identifies new and emerging technology as one of the big issues for transport systems, particularly 'how transport systems can adapt quickly to take advantage of new technologies to address many of the key issues facing the transport sector'. Connecting New Zealand sets out the overall objective of the government which has a focus on improving network efficiency and with the support of ITS, the transport network will help to achieve the government's broader goals.

The GPS recognises that improvements in technology can make a positive contribution and that 'there is considerable scope for innovation in the way that the land transport system is delivered'. The GPS makes reference to the findings of the Government's Intelligent Transport Systems Technology Action Plan 2014-18 (MoT 2014b) and that improvements have already been seen in the availability of real-time travel information through the greater use of GNSS and smart phones.

Accepting the important role of this technology, the GPS states that it can greatly support new travel demand management initiatives. Within the scope of the current GPS is the securing of these and other productivity improvements achieved through cost effective investment in existing technologies. The GPS recognises that technology will play an increasing part in managing network access and capacity over the coming decade. However, while the GPS 2015 provides for reporting on innovation and technology investment, it states that it 'does not endorse any specific form of technology in view of the speed of evolution' (MoT 2014c).

3.1.2 Technology and the delivery of efficient transport systems

With regards to funding and investment, the GPS has multi-class reporting lines, one of which is 'innovation and technology'. This is defined as 'investment in improved systems and associated

technology, including any pilot investments' with the aim to understand the benefits and costs associated with innovation and technology (MoT 2014c).

The MoT's (2014d) *Statement of intent* (SoI) envisages that ITS will offer New Zealand major opportunities to improve transport productivity, safety and efficiency in the medium to longer term. By advising the government on the ITS Technology Action Plan, the Ministry of Transport (2014b) expects that decisions can be made on how new technology can be applied in the New Zealand context.

The One Network Road Classification (ONRC) seeks to create a better and more transparent way to plan and invest in maintaining and operating New Zealand's road network and in turn help local government and the Transport Agency make better decisions based on the functions of our roads. To assist in achieving this, a number of performance measures have been established, one of which is travel time reliability. The ONRC states that being able to deliver on the predictability of travel time will 'rapidly improve as technology improves (smart phones, big data)' (The Road Efficiency Group 2014). Ensuring there is adequate travel information for road users so they are 'advised well in advance of issues affecting network performance and availability' has also been identified as one of the customer level of service outcomes in relation to the resilience performance measure.

The Transport Agency's statutory objective under the Land Transport Management Act is to 'undertake its functions in a way that contributes to an effective, efficient and safe land transport system in the public interest'. In order to achieve this, the Transport Agency aims to 'shape smart transport choices' and to 'incentivise and shape safe and efficient travel choices using a customer-focused approach' (NZ Transport Agency 2014a). Within a transport network that is fully integrated, the Transport Agency considers that customers should be able to make 'really smart choices about their driving, vehicles, route and timing' and will concentrate on 'assisting customers to make safe and efficient travel choices'.

Part of this objective is that 'transport users in key urban areas can readily access information on safe and efficient travel choice' and that by 2014/2015 'multi-modal information is available in the three major urban areas encouraging greater customer choice for mode, route selection and time of travel' (NZ Travel Agency 2014a). The overall net effect of customers making smarter transport choices will be a more efficient and effective transport network.

The *NZ Transport Agency position statement on intelligent transport systems* (NZ Transport Agency 2014b) identifies specific investment areas for ITS. High priority ITS investment areas that have relevance regarding crowdsourced information and technologies include:

- mechanisms for collecting quality data about the use of the network
- better-quality data to drive better operations, planning and investment
- more active network management
- mechanisms that enable the delivery of accurate information to travellers to promote smarter transport choices.

Crowdsourcing is explicitly identified in the position statement as a future information source for real-time and historic information about traveller and vehicle movements. However, the position statement does not explicitly identify the Transport Agency taking a lead role in developing crowdsourcing technologies – instead identifying the agency's role as:

- a facilitator and regulator of new wholesale quality information services
- a market follower in terms of progress and new developments
- encouraging new market opportunities (regarding information) and ensuring information quality.

3.2 High-level needs assessment

The literature review demonstrated that crowdsourcing can be used to inform a wide range of transport applications; however, not all applications are relevant in the New Zealand context nor will they meet the objectives of this research. To identify strategic transport needs for New Zealand, a number of policy and strategy documents have been consulted. In addition to strategic transport needs, the requirements for TIS have also been identified using previous Transport Agency research reports. A short summary of the objectives/needs from each document is provided below. Key themes across all these reports and documents have been identified and form the basis of the options scoping assessment (presented in section 3.3).

3.2.1 Transport needs

One of the specific research objectives was to determine the extent to which real-time or near real-time crowdsourced data is a valuable, trusted and safe means for the Transport Agency to collect data in the context of the agency's long term strategic outputs articulated in the 2010–13 SoI (NZ Transport Agency, 2010). There are four long-term outcomes for the transport sector as published in the SoI as follows:

- An efficient transport system that supports high levels of economic productivity, provides strong international connections for freight, business and tourism, and meets international obligations.
- A sustainable funding basis for transport infrastructure investments and use.
- A high-quality transport system for Auckland, the nation's largest economic hub.
- An accessible and safe transport system that contributes positively to the nation's economic, social and environmental welfare.

There is consistency between the earlier SoI (NZ Transport Agency 2010) and the more recently published SoI 2014–18 (NZ Transport Agency 2014), which has four long-term goals focused on improved efficiency, effectiveness, resilience and safety of the road network, enabling smart transport choices and maximising returns for New Zealand. The direction of the transport sector is articulated through eight impact statements which combined describe the desired long-term effect of the Transport Agency's work programme. Improved transport network efficiency achieved through the delivery of technological initiatives such as crowdsourcing has the potential to achieve all eight of the desired impacts (NZ Transport Agency 2010):

- *Better use of existing transport capacity* – existing capacity can be used more efficiently by reducing congestion and reducing the variability in travel times on the road network.
- *More efficient freight supply chains* – less variability in travel times on the network especially on national strategic freight routes.
- *A resilient and secure transport network* – the resilience of the network can be improved by reducing the variability in travel times on the road network.
- *Easing of severe urban congestion* – less congestion is a direct outcome of initiatives that improve the efficiency of the network.
- *More efficient vehicle fleets* – by reducing congestion on the road network the energy use per vehicle kilometre will also reduce.

- *Reductions in deaths and serious injuries from road crashes* – by reducing congestion and improving travel time reliability there is likely to be less driver impatience and risk taking which may lead to adverse safety outcomes.
- *More transport mode choices* – this desired impact can be largely supported by improved traveller information; however, it is noted that improved network efficiency can make public transport a more desirable mode choice.
- *Reduction in adverse environmental effects from land transport* – vehicles emissions can be reduced through more efficient use of the road network including less congestion (vehicle idling) and less otherwise unnecessary acceleration and deceleration arising from congestion.

Clearly there is a strong synergy between the Transport Agency's desired impacts and improved network efficiency outcomes with recurring themes being to reduce congestion, reduce variability in travel times and improve travel time reliability on the network.

These themes also align with network efficiency measures within the Ministry of Transport's (2012) Transport Monitoring Indicator Framework. This framework contains a large number of transport sector indicators that are monitored over time and is developed for transport sector groups as a tool for informing and evaluating transport-related policies and strategies.

3.2.2 Traveller information needs

High-quality information enables travellers to make informed travel choices regarding transport mode, route choice and travel time. The benefits of TIS therefore include both traveller satisfaction and improved network performance as demand for travel is spread throughout the day and across the transport network (Chang et al 2013; Raine et al 2014).

Chang et al (2013) have undertaken research into New Zealand road users' customer information needs and delivery systems for multimodal travel information. The authors identified information types sought by travellers, and prioritised these through an online survey. These results are summarised in table 3.1 with information needs classified as high, medium or low priority.

Table 3.1 Traveller information priorities (adapted from Chang et al 2013)

Priority	Information type
High	Road congestion information
	Location of road closures, roadworks and traffic incidents
	Anticipated travel times based on real-time updates
Medium	Weather conditions
	Road conditions
	Parking availability information
	Location of parking
Low	Next public transport service information (bus, train, ferry)
	On-board public transport (eg next stop information)
	Location of parking
	Location of points of interest, public toilets and rest areas
	Walking routes/facilities/journey times
	Cycling routes/facilities/journey times

Chang et al (2013) also identify high-quality real-time information as the highest priority for future improvements in providing traveller information in New Zealand. The top five real-time information types included network conditions (congestion, travel time) and road events (closures, road works and other incidents). Middle-ranked information types that could also be informed by crowdsourced data included real-time weather and road conditions, traffic cameras and parking availability. The highest-priority channels for delivering this information in New Zealand are websites, smartphone applications, radio and navigation systems (Chang et al 2015)

Public transport (PT) information (eg next service and next stop real-time information) ranked low, which may be indicative of the relatively low mode share of PT users compared to private vehicles amongst participants in the study. Active transport-related information, including walking and cycling routes, facilities and journey times also ranked lowly with no real-time active transport information types identified. Similarly to PT, this may be representative of the relatively low mode share of active transport modes.

Customer information needs also vary depending on location (Chang et al 2015):

- *In rural areas* – the highest priority information needs relate to whether particular routes are open during critical events (eg flooding, storms, snow and ice). PT information is a low priority.
- *In regional locations* – advanced notice of scheduled delays, parking information and delays due to unexpected weather events are highest priority, particularly for inter-city travel where alternative route options are limited.
- *In metropolitan areas* – drivers face more complicated information needs relating to unscheduled and scheduled delays as they arise. Avoiding congestion and optimising journey times are key issues for people in these locations.

3.2.3 Real-time information needs

One of the objectives of this research was to establish the value of real-time or near real-time crowdsourced data within a transport context. A number of the crowdsourced applications discussed in the literature review either do not collect data in real time, or data is not used in real time for an immediate response or update.

Real-time information is required to understand and respond to activities that have a relatively short lifespan, and usually require an immediate response by RCAs and road users. For example, a minor crash that blocks a lane discovered in real time using crowdsourced data would prompt the corresponding transport operations team to take steps to minimise the impact on the road network, while road users may be warned to avoid the area or possibly even provided with information suggesting alternative routes.

While all crowdsourced transport information could, in theory, be collected in real time or near real time, this is not always required within the corresponding application. For example, there is no need for real-time pothole identification if it takes several hours or days to respond to the problem. Similarly, tracking cyclists in real time serves little purpose as cycling routes are rarely affected by congestion or suffer from the extremes of travel time variability that other road users experience.

Historic information (information not collected in real time) is also often collected as part of a specific planning or data collection exercise, for example, the ‘Building our Footprints’ and ‘Share an Idea’ initiatives. Historic travel time information is also collected through ‘passive’ mechanisms such as GNSS sensors and Bluetooth tracking to update maps and create historic travel time profiles. This type of information collection does not require a RCA to respond in real time and can used for other purposes such as medium or long-term network monitoring and asset management.

One-off 'hack' crowdsourcing events are also excluded from this definition of real time, recognising that the key crowdsourced input is 'expertise', rather than data or information. It is recognised there may be a role for crowdsourcing events in the development of mobile data collection and/or traveller information distribution applications.

3.2.4 Other transport needs

Whilst addressing network efficiency and traveller information needs are the predominant focus of this research, there are other broader needs within the transportation industry that real-time crowdsourced data may address in part or full. Road safety, freight and asset management are examples of transportation fields that may benefit from the provision of real-time crowdsourced information and are discussed in more detail in chapter 4.

3.3 Real-time crowdsourcing scoping assessment

A selection of crowdsourcing applications was identified in the literature review, representing a cross-section of information collection technologies, applications and uses. Applications that have a real-time component have been assessed against the strategic transport and traveller information needs identified in section 3.2 (table 3.2 for passive applications and table 3.3 for active applications).

Active information collection options presented in table 3.3 have been categorised as either one-off event-based data collection, or continuous event reporting. This recognises the different quality, data collection and reporting needs associated with one-off versus on-going crowdsourcing projects. Both options support single or multi-modal information collection, depending on the application used and purpose of the project. A summary of the types of activity that a user could actively report is provided in Appendix D. Activities (or 'events') that users may report includes information specific to all or any mode of transport, and can include enforcement, safety, weather conditions, road conditions and closures, public transport capacity and quality of service, and pavement condition.

Although passive and active information collection applications are assessed separately, some applications (eg Waze and Tiramisu) combine both passive and active input elements. The distinction between the passive and active modes does not preclude a combination of both modes within the one application.

Table 3.1 Scoping assessment of passive information collection applications

Application	Description	Strategic transport needs	Traveller information needs	Other needs	Implementation factor
Private vehicle tracking	Applications collect in-vehicle travel time and speed data. They are used to improve real-time routing and provide accurate travel time predictions. Examples: Waze, Here-Drive, INRIX, Google Traffic	Improved network efficiency through real-time monitoring over a wide network. When combined with routing algorithms this supports better use of the existing transport capacity and improved travel time reliability.	Improved reporting of road congestion and expected travel times allows users to make informed real-time travel choices including routing and trip timing.	Network efficiency and traveller information benefits improve when combined with a routing algorithm that provides real-time, customised route updates to the user. Crowdsourced data could be used for network planning and monitoring purpose, for example, to identify congested areas and to measure the before and after impact of projects.	Requires a high number of users, although data gathered can be supplemented through other real-time network monitoring sources. A broad range of established and publically available products already used (eg Here Drive, Google Maps and Traffic). Some existing products not available in New Zealand or may not support data sharing.
Parking	Parking applications use vehicle and driver movement to predict parking space availability. Example: Parko	Improved vehicle circulation and congestion through dense urban areas as users spend less time circulating to find a car park. Supports efficient use of existing parking supply and road network.	Improved and customised parking availability information provided to the user to specify location of free (or becoming free) spaces.	Improved use of parking assets may reduce need for new parking stock. Could support demand-driven parking charging, enforcement and identifying parking needs.	Most suited to dense urban areas where on-street parking is at a premium. Requires a very high number of users for it to work. Currently not used in New Zealand. Established technologies provide similar information.
Public transport service tracking	Users on PT services are tracked to improve the reporting of arrival, departure and travel times. Examples: Tiramisu PT, Moovit	Improved prediction of arrival times and improved reliability of 'in-transit' times will improve PT as a choice of travel mode.	Provides high-quality, up-to-the-minute information on the proximity of the next PT service, alternative PT travel options, and in-transit arrival times.	Tracking network and operator performance. Better information provision may support increased PT mode share.	Crowdsourced data may provide limited benefit where vehicles are GNSS-enabled and provide real-time information feeds to VMS and via apps (eg Auckland, Christchurch). Existing applications for this purpose not widely used in New Zealand

Table 3.2 Scoping assessment of active information collection applications

Application	Description	Strategic transport needs	Traveller information needs	Other needs	Implementation factors
One- off event reporting	Collecting information on a one-off or infrequent event that could significantly disrupt transport services. These events may be expected (eg a sporting event, holiday congestion), or unexpected (eg natural disaster, transport system breakdown). Information can be collected across any or all modes (private vehicle, PT, active modes) Examples: Ushahidi (eg London tube strikes, Christchurch earthquake), social media sources	Road users receive real-time contextual information alongside official reports, influencing travel behaviour and minimising network impacts. Transport operation centres receive real-time contextual information regarding the functioning of the network and can respond accordingly.	Road users receive contextual information that may influence travel behaviour. Users understand the quality and spare capacity of the service they are expecting, and identify potential impediments that may affect route choice and travel time.	Safety benefits by alerting users and emergency services regarding real-time transport hazards. Asset management benefits if users are able to report issues with the transport network.	Can combine non-crowdsourced data in a single platform for richer traveller information. Short-term nature of information gathering means that quality control and moderation is more resource-intensive then ongoing event reporting. Moderators may be required to manually check and verify user reports.
Continuous event reporting	Applications that enable ongoing reporting of transport events, for example crashes, congestion, service quality etc. Information can be collected across any or all modes Examples: Waze, Tiramisu PT, Moovit, Cycle Atlanta, Collideosocpe, social media sources.	Road users receive contextual information that may influence travel behaviour. Transport operation centres receive real-time contextual information regarding the functioning of the network and can respond accordingly. Ability for active road users to report issues supports user engagement and encourages use of alternative modes.	Road users receive contextual information that may influence travel behaviour. Users understand the quality and spare capacity of the service they are expecting, identify potential impediments that may affect route choice and travel time.	Safety benefits by alerting users and emergency services regarding hazards. Asset management benefits if users are able to report issues with road network. Historic data can be used for network planning. Improve PT service quality and reporting of issues to operators. Elevate profile of PT as a mode.	Can combine other data into a single platform for richer traveller information. Automated data quality protocols can filter and process information, requiring less moderation than a one-off event. Effective method for engaging with small, yet engaged user-bases (eg cycling community).

4 Stakeholder engagement

The research team interviewed 24 key stakeholders during February and March 2015 to collect the stakeholder's views on a range of issues relating to the collection and dissemination of crowdsourced information.

The purpose of this consultation was to:

- Understand transport information needs, particularly focusing on real-time needs for customer information and network efficiency benefits.
- Understand how crowdsourced information would meet these needs and/or supplement existing information sources.
- Consider data quality needs, including accuracy and timeliness, required from a stakeholder and/or end user's perspective.
- Explore any unique knowledge regarding crowdsourcing information and applications.
- Explore stakeholder's concerns and experiences regarding:
  - privacy issues
  - safety issues
  - data and technology issues
  - integration with existing information sources and workflows
  - practicality of crowdsourcing applications
  - institutional or cultural barriers.

The stakeholders interviewed included representatives identified by the research team and project steering group. These were from the Transport Agency, Ministry of Transport, PT operations (Auckland and Christchurch), Auckland Transport, traffic operations centres (TOCs), transport user groups and technology firms. The response rate of stakeholders approached by the research team was 83% and a number of referrals were made to include other stakeholder representatives, demonstrating a high level of interest and engagement in this research topic.

The majority of stakeholder interviews were undertaken in person in Auckland, Wellington and Christchurch with two members of the research team present. At each interview, both research team members recorded notes on key points and themes which were compared and collated with all other responses for reporting purposes.

It is important to note that multiple interviews were held with some stakeholder organisations to receive a cross-section of views across the breadth of the organisation. In some instances, not all of the questions were raised with each stakeholder interviewee depending on the areas of knowledge and involvement of the respondent and on a couple of occasions due to time constraints. Where possible the stakeholder interviews were scheduled for one-hour duration.

This chapter of the research report summarises the outcomes and learnings arising from these interviews. Most of these themes reflect an anonymous and aggregated response to the questioning, although specific examples have been used on occasion to illustrate a particular point or highlight differences between organisations.

4.1 Overarching themes

Throughout the interviews, three overarching themes were expressed (explicitly or implicitly) by stakeholders:

- the role of crowdsourcing within ITS and TIS
- user empowerment
- organisational factors.

4.1.1 Integration of ITS, crowdsourcing and travel information delivery

A key theme expressed implicitly from the interviewees is that the concepts and practices of ITS, crowdsourcing and customer information delivery are interlinked and inseparable. Effective ITS require high quality data collection mechanisms, including road-side and in-car technologies. Human inputs through crowdsourcing can add value and context to ITS systems and help fill information gaps.

Crowdsourcing, by definition, is the collection (or 'pull') of information from crowd members, however it was noted by many stakeholders that transport crowd-based applications often combine both pull and push elements. For example, Facebook and Twitter as crowdsourcing mediums are useful for both receiving and responding/communicating user incident reports. The linkages between crowdsourcing and TIS are such that many stakeholders discussed TIS at length.

4.1.2 User empowerment

Stakeholders were highly aware of the benefits of *empowering* transport users to make informed decisions about their travel behaviour by providing timely, accurate and reliable travel information via the channels they are most comfortable using. Crowdsourcing is seen as an extension of this as it provides a means by which users can provide feedback about how the network is functioning and increases their level of engagement with the RCAs.

Users who are fully engaged with the transport system are also more likely to want to contribute to crowdsourced information gathering and less likely to need tangible incentives to do so. Good information also empowers RCAs to make informed and timely decisions about the operation of the transport network.

4.1.3 Organisational culture and direction

Many stakeholders discussed a perceived difference between some organisations' policy direction and expectations, and their business practices and risk management culture. Policy areas where this was identified included the Transport Agency's 'One Network' approach, use and supply of open data, emphasis on customer focus and the promotion of 'innovative' solutions, all of which bear relevance in the consideration of crowdsourcing applications.

Stakeholders were acutely aware of their organisations' expectations and policy direction in these areas; however, many felt the culture and bureaucracy of their organisation had not yet reached a point where they could be fully developed. Many of these areas were new to organisations and required some time to 'bed in' before being reflected in day-to-day business practice. Crowdsourcing is therefore seen as an area where organisations can push the limits and challenge current business practices.

4.2 Information needs

Stakeholders were asked to identify their existing information channels and information gaps, focusing on real-time information needs. The purpose of this was to understand where crowdsourcing applications have the greatest potential to deliver value to transport network operators and to road users through TIS.

4.2.1 Existing information sources/channels

4.2.1.1 Public transport

PT services in Canterbury and Auckland are tracked in real time using on-board automatic vehicle locating (AVL) technologies. The AVL systems use GNSS technology and actively monitor the location of each bus on the road network. This system generates huge amounts of data, some of which is published as a Google Transit Feed Service.

Information about incidents on the network that may affect the efficient and on-time running of PT services are received from bus drivers (often relayed via their operations staff), members of the public, TOCs and the NZ Police. Social media also has a prominent role in identifying incidents and issues relating to service quality and customer satisfaction. In Canterbury, many bus users use Facebook for reporting regarding issues such as late services, service updates, bus or bus stop maintenance issues. Twitter is the social media application most commonly used in Auckland for similar purposes.

In Auckland and Wellington, tap and go services collect information about where bus users board and alight, although this information requires extracting and post-processing and as such is not available in real time. In Christchurch boarding information is recorded using the same technology; however, no alighting information is collected.

4.2.1.2 Traffic operations centres (TOCs)

TOCs receive information from a wide range of sources to make informed network interventions and deliver traveller information to users. Travel time and congestion sources include GNSS and Bluetooth data feeds, although stakeholders noted that neither technology provides the complete picture of the road network. Commercial probe data sources can also be biased by the underlying composition of the vehicle fleets which contribute data.

Contextual information sources (the ‘eyes on the street’) include the NZ Police, transport operators, infrastructure managers, their own staff, PT operators, CCTV footage and some social media sources. Initial incident reports are usually provided by the NZ Police; however, this information is not always considered reliable (both in context and location).

Network contractors are generally considered by the TOCs to be the most trusted source of accurate information, but this information is received some time after the initial event has occurred (ie once an incident is being cleaned up). It was noted that issues identified by bus drivers and PT operators often do not get reported to the TOCs and there is plenty of scope to benefit from improved communications between regional councils and RCAs.

4.2.1.3 Other agencies/organisations

While PT operators and TOCs are considered to have the greatest real-time information needs, other stakeholders interviewed by the research team receive information from a range of sources including commercial travel time datasets, traffic surveys, call centres, websites and social media.

4.2.2 Specific information gaps

Stakeholders were asked to identify specific information gaps, real-time or otherwise, that may (or may not) be informed through crowdsourcing technology. Overall, the stakeholders identified real-time information as one of the most critical information gaps for network operations and customer information purposes. Table 4.1 is a summary of the breadth of information gaps identified by stakeholders.

Table 4.1 Specific crowdsourcing information gaps identified

Transport mode	Information gap
Public transport	Capacity/occupancy of services Customer satisfaction/quality of travel Where people get on and off the service Incidents affecting the PT network
Private and commercial vehicles	Origin/destination journey information identifying motorists' routes Traffic volumes Congestion Vehicle delays and travel time variability Monitoring real-time impacts of road works Differentiating between road work sites that are active/inactive Freight information Reporting dangerous drivers Rideshare (mass personal transport) applications
Walk/cycle	Information for tourists identifying cycle friendly routes Trip lengths (eg cycle vehicle kilometres travelled and travel times) Cyclist crash rates Buddy/social riding applications Cycling map updates Origin-destination information Pedestrian/cycle counts Cycle incident reporting[8] (eg crashes, near-misses, hazards)
Strategic/ multimodal	Planned event monitoring (for example sporting fixtures, community events, concerts) Incident reporting – expanding existing sources to include PT operators/drivers, parking officers, cyclists/pedestrians, heavy vehicle drivers etc Perceptions of the transport system (currently collected in travel surveys) Understanding how people react to events/disasters Throughput
Other	Requests for service (RFS) with actual location coordinates

While stakeholders were not specifically polled on the highest priority information gaps, those that were identified most frequently by stakeholders were:

- PT service occupancy and capacity (real-time)
- origin-destination across all modes, but particularly walking/cycling
- congestion in real-time

[8] Improved quality and quantity of data collection for real and perceived cycling safety, especially non-motor vehicle crashes, has also been identified as a high priority action by the Cycling Safety Panel (Cycling Safety Panel 2014)

- cycle incident reporting
- planned event monitoring for all modes in real time
- incident reporting in real-time.

4.2.3 Use of social media

Stakeholders were questioned on their use of social media channels and how they used information received via these channels.

Environment Canterbury primarily uses Facebook as a push-pull information channel with PT users, preferring Facebook over Twitter because it is better for facilitating two-way information exchange. Auckland Transport considers that there is more value in using Twitter rather than Facebook and this is primarily used to push information out.

Social media is used by other stakeholders to varying degrees as an information channel. Regional cycling user groups are active social media users. Many stakeholders identified social media as a valuable tool for both crowdsourcing and for the provision of traveller information, and that there is potential to improve the use of these applications to the benefit of the operation of travel networks.

4.2.4 Role of crowdsourcing in information collection

While discussing crowdsourcing applications in transport, stakeholders identified a number of themes regarding the use of crowdsourcing in collecting information. Overwhelmingly, stakeholders were very interested in the potential for crowdsourcing to add value to existing information collection channels by 'filling the gaps' in terms of both the coverage and quality of existing sources.

Discussions around the role of crowdsourcing primarily focused on 'active' data collection, in particular how the public could identify and provide contextual information about incidents and events which affect the operation of transport networks. Most stakeholders saw the greatest benefit in improving the number of 'eyes' on the network and empowering transport users to actively contribute information through crowdsourcing.

4.2.5 Role of traveller information systems

Stakeholders expressed concerns about the complex nature of sharing real-time, customised travel information (including crowdsourced information) with the public, particularly if crowdsourcing results in more alerts being generated and shared with users. The current government emphasis on improving real-time customer information channels was noted by many stakeholders, both at a strategic level and through the development of route-specific subscription-based notification systems (eg the 'On the Move' initiative).

Stakeholders also identified the need for a range of travel information channels, as different users have different communication preferences. Similarly, not all users would automatically be comfortable finding and using the appropriate applications or social media channels to submit crowdsourced information.

4.3 Information quality

4.3.1 Customer needs

Customers need access to reliable, timely and accurate travel time information. Stakeholders identified the Transport Agency and other transport operators as being the most likely 'trusted source' of information

for the general public. A clear theme also communicated by stakeholders is that while the accuracy of information is important, this does not mean that incomplete or untrusted information should be withheld provided it can be delivered with appropriate disclaimers or caveats regarding the source and confidence in the quality. Many stakeholders consider that it is better to have some information than none at all to empower users to make better informed travel decisions. This is consistent with the findings of Chang et al (2015) which found that users preferred to get fast information, and be updated regularly, even if it has not been fully verified.

4.3.2 In determining a response

A recurring theme in the stakeholder engagement interviews was identifying if and when traveller information should be issued to the general public following an incident or event on the network. The quality of the information received regarding the incident or event is pivotal to the decision to be made by TOCs and PT operators regarding notifying the public. Generally speaking, in those instances in which this scenario was discussed, stakeholders identified the following steps in determining whether a response should be posted:

1 Information or alert is received

2 The information is assessed – 'Do we need to investigate this further?'

3 The information is verified/validated through other sources (eg CCTV, Police, contractors)

4 Determine whether a response is required based on severity and likely network recovery time

5 The appropriate response is implemented (which may include travel alerts)

6 The situation is monitored until resolution.

When crowdsourced information is received at step 1, there must be sufficient confidence in this information to determine whether further investigation is needed, taking into account both the nature and severity of the incident reported, and the quality of the information (eg the number of reports). This is intrinsically linked to the importance of trust models to validate or verify the quality of the information source.

One stakeholder used the example of the challenge of comparing 10 verified reports of a train running two minutes late, to one unverified report of a train derailment. Some stakeholders used the terms 'threshold' or 'baseline' as a means of identifying an appropriate response to crowdsourced information. Crowdsourced information can also be used to provide context and verify alerts from other sources.

4.4 Privacy

There was interest among the stakeholders regarding privacy issues surrounding crowdsourcing technologies and information and most stakeholders were aware of public agencies' responsibilities under the Privacy Act, including the 12 privacy principles.

A key theme that arose was that while the Privacy Act must be complied with, it is a permissive act and privacy concerns should not be seen as a barrier in the development of innovative crowdsourcing applications. Privacy principles must be considered at the front-end and factored into the system design and architecture.

Stakeholders commented that the public is often highly critical of government agencies collecting personal information and that there could be great public interest in how crowdsourcing applications collect (or are perceived to collect) this type of information. It was noted by more than one stakeholder that the best

practice is to be open and honest with users about how and why information is being collected, and what it will be used for. This includes ensuring users provide 'informed consent' to the collection of this information, using privacy policy stated in simple-to-understand language and not hiding privacy among the terms and conditions.

A further recurring theme that emerged through the stakeholder discussions was that the general public may generally be more accepting of private organisations (eg Google) collecting their information than public agencies.

4.5 Safety

There was little feedback on safety issues relating to the use of devices in vehicles which was raised by the research team with some but not all stakeholders. The following points were noted:

- Mobile apps are a source of distraction for cyclists in addition to motorists.
- Consideration should be given to only using passive collection modes for in-car applications and not permitting active input.
- Any application that promotes the use of mobile devices in-car must be avoided and disclaimers (eg 'users shall not use this application while driving') is not sufficient mitigation.

4.6 Information collection, storage and retrieval

Stakeholders who were aware of their obligations under the OIA and LGOIMA made the following comments regarding information collection, storage and retrieval:

- Under the OIA/LGOIMA, all information is discoverable and can be requested.
- Information storage and collection needs (including private information) should be considered at the system design stage, not as an afterthought.
- Consideration should be given as to how large files and video data (eg dash camera video) should be retained and managed.
- There are no technological constraints to the collection, storage and retrieval of information – the key issue is likely to be the cost of implementing them. The security and privacy requirements of some technologies and applications may make them too expensive to deliver.

4.7 Incentives

The incentives that would encourage transport users to use a crowdsourcing application were discussed. Financial rewards (including prizes) were considered by most stakeholders to be infeasible due to funding restrictions and political/public perceptions about how public money is spent. Financial rewards were also perceived as unsustainable and unable to generate long-term engagement or encourage travel behaviour change.

Some stakeholders provided examples of how incentives had been used in influence travel choice or to promote the uptake of new travel systems. Auckland Transport noted that incentive schemes were used to encourage users to adopt the public transport 'HOP' smart card by offering discounted travel for users. Similar discounting/incentive schemes were considered to be less feasible for private vehicle users.

A number of stakeholders pointed out that if people see a clear benefit in sharing their information, they will naturally want to share this without the lure of additional incentives. People are naturally predisposed to help others and there is a spirit of reciprocity and public good among transport users. Gamification was also supported by some stakeholders to encourage contributions.

4.8 Institutional and cultural factors

4.8.1 'One Network' focus

Stakeholders recognised that the transport system involves multiple agencies, and users are often unaware of which agencies to contact regarding issues affecting the road network. The move towards a 'One Network' philosophy is recognised, but there are still challenges as each agency figures out how to apply this in practice. The perception of the 'One Network' approach was regarded as important for encouraging users to engage with transport providers and operations centres.

A single interface or 'brand' may be helpful in encouraging the general public to provide information across a range of transport networks and modes. A branded approach to information collection also helps distance the public agency from the application and can help reduce malicious/abusive content and 'cross-fire' on topical issues.

4.8.2 Innovation and risk

Stakeholders also identified innovation as an important area of consideration. While agencies desire to be 'innovative' at a strategic level, stakeholders noted that existing risk management and procurement business practices do not always support this. Crowdsourcing is inherently risky and therefore it was suggested by some stakeholders that third-party developers and the wider technology sector may be better placed to develop innovative crowdsourcing applications.

4.8.3 Political factors

Some stakeholder representatives noted that directly sourcing information from the public could be perceived as undermining the role of local government politicians as the 'voice' of the community. There was also concern about privacy concerns regarding tracking and surveillance becoming a sensitive political issue.

4.8.4 Role of 'hack' events

The relevance of transport 'hack' events such as OpenHack Christchurch and HACK:AKL in the development of crowdsourcing applications was discussed by some stakeholders. These events are perceived as a good way to generate innovation outside of a public agency setting; however, there are challenges in funding and developing sustainable solutions.

5 Evaluation of crowdsourcing approaches

Combining the outcomes of the literature review, needs assessment and stakeholder engagement, this chapter presents the results of two assessments:

- an evaluation of potential crowdsourcing information sources and their ability to meet the information needs identified through the desktop needs assessment and stakeholder engagement (section 5.1)
- a SWOT analysis of potential approaches the Transport Agency or other RCAs could initiate to crowdsource information (section 5.2).

The intention of this section is to provide direction to the Transport Agency and other stakeholders as to how the information needs can be met and what are the preferred crowdsourcing approaches, through a quantitative assessment of a broad range of internal and external influential factors.

5.1 Crowdsourcing data to meet information needs

An initial assessment of the extent to which passive and active crowdsourcing applications meet transport information needs was undertaken in chapter 3. Following stakeholder engagement, a wider range of potential information needs was identified, and the potential of social media sources was explored.

5.1.1 Crowdsourcing approaches

Four broad information sources have been identified for evaluation:

- **Social media monitoring** – using basic or advanced social media monitoring systems to view and potentially respond to social media alerts regarding transport network operations. The emphasis is on using existing social media accounts and services, or using out-of-the-box social media monitoring services (eg Hootsuite). This source requires active monitoring, interpretation and response to messages by a representative at a RCA or operations centre.
- **Social media mining** – using advanced social media data mining techniques to extract, locate, verify and aggregate data received from social media sources. In contrast to social media monitoring, data mining supports automation and techniques such as filtering, semantics and trust models to interpret and verify data sources.
- **Passive applications** – using an existing or new application to passively collect data regarding user location, movement and acceleration.
- **Active applications** – using an existing or customised application that enables users to actively report and locate events or incidents on the transport network or provide ratings of their experiences or the condition of assets.

The four approaches above are not exclusive or exhaustive. For example, social media can be monitored or mined using a combination of different tools and techniques and crowdsourcing applications can combine both passive and active sources. It is also acknowledged that social media is an active information source; however, it can be distinguished from other active information sources in that it utilises an existing platform and community, rather than requiring a bespoke data collection application.

5.1.2 Evaluation framework

The evaluation framework is a simplistic quantitative assessment of the extent to which the identified sources meet specific transport information needs. This evaluation was undertaken by the research team using a scale of 0 through 3:

- 3 – the source is best suited to fulfil the information need
- 2 – the source is suited to meet the information need in part or full
- 1 – the source may provide some data to fulfil the information need
- 0 – the source does not meet the information need.

In addition, a weighting of between 0 and 1 has been applied to demonstrate the relative importance of each of the information needs, as informed by the stakeholder consultation and strategic needs assessment. Higher priority needs, for example those that meet strategic needs and were widely supported by stakeholders, have been assigned a value of 1. Lower priority needs have been assigned a value of 0.5 if multiple stakeholders identified the corresponding need, or 0.2 if only one or two stakeholders highlighted the need.

For the purpose of the evaluation, it is assumed that if a passive or active application does not already exist then a suitable application could be designed and implemented to meet the information need.

5.1.3 Evaluation results

The transport information needs have been categorised as either real-time or non-real-time needs and evaluated separately in tables 5.1 and 5.2 and a list of transport modes provided alongside each need.

Table 5.1 Evaluation of real- time information needs

Information need		Mode	Social media monitoring	Social media mining	Passive applications	Active applications	Weight
High priority	Service occupancy/capacity	PT	1	1	0	2	1
	Congestion monitoring	PT, PV	1	1	3	3	1
	Travel times and variability	PT, PV	1	1	3	1	1
	Planned event monitoring	All	2	3	1	3	1
	Incident reporting	All	1	2	0	3	1
Lower priority	Reporting dangerous drivers	PT	1	1	0	3	0.2
	Parking availability/location	PV	0	0	2	1	0.2
	Weather/road information	All	1	2	0	3	0.2
WEIGHTED SCORE			6.4	8.6	7.4	13.4	

Abbreviations: CV = commercial vehicles, CY = cycling, PT = public transport, PV = private vehicles, WK = walking/other active modes

Table 5.2 Evaluation of other information needs

Information need		Mode	Social media monitoring	Social media mining	Passive applications	Active applications	Weight
High priority	Origin-destination/trip length	All	0	0	2	1	1
	Cycle incident reporting	CY	1	2	0	3	1
Lower priority	Customer satisfaction/perception	All	1	1	0	2	0.5
	Cycle and walking infrastructure updates (maps) and rating	CY, WK	1	1	1	3	0.5
	Throughput (counts)	All	0	0	0	0	0.2
	Requests for service	All	2	2	0	3	0.2
WEIGHTED SCORE			2.4	3.4	2.5	7.1	

Abbreviations: CV = commercial vehicles, CY = cycling, PT = public transport, PV = private vehicles, WK = walking/other active modes

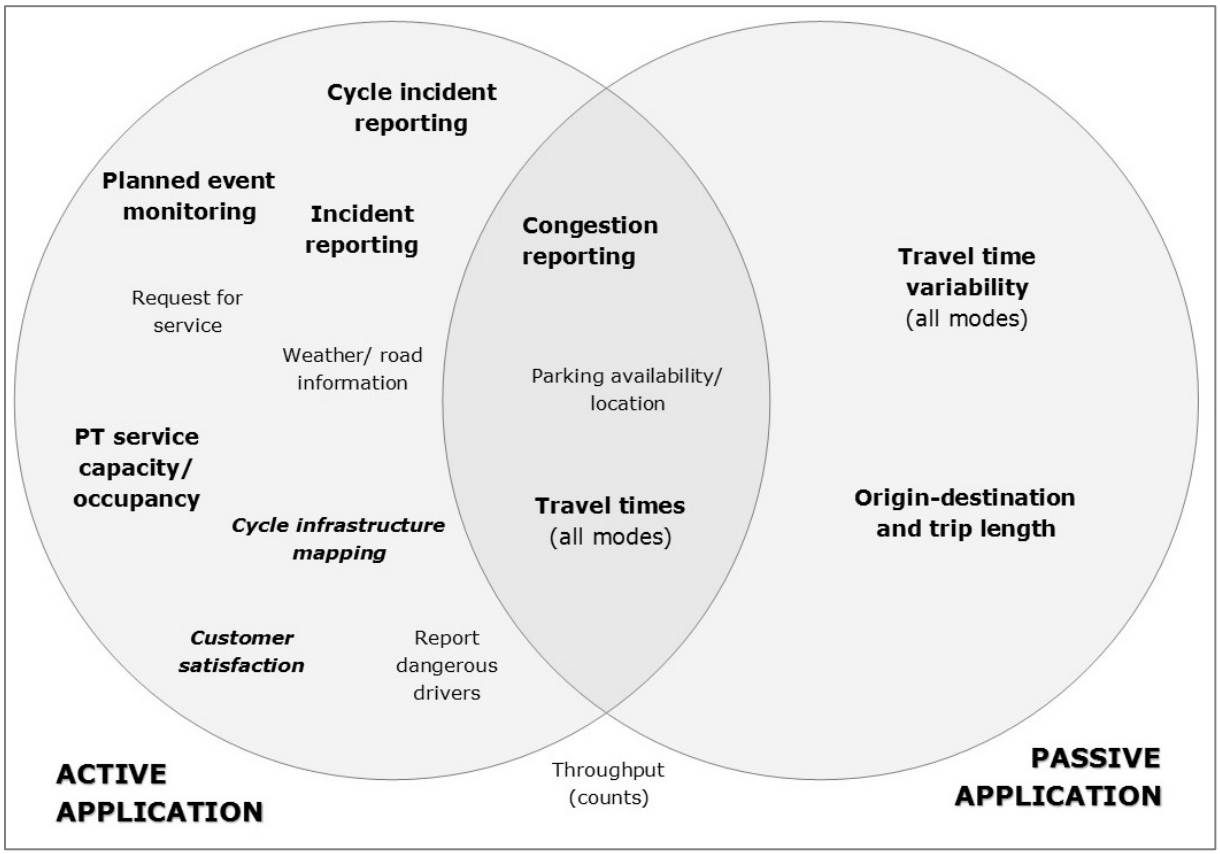
The results from the evaluation indicate that most information needs can be met through the use of either passive or active data collection applications, or by using a mix of social media approaches. The potential value of active crowdsourcing information sources (excluding social media) is evident. Passive crowdsourcing information sources only meet a small number of needs and some of this is currently met, in part, through commercial GNSS and Bluetooth sources.

There is also a role for social media in crowdsourcing information; however, the extent to which social media monitoring can be enhanced by data mining may be limited outside of event and incident reporting and will be sensitive to the sophistication of the mining method. In general, the information sourced through social media channels can also be met through other active crowdsourcing technologies.

While some social media sources support geotagging (adding spatial locations to posts), the combination of passive and active data sources meets all of the same needs and, in many cases, with a greater degree of suitability and customisation.

Figure 5.1 summarises the suitability of passive and active crowdsourced sources and travel information needs. High priority needs have been emphasised using bold text and social media sources have been combined to reduce complexity.

Figure 5.1 Suitability of crowdsourced information sources for different information needs



5.2 SWOT analysis of crowdsourcing approaches

SWOT analysis is a technique which supports strategic decision making. When applied to a problem or situation, SWOT analysis involves pinpointing four key areas (strengths, weaknesses, opportunities and threats) to assist decision makers in identifying the key facts of the issue/situation involved. Within SWOT analysis, strengths and weaknesses are internal factors that are controllable and can be acted upon. The opportunities and threats are external, largely uncontrollable and commonly identified as possible future events.

SWOT is traditionally used as a qualitative analysis tool, where each of the four factors as applicable to a particular issue or situation are documented on a matrix type structure as shown in figure 5.2. Qualitative SWOT analysis involves bringing together an informed group to brainstorm and complete the appropriate areas within the matrix as paragraphs or bullet points. The purpose of this analysis is to document existing knowledge on a given situation as an aid to decision making.

Figure 5.2 Typical SWOT matrix

	Helpful (to achieving the objective)	**Harmful** (to achieving the objective)
Internal origin (attributes of the organisation)	**S**trengths	**W**eaknesses
External origin (attributes of the environment)	**O**pportunities	**T**hreats

Quantitative SWOT analysis is less commonly used and involves adapting the traditional SWOT matrix to include numerical indexes to value inputs such as 'potential impact' and 'likelihood' to produce a graphical output in the form of a Cartesian chart. The quantitative approach has been applied for the analysis that follows.

5.2.1 Crowdsourcing approaches

Five high-level approaches to crowdsourcing information have been selected for SWOT analysis. These are proposed as potential strategies for the Transport Agency and other RCAs, and are as follows:

- **Approach A – Social media monitoring**. This approach involves focusing resources on monitoring existing and new transport-related social media feeds, either in isolation or using social media management systems. This approach assumes the social media feed is continuously monitored (in real time, where relevant to the application).
- **Approach B – Social media data mining.** This approach involves developing automated data mining capabilities, either through internal development or engaging with a third-party data mining service.
- **Approach C – Develop customised application(s).** In this approach, the Transport Agency would develop a customised web and/or mobile based application to source data for a particular information need.
- **Approach D – Procure crowdsourced data from third party providers**. In this approach, crowdsourced data is purchased or traded with existing third-party crowdsourcing application owners, for example Waze.
- **Approach E – Support private sector- led development of crowdsourcing platform.** This approach involves developing incentives to support the independent development of a crowdsourcing platform by the wider development/technology sector. Incentives may include making data more available,

publicising information needs to the technology sector and supporting hack events, competitions and other forms of engagement.

It is important to note that these five approaches reflect potential strategies which the Transport Agency may wish to initiate to crowdsource information. This is not an exclusive or exhaustive list of strategies, they are not mutually exclusive and there may be value in the Transport Agency and other stakeholders engaging in some or all of these approaches.

5.2.2 SWOT analysis and results

The crowdsourcing approaches have been assessed against a range of SWOT factors that have been identified through the literature review, strategic assessment and stakeholder feedback stage of the research. The strengths and weaknesses primarily address internal implementation matters. Each of the internal factors (strength and weaknesses) identified are measured on a scale -2 through +2 where:

- +2 is a strength
- +1 is a partial strength
- 0 is neither a strength nor weakness
- -1 is a partial weakness
- -2 is a weakness.

Opportunities and threats include a broad spectrum of external influences which impact upon the relative success of each approach. The opportunities and threats have been rated on a scale of -2 through +2, where:

- +2 is an opportunity
- +1 is a partial opportunity
- 0 is neither an opportunity nor threat
- -1 is a partial threat
- -2 is a threat.

A weighting factor has been applied to each factor to emphasise those considered more important for the overall evaluation. The most important factors have a value of ‘1’ and smaller values have been allocated to those factors considered to be less significant. These weightings are informed by the strategic needs assessment and evaluation stakeholder feedback reported in chapters 3 and 4 respectively.

The evaluation of each internal and external factor by the research team is somewhat subjective and informed by knowledge accumulated through the research to date. A summary of the SWOT analysis results is presented in table 5.3 and an explanation behind the rationale for scoring each factor is included in appendix C to this report.

Table 5.3 SWOT assessment of crowdsourcing approaches

Internal and external factors		Social media monitoring	Social media mining	Develop custom application	Procure data from third-party	Private sector leads development	Weight
Strength/Weakness	Control over data quality/validation	-2	-1	2	1	1	1
	Control over privacy issues	1	1	2	1	0	1
	Integration with current sources	0	0	2	2	1	0.5
	Adaptability/flexibility	1	0	0	0	0	0.5
	Initial investment (time/cost)	2	1	-2	-1	0	0.5
	Alignment with government strategic direction	1	1	0	1	2	0.2
	Ongoing investment (time/cost)	-1	-1	0	0	0	0.2
Opportunities	Meet real-time information needs	1	1	2	2	1	1
	Meet other information needs	1	1	2	2	1	0.5
	User empowerment/ engagement	2	2	1	0	1	0.5
	Cost-effective info source	2	1	0	0	1	0.5
	Collaboration public and tech sectors	0	1	0	1	2	0.2
Threats	Potential to damage reputation	-1	-1	-2	0	-1	1
	Bias in crowd input difficult to manage	-2	-2	0	-1	-1	0.5
	Requires high number of users	-1	-2	-1	-1	-2	0.5
	In-vehicle safety difficult to manage	-1	-1	0	-1	-1	0.5
	Prone to misuse/abuse	-1	-1	0	0	0	0.2
Strengths/weaknesses weighted score		0.5	0.5	4.0	2.7	1.9	
Opportunities/threats weighted score		0.3	-0.5	1.0	1.7	-0.1	

5.2.3 SWOT Cartesian plot

For each approach assessed, two scores have been calculated:

- strengths minus weaknesses
- opportunities minus threats.

These scores reflect two continuums, one for external factors (opportunities versus threats) and one for internal factors (strengths versus weaknesses). The continuum approach supports the balancing effects of many of the SWOT factors. In many cases, a strength in one factor will be offset by the weakness in the same criteria and for this reason strengths and weaknesses were scored on the same continuum in table 5.3. To test the sensitivity of the weightings applied to each factor, the results have been calculated with weightings removed – that is all factors have equal weighing in the evaluation. The two scores (weighted and un-weighted) for each approach are summarised in table 5.4.

Table 5.4 SWOT summary of scores

Approach	Weighted results		Un- weighted results	
	Strength – weakness	Opportunity - threat	Strength – weakness	Opportunity - threat
Social media monitoring	0.5	0.3	1.0	0
Social media mining	0.5	-0.5	0.5	-0.5
Develop customised application(s)	4.0	1.0	2.0	1.0
Procure crowdsourced data from third party providers	2.7	1.7	2.0	1.0
Support private sector-led development of crowdsourcing platform	1.9	-0.1	2.0	0.5

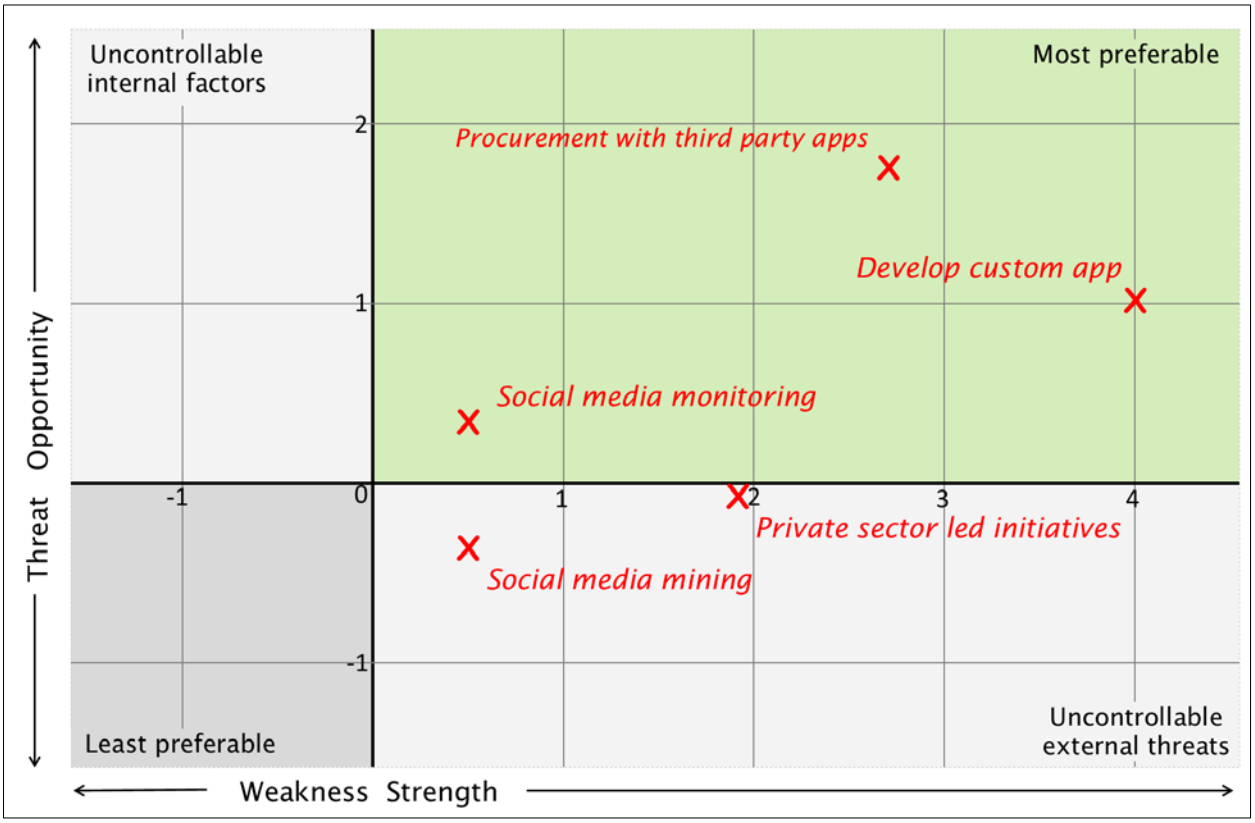
The sensitivity test results are grouped together more closely compared with the weighted results, indicating that those factors that attract a lower weighting are balanced by other factors attracting a higher value in the SWOT assessment. The two social media options have similar weighted and un-weighted scores. The three remaining approaches have generally lower 'strength-weakness' scores and similar 'opportunity-threat' scores when results are not weighted.

Figure 5.3 is a Cartesian graph of the weighted SWOT analysis outputs. The success of a particular approach is based on where it is located in a particular quadrant. The top- right quadrant ('strength and opportunity') is most preferable and the bottom-left quadrant ('weakness and threat') is least preferable. The 'middle ground' is represented in the top-left and bottom-right quadrants with a preference for the bottom-right quadrant. This is because the weaknesses represented in the bottom-right quadrant reflect controllable, negative influences, where as the top-right quadrant reflects the influence of uncontrollable external threats.

Three of the five approaches are situated in the upper-right quadrant of the Cartesian graph, with the 'procure crowdsourced data from third party providers' and 'develop customised application(s)' approaches clearly scoring higher in both internal and external factors. The 'social media monitoring' approach is also located in the upper-right quadrant with low positive scores for both internal and external factors. The 'support private sector-led development of crowdsourcing platform' and 'social media mining' options have slightly negative 'opportunity-threat' scores.

The benefits of social media mining over social media monitoring appear to be marginal. Data mining has performed poorly by comparison due to higher initial costs and a reduced level of adaptability and flexibility. There is also some uncertainty about the size of the social media user base is and how much additional information could be gleaned and analysed using this approach. This will be considered later in the real-world trial in chapter 6.

Figure 5.3 SWOT Cartesian map



The procurement of third-party crowdsourced data and Transport Agency-led application development options provide greater levels of control. With Agency-led application development, matters such as data quality validation, privacy control and data integration can be designed within the system architecture and user interface.

None of the five potential strategies in the SWOT analysis should be discarded outright. The analysis is subjective with some uncertainties and the outcome may change following the real-world trial phase of this research. This also reinforces a conclusion from the evaluation of information needs in section 5.1 wherein there is a role for both social media and other types of application (passive/active) in crowdsourcing information.

5.3 Summary of evaluation

The evaluation demonstrated that the use of both social media and customised applications, including passive and active crowdsourced information, potentially maximises the reach of the Transport Agency to deliver information needs for the supply of both traveller information and network efficiency outcomes.

The SWOT analysis demonstrated that the procurement of crowdsourced data from third party providers and the development of a customised application have the greatest potential benefits; however, social media sources also have some value and are suitable for certain applications, including planned event monitoring and incident reporting (eg unplanned and unforeseen incidents).

In both the information needs evaluation and SWOT analysis, there was some uncertainty regarding the practical application and potential of some crowdsourcing approaches, for example the usefulness of social media mining and the availability and suitability of third-party information sources.

6 Crowdsourcing trial

This chapter documents the development of the real-world trial, designed to test the findings from in the preceding chapters of the research report. To maximise the potential of the trial and address the research objectives, the aim of the trial was to:

- explore the potential for, and limitations of real-time (or near real-time) crowdsourcing applications
- determine whether crowdsourcing can meet high priority information needs
- test a data source or application that has not been used or previously tested in the New Zealand context
- explore the value of investing in social media mining relative to social media monitoring
- test an active crowdsource application through either the development of a customised solution or through procurement with a third party.

The research team considered a number of potential real-time crowdsourcing applications in consultation with the research owner and steering group. It was agreed that the trial would explore how crowdsourcing can be used to understand unplanned events on a road network.

6.1 Context

RCAs often have poor or limited visibility of real-time, unplanned events on many parts of the urban and rural road network. Improved visibility of these events, including weather conditions and road damage, improves an RCA's response and supports improved traveller information regarding the nature, scale and duration of the event.

The Queenstown-Lakes District was chosen as an appropriate area for the trial. The District has a mix of both urban and rural road networks, an engaged social media user base, and faces frequent winter weather events that affect key routes such as the Crown Range Road. The trial was undertaken between July and September 2015, targeting the winter months where unplanned, weather-related events occur more frequently.

The Queenstown-Lakes District Council (QLDC) uses several communication streams to send and receive information about road conditions, including:

- the QLDC Facebook page[9]
- the QLDC Twitter feed[10]
- The QLDC road conditions webpage[11]
- daily winter road report emails for registered users
- a text messaging service (from Twitter alerts).

QLDC also monitor their Facebook page, Twitter feed and service requests for public reports on road conditions. The QLDC communications team are responsible for responding to requests via these channels and issuing road condition emails and alerts.

[9] www.facebook.com/QLDCinfo

[10] https://twitter.com/QueenstownLakes

[11] www.qldc.govt.nz/services/transport/winter-driving/winter-road-conditions

6.1.1 Trial structure

The trial involved three separate activities:

1 the development and testing of a crowdsourcing web application and map viewer

2 social media monitoring

3 social media (Twitter) mining.

6.2 Web application development

Figure 6.1 provides an overview of the web application developed to collect, verify and display crowdsourced road reports. The development and implementation of the web application was undertaken by the research team in consultation with QLDC, CTOC and the Transport Agency.

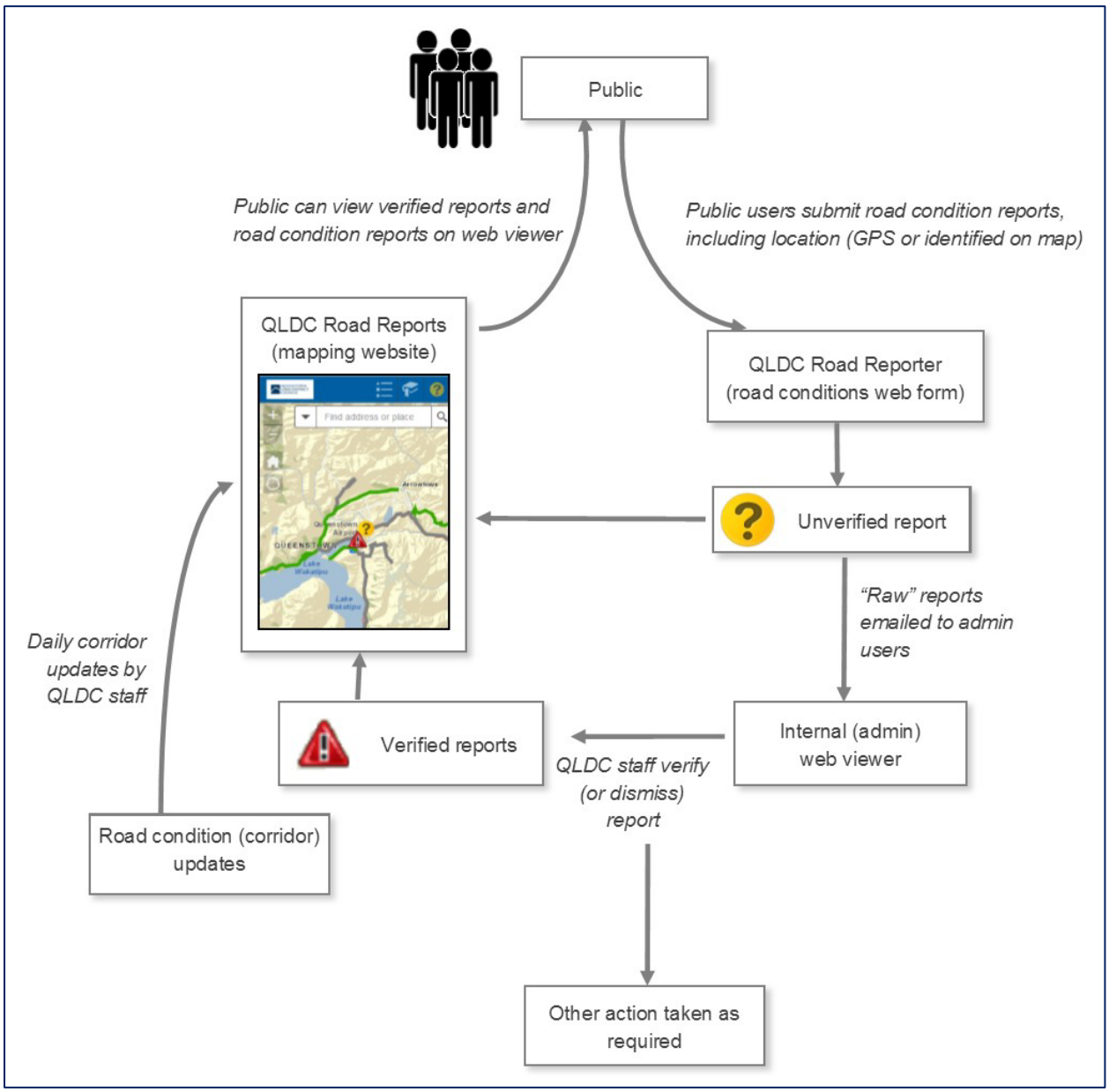
Three integrated components were designed to deliver the web application, utilising the ArcGIS platform:

1 a public reporting form

2 a public road conditions map

3 an internal administration web viewer for editing and verifying reports.

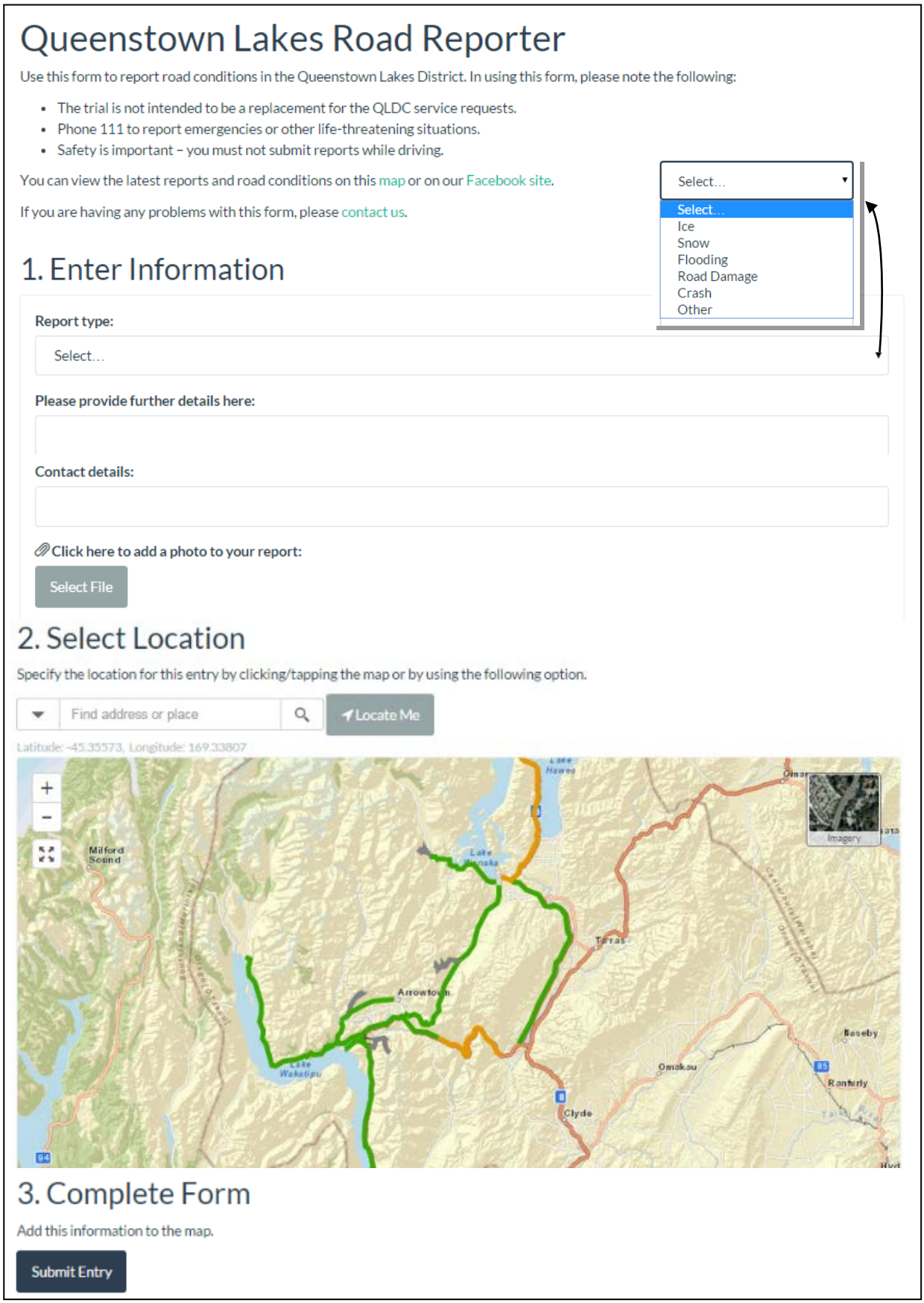
Figure 6.1 Crowdsourcing web application framework



6.2.1 Reporting form

A web-based reporting form, the 'QLDC Road Reporter was developed to collect public reports on road conditions. Figure 6.2 is a screenshot of the reporting form.

Figure 6.2 Screenshot of the submission form used by trial participants



The specifications for the form include:

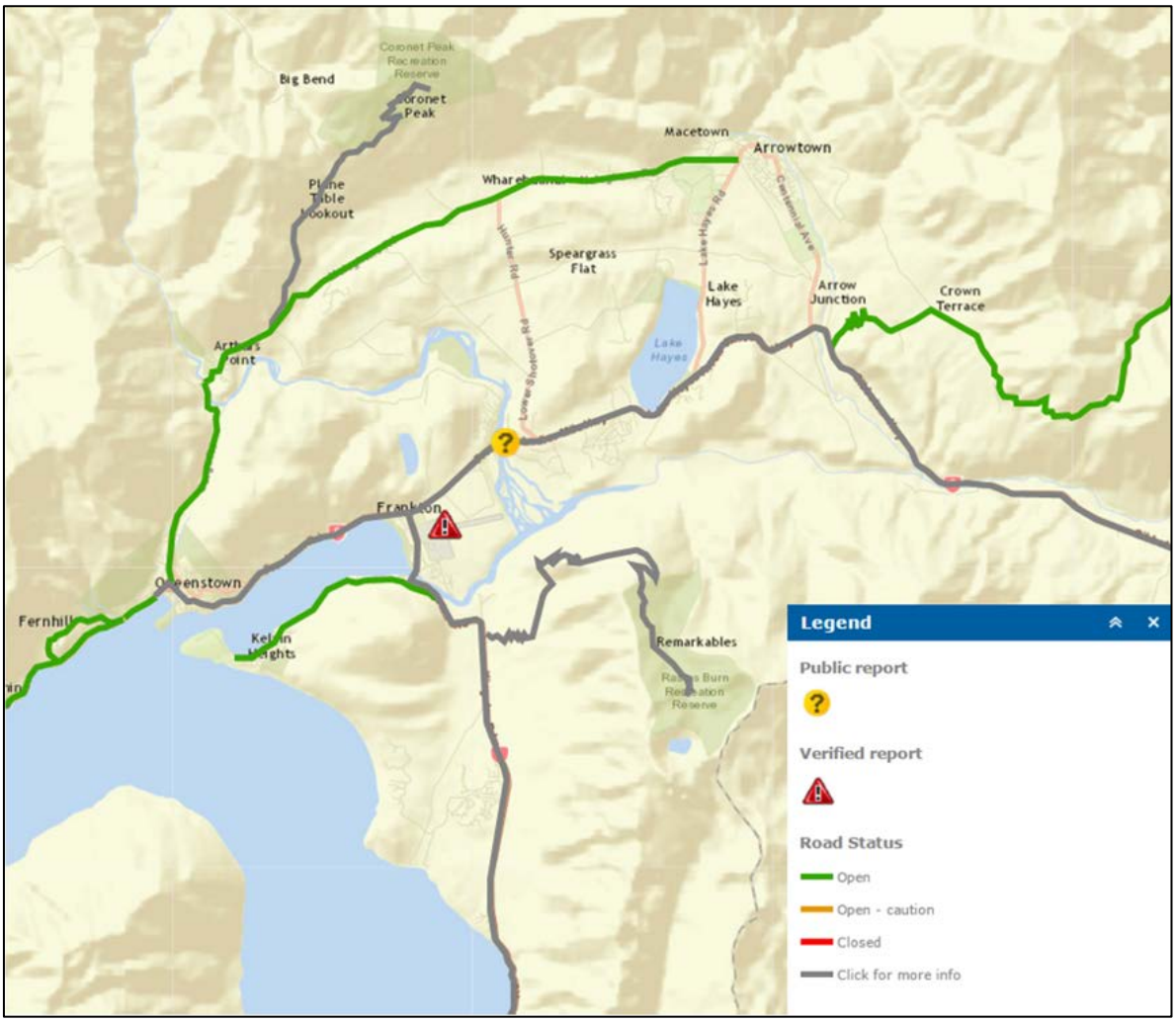
- a disclaimer addressing appropriate and safe in-vehicle use
- optimisation for display on desktop, mobile and tablet devices
- customised reporting categories (ice, snow, flooding, road damage, crash, other)
- options for users to provide additional comments and add attachments (eg photos)
- a map and address panel for locating the report, with an option to use the device's location service
- a link to the road conditions map.

Once a report was submitted, it appeared on both the public-facing and internal web maps. An email alert was also sent to the trial administrators.

6.2.2 Public road conditions/reporting map

Submitted reports were displayed on a public-facing road conditions webmap, shown in figure 6.3. Unverified reports were displayed as question marks and verified reports were displayed as exclamation marks.

Figure 6.3 Screenshot of the 'QLDC Road Reports' road conditions map



The web application administrators used an internal-facing webmap to validate and monitor reports. The validation of public reports included checking the report against other sources of information (for example contractor and Police reports). Validated reports could also be pushed out through the Council's social media channels and/or passed on for further action (for example sending a request through to the roading contractor). A process for removing closed events was also provided.

Public reports were stored in a central database and recorded the date and time received, the administrator username, and the date and time the report was verified, closed or removed. The collection, storage and privacy of personal information was considered throughout, including minimal personal information collection and password security for administrators.

6.2.3 Web application roll-out

The web application trial was active over a three-month period between July and September 2015.

6.2.3.1 Closed trial

The web application was launched to a closed group of users to ensure the administrators were not overwhelmed with user reports. Only the web-based reporting form was available to submitters at this time. Approximately 30 people were recruited for the closed trial using one of the following methods:

- an invitation to active social media users or members of the public known to the Council/Transport Agency as potentially reliable information sources
- posts on the Transport Agency and QLDC Facebook pages and QLDC and Transport Agency Otago/Southland Twitter feeds
- a message added to the bottom of the QLDC daily road email updates
- internal promotion to QLDC/NZTA staff and wider networks.

6.2.3.2 Open trial

The trial was opened to the general public three weeks after the closed trial commenced. The road reports webmap was launched at the same time to help build interest in the trial and provide instant feedback for users who made submissions.

A message was posted on the QLDC Facebook page on 17 August 2015 containing an explanation of the trial and links to the map and form. The trial was also promoted in the August 2015 QLDC newsletter 'Scuttlebutt'. Following the initial stages of the open trial, further messages promoting the trial were posted on the QLDC Facebook page prior to, or during, major weather event when travel disruption was possible.

6.3 Social media monitoring

While the web application trial was underway, the research team monitored the QLDC Facebook site using Facebook Analytics to explore the number of page views and reports submitted by users over the same period. This information was used to compare the level of engagement and number and type of crowdsourced reports received via Facebook and web application (chapter 7).

6.4 Social media (Twitter) mining

Mining of social media channels was investigated by:

- researching the publicly available data being generated from these sites
- researching the cost of obtaining data from these sites which is not publicly available
- quantifying the likely amount of data that could be extracted from these sites
- assessing the value of the data extracted from these sites.

Several social media sites were considered for mining. The social media site with the greatest usage in New Zealand is Facebook, with approximately 54% of New Zealanders using this site (Adcorp 2013). Facebook has very restricted data access and does not release data to the general public, making it unavailable for the purposes of this research. The next highest social media sites by usage were WordPress (23%), Tumblr (17%), LinkedIn (16%) and Twitter (8%). WordPress, Tumblr and Linkedin were deemed to be of little value to the trial due to the nature of their content. Twitter can provide real-time data (Tweets) via an API and sells historic datasets upon request. Twitter was therefore considered the best option for testing social media mining methods.

6.4.1 Twitter mining trial

A live stream of Twitter data is freely available through a streaming public API that returns a small, random sample (~1%) of all public Tweets. Access to a greater sample of Twitter data or the full Twitter stream, also known as the 'Firehose', can be purchased through services such as GNIP[12].

Twitter data is provided in JSON (JavaScript Object Notation) format that can be collected and processed using scripting languages. Collected tweets can also be filtered using location and keyword(s). Tweets can be geotagged (linked to a geographic location from which they were sent), providing a spatial location which can be mapped. Only approximately 2% of tweets are geotagged which limits the amount of spatial data that can be retrieved from the API feed.

The public Twitter API was tested to determine if any useful data related to road conditions could be retrieved. This was not undertaken during the trial period and could not be completed retrospectively at the conclusion of the trial. Instead, the public Twitter API was tested over a one-week period beginning 21 October 2015. Tweets generated in New Zealand and containing the keywords 'snow', 'ice', 'accident', 'crash' or 'slip' were extracted for further analysis (chapter 7).

[12] www.gnip.com

7 Trial results and analysis

At the completion of the trial, the research team analysed and compared the level of public engagement for the web application and QLDC Facebook site. A follow-up survey of Queenstown residents was also undertaken. Similarly, QLDC staff were interviewed for their perspectives of the trial.

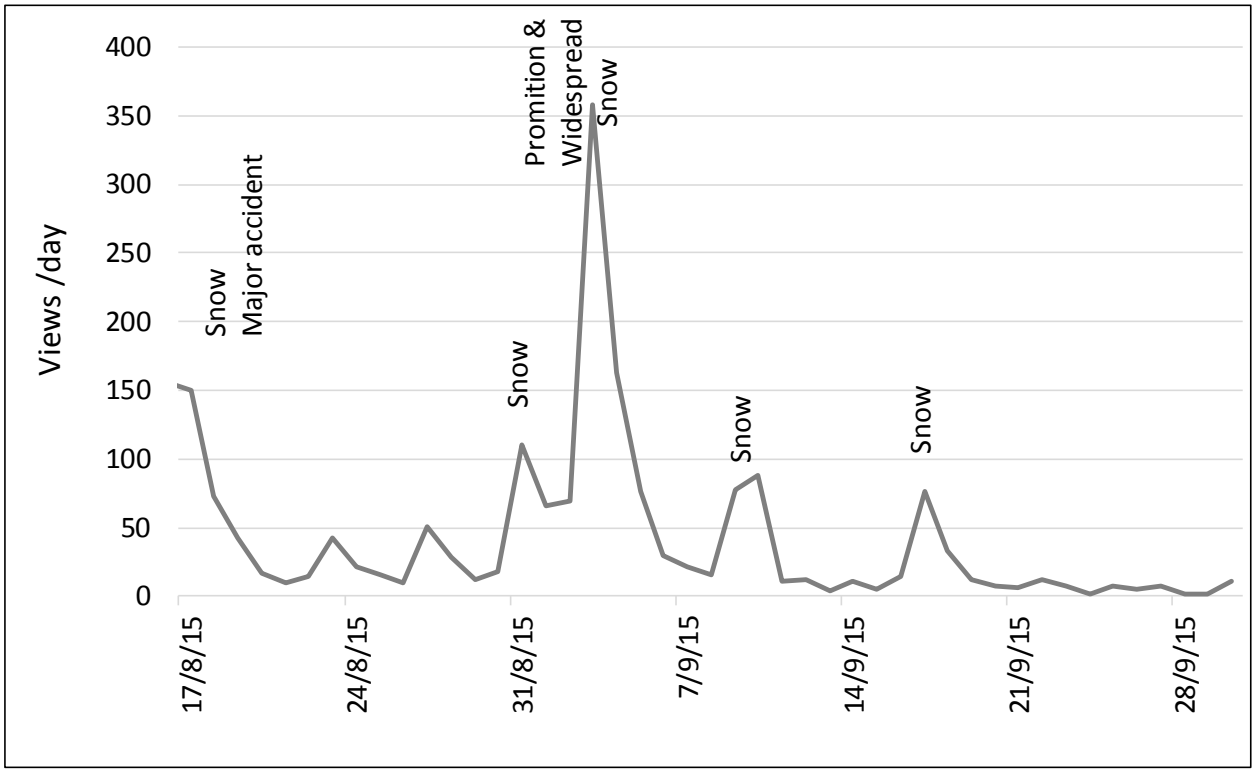
This chapter presents the results of the trial and the trial follow-up. A separate discussion of the results of the social media (Twitter) mining trial are provided at the end of this chapter.

7.1 Trial results

7.1.1 Web application

Figure 7.1 displays the number of views per day for the road conditions mapping website during the trial period, demonstrating that site viewership peaks occurred when the application was promoted or when there was a severe weather event.

Figure 7.1 Road conditions web application views per day



Over the duration of the trial, approximately 2,000 views of the web application were recorded. At its peak, the web application attracted 358 views in one day (4 September 2015). This peak coincided with the trials promotion on the QLDC Facebook page and a widespread snowfall event.

Only five crowdsourced reports were submitted via the reporting form during the trial period. Most of these were ice or snow reports. One report identified a number of potholes on State Highway 6 through Kawarau Gorge, which the reporter considered a danger to vehicles. This report was not received through any other communication channel and enabled QLDC to respond quickly by sending contractors out to assess the damage.

A likely reason for the low number of reports is the start date of the trial – at a time after a number of significant snowfall storm events had already passed and the novelty of reporting weather-related events would have dissipated.

7.1.2 Facebook monitoring

Figure 7.2 shows the daily total reach of the QLDC Facebook page over the same period as figure 7.1. 'Daily total reach' measures the number of Facebook users who have seen or accessed the QLDC Facebook page. The largest spike in daily total reach (~15,000 views) was also on 4 September 2015 during the major snowfall event.

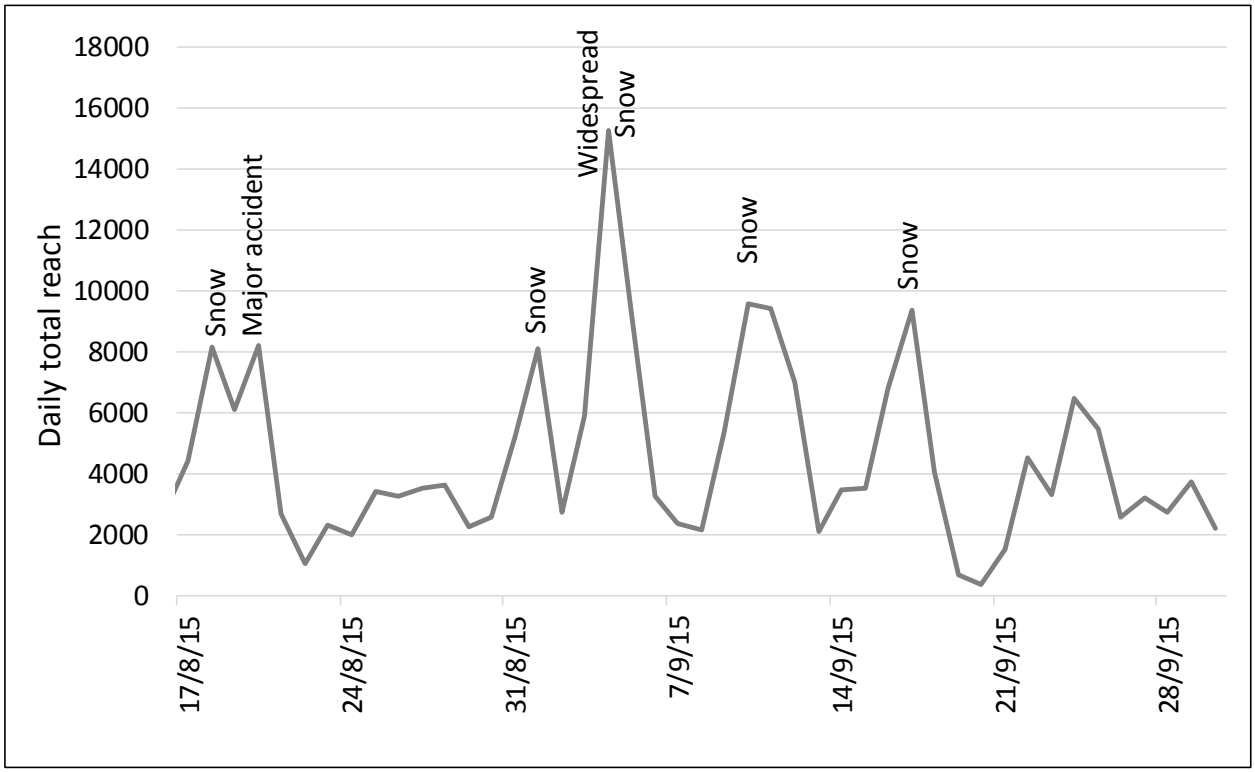


Figure 7.2 Daily total reach of QLDC Facebook page (17 August 2015 to 30 September 2015)

As well as being an 'information push' channel, the QLDC Facebook page performed a sharing function whereby page users would tag friends in a post about road conditions if they thought the information was relevant. Out of the 24 specific public comments about road conditions on the QLDC Facebook page, 10 were submitted during the snow event which began on 18 July 2015.

Questions about road conditions on the QLDC Facebook page were answered by QLDC or, in some instances, other Facebook users. Facebook users also often posted questions for information on road conditions that they could have found published elsewhere, for example the Transport Agency state highway conditions website. When this occurred, the QLDC Facebook administrators would either post an update, or direct the user to the Transport Agency website.

7.1.3 Comparison of crowdsourcing methods

Table 7.1 summarises the number and general content of reports received through the web application and Facebook monitoring.

Table 7.1 Comparison of reports submitted through the crowdsourcing application and Facebook

Crowdsourcing stream	Monitoring period	Number of reports submitted	Number of views during the monitoring period	Type of reports submitted
QLDC Facebook page	1 July 2015 – 30 Sept 2015	24	48,781[13]	Reports of ice/snow Reports of crashes/current events Weather conditions Confirmation of road conditions Weather conditions
Web application	27/ July 2015 – 30 September 2015	5	1,912	Reports of ice/snow Reports of wet roads Weather conditions Poor road condition

When comparing the type of information submitted to Facebook and the web application, the content and detail of reports were similar. The large number of views relative to the number of reports received indicates that most users were using these channels to receive information rather than provide it.

7.2 Survey results

A survey was undertaken at the end of the trial period to understand the public response to the trial. The survey was prepared using SurveyMonkey and approved by the Transport Agency legal team.

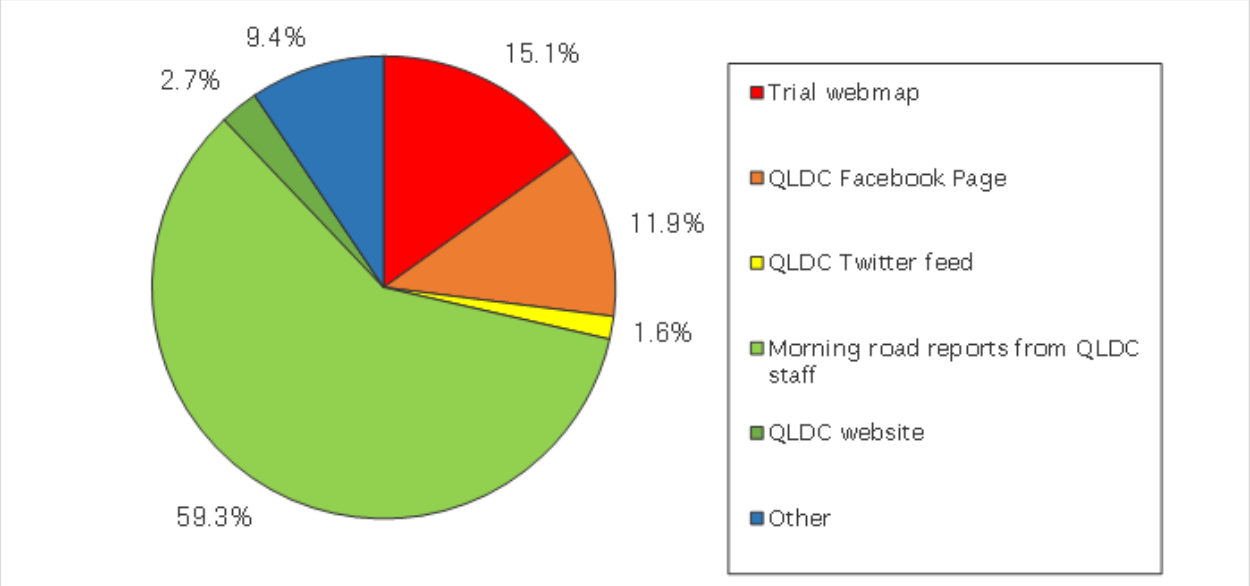
The survey was publicised via an email to the QLDC winter road reports email group and through a post on the QLDC Facebook page and as such identified existing engaged users rather than being a random sample of the travelling public. For this reason, there was a strong element of bias in the results as most survey participants currently use Facebook and/or the email facility to receive traveller information and to provide road reports. Subsequently no statistical inferences have been made around the survey results in the following sections. The survey was completed by 391 participants. The key results from the survey are summarised in the sections below. Detailed results are provided in appendix D.

7.2.1 Receiving road information

Figure 7.3 displays the preferred methods of receiving road information. Most respondents indicated they preferred the morning road report email from QLDC staff. This is to be expected as many respondents had been contacted via this mailing list. Smaller proportions of respondents indicated they preferred the trial web application (15%), QLDC Facebook page (12%), or other methods (9%).

[13] The number of people who engaged with the QLDC Facebook page by either clicking on something, commenting, or creating something on the page.

Figure 7.3 Preferred methods for receiving road information

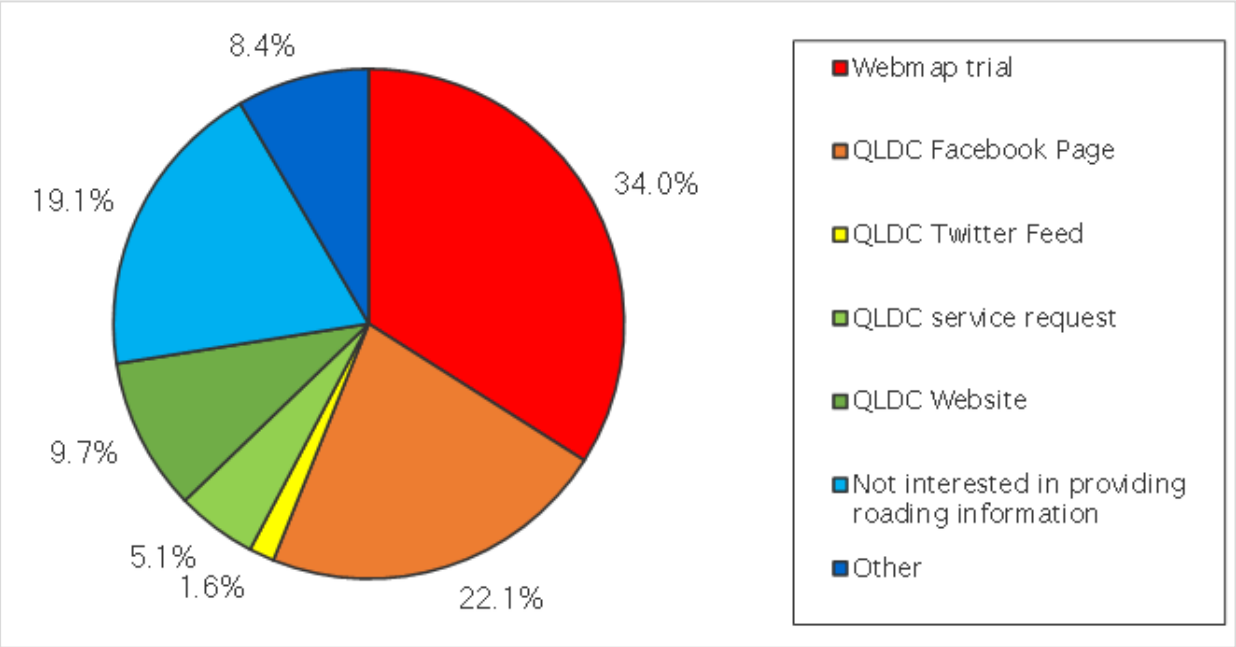


7.2.2 Provision of road information

The survey confirmed that the majority of respondents did not provide any road-related information to QLDC (88%). Of those who did submit information, most preferred the QLDC Facebook page, direct communication with QLDC staff, or a different communication channel. There were also some comments from respondents indicating they would have been happy to contribute road reports if they had known they could.

Figure 7.4 displays the preferred methods for reporting road conditions. The highest proportion (34%) preferred to provide information through a tool similar to the web application developed for the trial. The second preferred option was providing information through the QLDC Facebook page (22%).

Figure 7.4 Preferred methods for reporting road conditions



Nineteen percent of respondents said they had no interest in providing information about road conditions. These respondents were questioned to determine what measures might encourage them to submit this information in the future. Responses included having a dedicated mobile application or mobile text report number.

7.2.3 Response to the web application

Forty-six respondents (13%) indicated they were aware of the web application before taking the survey. Of these, over half discovered the trial through the QLDC Facebook page (35%) or word-of-mouth (20%).

Of those who were aware of the application, 27 said they had not seen anything worth reporting, four indicated they had submitted something using the application and four indicated they used a different method to report road conditions. Six respondents provided other reasons for not submitting reports, including that it was too difficult to use or that they were not used to working with technology.

7.2.4 Demographics

Respondents were asked to identify their age range, how long they had lived in the district (if at all), and the home and work suburb or locality. It was found that the average age of respondents was approximately 50 years old and almost three quarters (73%) had lived in the area for more than two years. The majority of respondents (68%) worked in the main employment centres (Queenstown, Frankton, Arrowtown or Wanaka). Fourteen percent worked outside of the district and 9% were either unemployed or working from home.

Most respondents (79%) lived outside the main Queenstown-Frankton urban area – 33% of respondents in Wanaka and adjacent areas accessible from Queenstown via the Crown Range Road, and 11% of respondents lived outside the Queenstown-Lakes District. Overall, the responses indicated that a large proportion of respondents travelled a considerable distance to work and subsequently could be reliable sources of real-time road condition reports.

7.3 Feedback from QLDC

The QLDC staff responsible for administrating the trial website and Facebook page were interviewed following the trial for their perspectives. The key findings from these interviews are reported below.

7.3.1 Verification of reports

The uniqueness and severity of each crowdsourced report dictated the amount of verification required by QLDC administrators. For example, isolated reports of ice would not be followed up if ice is a widespread problem in the district at that particular time. If a significant isolated report was received, for example a crash, verification would generally be sought before issuing a public notification. Public reports can be verified in a number of ways, including checking:

- with other QLDC staff members who had driven the affected road
- with contractors or Police
- other social media channels for similar reports.

Certain members of the public known to QLDC staff were considered reliable sources of information and their reports generally considered trustworthy. QLDC staff also monitored conversations on other popular

local Facebook groups such as 'Queenstown Whinge'[14] and 'Queenstown Buy Sell or Swap'[15] to find or validate public reports.

The QLDC staff felt that if the number of crowdsourced reports through the web application and Facebook page increased, they would need formalised procedures for verification and response, particularly if the information was from an unknown user.

7.3.2 Internal policies and procedures

In terms of communication policy, Facebook and other crowdsourced social media information are treated the same as any other communication channel by QLDC. Public posts on the QLDC Facebook page are automatically viewable by other Facebook users, some of whom would then occasionally post a response to provide additional discussion or context. These conversations are monitored by QLDC and the page administrators moderate these when necessary.

7.3.3 Usefulness of crowdsourced information

QLDC staff stated that road-related events could be missed if public reporting did not occur. In addition to identifying events as they occur, the significance or context of an event can also be discovered through these reports.

If there were an increase in the amount of crowdsourced information being received, QLDC would likely treat it the same way as at present, but possibly with some policy around how to filter the incoming information for severity and value. More reports would help to provide a clearer picture of exactly what was occurring on the district's roads in real time.

The administrators also highlighted the 'community' function of the QLDC Facebook page. Rather than being viewed simply as an information 'push' channel, the public felt comfortable sharing information and intelligence with QLDC and each other. QLDC administrators acknowledge user comments as much as possible to show their input is valued, and to provide timely responses to questions.

7.3.4 Usefulness of web application

QLDC staff noted that if the trial had been conducted during June and July when the worst winter weather occurred, there might have been greater awareness of the web application. Harnessing public interest in the first major winter weather events would have resulted in increased use of the application during subsequent minor weather events.

The web application itself was considered easy to use by QLDC staff but they thought the style of the reporting form might have been a deterrent to users who were more familiar with the existing service request and Facebook channels. Better integration between the web application and Facebook page could have delivered greater value.

The staff found the web application easy to administer; however, having to manually update road condition for each corridor on the map took time. If continued in the future, this process would need to be simplified.

[14] https://www.facebook.com/groups/queenstownwhinge/ (~6,500 members)

[15] https://www.facebook.com/groups/556109064463545/ (~5,500 members)

7.4 Social media mining

During the week-long trial of the Twitter API, some tweets were extracted that matched the query parameters but no specific information relating to road conditions was found. For example, a keyword match against the word ‘ice’ included Tweets on ‘ice hockey’ and ‘ice cream’. As the free Twitter API only includes a subset of all Twitter posts it is possible that Tweets regarding road conditions were posted during the trial period but not included in the public API feed. Of all the Tweets received, only a small percentage (~2%) were geotagged with a spatial location.

7.4.1 Discussion of data mining results

The Twitter mining trial did not generate any useful content due to the small subset of data available through public API and the relatively basic filtering methods that were applied. Future results would therefore be improved by:

- purchasing a larger subset of data
- using more sophisticated filtering techniques such as wordset algorithms and semantic analysis.

Access to a larger subset of data is only possible through paid services such as GNIP. Any decision to invest in this service would need to be weighed against the potential value of the data that can be returned. This would be most successful if applied on a dense urban environment where a greater density of tweets is available. The filtering of tweets could also be enhanced by using word set algorithms (Fu et al 2015) or other advanced filtering techniques.

The low rate of geotagging makes it difficult to locate the source of a tweet and it is possible that the location of the tweet does not related to the road event reported. For example, a user may wait until they have arrived at work to post an image of a fallen tree they discovered during their commute to work. Therefore, it would be almost impossible to assess spatial precision or aggregate tweets based on geotagged location data.

7.5 Summary of trial results

Although the success of the trial was limited by the start date (after winter had begun), the overall response to the web application and reporting form suggest there is value in continuing the trial in following winters or expanding it to other areas.

From the analysis of results, including the survey and staff interviews, the following matters were identified as areas for future enhancements, specifically for application in the Queenstown-Lakes District:

- Undertake a survey or focus group with key stakeholders to better understand local user needs for both reporting and receiving road related information.
- Redesign the map and public reporting interface to better meet identified user requirements and to improve usability.
- Ensure the web application and reporting form is easy to access from existing reporting channels, for example by creating permanent links from the QLDC Facebook page, morning road reports email and QLDC website.
- Clearly communicate the type of information QLDC would like reported.

- Target users who regularly travel particular routes (such as the Crown Range Road) whose reports can be trusted without substantial verification. This could include commuters, commercial operators and QLDC employees.

General findings from the trial that have specific relevance to the objectives of this research are:

- Both the custom web application and Facebook show potential for sourcing information on rural road networks where coverage by existing information sources is poor.
- The QLDC experience with Facebook supports user engagement and two-way communication. This is fostered by:
  - establishing a reputation for timely responses to user comments/questions
  - providing timely and trusted travel information, even where details are incomplete
  - an engaged user community
  - dedicated resources for managing social media channels.
- The crowdsourcing platform should be closely integrated with existing traveller information channels to make it easy for users to report road incidents, and to provide instant gratification for users who see their reports are being acknowledged.
- The trial demonstrated how the Transport Agency and local authorities can work together to develop and launch a crowdsourcing and customer information project.

8 Discussion

This chapter collates and summarises the key learnings across all stages of the research. These are formatted as a series of questions, each addressing specific research objectives and other matters raised by the steering group during the research process.

8.1 Do strategic, legal and policy frameworks support crowdsourcing in New Zealand?

This question addresses the high-level frameworks and policies that support (or hinder) a crowdsourcing ecosystem at the Transport Agency and other organisations responsible for delivering transport services.

8.1.1 Transport and technology

The Ministry of Transport and the Transport Agency provide a clear mandate for the use of technology in transport through key strategic documents including the *Intelligent transport systems action plan 2014–2018* (MoT 2014b) and the *NZ Transport Agency position statement on intelligent transport systems* (NZ Transport Agency 2014b). Similarly, the 'One Network' philosophy of the Transport Agency supports the receipt of crowdsourced information from a range of sources with no boundary between different road types and transport modes. The broader data management goals of the New Zealand government also support the use and sharing of data to drive social and economic value.

Despite clear policy on the use of technology in transport there is no specific strategic direction for the role of crowdsourced information in New Zealand's transport system. However, generic goals for improving information collection and sharing suggest that crowdsourcing has a role to play in improving customer information and supporting smarter transport choices.

The stakeholder engagement revealed some disconnect between the strategic directions of their organisations, and how this is applied in practice. For example, innovative crowdsourcing applications are generally considered risky and existing risk management and procurement practices can be barriers to their use. The Transport Agency, in commissioning this research, has indicated a desire to provide leadership in this area. Specific recommendations on applying this in practice are provided in section 8.4 (role of the Transport Agency) below.

8.1.2 Privacy

The literature review and stakeholder revealed that nothing in the Privacy Act 1993 restricts public agencies from undertaking crowdsourcing activities. Social media platforms have become ubiquitous and users generally understand and accept privacy conditions and controls. The Department of Internal Affairs has also published a number of documents setting a clear direction for the use of third-party social media sites such as Twitter and Facebook. For custom applications developed by government, the privacy principles can be built into the system design and terms of use. Third-party applications, particularly those developed overseas, can raise additional privacy concerns that need to be addressed, particularly if a government agency actively encourages public to use them.

Consultation with the Transport Agency's legal team revealed an interest in privacy matters across all merging ITS platforms and applications, including crowdsourcing. For this reason, a specific recommendation arising from this research is that a clearer process for undertaking privacy reviews of potential ITS applications is established.

8.1.3 Safety

There are clear legislative restrictions on the use of mobile devices in vehicles that will affect the types of crowdsourcing applications that can be used or promoted in New Zealand. Investigations into safety issues associated with in-vehicle devices and user interfaces have been identified in the ITS Action Plan (MoT 2014b) and the Ministry of Transport is leading this project, which may result in changes to legislation. In the meantime, any application that enables users to actively support reports while at the control of a vehicle is clearly in breach of the law and should be carefully reviewed before public release.

8.1.4 Other matters

Government policy and direction in areas such as social media use, web application development, open data principles, licensing and copyright all provide some direction on matters that need to be considered when undertaking crowdsourcing – these are discussed in detail in the literature review (chapter 2).

8.2 How can crowdsourcing deliver value?

This question specifically focuses on identifying where crowdsourcing can deliver the greatest value in terms of filling real-time or near real-time information gaps, focusing on:

- information gaps that cannot be filled using existing technologies and data sources
- information gaps where crowdsourcing is the best or only option in terms of quality, coverage or cost.

Other matters that drive value from crowdsourcing are also discussed: the role of TIS, the potential to fill other (non-real-time) information gaps and driving an open data ecosystem.

8.2.1 High priority information needs

Information needs that could potentially be met through crowdsourced data were identified in the literature review and stakeholder engagement, and refined through the SWOT analysis. A summary of the highest priority needs is discussed below. Reporting dangerous drivers, parking availability and weather/road information was also considered as a lower priority information need.

8.2.1.1 Service occupancy/capacity (PT)

The inability to measure service occupancy was a major information gap identified by PT stakeholders. While Wellington and Auckland have 'tap-and-go' systems that could, in theory, provide real-time occupancy information, the current systems are not designed for this. A simple-to-use crowdsourcing reporting application could provide some visibility on capacity. To support user engagement, this type of application should be combined in conjunction with the development of PT traveller information applications and/or smartphone bus pass systems – empowering users by enabling them to view real-time occupancy reports.

8.2.1.2 Congestion reporting

Speed, travel time and congestion information is currently supplied through passive/probe-based collection including GNSS and Bluetooth commercial systems. These systems can already quantify congestion in terms of travel time and delay, can have variable coverage, and are less effective for understanding the context in which the congestion has developed. An active crowdsourcing application could fill this information need, either through a custom-application or a third-party service. Safety concerns, however, could be a major limitation to this service.

8.2.1.3 Monitoring events

Crowdsourced information for one-off or infrequent major events, including planned and unplanned events (eg a major sporting event or natural disaster), is particularly effective as:

- the event results in unusual or unique conditions that people are more likely to report on
- people can directly see the value in sharing critical real-time information with other transport users.

Specific considerations in this application include the ease in which the application can be deployed and accessed and providing semi-structured reporting options that take into account the range of incidents occurring on transport networks. Active applications and social media are likely to be the most effective for this purpose, particularly when coupled with customer information and updates.

8.2.1.4 Incident and hazard reporting

Real-time information on incident and hazards are a particular issue on rural roads, where the coverage provided by other information channels (eg passive data sources or CCTV) is limited or non-existent. This type of reporting was applied in the QLDC trial, where users were provided with an application for reporting incidents such as crashes, road conditions (eg potholes) or weather conditions. A custom application is the best opportunity for this, although Waze does offer some limited incident reporting data. Social media monitoring and mining are also feasible in certain environments (see section 8.3). Safety concerns regarding users interfacing with technology would need to be effectively addressed as far as possible.

8.2.1.5 Travel times and variability

Travel time reporting in major urban areas and on the regional state highway network is increasingly provided through GNSS and Bluetooth sensor technologies. These technologies generally report average travel times in a near real-time sense and have limited capability to report upon travel time variability or reliability.

There may be a secondary role for active crowdsourced information to provide a measure of the variability in travel times experienced by road users; however, the data generated through existing technologies could be enhanced to calculate and share travel time variability measures through traveller information channels. As for congestion reporting, active crowdsourcing reports would be more effective in providing context to explain why travel times may be outside of the norm. In most instances the ‘why’ will relate to an incident on the network such as a crash, breakdown or road works.

8.2.2 Crowdsourcing and traveller information systems

Feedback from stakeholders and the trial results suggest that crowdsourcing applications deliver greatest value when coupled with a traveller information service. The coupling of traveller information and crowdsourcing is one of the major recommendations for this project. This is effective for the following reasons:

- Users who submit information receive gratification by seeing their report publicised.
- The crowdsourcing collection method is promoted directly or indirectly through the service interface.
- The service creates awareness of the value of crowdsourced reports for other transport users.

For example, adding road status information in addition to public reports on the QLDC road reports webmap, in conjunction with a link to the reporting form, improved awareness and accessibility of the crowdsourcing trial. Similarly, the use of Facebook by QLDC staff for the timely reporting of road conditions encouraged other Facebook users to provide additional detail and reporting.

The decision to combine 'push' and 'pull' information in a single platform is critical as it will impact on the systems and mediums used to interact with the crowd. Users of transportation data principally want a way to consume information about the transport network, identifying opportunities or impediments that may affect their journeys, both positively or negatively. If transport agencies want to engage the crowd to contribute data, then it is recommended data collection and information delivery mechanisms are integrated. In the case of a bespoke application, a single screen should ideally display information and allow users to modify existing content or create new content. The interface must also be simple – easy to use and understand. Given that contributors of real-time traveller information are limited by time, opportunity and safety constraints, it is logical for future work to lift engagement through simplification of the interface between the agency and the crowd.

8.2.3 Crowdsourcing other information needs

While the explicit purpose of this research is to investigate the potential of crowdsourcing to meet real-time or near real-time information, a number of potential non-real-time crowdsourcing applications were identified that warrant further investigation:

- Highest priority:
  - origin-destination and trip length (all modes)
  - cycle incident reporting
- Lower priority:
  - customer satisfaction/perception (all modes)
  - cycle infrastructure updates (maps) and ratings.

In addition, real-time information from crowdsourcing applications can be collected and stored for identifying trends and changes over time. This information can then be used for medium and long-term planning purposes, including planning assessments and strategic planning projects. Most outputs of this research, including recommendations, are equally relevant for historic data collection purposes as they are for real-time or near real-time applications.

8.2.4 Delivering value through improved data sharing

NZ Transport Agency research report 572 'Detailed customer requirements of travel information services' (Chang et al 2015) provides specific recommendations as a series of tasks for improving data standards and exchange. These support the findings of this research, particularly regarding the potential integration of crowdsourced information, which is relevant above all in supporting private sector-led development of crowdsourcing platforms (as assessed in the SWOT analysis). The research team supports these recommendations.

In line with open data principles, crowdsourced data should be made public and reusable through a creative commons licence. An agency would need to establish a policy about how this data is to be held following collection, and how it is to be re-released without impacting on the privacy of individual contributors.

8.2.5 Maximising value through engagement and design

Successful crowdsourcing applications require an easy-to-use interface and an engaged user base. A number of methods for ensuring initial and ongoing engagement have been identified through the

literature review, stakeholder engagement and the trial. In addition to building crowdsourcing into a traveller information system (see section 8.2.2), user engagement can be supported by:

- targeted marketing
- levering existing social media sites and user groups
- designing with ease of use in mind
- post-implementation review.

For future projects, it is recommended that user feedback is sought throughout the design and development phase. This may include a mix of active engagement (eg focus groups), usability testing and beta testing.

8.3 Which approaches are best suited to meeting real-time information needs?

A number of approaches to crowdsourcing have been reviewed in this research, including:

- passive and active approaches
- social media
- existing applications and third-party data sources, including commercial and community-based applications, and
- custom-developed applications.

As determined through the literature review, SWOT analysis and Queenstown-Lakes trial, each approach has strengths and weaknesses depending on the audience, application and location context. Given the range of information gaps crowdsourcing can potentially fill, there is no one-size-fits-all engagement approach that can be recommended from this research. For any particular application considered in the future, a thorough review of similar applications developed overseas is recommended, using this research as a starting point. The types of crowdsourcing approaches and their suitability for different situations are summarised below.

8.3.1 Social media monitoring

Social media monitoring involves using existing social media channels (eg Twitter and/or Facebook) to receive travel information alerts from the public. Learnings from the real-world trial indicate that to be successful, agencies considering social media as a crowdsourcing tool need to have:

- an established reputation for timely responses to user comments/questions and providing timely and trusted travel information (even when details are incomplete)
- an engaged user community
- dedicated resources for managing social media channels.

Social media monitoring is considered to be best suited for understanding real-time traffic conditions, particularly providing context regarding unusual or unique events. This method is not suited to applications or locations where large amounts of feedback are received, or where there are insufficient resources to manage the social media account appropriately. In these situations, social media mining tools may be helpful.

8.3.2 Social media mining

Social media mining involves using data analytics to mine social media channels for useful transport-related information. It was found in the trial that based on usage and engagement levels, Twitter is the most suitable candidate to be mined in New Zealand using existing tools and APIs. This approach was largely unsuccessful in the real-world trial due to a lack of reports, low rate of geo-located tweets and the simplistic approach to filtering that was applied.

The data mining component of the trial had a higher likelihood of being successful in a dense urban environment (eg Auckland) where the number of reports are higher, and if more sophisticated filtering/mining techniques were utilised such as those trialled by Fu et al (2015).

Applied successfully, this approach has the potential to pick up a broad range of network incidents across multiple Twitter accounts with minimal active oversight, but should ideally supplement rather than replace social media monitoring.

8.3.3 Custom-developed applications

Custom applications can be developed to suit a range of transport networks and audiences, including alternative modes. The type and complexity of application can range from device friendly websites (as used in the real-world trial), off-the-shelf applications or templates such as Ushahidi, to bespoke smartphone applications. Custom applications should be closely coupled with traveller information services to drive user engagement and deliver value. Integration with other traveller information channels and social media sites is also desirable to improve coverage and promote engagement.

Users of existing Transport Agency tools such as the Zero Harm app provide crowdsourced information relating to transport networks on occasion and have the potential to be increasingly valuable to RCAs as they become more widely established. It is recommended that the zero harm app be monitored and reporting integrated with other channels where possible. Other initiatives and emerging technologies developed within the Transport Agency could be assessed to unlock further benefits in crowdsourcing traveller information.

8.3.4 Third-party applications

Currently, there are no real-time third-party applications that have sufficient coverage to provide useful active crowdsourced information for supporting traveller information or network efficiency. This could readily change as new applications are developed, grow, or are made available in New Zealand. It is recommended that the Transport Agency continue to monitor the evolution and growth of third-party applications in New Zealand, and keep abreast of overseas initiatives.

8.3.5 Sector-led initiatives

In terms of delivering innovation, private sector companies are better placed to deliver novel (and potentially risky) crowdsourcing applications. Improving data exchange and quality standards are recommended to drive private sector investment in crowdsourcing and transport application development (see section 8.2.4).

8.4 How can the NZ Transport Agency support crowdsourcing?

The research provided an insight into the potential benefits of crowdsourcing for transport operations and customer information, and demonstrated the value of different approaches to collecting this data. Any transport body (including RCAs and public transport providers) can undertake crowdsourcing to meet the specific needs and information gaps in their organisations. This question explores the role of the Transport Agency in creating a technical framework and strategy for crowdsourcing initiatives both within the agency and in the wider New Zealand transport industry.

The *NZ Transport Agency position statement on intelligent transport systems* (NZ Transport Agency 2014b) identifies mechanisms for collecting quality data about the use of the network as a priority for investment. While non-specific about crowdsourced information, the position statement identifies the role of the Transport Agency as:

- a facilitator and regulator of new wholesale information services
- a market follower in terms of new ITS developments, including smart phones
- encouraging new market opportunities (regarding information) and ensuring information quality.

A focus on high-quality information is noted throughout the position statement and therefore a key role for the Transport Agency is to provide leadership in terms of establishing 'trust' or quality ratings for crowdsourced information. This applies to crowdsourced information collected by the Transport Agency, or information procured from third-party providers. This research will provide a framework for developing trust models and validation policies for future crowdsourcing initiatives, noting that this may not be required for all applications.

The position statement also encourages sector-led ITS development and investment as a means of mitigating the risk of overinvestment in technologies and information services, and directs the Transport Agency to take a 'market follower role'.– monitoring global trends in terms of new ITS developments and investment practices. As crowdsourcing technologies and applications continue to develop, the best strategy for following international best practice will be to establish relationships with leading international transport bodies and academia, and join relevant crowdsourcing-focused ITS forums.

Finally, the Transport Agency can support private sector crowdsourcing and traveller information applications by supporting open data practices, as discussed in section 8.2.4.

The Transport Agency can best support transport crowdsourcing initiatives by:

- monitoring international best practice
- building and maintaining relationships with leading transport agencies and academia involved in crowdsourcing projects
- fostering relationships with TOCs and RCAs to understand information needs, gaps and local crowdsourcing projects
- supporting open data practices
- providing leadership on how to establish trust and the measurement of crowdsourced information quality.

8.5 Can crowdsourced data/information be trusted?

Given the range of crowdsourcing approaches and applications, there is no one-size-fits all model to validating data and understanding trust. An extensive discussion on how to model 'trust' in crowdsourced information is provided in the literature review, and these findings were supported through the stakeholder consultation.

The first step in determining a validation process for a crowdsourcing application is to determine a threshold for human versus automated verification. This threshold depends on a number of factors, including:

- the resources available to the project
- the number of reports expected
- the format and engagement platform in which the information is collected (eg structured or unstructured reports; social media versus custom applications)
- who will view the information and how will it be used
- whether sufficient personal information can be collected to establish user reputation or likely quality of report
- how spatial location will be recorded.

Once an agency makes a series of important policy decisions about the way it collects and delivers data, the mediums for this and the threshold for each type of event that requires human verification, then a large body of crowdsourced data can be assessed for trustworthiness through automated and algorithmic models.

A trust model for real-time crowdsourced traveller information would need to account for spatial and temporal considerations as a baseline, and preferably include data author information. An assessment of semantic content would be a valuable addition to any trust model, although this would not necessarily be required to proceed in the situation of a specifically developed bespoke application.

It is recommended that trust and validation matters are considered early in any crowdsourcing project, although the exact specifications or policy for this may need to be developed over time due to uncertainty. In the trial, for example, the number of reports expected was unknown and did not warrant the development of a trust model or policy.

9 Conclusion and recommendations

This report explores the potential for crowdsourced information and technologies to support traveller information and network efficiency outcomes across transport networks. This section discusses the key conclusions arising from the research and presents a summary of recommendations for further researchd.

9.1 Conclusions

The objectives of this research have been addressed through the discussion presented in chapter 8. The objectives were to determine:

- the extent to which New Zealand's legal and government policy framework supports crowdsourced data as a means of collecting data for use by government
- the value of real-time or near real-time crowdsourced data within a transport context, and to what extent can crowdsourced data can contribute to better service delivery and outcomes for the travelling public and the Transport Agency
- the role of the Transport Agency in creating the technical framework and strategy required to support the practical implementation of a crowdsourced ecosystem.

The research concludes that the Ministry of Transport and the Transport Agency provide a clear mandate for the use of technology in transport through their respective strategic documents and the broader data management goals of the New Zealand government also support the use and sharing of data to deliver social and economic value. There is however, no clear direction regarding the role of crowdsourcing information in New Zealand's transport system but in commissioning this research the Transport Agency is demonstrating leadership in this area.

The New Zealand government's legal framework does not restrict public agencies from crowdsourcing information, with the Privacy Act 1993 being the applicable legislation in this regard. Legislative restrictions on the use of mobile devices in vehicles will affect the use and type of crowdsourcing applications that can be promoted or used in New Zealand and at the time of writing the Ministry of Transport is working towards legislation to address this matter. Government policy and direction in areas such as social media use, web application development, open data principles, licensing and copyright all provide some direction on matters that need to be considered when undertaking crowdsourcing.

A broad range of applications of real-time or near real-time crowdsourced information was specified in the research with public transport service occupancy/capacity, congestion reporting, planned event monitoring and incident and hazard reporting identified as high priority information needs. Other real-time needs that could be met through crowdsourcing included reporting dangerous drivers, parking availability and weather/road information reporting, with a number of non-real-time information needs also identified that would support traveller information for all modes of transport.

Crowdsourcing applications are likely to deliver greatest value when coupled with a traveller information service. This is effective as users who submit information receive gratification by seeing their report publicised. The crowdsourcing collection method is promoted directly or indirectly through the service interface, and the service creates awareness of the value of crowdsourced reports for other transport users.

A number of approaches to crowdsourcing have been reviewed in this research including passive and active approaches, social media, existing applications and third-party data sources, including commercial and community-based applications and custom-developed applications.

As determined through the literature review, SWOT analysis and real-world trial, each approach has strengths and weaknesses depending on the audience, application and location context. Given the range of information gaps crowdsourcing can potentially fill, there is no one-size-fits-all engagement approach that can be recommended from this research but instead an open minded approach is recommended to embrace a range of initiatives where there is a valuable level of engagement.

The Transport Agency can best support transport crowdsourcing initiatives by:

- monitoring international best practice
- building and maintaining relationships with leading transport agencies and academia
- fostering relationships with TOCs and RCAs to understand information needs, gaps and local crowdsourcing projects
- supporting open data practices
- providing leadership on how to establish trust and the measurement of crowdsourced information quality.

A key role for the Transport Agency is to provide leadership in terms of establishing 'trust' or quality ratings for crowdsourced information. Given the range of crowdsourcing approaches and applications, there is no one-size-fits all model to validating data and understanding trust. Once an agency makes a series of important policy decisions about the way it collects and delivers data, the mediums for this and the threshold for each type of event that requires human verification, then a large body of crowdsourced data can be assessed for trustworthiness through automated and algorithmic models. Trust and validation matters should be considered early in any crowdsourcing project, although the exact specifications or policy for this may need to be developed over time due to uncertainty.

9.2 Recommendations

Specific recommendations for the Transport Agency towards the implementation of this research and future study are as follows:

1 Establish a formal process for reviewing privacy issues before a crowdsourced application is developed, or where crowdsourced data will be sourced from a third-party supplier. This process can be equally relevant for other ITS applications where personal data is collected, including the movements of individual transport users. This process shall include:

   a privacy statements and compliance with the Privacy Act 1993

   b security testing to ensure there will be no loss or misuse of personal information

   c a legal opinion regarding the re-release of data to the wider public to address licensing, data ownership and the requirements of the OIA and Privacy Act 1993.

2 Consult/collaborate with the Ministry of Transport regarding potential safety issues of any proposed crowdsourcing application or service being undertaken by the NZ Transport Agency or a RCA, including:

   a whether the proposed safety measures are appropriate for the context and user type

   b whether the application will put users in breach of existing or proposed legislation.

3 Prioritise identified high priority crowdsourcing information gaps for further scoping or development of pilot projects.

4 Re-run the QLDC crowdsourcing trial over next winter, taking into account the recommended enhancements discussed in section 7.5 of this report.

5 Consider potential crowdsourcing options when developing TIS (and vice versa) to integrate crowdsourcing data collection and traveller information provision into a single easy-to-use platform

6 Investigate potential for crowdsourcing in meeting non-real-time needs, especially regarding public transport, walking, cycling, ridesharing and other emerging applications.

7 Ensure crowdsourced data is made public and reusable through a creative commons licence, taking into account the privacy of individual contributors (as required under the Privacy Act).

8 Continue to follow social media mining trials undertaken elsewhere and consider piloting this technique on a dense urban network

9 Monitor the usage of third-party applications (and any new applications that arise) to determine future potential.

10 Build upon the findings of this research by assembling a guidance document for the industry to promote 'best practice' for crowdsourcing data to support transport objectives. A guidance document should provide direction on how to derive value from existing social media channels and address the validation of crowdsourced information.

10 References

Adams, B and G McKenzie (2013) Inferring thematic places from spatially referenced natural language. Pp201–221 in *Crowdsourcing geographic knowledge: volunteered geographic information (VGI) in theory and practice*. D Sui, S Elwood and Goodchild, M (Eds). New York: Springer.

Adcorp (2013) *Social media statistics August 2013, Australia and New Zealand*. Accessed 1 December 2011. www.adcorp.co.nz/news-blog/Social-Media-Statistics-August-2013

Ali, K, D Al-Yaseen, A Ejaz, T Javed and HS Hassanein (2012) CrowdITS: Crowdsourcing in intelligent transport systems. In *IEEE Wireless Communiciations and Networking Conference,* pp3345–3349.

Aragó, P, L Díaz and J Huerta (2011) A quality approach to volunteer geographic information. *7th International Symposium on Spatial Data Quality*, Coimbra, Portugal.

Astleitner, H (2004) Multimedia elements and emotional processes. *E-Journal of Instructional Science and Technology 7*, no.2.

Birrel, E (2015) Integrating third party crowdsourced data [T3 Intelligent Transportation Systems Training Webinar], *Using crowdsourced data from social media to enhance TMC operations*. Intelligent Transportation Systems Joint Program Office, United States Department of Transportation.

Blum, B (2012) *App steers you to parking spots*. Accessed 12 December 2014. www.israel21c.org/technology/app-steers-you-to-parking-spots

Center for Automotive Research (2015) *Crowdsourcing transportation systems data*. Report prepared for the Michigan Department of Transportation, 66pp, sourced from: www.cargroup.org/?module=Publications&event=View&pubID=116

Chan, M and A Singhal (2013) The emotional side of cognitive distraction: implications for road safety. *Accident Analysis and Prevention 50*: 147–154.

Chang, J, C Morahan, G Rive, J Thomas and C Crooks (2013) Customers' requirements of multimodal travel information systems. *NZ Transport Agency research report 540*. 177pp.

Chang, J, C Jones, K Mora, G Rive and J Beetham (2015) Detailed customer requirements of travel information services, and the effectivenesss of current channels, *NZ Transport Agency research report 572*. 237pp.

Chariot (2015) Welcome to Chariot [video]. *www.youtube.com/watch?v=bLE8pU47K90*

Christchurch City Council (2011) *Share an idea*. Accessed 8 December 2014. www.ccc.govt.nz/homeliving/civildefence/chchearthquake/ShareAnIdea.aspx

Coleman, D, Y Georgiadou and J Labonte (2009) Volunteered geographic information: the nature and motivation of produsers. *International Journal of Spatial Data Infrastructures Research 4*, no.1: 332–358.

Collideoscope (2014) *Find road collisions and near misses near you*. Accessed 11 December 2014. http://collideosco.pe

Controller and Auditor-General (2013) *Learning from public entities' use of social media*. Discussion paper from the Office of the Auditor-General. 52pp.

Cooper, JM, H Ingebretsen and DL Strayer (2014) *Mental workload of common voice-based vehicle interactions across six different vehicle systems*. Report prepared for the AAA Foundation for Traffic Road Safety, Washington DC, USA.

Corsar, D, P Edwards, CC Bailie, M Markovic, K Papangelis and JD Nelson (2013) GetThere: A rural passenger information system utilising linked data and citizen sensing. In *Proceedings of the International Semantic Web Conference (Posters & Demos),* 2013, pp85–88.

Cycle Atlanta (2014) *Cycle Atlanta – mapping the ride to better Atlanta*. Accessed 12 December 2014. http://cyclatlanta.org

Cycling Safety Panel (2014) *Safer journeys for people who cycle*. Cycle Safety Panel final report and recommendations. www.saferjourneys.govt.nz/assets/Panel-Report-Safer-cycling.pdf

Denne, TR, A Irvine and C Sweetman (2013) Blueprint for a best practice measurement indicator set and benchmarking. *NZ Transport Agency research report 522.* 147pp.

Department of Internal Affairs (DIA) (2011a) *Social media in government: high-level guidance*. https://webtoolkit.govt.nz/files/Social-Media-in-Government-High-level-Guidance-final.pdf

Department of Internal Affairs (DIA) (2011b) *Social media in government: hands-on toolbox*. https://webtoolkit.govt.nz/files/Social-Media-in-Government-Hands-on-Toolbox-final.pdf

Department of Internal Affairs (DIA) (2012) *Social media in government: how to handle a mishap.* https://webtoolkit.govt.nz/files/Social-Media-in-Government-How-to-Handle-a-Mishap-v1-0.pdf

Doan, A, R Ramakrishnan and AY Halevy (2011) Crowdsourcing systems on the world-wide web. *Communications of the ACM 54*, no.4: 86–96.

Du, H, N Alechina, G Hart, M Jackson, S Anand, J Morley, D Leibovici and M Ware (2012) Geospatial information integration for authoritative and crowd sourced road vector data. *Transactions in GIS 16,* no.4: 455–476.

Elley, R (2014) *Building our footprints competition results*. Accessed 8 December 2014. http://canterburymaps.govt.nz/building-our-footprints-competition-results

Elwood, S, M Goodchild and D Sui (2012) Prospects for VGI research and the emerging fourth paradigm. In D Sui, S Elwood and M Goodchild (2013) *Crowdsourcing geographic knowledge: volunteered geographic information (VGI) in theory and practice*. New York: Springer.

Filippi, F, G Fusco and U Nanni (2013) User empowerment and advanced public transport solutions. *Procedia – Social and Behavioural Sciences 97*: 3–17.

FixMyStreet (2014) *Fix my street*. Accessed 11 December 2014. *www.fixmystreet.org.nz*

Florida Department of Transportation (FDoT) (2014) The Waze connection, *Florida Department of Transportation's Traffic Engineering and Operations Newsletter* August 2014, pp1–2. Accessed 29 November 2014. www.dot.state.fl.us/trafficoperations/Newsletters/2014/2014-Aug.pdf.

Fu, K, R Nune and JX Tao (2015) Social media data analysis for traffic incident detection and management. *Transportaiton Research Board Annual Meeting 2015 paper #15-4022*, Washington DC.

Future Christchurch (2013) Share an Idea scoops another international award. Accessed 8 December 2014. *www.futurechristchurch.co.nz/news/international-award-for-share-an-idea*

Gilbert, C (2012) *Request for information – Share an Idea*. Letter to Layton Duncan in response to a request under the Local Government Official Information and Meetings Act 1987, 17 December 2012.

https://fyi.org.nz/request/670/response/3459/attach/html/3/Letter%20to%20Layton%20Duncan%2017.12.12%20Request%20for%20Information%20Share%20an%20Idea.PDF.pdf.html

Goodchild, M (2007) Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0. *International Journal of Spatial Data Infrastructures Research 2*: 24–32.

Grant-Muller, SM, A Gal-Tzur, E Minkov, S Nocera, T Kuflik and I Shoor (2014) The efficacy of mining social media data for transport policy and practice. *Transportaiton Research Board Annual Meeting 2014 paper #14-1716*, Washington DC.

Hackakl (2014) Hackakl. Accessed 27 November 2014. *http://hackakl.org.nz/*

Haklay, M, S Basiouka, V Antoniou and A Ather (2010) How many volunteers does it take to map an area well? The validity of Linus' Law to volunteered geographic information, *The Cartographic Journal 47,* no.4: 315–322.

Harris, CG (2011) Dirty deeds done dirt cheap – a darker side to crowdsourcing. *IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing.*

Hawkins, AJ (2015) *Uber and Lyft will be the subjects of an environmental impact study*. www.theverge.com/2015/11/13/9730458/uber-lyft-environment-impact-cost-NRDC-Berkeley-study

Heipke, C (2010) Crowdsourcing geospatial data. *ISPRS Journal of Photogrammetry and Remote Sensing 65*: 550–557.

Here (2014) *Android fans rejoice: HERE for Samsung Galaxy Smartphones. A*ccessed 17 December 2014. *http://360.here.com/2014/08/29/here-samsung-galaxy-android/*

Hopkin, J, SD Ball, A Hannay, R Hutchins, D Palmer, S Rahman, D Naberezhnykh and D Longhi (2014) Pilot trial of a transport crowdsourcing smartphone app. *Transport Research Laboratory published project report PPR719*.

Howe, J (2009) *Crowdsourcing: why the power of the crowd is driving the future of business*. New York: Crown Business.

Ishiko, H, K Nkano and R Zheng (2013) Evaluation of safety of navigation systems. In *Proceedings of the 20th ITS World Congress*, Tokyo, 14–18 October 2013.

Jamson, SL (2013) What impact does legislation have on drivers' in-vehicle use of nomadic devices? *Transportation Research Part F: Traffic Psychology and Behaviour 16*: 138–150.

Johnson, PA and RE Sieber (2013) Situating the adoption of VGI by government, in D Sui, S Elwood and M Goodchild (2013) *Crowdsourcing geographic knowledge: volunteered geographic information (VGI) in theory and practice*. New York: Springer.

Kalbasi, R, K Janowicz, F Reitsma, L Boerboom and A Alesheikh (2014) Collaborative ontology development for the geosciences. *Transactions in GIS 18*, no.6: 834–851.

Kanhere, S (2011) Participatory sensing: crowdsourcing data from mobile smartphones in urban spaces. *12th International Conference on Mobile Data Management,* 2011.

Kershi, JJ and J Clark (2012) *The GIS guide to public domain data.* Redlands, California: Esri Press.

McNamara, T (2011) *eq.org.nz – the power of Ushahidi*. IITP Techblog, 18 March 2011. Accessed 19 December 2014. www.techblog.co.nz/94-eqorgnzThePowerofUshahidi

Marshall, P (2014) *How to build a crowdsourced traffic safety network*. Accessed 9 December 2014. http://gcn.com/articles/2014/02/11/crowdsourced-traffic.aspx

Mashhadi, A and L Capra (2011) Quality control for real-time ubiquitous crowdsourcing. In *Proceedings of the 2nd international workshop on Ubiquitous Crowdsourcing*. ACM, 2011.

Miller, L (2015) *Specialised applications* [T3 Intelligent Transportation Systems Training Webinar]. *Using crowdsourced data from social media to enhance TMC operations*. Intelligent Transportation Systems Joint Program Office, United States Department of Transportation.

Ministry of Transport (MoT) (2010) *Safer journeys. New Zealand's road safety strategy 2010–2020.*

Ministry of Transport (MoT) (2012) *Transport indicators*. Accessed 18 February 2015 www.transport.govt.nz/ourwork/tmif

Ministry of Transport (MoT) (2014a) *Diverted attention 2014.* www.transport.govt.nz/assets/Uploads/Research/Documents/Attention-diverted-by-2014.pdf

Ministry of Transport (MoT) (2014b) *Intelligent transport systems technology action plan 2014–18*. www.transport.govt.nz/assets/Uploads/Our-Work/Documents/ITS-technology-action-plan-2014pdf

Ministry of Transport (MoT) (2014c) *Government policy statement on land transport 2015/16 – 2024/2025*. www.transport.govt.nz/assets/Uploads/Our-Work/Documents/GPS-2015.pdf

Ministry of Transport (MoT) (2014d) *Ensuring our transport system helps New Zealand thrive. Statement of Intent 2014–2018.* www.transport.govt.nz/assets/Uploads/About/Documents/MoT-SoI-2014-2018.pdf

Misra, A, A Gooze, K Watkins, M Asad and CA Le Dantec (2014) Crowdsourcing and its application to transportation data collection and management. *Journal of the Transportation Research Board 241*. Washington DC: Transportation Research Board of the National Academies, pp1–8.

Mooney, P, P Corcoran and AC Winstanley (2010) Towards quality metrics for OpenStreetMap. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 514–517.

Moovit (2014) *Moovit app*. Accessed 10 December 2014. *www.moovitapp.com*

NV Interactive (2014*) Consultation reinvented*. Accessed 8 December 2014. *www.nvinteractive.co.nz/our-work/share-an-idea*

New Zealand Government (2014a) *New Zealand Government web toolkit*. Accessed 27 November 2014. https://webtoolkit.govt.nz

New Zealand Government (2014b) *Web usability standard 1.2*. Accessed 11 December 2014. https://webtoolkit.govt.nz/standards/web-usability-standard-1-2

New Zealand Government (2014c) *Terms of use*. Accessed 11 December 2014. https://webtoolkit.govt.nz/guidance/legal-and-policy/terms-of-use/

New Zealand Transport Agency (2014a) *NZ Transport Agency Statement of intent 2014–18*. www.nzta.govt.nz/resources/statement-of-intent/docs/SOI-2014-18-web.pdf

New Zealand Transport Agency (2014b) *NZ Transport Agency position statement on intelligent transport systems: responding to the opportunities*. www.nzta.govt.nz/assets/resources/intelligent-transport-systems/docs/nz-transport-agency-intelligent-transport-systems.pdf

New Zealand Transport Agency (2014c) *Zero harm. New Zealand Transport Agency strategic plan 2014–2020.* https://hip.nzta.govt.nz/__data/assets/pdf_file/0017/52424/Zero-Harm-Strategic-Plan-2014-2020.pdf

Nunes, AA, T Galvao and J Falcao e Cunha (2014) Urban public transport service co-creation: leveraging passenger's knowledge to enhance travel experience. *Procedia – Social and Behavioural Sciences 111*: 577–585.

OpenHack Christchurch (2013) OpenHack Christchurch. Date accessed 27 November 2014. *http://openhackchristchurch.splashthat.com/*

Owens, JM, SB McLaughlin and J Sudweeks (2011) Driver performance while text messaging using handheld and in-vehicle systems. *Accident Analysis & Prevention 43*, no.3: 939–947.

Parko (2014) *Parko*. Accessed 15 December 2014. www.parco.co.il

QSR International (2011) *106,000 ideas: NVivo's role in the plan to rebuild Christchurch's central city*. Accessed 18 December 2014. www.qsrinternational.com/news_whats-new_detail.aspx

Raine, J, A Withill and M Morecock Eddy (2014) Literature review of the costs and benefits of traveller information projects. *NZ Transport Agency research report 584*. 61pp.

Rayle, L, S Shaheen, N Chan, D Dai and R Cervero (2014) *App-based, on-demand ride services: comparing taxi and ridesourcing trips and user characteristics in San Franscisco.* University of California-Berkeley Transportation Sustainability Research Center. http://tsrc.berkeley.edu/node/797

Rickards-Simms, T (2014) Distracted driving and in-vehicle devices: literature review. Unpublished report, Ministry of Transport. 47pp.

Roose, K (2013) *Did Google just buy a dangerous driving app? A*ccessed 19 December 2014. http://nymag.com/daily/intelligencer/2013/06/did-google-just-buy-a-dangerous-driving-app.html

Sachtleben, A (2014) *Hackakl winners try to prise Aucklanders out of their cars*. Accessed 27 November 2014. http://idealog.co.nz/tech/2014/05/hackakl-winners-try-prise-aucklanders-out-their-cars

Samsung (2014) *Here BETA*. Accessed 27 November 2014. http//apps.samsung.com/earth/appquery/appDetail.as?COUNTRY_CODE=NZL&appId=com.here.app.maps

Severinsen, J (2014) *Establishing trust for crowdsourced geographic information.* Unpublished Masters Thesis, University of Canterbury.

Smith, D, D Harris, S Clarke and P Kapugama (2014) Indentify the uses of emerging sources of digital data to assess the efficiency of the state highway network. *NZ Transport Agency research report 559*.

Sodnik, J, C Dicke, S Tomazic and M Billinghurst (2008) A user study of auditory versus visual interfaces for use while driving. *International Journal of Human-Computer Studies 66*, no.5: 318–332.

Sorenson, J (2010) One in every crowd. *Traffic Technology International 2*: 50–56.

Souliotis, N, A Tsadimas and M Nikoaidou (2014) Real-time information about public transport's position using crowdsourcing. In *Proceedings of the 18th Panhellenic Conference on Informatics*, New York, pp 1–6.

Starbird, K, J Maddock, M Orand, P Achterman, RM Mason (2014) Rumors, false flags, and digital vigilantes: misinformation on Twitter after the 2013 Boston marathon bombing, *Conference 2014 Proceedings:* 654–662.

State Services Commission (2010) *New Zealand Government open access and licensing framework (NZGOAL) version 1.* https://ict.govt.nz/assets/Uploads/Documents/NZGOAL.pdf

Steinfield, A, J Zimmerman, A Tomasic, D Yoo and R D Aziz (2011) Mobile transit information from universal design and crowdsourcing. *Transportation Research Record 2217*: 95–102.

Stewart, A (2015) *Uber to launch in Christchurch*, The Press, 15 May 2015. www.stuff.co.nz/business/industries/68543925/Uber-to-launch-in-Christchurch

Strava (2014a) *Strava*. Accessed 18 December 2014. www.strava.com/

Strava (2014b) *Strava metro*. Accessed 18 December 2014. http://metro.strava.com/

Street Bump (2014) *Street bump*. Accessed 11 December 2014. www.streetbump.org

The City of Atlanta (2011) *Cycle Atlanta: phase 1.0*. Atlanta, USA: Department of Planning and Community Development http://documents.atlantaregional.com/lci/2012Applications/Innov_Atlanta_CycleAtlanta.pdf

The Road Efficiency Group (2014) *One network road classification performance measures guidance document*. www.nzta.govt.nz/projects/road-efficiency-group/docs/onrc-performance-measures-guidance-document.pdf

The World Bank (2012) *China: Online platform engages the public in urban transport planning*. Accessed 27 November 2014. www.worldbank.org/en/news/feature/2012/05/24/China-Online-Platform-Engages-the-Public-in-Urban-Transport-Planning0

Tiramisu (2014a) *Tiramisu the real-time bus tracker*. Accessed 27 November 2014. http://tiramisutransit.com

Tiramisu (2014b) *Tiramisu informed consent and privacy policy*. Accessed 27 November 2014. www.tiramisutransit.com/t/consent

TomTom (2014) *TomTom map share*. Accessed 15 December 2015. www.tomtom.com/en_nz/maps/map-share

Trame, J and C Keßler (2011) *Exploring the lineage of volunteered geographic information with heat maps*. Hamburg: GeoViz

Tsimhoni, O, D Smith and P Green (2004) Address entry while driving: speech recognition versus a touch-screen keyboard. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, w*inter 2004, 46*: 600–610.

Wihbey, J (2014) The challenges of democratizing news and information: examining data on social media, viral patterns and digital influence. *Shorenstein Center on Media, Politics and Public Policy Discussion Paper Series*, #D-85.

Wilson, N, G Thomson, N Starkey and S Charlton (2013) Persisting mobile phone use while driving and possible solutions for New Zealand. *Journal of the New Zealand Medical Association 26*, no.1384: 138–140.

Ünal, AB, D de Waard, K Epstude and L Steg (2013) Driving with music: Effects on arousal and performance. *Transportation Research Part F: Traffic Psychology and Behaviour 21*: 52–65.

Ushahidi (2014) *Ushahidi – your tools for changing the world*. Accessed 19 December 2014. www.ushahidi.com/product/ushahidi/

van Nes, N, C Michiel, A Knapper and S Wesseling (2012) Manual interactions with ITS while driving: naturalistic driving observations on mobile phones and navigations systems. In *Proceedings of the 19th ITS World Congress*, Vienna, Austria, 22–26 October 2012.

Waze (2014a) *Waze – privacy policy*. Accessed 27 November 2014. www.waze.com/legal/privacy/

Waze (2014b) *Waze help – flagged accounts*. Accessed 27 November 2014. https://support.google.com/waze/answer/6086829?hl=en&ref_topic=6024556

Ye, M, K Janowicz, C Mulligann and W Lee (2011) What you are is when you are: the temporal dimension of feature types in location-based social networks. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems,* November 2011, 102–111.

Yin, J, A Lampert, M Cameron, B Robinson and R Power (2012) Using social media to enhance emergency situaiton awareness, *IEEE Intelligent Systems 27*, no.6: 52–59.

Zhang, Y, B Reimer, B Mehler, J Dobres, S Pala, L Angell and S Ifushi (2013) An evaluation of the visual demands of portable telematics in young adult drivers. In *Proceedings of the 20th ITS World Congress*, Tokyo, 14–18 October 2013.

Zibreg, C (2010) *Google releases Open Spot, a crowdsourced parking spot app for Android*. Accessed 10 December 2014. www.geek.com/android/google-releases-open-space-a-crowdsourced-parking-spot-app-for-android-1269214

Zimmerman, J, A Tomasic, C Garrod, D Yoo, C Hiruncharoenvate, R Aziz, NR Thiruvengadam, Y Huang and A Steinfeld (2011) Field trial of Tiramisu: crowd-sourcing bus arrival times to spur co-design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 1677–1686

Appendix A: Crowdsourcing application reviews

A number of real-world crowdsourcing applications have been selected and reviewed to demonstrate how crowdsourcing is applied in the transportation sector. This is not an exhaustive list of available applications and the information presented in this appendix has been drawn primarily from web-based searches gathered primarily over November/December 2014. Where possible specific information about each application has been sourced directly from the applications website; however, in some instances this information has been sourced from third party sites including technology reviews and news websites. The research team and the Transport Agency have not approached any of the application owners directly to verify that this information is correct and current. The sources used for each application are listed under references in each table.

Waze	
Owner	Google
Country	Initially Israel, now globally available.
What it is	A community-based traffic and navigation app that uses crowdsourced information to provide alerts on traffic incidents and congestion.
How it works	 After typing in their destination address, users drive with the app open on their phone to passively contribute traffic and other road data. They can also take a more active role by sharing road reports on accidents, police traps, or any other hazards along the way, alerting other users in the area. Waze is also home to an active community of online map editors who ensure that the data in their area is as up-to-date as possible.
Data currency	Real time and historic.
How data is collected	A combination of passive collection (vehicle speed and location from GNSS) and active collection (user entry of traffic incidents).
How the data is used	Updates live maps showing traffic incidents. A data sharing arrangement has been established with at least one government authority – Florida Department of Transportation (FDoT), providing two-way access to traffic information. This has improved FDoT's coverage, particularly on interstate highways.
Safety	Waze disables text input when it detects the car is in motion (the only way to override the block is for the user to tell Waze they are riding in the passenger seat). It has also added a hands free option that allows drivers to input alerts with voice commands.
Privacy	Waze gives the user full control over their privacy settings.
Data validation	To ensure the currency and accuracy of the real time road information, Waze monitors irregular use of their products to automatically detect bots, spam and other types of abuse. Users reporting a very high number of reports and/or sending reports that have been flagged as unhelpful by other users will be identified and unable to send further reports.
Incentives	Improved routing using real time traffic and road information. The user can earn points as they contribute road information and climb the ranks in their user community.
Link	www.waze.com
References	Birrel (2015); Roose (2013); Waze (2014a); Waze (2014b)

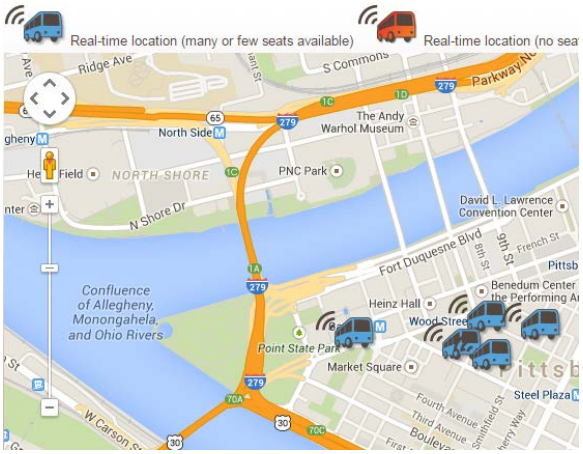
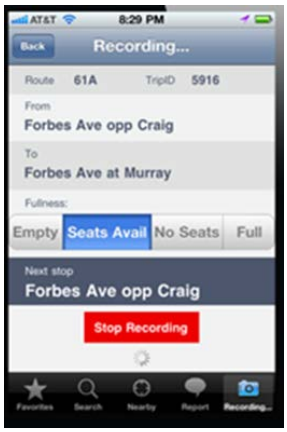
UDOT citizen reporter	
Owner	Utah Department of Transportation (UDOT), USA
What it is	A smartphone application that allows citizens to report road and weather conditions on selected roadways. The primary purpose of the application is to improve the frequency and accuracy of weather and road reports.
How it works	The reporting app allows registered users to report weather conditions (eg clear, overcast, rain, snow, fog) and road conditions (eg dry, wet, slushy) on a selected list of roadways. Users can also provide comments. Citizen volunteers are trained to use the app before being given access to it. 
Data currency	Real time, although the data stamp can be changed to allows users to submit near real-time reports.
How data is collected	Users manually input reports. Location is determined using a drop-down list of selected roadways, rather than relying on location-based sensors.
How the data is used	The information received supplements existing road and weather reports that are submitted by contractors, road weather information stations and UDOT traffic cameras. The crowdsourced information is compiled with these sources to determine road surface condition. This is passed on to traffic operation centres and to the public.
Safety	Safety matters (eg not reporting while driving) are discussed in the training programme. Warnings are also provided when signing into the application interface.
Privacy	Only an email address is required to undertake training and establish an account.
Data validation	Multiple inaccurate reports will cause a user's account to be flagged and incoming data no longer used.
Incentives	The application is bundled up with the UDOT traffic smartphone application – users can (indirectly) see how their feedback improves traffic condition reporting.
Link	www.udottraffic.utah.gov/CitizenReporting
References	US Department of Transportation (2013); Miller (2015)

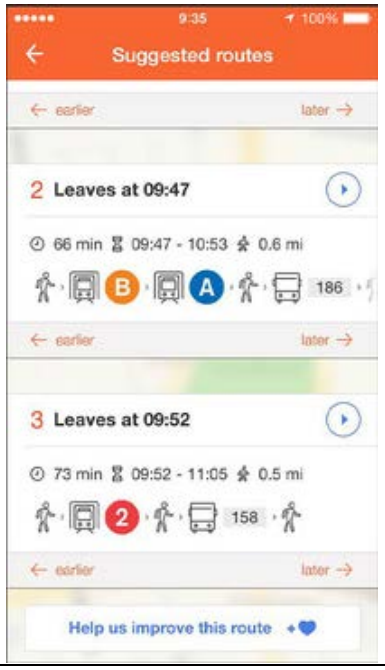
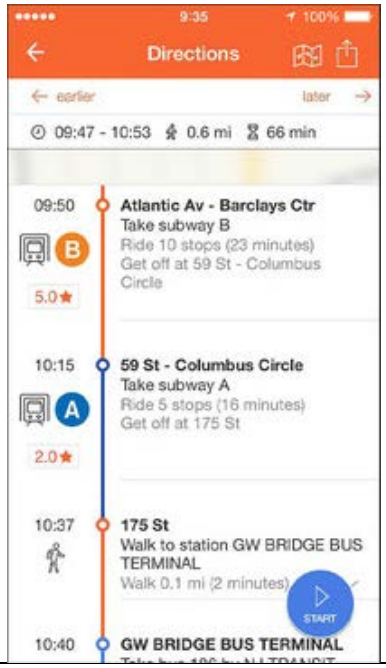
TomTom map share	
Country	Available worldwide with regional maps. Including New Zealand.
What it is	Map share allows TomTom customers to make improvements to their map directly on their TomTom navigation device and share these with the Map Share Community.
How it works	Users make improvements to the maps on their device. These improvements show some of the more dynamic and frequent changes on the road including: • changed speed limits • new street names • blocked roads • new traffic directions • altered turn restrictions These changes are uploaded to the Map Share Community and once checked by TomTom moderators, they are bundled up as free map updates to the wider Map Share Community. 
Data currency	Daily.
How data is collected	Users have two options for actively entering data. They can enter reports online which are then added to the maps by TomTom, or they edit their own map on their device and when they connect to the internet the changes are uploaded to TomTom's Map Share Community. The user can opt to allow the passive collection of anonymous usage statistic's through their user settings. This means that all of the TomTom tracks will be sent anonymously to TomTom each time the device is synchronised with TomTom Home.
How the data is used	The map data is used to improve the maps used for navigating with the TomTom device. The track data contains coordinates, speed and directions enabling speed profile databases to be built.
Safety	TomTom does not allow navigation to be edited while the vehicle is moving. This can be disabled to allow passenger changes while in transit through device settings.
Data validation	When a user Joins the TomTom Map Share community, they nominate the 'level of trust' for the map changes received. This gives the user control over map corrections received, for example, opting to receive only changes that have been verified by TomTom.
Incentives	More up-to-date maps and therefore better routing, less delays.
Link	www.tomtom.com/mapshare
References	Private (2008); TomTom (2014)

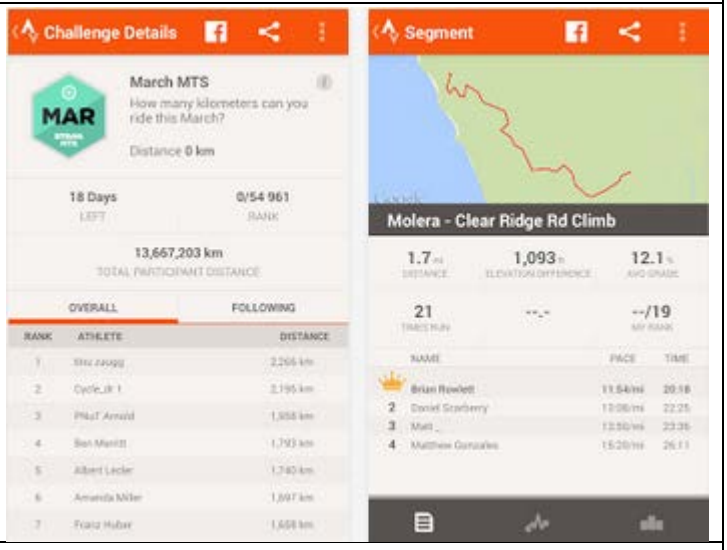
HERE Drive	
Owner	Nokia
Country	Worldwide including New Zealand.
What it is	HERE Drive is a routing and navigation application.
How it works	Here can be used online or offline. If the maps for the app are downloaded then the user can be offline and move around the map, find places, plan routes and navigate. Being offline means there is no mobile data usage and the app moves faster. Being online however, allows the user to get current information about traffic congestion and allows them to view satellite maps. 
Data currency	Real-time if online.
How data is collected	Active and passive. Drive information including user drive speed, location and direction are collected using a combination of GNSS and accelerometer data.
How the data is used	Data collected helps Here to determine current traffic congestion, preferred routes etc.
Safety	Disclaimer when loading app on device: 'Your first consideration while using the app in traffic has to be safety.'
Privacy/ownership	No exact location or other sensitive data is collected. Nokia do not on-sell user details to third parties. Nor do they keep a permanent record of where the user has been.
Incentives	Improved routing.
Link	https://help.here.com/wp8/drive/
References	Here (2014); Samsung (2014)

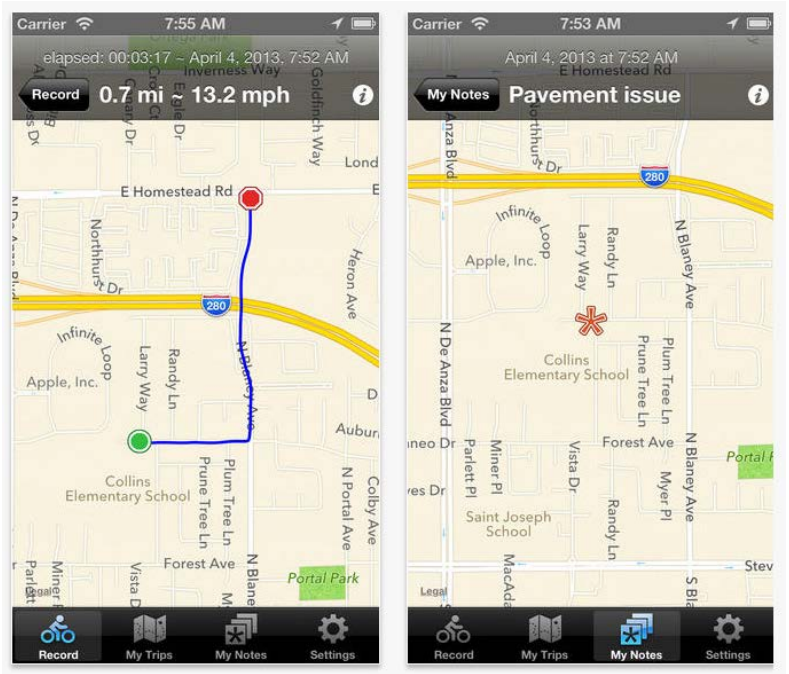
Street Bump	
Owner	City of Boston (USA)
What it is	Street Bump is a crowdsourcing project that allows residents to help improve their neighbourhood streets through use of the Street Bump mobile app. The app allows the users to passively collect road condition data while they drive.
How it works	 Street Bump is a free application that runs in the foreground and utilises a smartphone's accelerometer and GNSS sensors to detect 'bumps" in the road. Once installed, users choose to 'Record a Trip" to start recording bumps. Once completed, the "End Trip and Upload' button is clicked to submit the trip or the 'Cancel" button can be clicked to delete the trip.
How the data is used	Boston analyses the data coming in, and if multiple reports are being made for the same location, then the city will inspect the obstacle and assign it to a queue for short-term repair or record its location to assist with their long-term repair planning.
Safety	No user entry needed except at the beginning and end of the trip.
Data validation	Relies on multiple reports of the same bump before it is flagged to be checked.
Incentives	Targeted road repairs to help improve the roads travelled. Users can earn points for the number of fixed bumps that they have reported.
Other comments	In partnership with New Urban Mechanics, Connected Bits designed and developed the app, collaborating with global design company IDEO and building upon research by Professor Fabio Carrera. The City of Boston will make the app freely available so others can use and build on the project.
Link	www.streetbump.org
Reference	Street Bump (2014)

Parko	
Owner	Parko (Israel)
What it is	A crowdsourcing solution for parking problems. It is a mobile application that provides crowdsourced, real-time alerts on parking spaces that are about to be vacated, and connects motorists with available (or about to become available) parking spaces.
How it works	 Parko uses an algorithm along with data from the GNSS and accelerometers within the mobile device to know when the user is walking to their car and moving from the parking space. Users can use the detailed map viewer to view live and typical parking availability information for any day of the week, and also view parking spaces in their vicinity that are about to become available. The app includes paid parking locations and centres.
Data currency	Real time.
How data is collected	Passive. GNSS and accelerometers to determine user proximity to vacancies and alert the user.
Incentives	Users can earn prizes, coupons and real money by sharing parking spaces with others.
Link	www.parko.co.il
References	Parko (2014); Parko – mobile application to help you find parking in Tel Aviv, Los Angeles, and more (2013)

Tiramisu	
Owner	Carnegie Mellon University and Tiramisu Transit LLC.
Country	USA.
What it is	Tiramisu is part of a research study that allows users of public transportation to engage with the PT authority through sharing real-time ride information and reporting problems in the PT infrastructure.
How it works	Tiramisu processes incoming GNSS traces and generates real-time arrival time predictions for buses. It focuses on crowdsourcing information about bus location and bus load, predicting the arrival time of buses and providing a convenient platform for reporting problems and positive experiences within the PT system.  
Data currency	Real-time and historic data.
How data is collected	A combination of passive collection (vehicle speed and location from GNSS) and active collection (user entry of suggested service improvements).
How the data is used	Updates live maps showing the occupancy of services. Provides a list of reports by users on a website.
Incentives	Improved bus services by providing a mechanism for users to give feedback about the services to the transit authority.
Link	www.tiramisutransit.com
References	Steinfield et al (2012); Tiramisu (2014a); Tiramisu (2014b); Zimmerman et al (2011)

Moovit	
Owner	Developed by Israeli start-up Tranzmate
Country	More than 400 cities worldwide, including Auckland, Christchurch and Wellington.
What it is	A real-time multi-modal public transit app for routing and bus tracking.
How it works	Helps users find the fastest, least crowded public transit route by tracking the movements of its users and matching this against official transit data.  
Data currency	Real-time data.
How data is collected	Location and user input data is collected.
How the data is used	In addition to supporting real-time bus tracking, the data collected is used by Moovit to understand usage and enhance the user experience. Anonymous data is also provided to transit agencies, municipalities, government offices and/or other third parties which may have an interest in information for research, public transit planning and development purposes.
Privacy/ownership	User generated content is the sole and exclusive property of the user, subject to the applicable copyright law. Personal information is not shared with third parties except as required by law.
Incentives	Access real-time public transport information.
Link	www.moovitapp.com
References	Moovit (2014); MPA 2014 pitches – Moovit, 2014

Strava	
Country	Available worldwide and based in USA.
What it is	Strava is a community of runners and cyclists from all over the world. It allows users to connect and compete with each other via mobile and online apps.
How it works	Strava lets the user track their rides and runs via their device. The GNSS sensor within the user's device allows the user to analyse and quantify their performance. 
Currency	Real time.
How data is collected	Active and passive. Once the user has created an account on the Strava website, they record their activity and upload it to their Strava account. They can then compare their data against themselves, friends and others.
How the data is used	Individuals use Strava to find the most popular and competitive segments to run/ride. Users can run or ride a segment (specific section of road or trail) and compare their effort against past efforts, as well as other athletes who've run or ridden the same segment. Strava Metro is a data service that Strava uses to sell location tracked activities that are uploaded by users. The data has all identifiable user information removed before sale so that it just shows trends.
Privacy/ownership	Strava allows any individual activity to be private. A privacy zone perimeter can be created around any address. Profiles are viewable only by signed in Strava members, and last names can be abbreviated to provide further anonymity.
Data validation	Strava can cancel user accounts for breaches of protocol.
Incentives	Motivation for training. Compare fitness to others.
Link	www.strava.com
References	Strava (2014a), Strava (2014b)

Cycle Atlanta	
Owner	City of Atlanta (USA)
What it is	A smart phone app for recording bicycle trips. There is also an interactive map that allows users to see where other cyclists are riding. Users can view maps of their rides, track time, distance, and average speed, and share maps of favourite biking routes with friends.
How it works	The app uses a mobile device's location service to record routes in real-time, allowing the City of Atlanta to know which routes cyclists prefer. The app allows users to report problems along their route such as potholes and obstructed bike lanes. 
Data currency	Real time.
How data is collected	Passive data using a device's location service and active data as users enter problem reports into the app.
How the data is used	Transport planners at the City of Atlanta use the data to make strategic improvements to cycle infrastructure, as a means of making cycling in Atlanta more pleasant and encouraging those who do not currently bike to give it a go. This app is part of an ongoing project being conducted to promote sustainable growth in the Atlanta metropolitan area over the next fifty years. The Cycle Atlanta smart phone app will aid in collecting important and necessary data on the corridors highlighted in the map.
Link	www.cycleatlanta.org
References	Cycle Atlanta (2014); The City of Atlanta (2011)

Share an Idea	
Owner	Christchurch City Council, New Zealand
What it is	Share an Idea was a community public engagement campaign, enabling the community to share ideas for the redevelopment of the Christchurch Central City following the earthquakes of 2010 and 2011.
How it works	Using specialist software (NViVo), content collected through the Share an Idea initiative could be analysed in greater detail than could be achieved through manual data cleaning and coding processes. This helped the council identify and prioritise themes and visualise key themes through word trees and tag clouds. 
Data currency	Six weeks from May to June 2011.
How data is collected	Actively via the web application.
How the data is used	The information collected with Share an Idea was used to help inform the draft Central City Plan which would go on to guide the redevelopment of the Central City following the February 2012 earthquake.
Privacy/ownership	Private information is scattered throughout the information provided. A separate project is underway to remove this information so the raw data can be made available to the public.
Data validation	Data captured via web portal was manually checked before comments posted online.
Incentives	Allows the public to have a say in how the City should be redeveloped.
References	Christchurch City Council (2011); Future Christchurch (2013); NV Interactive (2014); QSR International (2011)

Building our Footprints	
Owner	Collaboration between LINZ, Canterbury University, Environment Canterbury, Selwyn and Waimakariri District Councils, and Christchurch City Council (New Zealand).
What it is	A crowdsourcing competition open to Canterbury high school students, with participants required to accurately map building footprints. The competition was run by Land Information New Zealand (LINZ), Environment Canterbury (ECan) and Canterbury University from 28 July-29 August.
How it works	The students involved in the competition used the Canterbury Maps web application hosted by ECan to digitise building footprints. While capturing the shapes, participants had the ability to tag them with information about the types of buildings they were (residential/ commercial/ backyard sheds etc). During the competition Canterbury high school students mapped nearly 20,000 building footprints in the Selwyn, Waimakariri and Christchurch local authority areas making a significant contribution to the rebuild. 
Data currency	One-off project.
Data collection and use	Data is entered via the custom web application. The data was used by the Council to inform a building envelope layer.
Privacy/ownership	All data collected through the competition will be released under a creative commons license so it can be reused by interested parties, including the Local Authorities and OpenStreetMap – a free crowdsourced world mapping service.
Data validation	In the Building our Footprints project, the quality of each crowdsourced building footprint was modelled using information about the author of the data, elements of the data's spatial precision, and the provenance, or temporal quality of the feature.
Incentives	Prizes for individuals who captured the most buildings.
Link	www.canterburymaps.govt.nz/buildingourfootprints
Reference	Elley (2014)

Ushahidi	
Owner	Open source software under the GNU Lesser General Public License (LGPL).
Country	Nairobi, but available for wider use worldwide.
What it is	Ushahidi provides a range of open source resources available for free download that can then be customised for specific applications.
How it works	The Ushahidi software includes: **Ushahidi Platform.** This is open source software that can be installed on a server for information collection, visualisation and interactive mapping. **Ushahidi Mobile App.** This app allows viewing and creation of incident reports. It supports loading of multiple deployments at one time, quick filtering through reports, exploring incident locations on the map, viewing report photos, news articles and media as well as sharing reports via email, SMS or Twitter. **Crowdmap** is a hosted service for mapping anything on the web, focused on a more social mapping experience.
Example applications	**Example 1:** Beijing Transport Research Center (BTRC) completed a pilot project with the aim of using a crowdsourcing methodology to change how urban transport improvements are planned and delivered. Ushahidi was used to develop the online urban transport public feedback portal, and this portal was designed to allow anyone to report issues on cycling and walking infrastructure via the web, smart phone apps, SMS or social media. The output reports could be aggregated and reported to the Beijing municipal government.  **Example 2:** Christchurch Recovery Map was a website developed following the Christchurch earthquake of February 2011. This website enabled crowdsourced information to be collected from tweets, emails, SMS messages and web form submissions and disseminated alongside third party information including official information. It was built by a team of volunteers using Ushahidi software and provided a central repository of relevant information.
Link	www.ushahidi.com
References	The World Bank (2012); Ushahidi (2014); McNamara (2011)

HackAKL	
Supporters	Auckland Transport in conjunction with AUT University and Propellerhead.
What it was	A collaborative event run in May 2014 that brought together over 300 people interested in utilising open government data to help develop new applications.
How it worked	Auckland Transport published a beta application programming interface (API) that opened up the organisation's data, including real-time bus information, geo-coding, roading and congestion data. Teams of developers were then empowered to work on developing applications based on the data provided by Auckland Transport. The supreme winner was an app called 'My AT 2.0' that integrated real-time bus tracking, travel history statistics with levels and rewards based on the distance users travel on the network.
Link	www.hackakl.org.nz
References	Hackakl, 2014; Sachtleben, 2014

OpenHack Christchurch	
Supporter	Land Information New Zealand.
What it was	A collaborative event run in December 2013 that involved backend, frontend and mobile developers to work together to build a journey planning application for Christchurch using road closure/restrictions and traffic congestion and flow data.
How it worked	The event was built around a number of key activities: • Take an XML data feed and wrap a RESTful API around it and expose it with suitable end-points. • Deploy a routing service using Open Street Map/other and integrate directional road closures and traffic congestion data so that web/mobile apps can request the best route to get from Point A to Point B. • Work with/provide feedback on ECAN's routing service • Develop a prototype journey planning application that makes it easy for road-users to avoid road closures and congestion.
Reference	OpenHack Christchurch (2013)

Appendix B: Events that can be collected using active user data collection

The following table presents a list of incidents and events that could be identified through active crowdsourcing of information.

Mode	Input category	Type of input
Private vehicle	Enforcement	Red-light camera
		Speed camera (fixed or mobile)
		Police checkpoint
	Weather conditions	Flood/standing water
		Snow and Ice
		High wind
		Fog
	Road conditions	Vehicle breakdown
		Road damage (eg pot-hole, washout)
		Crash
		Wandering animals/stock
		Signals not working
		Missing/damaged signage
	Road closure or restriction	Congestion (eg mild, moderate, severe, standstill)
		Event
		Road works
		Changed road layout
Public transport	Occupancy	Full, half-full, empty
	Quality of service	Rating (eg 1–5 scale re quality of service)
		Graffiti/damage
		Passenger or driver behaviour
		Service reliability
Walking, cycling	Pavement/trail condition	Surface quality
		Connections (missing or directions required)
		Obstructions
		Hazards
	Safety	Near-miss
		Crash
		Perception of safety (eg unlit walking tracks or dangerous intersections)

Appendix C: SWOT assessment rationale

An overview of the rationale behind the quantitative assessment of the strengths, weaknesses, opportunities and threats in the SWOT analysis is presented based upon the engagement approaches described in detail in chapter 5 of the main body of the report. These are as follows:

- Approach A – Social media monitoring.
- Approach B – Social media data mining.
- Approach C – Develop customised apps and websites.
- Approach D – Procure crowdsourced data from third party providers.
- Approach E – Support private sector-led development of crowdsourcing platform.

Each of the internal factors (strength and weaknesses) identified are measured on a scale -2 through +2 where:

- +2 is a strength
- +1 is a partial strength
- 0 is neither a strength nor weakness
- -1 is a partial weakness
- -2 is a weakness.

The opportunities and threats have been rated on a scale of -2 through +2, where:

- +2 is an opportunity
- +1 is a partial opportunity
- 0 is neither an opportunity nor threat
- -1 is a partial threat
- -2 is a threat.

The rationales for assigning the values to each factor are presented in table C.1 for strengths and weaknesses and table C.2 for opportunities and threats.

Table C.1 Strengths and weaknesses assessment rationale

Strengths/weaknesses	Assessment rationale
Integration with current sources	The Transport Agency uses social media feeds already, but it is difficult to integrate these with some quantitative sources so this is considered to be neither a strength nor weakness. Data collection applications can be designed to integrate with other sources although it is considered that there is more control over integration with the Transport Agency development or procurement compared to private sector development.
Control over data quality/validation	There are limits on the extent to which information quality can be assured from social media sources. Data mining techniques have the potential to improve the accuracy of data through careful filtering associated metadata. Trust models can be designed into apps so this is a strength particularly for any customised development.
Control over privacy issues	Privacy issues are generally accepted to be effectively managed within social media options as they are open platforms with accepted terms and conditions. Control over privacy

Strengths/weaknesses	Assessment rationale
	matters can be designed into applications, especially if they are developed with direction from the the Transport Agency.
Adaptability/flexibility	Social media sources can change over time, however the use of conventions such as 'hashtags' provides some flexibility in this regard. Data mining methods may require periodic revision. Adaptability can generally be designed into any application so it is considered to be neither a strength or a weakness.
Alignment with government strategic direction	All options generally align with the government's ITS strategy and the Transport Agency's position statement on ITS. Some options provide more opportunities to work in collaboration with the development community than others.
Initial investment (time/cost)	There is very little set-up required for social media options, especially as the Transport Agency and other RCAs already use social media channels including for the provision of traveller information. More resource is required to scope, design, develop and facilitate internal development of solutions, but comparatively this is neither a strength or weakness for third party options.
Ongoing investment (time/cost)	The extent of ongoing investment can be managed in the architecture development but the ongoing monitoring requirements of social media will demand ongoing resource.

Table C.2 Opportunities and threats assessment rationale

Opportunities/threats		Assessment rationale
Opportunities	Meet real-time information needs	Refer to table 5.1 in the main body of the report for details. There may be less value in private sector led initiatives as the Transport Agency does not have the same level of intervention and control.
	Meet other information needs	Refer to table 5.2 in the main body of the report for details. There may be less value in private sector led initiatives as the Transport Agency does not have the same level of intervention and control.
	User empowerment	A Transport Agency platform has the potential to directly target and empower users to make smart travel choices and increase their level of engagement with the Transport Agency and other RCAs. This empowerment/engagement is less likely to eventuate through third party strategies. As social media is a part of many New Zealanders everyday life there is an opportunity to increase the extent of engagement through these channels.
	Cost-effective info source	Social media options are comparatively cost effective relative to the other options. The costs may be less in private sector led initiatives compared to the internal customised development option.
	Collaboration public and tech sectors	Private sector led initiatives support a collaborative environment, and the procurement and data mining options involve a limited level of engagement with the development community.
Threats	Bias in crowd input difficult to manage	Social media sources can be bias due to the composition of the user base, and there is little ability to understand and manage this bias. Any bias can be understood and accounted for through the design architecture of applications subject to careful management of privacy matters but may be less manageable for private sector led initiatives.
	Requires high number of users	Social media mining requires a high number of users to fully develop semantic understanding, to develop trust models and to automate the processing of contextual information. This is generally less of a risk for other types but there is an opportunity to promote crowdsourcing to get a larger user base with customised and procurement strategies.

Opportunities/threats		Assessment rationale
	Prone to misuse/abuse	This is a risk with social media feeds as there is little control over misuse and abuse and subsequently the feeds will require moderation. Quality control and user input functionality can be carefully designed to mitigate this risk on other platforms.
	In-vehicle safety difficult to manage	There are no inherent safety controls on the use of social media platforms in vehicles however custom applications can be designed to take into account safety factors. Safety enhancements and training can be designed into customised apps.
	Potential to damage reputation	Any Transport Agency branded development carries a greater risk of damaging the agency's reputation if things go wrong. The agency can be distanced from this through third party engagement depending on the level of endorsement or perceived involvement.

Appendix D: Real- world trial survey results

This appendix presents the detailed results of the user survey undertaken at the completion of the real-world trial.

Table D.1 During the past winter, did you receive reports about road conditions through any of the following channels? (total responses = 392)

	Percentage	Count
QLDC Facebook page	43%	168
Morning road reports from QLDC staff	85%	334
QLDC Twitter feed	7%	26
QLDC website	15%	57
NZ Transport Agency website/Facebook page	12%	45
I didn't receive any reports about road conditions	0%	1
Other	9%	37

Respondents were asked if they had any suggestions for improving the road reports which QLDC issue across multiple communication streams. 89% of respondents either skipped the question or indicated they were satisfied with the reports being issued. The list below summarises suggestions from respondents who did provide feedback. The suggestions are ordered from most stated to least stated.

- Earlier reporting (as some respondents begin commuting before 6:30am) [10 respondents]
- Consistent report times [4 respondents]
- Text alerts with details of specific road events [4 respondents]
- Information about the state highways/major local roads in the wider region [4 respondents]
- Photos of conditions [4 respondents]
- More informative/specific information during major winter events [3 respondents]
- More updates during the day (and before the evening commute) [3 respondents]
- Skifield road conditions specified on the page rather than a link [2 respondents]
- Web cameras (particularly on Crown Range Road) [2 respondents]
- Faster alerts when road conditions change [1 respondent]
- Really obvious link on QLDC homepage to road information [1 respondent]
- Link in email to an interactive map with regular updates [1 respondent]
- More promotion of the page (particularly to international drivers) [1 respondent]
- Name the routes (eg Kelvin Heights to Frankton) [1 respondent]
- Include the next day's weather forecast [1 respondent]
- State out-of-the-ordinary information before more general information in email [1 respondent]
- Have a different field for skifield closures in the morning road report emails [1 respondent]

Table D.2 What would be your preferred method for receiving roading information? (total responses = 370)

	Percentage	Count
A road reporter tool similar to the trial recently launched by QLDC and NZ Transport Agency	15%	55
QLDC Facebook page	12%	44
QLDC Twitter feed	2%	6
Morning road reports from QLDC staff	59%	220
QLDC website	3%	10
Not interested in receiving roading information	0%	0
Other (please specify)	9%	35

Other communication methods which respondents specified they would prefer included:

- a text messaging service
- an app
- variations of email services
- More FM
- the reporter tool linked to Facebook
- live tweets from contractors and drivers
- a single platform with all crowdsourced information

Table D.3 During the past winter, did you submit reports to QLDC about road conditions through any of the following channels? (total responses = 386)

	Percentage	Count
QLDC Facebook page	6%	22
QLDC Twitter page	1%	2
Direct communication with QLDC staff	4%	16
QLDC service request	1%	3
I didn't submit any reports about road conditions	88%	338
I submitted reports through another channel (please specify)	4%	14

The communication channels which respondents specified other than options in the survey included:

- the Upper Clutha trading post Facebook page
- Wanaka Police
- the Wanaka Sun Facebook page and Twitter feed
- personal Facebook posts
- the Queenstown trading Facebook page
- More FM.

There were also some comments which stated they would have been happy to contribute road reports but did not know they could submit reports to QLDC.

Table D.4 What is your preferred method for providing reporting road conditions to QLDC/other road users? (total responses = 370)

	Percentage	Count
A road reporter tool similar to the trial recently launched by QLDC and NZTA	34%	125
QLDC Facebook page	22%	82
QLDC Twitter feed	2%	6
QLDC service request	5%	19
QLDC website	10%	36
Not interested in providing roading information	19%	71
Other (please specify)	8%	31

The methods which respondents specified other than the options in the survey included:

- text, email or phone
- Queenstown trading Facebook page
- an app
- other community Facebook pages.

Respondents who indicated they were not interested in providing road information were asked if there was anything else that could be done to encourage them to provide information. Some of this feedback included having an app or text report number. There were also some comments which stated they would have been happy to contribute road reports but did not know they could submit reports to QLDC.

Table D.5 How did you find out about the trial website? (total responses = 46)

	Count
QLDC Facebook page	16
Contacted directly by QLDC or the NZ Transport Agency	3
QLDC newsletter	5
QLDC or NZ Transport Agency Otago/Southland Twitter feeds	0
QLDC daily road email updates	7
Word of mouth	9
Other (please specify)	6

The methods which respondents specified other than the options in the survey included:

- the Transport Agency website
- direct contact with QLDC staff
- the Wanaka Sun website, Facebook page and Twitter feed.

Table D.6 Did you submit reports on the trial website? (total responses = 46)

Response	Count
Yes	4
No - I didn't have any interest.	5
No - I didn't see anything worth reporting.	27
No - I used another method to inform QLDC of road conditions.	4
No - The site didn't work on my device.	0
No - It was too difficult to use.	1
Other (please specify)	5

Reasons which respondents specified other than the options in the survey included:

- not good with technology
- someone else at home/work does the reports
- only heard about it indirectly and didn't know how to access it
- I didn't know I could provide reports with the web application.

Table D.7 Which age bracket do you fit into? (total responses = 349)

Younger than 18	18 to 24	25 to 34	35 to 44	45 to 59	60+
0%	1%	11%	20%	40%	27%

Table D.8 How long have you lived in the Queenstown- Lakes District? (total responses = 348)

Less than one year	1 to 2 years	2 to 5 years	5 to 10 years	More than 10 years	I do not live in the Queenstown- Lakes District
7%	9%	18%	16%	39%	10%

Table D.9 What is you closest work/home suburb (selecting the closest town/suburb to your workplace)?

Suburb	Work suburb (total responses = 335)	Home suburb (total responses = 337)
Queenstown	30%	9%
Fernhill	1%	5%
Kelvin Heights	0%	2%
Glenorchy	1%	2%
Frankton	18%	7%
Lake Hayes Estate/Shotover Country	1%	7%
Arrowtown	5%	15%
Wanaka	15%	24%
Lake Hawea	3%	4%
Albert Town	1%	4%
Jacks Point	0%	3%
Kingston	0%	1%

Suburb	Work suburb (total responses = 335)	Home suburb (total responses = 337)
Makarora	0%	0%
Luggate	0%	1%
Arthurs Point	1%	4%
Outside Queenstown Lakes District	14%	11%
I don't travel to work	9%	0%

Appendix E: Glossary

API	application programming interface
AT	Auckland Transport
AVL	automatic vehicle location (technology)
CCTV	closed circuit television
CGI	crowdsourced geographic information
CTOC	Christchurch Transport Operations Centre
DIA	Department of Internal Affairs
GNSS	global navigation satellite system
GPS	Government Policy Statement on Land Transport
ITS	intelligent transport systems
LGOIMA	Local Government Official Information and Meetings Act 1987
MOT	Ministry of Transport
NZGOAL	The New Zealand Government Open Access and Licensing framework
OIA	Official Information Act 1982
OSM	OpenStreetMap
ONRC	One Network Road Classification
PT	public transport
QLDC	Queenstown–Lakes District Council
RCA	road controlling authority
RFS	request for service
SMS	short message service
SWOT (analysis)	SWOT analysis involves pinpointing four key areas (strengths, weaknesses, opportunities and threats) to assist decision makers in identifying the key facts of the issue/situation involved
TIS	traveller information systems
TLA	territorial local authority
TOC	traffic operations centre
UDOT	Utah Department of Transport
VMS	variable message sign