A commonly given reason for not using public transport in rural areas is lack of convenience, which includes inadequate passenger information. Real-time passenger information (RTPI) systems use a variety of information sources to provide passengers with information such as estimated vehicle arrival times and notifications of delays or cancellations. Few RTPI systems exist in non-urban areas, for a variety of reasons, including lack of infrastructure for obtaining and providing real-time information.

**Problem**

In the Informed Rural Passenger Project we are developing GetThere, a RTPI system for rural areas. GetThere uses data from multiple sources (including observations such as vehicle location and occupancy level crowd sourced from users), and delivers information via multiple channels (smartphone app, SMS).

**GetThere**

**Data Quality**

Given the open nature of the data, issues around quality and trust naturally arise. We have developed a quality assessment model and associated reasoning service to evaluate data used by the system.

**Disruptions**

Users can report travel disruption, such as road closures and roadworks. These are described using the travel disruption ontology and linked to the infrastructure dataset. This allows, for example, the quality assessment service to consider any disruption that may affect a particular bus route and the location observations associated with it.

**Deployment**

Working with various stakeholders in the Scottish Borders, UK, we are conducting an extended study, deploying the GetThere system on eight First Group routes, investigating changes in travel behaviour as a result of using the system. User feedback gathered through interviews and focus groups will be used to guide future development of the system.

**Discussion**

We use open datasets to provide domain data where possible. This has presented various challenges including dealing with information that may be out-of-date, and integrating heterogeneous data sources. Provenance and quality assessment are used to address the former, while linked data is used to address the latter. Although this has required converting some existing datasets, it has brought benefits in terms of simplified integration and maintenance.