Queensland Government

Queensland **Transport**

Developing a transport system data architecture for strategic planning and performance measurement

Alan Beale

Data Architect

Transport Information Systems

Integrated Transport Division

Queensland Transport

Chris Oldham

Information Manager (Transport Systems)

Transport Information Systems

Integrated Transport Division

Queensland Transport

30 July 2006

Contents

1. Introduction 3

2. Scope 3

3. Background 4

4. Data architecture—what is it? 5

5. The Data Architecture – how it fits 6

6. Data architecture—what can it do? 6

7. Meeting Queensland Transport business goals 7

8. Transport Systems Information goals 7

9. Building the architecture 8

9.1 Compiling the attributes 9

9.1.1 Abstracting the attributes from the business plan (TCP) 9
9.1.2 Testing the attributes against the business plan (TCP) 10
9.1.3 Scoping and clarifying the attributes under user consultation 11

9.2 Building the activity dimension using the compiled attributes 11

9.3 Building the data analysis dimension using the activity dimension templates 12

10. Carrying out the data assessment 14

11. The summary version 14

12. Meeting data collection cycles 15

13. Linking users to the data architecture 16

14. The Benefits of the data architecture 16

15. The Future 17

16. Conclusion 17

17. Appendix 18

18. References 22

1. Introduction

The larger an organisation or business becomes the more critical is its need to maximise efficiencies in data management. The data architecture described in this paper addresses aspects of this need by providing a methodology that maps certain business needs against data-gathering activities. It offers a foundation for strategic planning and performance measurement, firstly, through applying specific attributes that indicate what data needs to be sourced, and secondly through built-in measures that rate the adequacy of existing data and datasets. The methodology thus highlights areas where collected data may be of unacceptable quality; or has been unnecessarily duplicated or is missing, and exposes datasets that may be inaccessible. Such findings can subsequently lead to specific remedial strategies. These include efforts to source the exact data required, and initiatives that result in easier access to datasets that may otherwise require prohibitively long lead times to access.

The data architecture is therefore able to support efficient decision making and performance monitoring through targeted data collection and distribution.

Created by the Transport Systems Information unit at the Queensland Department of Transport (QT), the architecture enhances the Department's stewardship of the state transport system by providing a framework that guides and drives data collection. It does not model QT systems; however, it does provide a structure and procedures by which all data-reliant divisions within the Department benefit.

2. Scope

At present, QT is developing a full Information Management program aimed at defining and modelling all data and information systems across the organisation. The data architecture as outlined in this paper is a subset of this activity and developed outside this scope. It does not model existing QT systems, instead, it takes a pro-active approach leading data collection and fulfilling Queensland Transport's mission to 'develop, lead and manage transport in Queensland'.

To date there has been little published on data sourcing methodologies—especially for the transport industry. Mostly, the term data architecture has been used to describe existing business system relationships and the flow of data through an organisation—particularly focussing on information and communication systems. These information and system architectures are not within the scope of this paper.

In this paper, the term data architecture refers to a framework that maps an organisation's high level business goals and decision-making processes against its data gathering activities. This type of data architecture or data collection framework is designed to lead an organisation's data sourcing activities—and is recommended for organisations such as government transport bodies that rely heavily on surveys, statistical analysis and business modelling. The data managed by this high-level data architecture focuses on strategic planning and performance measurement information that has many uses for many users. It does not cater specifically to project data needs, although it will undoubtedly assist projects to some degree.

3. Background

Queensland Transport stores its planning and performance data in around 200 datasets ranging from databases through spreadsheets to reports located within various branches across the department. Some information within these datasets is replicated across more than one dataset and other strategic information is missing (PAS, 2003)—indicating inefficiencies in data collection.

Data at QT is collected in both centralised, common acquisition as well as decentralised independent acquisition by its various branches. As QT and the Department of Main Roads come under the same portfolio, there is also a certain amount of data sharing across both departments, although each department gathers some data independently of the other. Data is also sourced from external agencies such as Queensland Rail, the Australian Bureau of Statistics and the Brisbane City Council. Of course, these external bodies also conduct their data gathering activities independently of QT.

Until the present, data collection and data management within QT has taken place without an overarching framework, which at times may have led to data duplication. Additionally dataset access difficulties, data gaps, and dataset under-utilization have been areas of concern.

Consequently, this raises the need for a methodology that will guide efficient data gathering and assess what data is available, appropriate and accessible—across the department. Such a methodology or data architecture provides a basis for QT's on-going data gathering program.

In 2003 a report was commissioned by the Performance and Information Strategy Team (P&IST - a team of QT and MR senior officers) to examine and make recommendations concerning the apparent data gaps within the transport portfolio. This report identified certain data shortfalls and recommended that budget and resources be allocated to addressing this area. The executive agreed to these proposals, and charged Transport Information Systems Branch (TIS) with the task of sourcing information to address these known data gaps.

As this work was progressing, in consultation with the P&IST, it became apparent that a framework was necessary to describe existing and potential gaps. Additional pressures, such as urban congestion, was changing the potential priorities of information needs and this needed to be described in terms that the business could direct resources to best effect. To address this need, TIS developed a framework, or data architecture, to provide a platform for ongoing business analysis.

4. Data architecture—what is it?

The data architecture is a methodology that provides several multi-dimensional high-level views of an organisation's strategic planning and performance measurement data. Such views offer multiple assessments that reveal at any given time, data condition—in the light of current corporate goals and objectives.

The architecture consists of four key elements:

1. A set of data attributes which are labels that represent categories of information—for instance the attributes 'volume of activity', 'user perception' and 'transport emissions' represent many fields in each dataset. (See Appendix for a list of attributes and their definitions.)
2. An activity or modal dimension—consisting of one layer for each mode of transport
3. A data analysis dimension consisting of industry-standard statistical analysis parameters.
4. A 'lightweight' dimension that summarises the above.

The architecture is derived in three steps by:

5. Abstracting the data attributes from the corporate plan
6. Scoping the attributes under user consultation
7. Building the framework whereby the attributes can be assessed against the datasets for each mode using the analysis parameters

5. The Data Architecture – how it fits

The Data Architecture fits snugly into QT's existing data program by providing data assessments on demand.

Once the data assessment phase completes, the results indicate what data collection strategies will be required for the next data program. The data program is completed when all data specified by the program has been gathered. The datasets can then be accessed by users for planning and performance measurement.

The data analysis cycle commences again when user feedback and data collection cycles then trigger the next data assessment program.

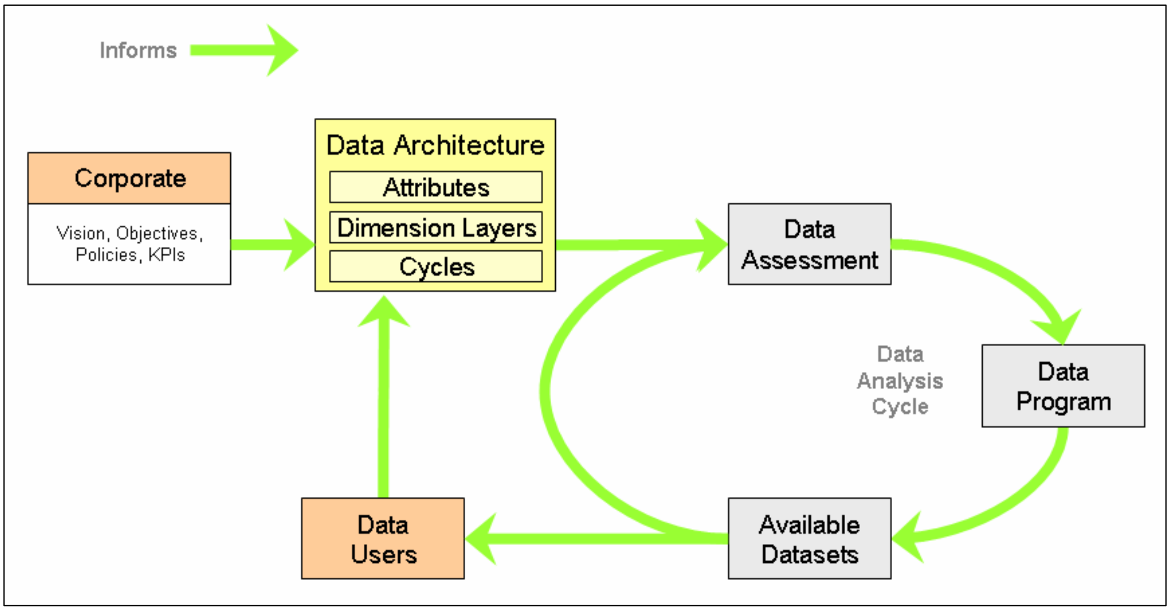


Figure 10: Drivers acting upon the Data Analysis Cycle

6. Data architecture—what can it do?

The data architecture is a framework by which an organisation can manage its data gathering and data identification activities. It is a multi-faceted system which is based on business goals and objectives, and is able to carry out the following:

1. Specify what data needs to be collected for strategic planning and program planning and budgeting
2. Specify what data needs to be collected for performance monitoring and review
3. Specify how often the data needs to be collected
4. Assess collected data against each business activity, viewed by parameters such as:
    - the relevance of the data to the user
    - the usefulness of the data measured against user requirements
    - the currency (age) of the data
    - the level of detail of the data

- how accessible the data is within the organisation
- the quality or reliability of the data based on sample size, method of collection and so on
- the coverage of the data and
- the frequency of collection

5. Present the above information in a structured and consistent way that is easily understood
6. Adapt to user needs and on-going research and development
7. Incorporate new assessment perspectives as the need arises
8. Is not restricted by business size, activity or organisational structure

7. Meeting Queensland Transport business goals

Under the Transport Portfolio, Queensland Transport strives to meet government outcomes and its own vision of responsible and efficient transport planning and management.

Government outcomes that determine QT strategic planning can be summarised as:

1. Build Queensland's economy
2. Strengthen Queensland communities and
3. Protect Queensland's environment

Every year, QT and Main Roads updates the Transport Coordination Plan (TCP), which addresses the above government outcomes and specifies the Portfolio's goals for transport planning across the state for the following 20 years.

The plan states its own purpose on page 8, to 'guide the planning and operations of the Transport Portfolio' (TCP, 2005).

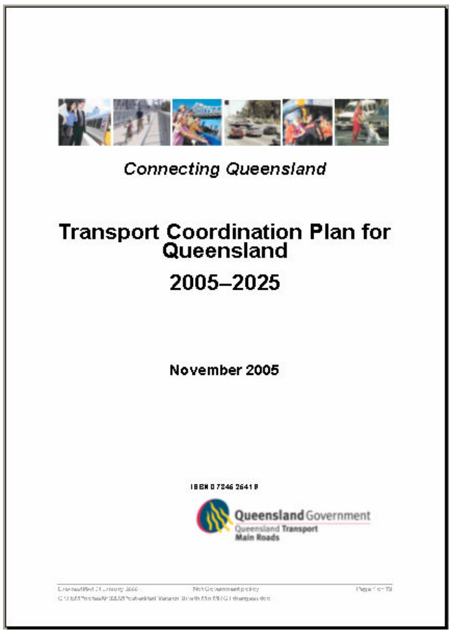


Figure 1: Transport Coordination Plan

Within this plan, the two transport bodies detail in broad terms their functions and the measures by which they will assess their performance. The TCP provides all divisions within QT and Main Roads a clear understanding of what objectives they need to meet. Indirectly, it points out what data needs will also arise.

8. Transport Systems Information goals

The Transport Systems Information (TSI) unit located within the Integrated Transport Planning division of QT is mandated to identify and satisfy strategic data needs. The unit has a diverse spread of clients—other branches within QT and Main Roads that depend on people movement data and freight data.

In carrying out its data identification activities, the TSI unit must have a clear picture of all data available to the organisation, and what condition it is in. To identify data gaps, the TSI unit must be across current user needs, and be able to assess available data to see whether it meets such needs.

Additionally, while making its data assessments, the TSI unit strives to meet best practice standards, as stated by the following:

- Ensure that collected data is good, clean and timely
- To locate or source data that supports transportation policy analysis
- Ensure that the data is easily accessible so that more people within QT can use it to benefit their clients and other stakeholders
- To carry out data identification, gathering and analysis activities in a way that reflects the Queensland Transport corporate vision and mission, namely: 'to provide better transport for Queensland and to develop, lead and manage transport in Queensland which is safe, secure, efficient, inclusive, ecologically sustainable and promotes a strong economy' (QTSP, 2005).

Anticipating the future, the unit also sees itself working towards fulfilling these goals:

- Having a small team of data analysts that will watch incoming data to discover emerging trends as they unfold
- Be able to seize relevant results and circulate them, so more information will be developed and delivered to those who can make transportation better
- That everyone who is starting a planning effort or policy study related to transportation, can come to TSI and access data and information that will assist their work

With the foregoing in mind, it is essential the TSI unit manages its data identification and collection activities in a thorough, structured manner that accommodates these principles.

The data architecture outlined in this paper, addresses these goals—both current and future.

9. Building the architecture

So how does the data architecture link the QT strategic business needs against its numerous datasets—each containing their own complex spread of data fields?

The data architecture is based on the QT business interests as embodied in the details provided by the Transport Coordination Plan.

The plan provides a blueprint for QT's transport-related goals and activities for the following 20 years. It focuses on people movement, freight movement and integrated transport planning which are QT's three foundation transport interests.

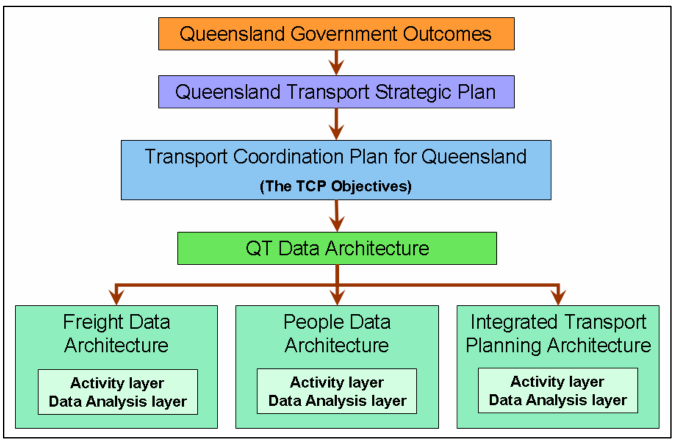


Figure 2: Overview of Data Architecture

Consequently, the data architecture reflects these interests by being split into a Freight Data Architecture, a People Data Architecture and an Integrated Transport Planning Architecture (see Figure 2). In this discussion, we are featuring examples from the Freight and People Data Architectures.

So how are the three architectures constructed? All architecture is constructed by carrying out the following steps:

1. Compiling the data attributes, which is achieved by:
2. Abstracting the attributes from the business plan (TCP)
3. Testing the attributes for relevance against the TCP
4. Scoping the attributes under user consultation (Figure 5)
5. Building the activity or modal dimension using the compiled attributes
6. Building the data analysis dimension using the activity layers with the attributes
7. Building the summary dimension

These steps are explained in more detail below.

9.1 Compiling the attributes

9.1.1 Abstracting the attributes from the business plan (TCP)

The Transport Coordination Plan provides a comprehensive outline of both Queensland Transport and Main Roads goals and objectives. The TCP is based on the QT and Main Roads strategic plans which are based on Queensland Government Outcomes. It is therefore an excellent foundation upon which the data architecture can be structured.

The TCP Objectives

1. Make the most of the existing transport system.
2. Invest in Queensland's transport system.
3. Keep the system working well.
4. Get more people walking, cycling and using public transport.
5. Support regional and remote communities.
6. Move freight efficiently.
7. Make transport safer.
8. Make transport more secure.
9. Care for our natural and built environment.
10. Integrate transport planning and land-use planning.

Figure 3: The TCP Objectives

So the initial step is to identify from the TCP what the data needs will be. This is achieved by abstracting individual attributes from the Plan's stated goals, objectives and performance measures.

Specific details regarding transport and transport-related activities are outlined within each of the 10 TCP Objectives. From these and the associated policy priorities and performance indicators, it is possible to get an understanding of what data QT requires for planning, decision making and performance assessment.

By abstracting key concerns and objectives from the TCP, it is possible to identify different sets of attributes that can be used for QT's major interests, such as the freight transport architecture, the people-movement architecture or the integrated transport planning architecture.

9.1.2 Testing the attributes against the business plan (TCP)

A close examination of the 2005-2025 TCP reveals 42 attributes that represent the key themes of the TCP that can be built into a structure for assessing the Queensland transport system (and perhaps many other large transport systems).

Of these 42 attributes, 39 are applicable to people transport and 36 for freight. The attributes cover traditional requirements such as volume of activity, activity type, speed and environmental impact. However the TCP reveals a number of other essential data attributes that are not necessarily being gathered, such as user perception, security, planned development, climate and emerging industries.

	A	B	C	D	E	F	G	H	I	J	K
1	PEOPLE ARCHITECTURE ATTRIBUTES ARISING FROM THE TCP										
2		TCP Objectives									
3		1 Make the most of the existing transport system	2 Invest in Queensland's transport system	3 Keep the system working well	4 Get more people walking, cycling and using public transport	5 Support regional and remote communities	6 Move freight efficiently	7 Make transport safer	8 Make transport more secure	9 Care for our natural and built environment	10 Integrate transport planning and land-use planning
4	ATTRIBUTES										
5	Speed/Travel Times	✓									
6	Speed variability	✓									
7	Access	✓	✓		✓	✓					✓
8	Infrastructure Projects	✓	✓	✓	✓	✓		✓			✓
9	Condition of infrastructure	✓	✓	✓	✓	✓		✓			✓
10	Infrastructure and services costs		✓	✓	✓	✓		✓			✓
11	Transport management technologies	✓		✓				✓	✓		
12	Volume of Activity	✓	✓	✓	✓	✓			✓		✓
13	Occupancy rates	✓		✓	✓	✓					
14	Incidents	✓		✓				✓	✓		
15	Accessibility/Services available	✓	✓	✓	✓	✓			✓		✓
16	User Perception of Transport System	✓	✓	✓	✓	✓					✓
17	Community cost			✓	✓	✓					
18	User cost			✓	✓	✓					
19	Intermodal Times	✓			✓						
20	User Perception Benefits PT, walk, cycle				✓						
21	Planned Government Development	✓	✓	✓		✓				✓	✓
22	Planned Land Development	✓	✓	✓		✓				✓	✓
23	Emerging Industries	✓	✓	✓		✓					✓
24	Land Valuations	✓	✓	✓		✓					✓
25	Regional Population	✓	✓	✓		✓					✓
26	Flood Regions					✓					✓
27	Fatalities							✓			
28	Injuries							✓			
29	Accidents	✓						✓			
30	High-risk behaviours							✓			
31	Vehicle Management Technology							✓			
32	User Perception Safety							✓			

Figure 4: Testing attributes for the People Architecture against the TCP

Now that we have abstracted the attributes from the TCP, how can we be certain the attributes are appropriate? One way to confirm their relevance is by crosschecking the attributes from each TCP Objective against other Objectives. To achieve this, a spreadsheet is used to match the abstracted attributes against each of the 10 TCP Objectives (see Figure 4). The result shows the attributes initially selected from the TCP Objectives have a high degree of relevance across many of the other Objectives, meaning that any datasets matching these attributes will be of interest to several or many branches within QT. This gives us a good indication that the abstracted attributes are appropriate.

The next step involves user consultation to scope, clarify, define and finalise the attributes.

9.1.3 Scoping and clarifying the attributes under user consultation

While the attributes are based on the TCP and have corporate relevance, we need to get an idea of the range of each attribute. Additionally, it needs to be asked of each attribute, 'What business drivers are causing us to seek the attribute of this data? Is it for performance measurement? Is it for planning and modelling? Or is it for evaluation and impact analysis?'

Only the users can answer these questions. So the next step in compiling the attributes is to meet with the users, and ensure the range and reason for seeking each attribute. By conducting a user consultation process, we explore the depth of each attribute, and we extend and add value to the attributes.

The user consultation process also provides an opportunity to clarify attribute meanings and add other attributes not identified in the TCP.

9.2 Building the activity dimension using the compiled attributes

Once the attributes have been settled with the users and any other stakeholders, it is possible to build a matrix for each activity or (in the case of Queensland Transport), each mode of transport, such as for people: car, bus, taxi, sea etc.

In the sections above, we mentioned that separate sets of attributes could be abstracted from the QT corporate plan for each of the QT principal interests, such as the freight architecture or people movement architectures. Any modifications to the attributes through the user consultation process should not compromise the distinctiveness of the attribute sets for each interest as these will provide the basis for the analyses of the interests.

So each interest has its own set of data attributes. In other words, the freight architecture has its own set of data attributes; the people movement architecture has its own set of attributes and so on.

Taking one interest at a time, the attributes unique to the interest, are imposed upon each mode of transport or activity.

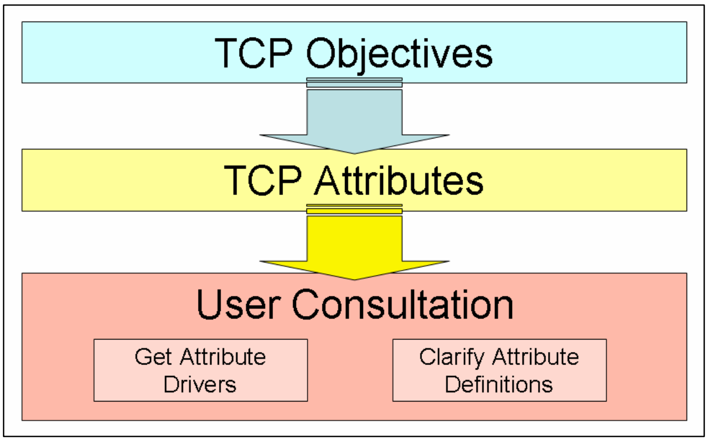


Figure 5: User consultation process

At QT, the architecture has been implemented using a spreadsheet platform. While the architecture is an abstract concept, it could however be managed using a database with an online front end.

Each mode or activity is now presented on a separate worksheet with the same attributes repeated on each worksheet (see Figure 6).

So, if we use the freight architecture as an example, we now have six worksheets—one for each mode: light commercial, truck, rail, sea, air and cycle. On each sheet, the X axis lists the compiled attributes, and the Y axis lists the datasets. Each dataset can then be compared with each attribute for each mode of transport.

The activity sheets represent the first dimension in the data architecture. Each worksheet within this dimension represents a mode of transport, and each provides a template for the various assessments in the next dimension—the data analysis dimension.

FREIGHT TRANSPORT - ASSESSMENT OF DATA SOURCES AGAINST COMPILED ATTRIBUTES

Architecture for Freight Mode:
Car/Light commercial

Strategic Information/Source	Source	Speed/ Travel Time	Speed Variability	Access	Infra-structure Projects	Condition of infra-structure	Infra-structure and services costs	Transport mngmnt tech-nologies	Volume of Activity	Incidents	Accessibility /Services available	User Perception of Transport System	Planned Govt Dvlpmnt	Planned Land Dvlpmnt	Emerging Industries	Land Valuations	Reg Popu
EXISTING DATA SOURCES																	
INFRASTRUCTURE																	
ARMIS	DMR																
Rail and Sea Infrastructure	QT RPF, MSQ																
TRAILS (Qld Registration & Licencing Information)	QT																
Commercial Vessel Licensing & Rego	QT																
ECONOMY																	
Economic Statistics	ABS, OESR																
Industry Statistics	ABS, OESR																
ACTIVITY																	
Brisbane Strategic Transport Model (BSTM)	QT																
SEQ Freight Survey ('95) & Metro Freight Study ('93)	QT																
Input / Output Freight Generation - SEQ	QT																
Brisbane Truck Origin - Destination Matrix	QT																
Gateway Upgrade Project (GUP)	DMR																

Figure 6: Freight: Car/Light commercial spreadsheet showing attributes and datasets

Currently there are eight modes for the people architecture: bus, car, taxi, rail, sea, air, bicycle and walking. There are six modes for the freight architecture: light commercial, truck, rail, sea, air and bicycle.

The next step now involves creation of the data analysis dimension.

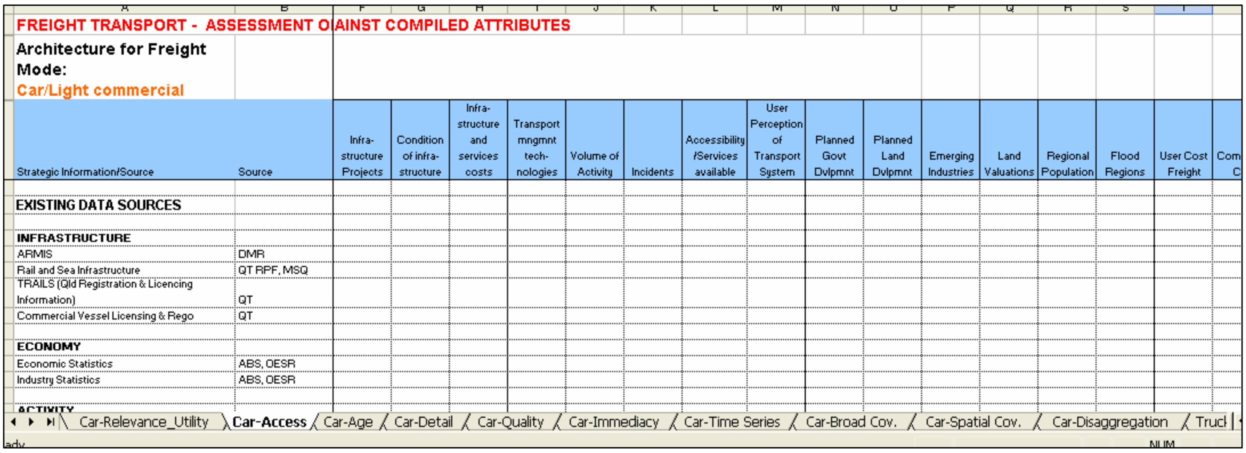
9.3 Building the data analysis dimension using the activity dimension templates

The quality of each dataset is analysed by adding assessments for each mode or activity. In the above section we have set up templates for each mode which will now be expanded for the data analysis dimension.

While the activity sheets above could be used to provide a 'first-cut' assessment of the datasets against each attribute, there are still many unanswered questions about the data. According to the United States Bureaus of Transportation Statistics, there are six goal areas that reflect the key attributes of data and analysis that must be addressed if gathered data is to meet desired outcomes. These are: relevance, quality, timeliness, comparability, completeness, and utility (BTS, 2005).

At QT, we have built similar assessments into the data architecture, and added a few others. For each of the above activities or modes of transport (discussed in section 9.2), separate worksheets are allocated for assessment of each attribute in relation to each mode, such as accessibility (access), relevance (to the user), currency (age) and detail (level of detail) etc. (See Figure 7).

As illustrated in Figure 7, one sheet is used for each assessment of the Car/Light commercial freight mode. Each mode has its own set of assessments.



FREIGHT TRANSPORT - ASSESSMENT O AINST COMPILED ATTRIBUTES

Architecture for Freight Mode:
Car/Light commercial

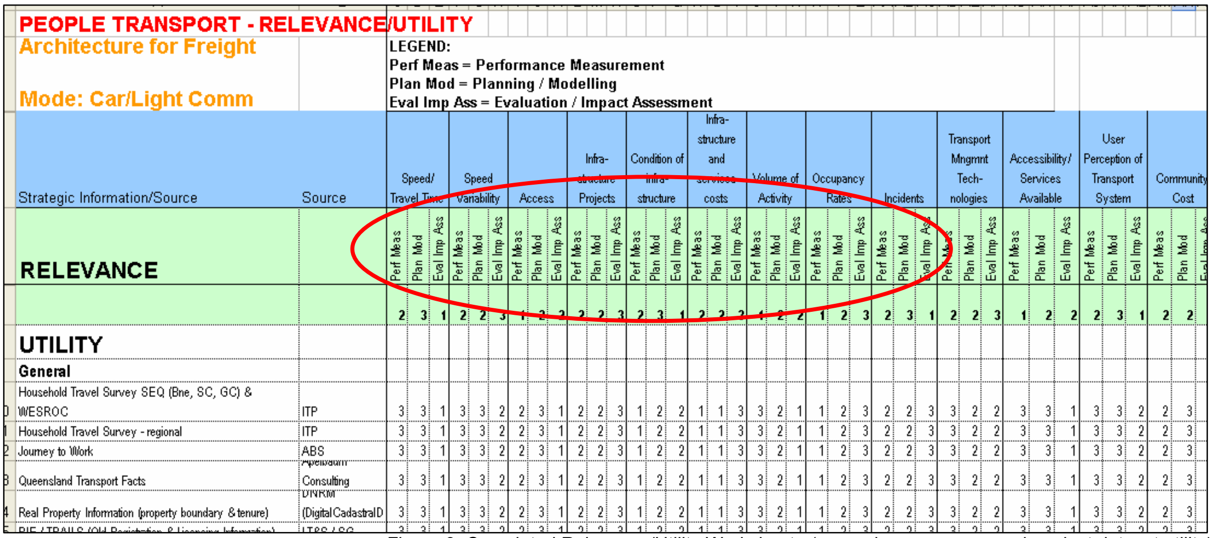
Strategic Information/Source	Source	Infra-structure Projects	Condition of infra-structure	Infra-structure and services costs	Transport mngmnt tech-nologies	Volume of Activity	Incidents	Accessibility /Services available	User Perception of Transport System	Planned Govt Dvlpmnt	Planned Land Dvlpmnt	Emerging Industries	Land Valuations	Regional Population	Flood Regions	User Cost Freight	Com C
EXISTING DATA SOURCES																	
INFRASTRUCTURE																	
ARMIS	DMR																
Rail and Sea Infrastructure	QT RPF, MSQ																
TRAILS (Qld Registration & Licencing Information)	QT																
Commercial Vessel Licensing & Rego	QT																
ECONOMY																	
Economic Statistics	ABS, OESR																
Industry Statistics	ABS, OESR																
ACTIVITY																	

Car-Relevance_Utility / Car-Access / Car-Age / Car-Detail / Car-Quality / Car-Immediacy / Car-Time Series / Car-Broad Cov. / Car-Spatial Cov. / Car-Disaggregation / Truck

Figure 7: Example of the Car/Light commercial mode under the Freight Architecture: the tabs at the bottom of the picture indicate each assessment worksheet

The QT Data Analysis dimension is comprised of the following layers:

- The Access worksheet: provides a rating on how accessible the attribute is within the dataset. For example, is this information readily available or does the owner require authorisation/payment or some other inducement to release the information?
- The Age worksheet: rates the attribute's currency. Is this old data that is no longer useful or is it current?
- The Detail worksheet: assesses the level of detail associated with the attribute. Is the detail poor, mediocre or good?
- The Quality worksheet: assesses the quality of the data based on sample size etc.
- The Immediacy worksheet: assesses whether the data is useful for immediate use
- The Time series worksheet: assesses whether the data has been collected long enough and frequently enough to be of use
- The Broad coverage worksheet: assesses how well the data meets this aspect of coverage
- The Spatial coverage worksheet: assesses the degree of this aspect of coverage
- The Disaggregation worksheet: assesses the level of disaggregation, and
- The combined Relevance/Utility worksheet: features a Relevance row which is completed under user consultation. It allows users to specify the degree of relevance of each attribute against three uses: Performance Measurement, Planning/Modelling and Evaluation/Impact Analysis. User response also provides an indication of which collection cycles the data attribute falls under (see Section 10).
- The remainder of the sheet (the Utility section) is used to rate the datasets against these uses. This means that by comparing scores entered in the Relevance row against scores entered in the Utility section, it is possible to see how the data attributes and datasets are performing against user requirements (see Figure 8).



PEOPLE TRANSPORT - RELEVANCE/UTILITY																				
Architecture for Freight Mode: Car/Light Comm		LEGEND: Perf Meas = Performance Measurement Plan Mod = Planning / Modelling Eval Imp Ass = Evaluation / Impact Assessment																		
Strategic Information/Source	Source	Speed/ Travel Time			Speed Variability			Access			Infra-structure Projects			Condition of Infra-structure			Infra-structure and services costs			
RELEVANCE		Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass	
		2	3	1	2	2	3	1	2	2	2	2	3	2	3	1	2	2	2	
UTILITY																				
General																				
Household Travel Survey SEQ (Bne, SC, GC) & WESROC	ITP	3	3	1	3	3	2	2	3	1	2	2	3	1	2	2	1	1	3	
Household Travel Survey - regional	ITP	3	3	1	3	3	2	2	3	1	2	2	3	1	2	2	1	1	3	
Journey to Work	ABS	3	3	1	3	3	2	2	3	1	2	2	3	1	2	2	1	1	3	
Queensland Transport Facts	Apelbaum Consulting	3	3	1	3	3	2	2	3	1	2	2	3	1	2	2	1	1	3	
Real Property Information (property boundary & tenure)	DNRM (Digital Cadastral D	3	3	1	3	3	2	2	3	1	2	2	3	1	2	2	1	1	3	
RIE / TRAILS (Qld Registration & Licensing Information)	LTSS / SG	3	3	1	3	3	2	2	3	1	2	2	3	1	2	2	1	1	3	

Strategic Information/Source	Volume of Activity			Occupancy Rates			Incidents			Transport Mngmnt Tech-nologies			Accessibility/ Services Available			User Perception of Transport System			Community Cost		
RELEVANCE	Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass	Perf Meas	Plan Mod	Eval Imp Ass
	1	2	2	1	2	3	2	3	1	2	2	3	1	2	2	2	3	1	2	2	
Household Travel Survey SEQ (Bne, SC, GC) & WESROC	3	2	1	1	2	3	2	2	3	3	2	2	3	3	1	3	3	2	2	3	
Household Travel Survey - regional	3	2	1	1	2	3	2	2	3	3	2	2	3	3	1	3	3	2	2	3	
Journey to Work	3	2	1	1	2	3	2	2	3	3	2	2	3	3	1	3	3	2	2	3	
Queensland Transport Facts	3	2	1	1	2	3	2	2	3	3	2	2	3	3	1	3	3	2	2	3	
Real Property Information (property boundary & tenure)	3	2	1	1	2	3	2	2	3	3	2	2	3	3	1	3	3	2	2	3	
RIE / TRAILS (Qld Registration & Licensing Information)	3	2	1	1	2	3	2	2	3	3	2	2	3	3	1	3	3	2	2	3	

Figure 8: Completed Relevance/Utility Worksheet: (user relevance compared against dataset utility)

10. Carrying out the data assessment

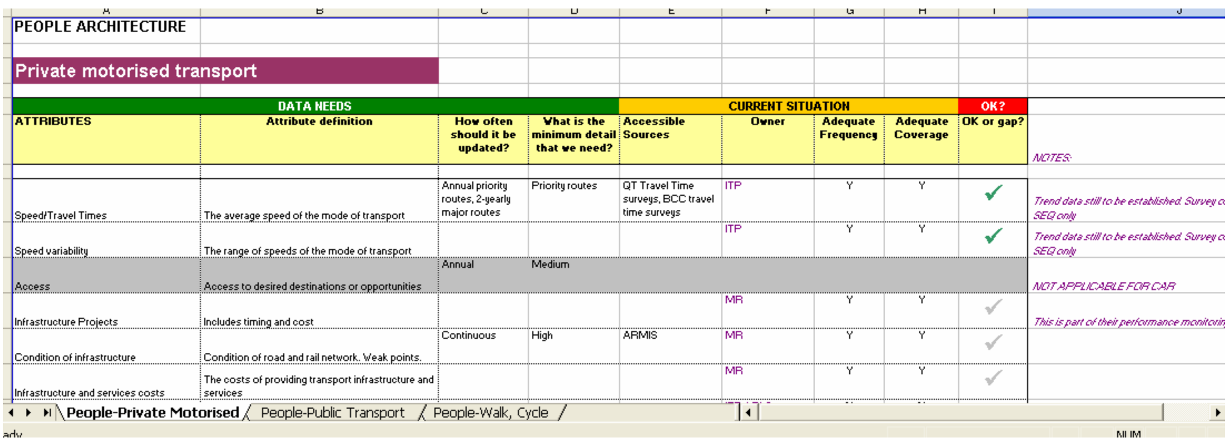
At this point we have the completed worksheets which can now be used to assess the datasets. Commencing with the Relevance/Utility worksheet, a number is scored in each intersecting cell. If the dataset matches the attribute for the particular mode under the selected assessment, then a high score is entered in the intersecting cell. If the attribute offers a poor match, a low score is entered.

As mentioned above, the Relevance/Utility worksheet is partially completed under user advice. So when complete, it provides a good indication of user data needs. (See Figure 8.) By comparing the scores in the Relevance section with the scores in the Utility section, any significant deficiencies (that is, real data gaps) are revealed. Based on the conclusions drawn from this comparison, a thorough assessment of the significant deficiencies can be made by completing and viewing the other sheets in the data analysis dimension.

When all sheets have been scored, it is evident where data may be duplicated, missing, or of poor quality. It is then easy to see where potential gaps may exist.

11. The summary version

Once the foregoing has been completed, the information revealed by the architecture can be condensed into a 'lightweight' version for summary purposes. A sample is provided:



PEOPLE ARCHITECTURE									
Private motorised transport									
DATA NEEDS				CURRENT SITUATION				OK?	
ATTRIBUTES	Attribute definition	How often should it be updated?	What is the minimum detail that we need?	Accessible Sources	Owner	Adequate Frequency	Adequate Coverage	OK or gap?	NOTES:
Speed/Travel Times	The average speed of the mode of transport	Annual priority routes, 2-yearly major routes	Priority routes	QT Travel Time surveys, BCC travel time surveys	ITP	Y	Y	✓	Trend data still to be established. Survey c SEQ only
Speed variability	The range of speeds of the mode of transport				ITP	Y	Y	✓	Trend data still to be established. Survey c SEQ only
Access	Access to desired destinations or opportunities	Annual	Medium						NOT APPLICABLE FOR CAR
Infrastructure Projects	Includes timing and cost				MR	Y	Y	✓	This is part of their performance monitorin
Condition of infrastructure	Condition of road and rail network. Weak points.	Continuous	High	ARMIS	MR	Y	Y	✓	
Infrastructure and services costs	The costs of providing transport infrastructure and services				MR	Y	Y	✓	

People-Private Motorised / People-Public Transport / People-Walk, Cycle

12. Meeting data collection cycles

An important question facing any data collection program is 'How often data should be collected?'

The data architecture helps to answer this question by firstly addressing an organisation's corporate goals then by separating its planning data from its performance measurement data. As illustrated, each subset of data must be gathered according to its own cycle.

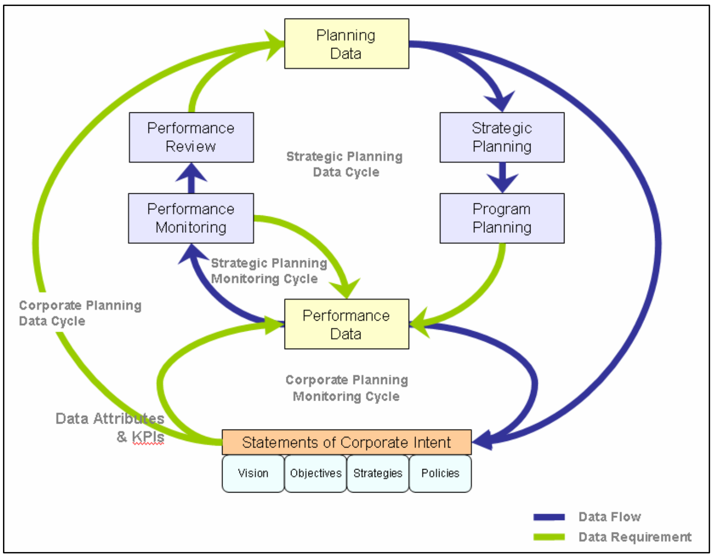


Figure 9: Data collection cycles

The length of the cycle is determined in consultation with users; however we know that there are four cycles, each with varying lengths:

Corporate planning data cycle: likely to be the longest cycle as corporate planning is usually timed over many years—for example the Transport Coordination Plan covers 20 years so data could be collected around once every five years

Corporate planning monitoring cycle: contained within the corporate planning data cycle. This cycle can run many times within one cycle of the corporate planning data cycle for monitoring corporate planning so an ideal collection rate may be once annually or once every two years.

Strategic planning data cycle: likely to run more frequently than the corporate planning data cycle as strategic planning cycles are usually of shorter duration than corporate planning cycles so this data could be collected around every third year.

Strategic planning monitoring cycle: contained within the strategic planning data cycle. This cycle can run many times within one cycle of the strategic planning data cycle for monitoring strategic planning, so this data could be collected once annually.

Undoubtedly, there will be exceptions to the above, so user consultation will identify any requirements that would justify such exceptions.

13. Linking users to the data architecture

The architecture is a disinterested model—designed to meet the data needs of any user within an organisation irrespective of organisational structure. It simultaneously addresses user needs and corporate planning requirements independently of department, branch or unit assignment. Its independent application is for two reasons; one being that organisational structure changes from time to time, and any data architecture model has to sit over the top of such organisational changes.

A second reason is that it allows the architecture to be transportable from one organisation to another. This fact alone supports the robustness of the architecture.

However, given a particular organisational structure, the architecture provides a convenient framework for mapping data activities against organisational units as illustrated below. (Figure 11).

PEOPLE ARCHITECTURE LINKS: ATTRIBUTES TO USERS	Maritime Safety Queensland	Rail, Ports and Freight	Land, Transport and Safety	Corporate Governance	Services Group	Integrated Transport Planning	Passenger Transport	Infrastructure Program	TransLink	Main Roads	Intelligent Transport System Development	Road System and Engineering	
ATTRIBUTES													
Speed/Travel Times	✓												
Speed variability	✓												
Access	✓	✓		✓	✓					✓			✓
Infrastructure Projects	✓	✓	✓	✓	✓		✓			✓		✓	✓
Condition of infrastructure	✓	✓	✓	✓	✓		✓			✓		✓	✓
Infrastructure and services costs		✓	✓	✓	✓		✓			✓		✓	✓
Transport management technologies	✓		✓				✓	✓			✓	✓	
Volume of Activity	✓	✓	✓	✓	✓			✓		✓	✓		✓
Occupancy rates	✓		✓	✓	✓								✓
Incidents	✓		✓				✓	✓			✓	✓	
Accessibility/Services available	✓	✓	✓	✓	✓			✓		✓	✓		✓
User Perception of Transport System	✓	✓	✓	✓	✓					✓			✓
Community cost			✓	✓	✓								✓
User cost			✓	✓	✓								✓
Intermodal Times	✓			✓									✓
User Perception Benefits PT, walk, cycle				✓									✓
Planned Government Development	✓	✓	✓		✓				✓	✓			
Planned Land Development	✓	✓	✓		✓				✓	✓			
Emerging Industries	✓	✓	✓		✓					✓			
Land Valuations	✓	✓	✓		✓					✓			
Regional Population	✓	✓	✓		✓					✓			
Flood Regions					✓					✓			

Figure 11: Spreadsheet linking attributes to users

14. The Benefits of the data architecture

By integrating the data architecture into QT's data gathering program, greater efficiencies are achieved:

- Unnecessary data duplication is avoided and real data gaps are readily exposed.
- When the data analysis cycle under the data architecture runs in a timely manner, users can expect data that is more up-to-date and relevant to business needs.
- Because the data architecture also identifies the level of dataset access, pro-active strategies can be put in place that will remove access barriers before data is requested.
- There is less opportunity for datasets to be 'lost' or under-utilized.
- Changing data requirements are readily catered for.

15. The Future

The QT data architecture will evolve over time adapting to its changing environment. Minor on-going modifications will be required as a result of:

- Changing government policies causing changes in QT's own policies and data needs
- Consultation with users and other stakeholders which will help to refine the data attributes (see Figure 8).
- Research and development which will require the addition of new attributes into the data architecture

The architecture has the potential to increasingly drive data gathering so that datasets should conform more and more to the architecture. Eventually, the architecture could be used to specify groups of datasets required by each branch to perform its function.

16. Conclusion

The Data Architecture has been designed to be flexible enough to scrutinize every imaginable dataset characteristic. It has the scope to provide, in many cases more information than is actually needed, however this is the beauty of the architecture in that it can deliver the goods when they're needed.

The separate worksheets for each mode ensure that all datasets are subjected to a range of assessments that will fulfil all QT strategic planning and performance measurement needs.

The use of spreadsheets, rather than a more formalised structure such as a database, allows easy access to the structure. As the architecture evolves then changes need to be applied to the structure. The usefulness of the architecture is a key factor, and therefore the form that it should take is still evolving. The flexibility that spreadsheeting gives us is highly rated.

In the case of the People Architecture, there are eight modes, with 39 attributes, subjected to 9 single assessments and one six-way assessment—that amounts to (8 x 39 x (9 + 6)) 4,680 possible different ways of looking at each dataset related to people transport.

In conclusion, the data architecture outlined in this paper provides a thorough, structured methodology that is capable of assessing existing data and guiding decision-making and planning for on-going data collection. It has the capacity to embrace whatever future data needs will arise irrespective of organisational changes.

17. Appendix

Attributes based on TCP Objectives (Obj)

Obj	Attribute	Freight	People
1	**Speed/Travel Times**	The average speed of the mode of freight transport	The average speed of the mode of people transport
1	**Speed variability**	The range of speeds of the mode of freight transport	The range of speeds of the mode of people transport
1	**Access**	The user's ability to send or receive freight to or from desired destinations or opportunities	The user's ability to travel to or from desired destinations or opportunities
2	**Infrastructure Projects**	Details of past, present and planned infrastructure projects related to freight transport—includes timing and cost	Details of past, present and planned infrastructure projects related to people transport—includes timing and cost
2	**Condition of infrastructure**	The condition of road, rail and air infrastructure as related to freight transport—includes any weak points.	The condition of road, rail and air infrastructure as related to people transport—includes any weak points.
2	**Infrastructure and services costs**	The costs of providing freight transport infrastructure and services	The costs of providing people transport infrastructure and services
3	**Transport management technologies**	The details of new and existing freight transport management technologies	The details of new and existing people transport management technologies
3	**Volume of Activity**	The number of occurrences of any activity related to freight transport in any region. This would mostly to refer to each mode	The number of occurrences of any activity related to people transport in any region. This would mostly to refer to each mode.
3	**Occupancy rates**	Not applicable	The rates of occupancy of each mode of transport
3	**Incidents**	Incidents that disrupt the transport system and that would affect freight transport. Including but not limited to accidents.	Incidents that disrupt the transport system and that would affect people transport. Including but not limited to accidents
3	**Accessibility/ Services available**	User ability to access and use the freight transport system—especially relevant to remote areas and communities. Also includes services for people with a disability and shift workers.	User ability to access and use the people transport system—especially relevant to remote areas and communities; also includes services for people with a disability and shift workers

3	**User Perception of Transport System**	The user view of how well all modes of the transport system appear to be running	The user view of how well all modes of the transport system appear to be running
4	**Community cost**	Not applicable	Costs to the community for people transport infrastructure and services
4	**User cost**	Not applicable	The cost to the user across all modes of transport
4	**Inter-modal Times**	Not applicable	Waiting times when linking from one transport mode to the next
4	**User Perception Benefits**	Not applicable	User view of the benefits of travel by public transport, cycling or walking
5/10	**Planned Government Development**	Local, state and federal government infrastructure planning that relates to freight transport	Local, state and federal government infrastructure planning that relates to people transport
5/10	**Planned Land Development**	Local, state or federal government land planning that relates to freight transport	Local, state or federal government land planning that relates to people transport
5/10	**Emerging Industries**	Planned industrial development and emerging industries that would place additional demands on the freight transport system	Planned industrial development and emerging industries that would place additional demands on the people transport system
5/10	**Land Valuations**	Land valuations for freight transport infrastructure planning	Land valuations for people transport infrastructure planning
5/10	**Regional Population**	Regional population figures for freight transport planning	Regional population figures for people transport planning
5/10	**Flood Regions**	Flood data for all-weather freight transport infrastructure planning and development	Flood data for all-weather people transport infrastructure planning and development
6	**User Cost Freight**	The cost to the user for freight across all modes of transport	Not applicable
6	**Community cost**	Costs to the community for freight transport infrastructure and services	Not applicable
6	**User Perception Freight**	User view of costs, efficiency, incidents and other aspects of freight transport across all modes	Not applicable
6	**Inter-modal Times**	Freight waiting times when linking from one mode to the next	Not applicable
7	**Accidents**	All data concerning accidents across all modes of transport that would affect freight transport	All data concerning accidents across all modes of transport that would affect people transport
7	**Fatalities**	Not applicable	Data relating to fatalities in the transport system

7	**Injuries**	Not applicable	Data relating to injuries in the transport system
7	**High-risk behaviours**	Drug, alcohol, speed, fatigue data affecting the freight transport industry	Drug, alcohol, speed, fatigue data affecting the people transport industry
7	**Vehicle Safety Technology**	All data concerning implementations of in-vehicle technology for freight transport	All data concerning implementations of in-vehicle technology for people transport
7	**User Perception Safety**	User view of safety issues affecting freight transport across all modes of transport	User view of safety issues affecting people transport across all modes of transport
8	**Security Risks**	Security risks across all modes of transport that would affect freight mode choices and the transport of freight	Security risks across all modes of transport that would affect mode choices
8	**Dangerous Goods Incidents**	Dangerous goods incidents that would affect freight-mode choices and the delivery of freight	Dangerous goods incidents that would affect mode choices
8	**Security solutions - other organisations**	Security solutions developed or in development by other government or private organisations that would benefit the freight transport system	Security solutions developed or in development by other government or private organisations that would benefit the people transport system
8	**User Perception Security**	User view of how secure each mode of transport and how it may affect freight transport mode choices	User view of how secure each mode of transport and how it may affect transport mode choices
9	**Transport Emissions**	Transport emissions across all modes of freight transport	Transport emissions across all modes of people transport
9	**Energy Consumption**	Energy consumption across all modes of freight transport	Energy consumption across all modes of people transport
9	**Noise**	Noise created by all modes of freight transport	Noise created by all modes of people transport
9	**Environment-ally-Friendly Vehicles**	Data on environmentally-friendly freight transport vehicles	All data on environmentally-friendly people transport vehicles
9	**Heritage Sites**	Data on heritage sites and environmentally fragile regions that would affect infrastructure planning for freight transport	All data on heritage sites and environmentally fragile regions that would affect infrastructure planning for people transport
9	**Waterway Pollution**	Data on waterway pollution caused by freight transport	All data on waterway pollution caused by people transport
9	**User Perception Environment**	User view of the environmental aspects of freight transport—includes noise, pollution complaints	User view of the environmental aspects of people transport—includes noise, pollution complaints

Note: although many of the attributes appear common to both the freight and people architectures, subtle differences from each perspective demands that both architectures have their own set of attributes.

18. References

BTS, *A Strategic Plan for Transportation Statistics (2000-2005),* US Bureau of Transportation Statistics, Washington, March 2000.

PAS, *Framework for Decision Making & Strategic Information Management*, Policy Appraisal Services Pty Ltd, Brisbane, February 2003.

QTSP, *Queensland Transport Strategic Plan 2005-09*, Queensland Transport, Brisbane, 2005

TCP, *Transport Coordination Plan for Queensland 2005—2025*, Queensland Transport, Brisbane, November 2005.