

# Buffalo City Municipality Traffic Safety Plan

## PROBLEMS RELATING TO THE CLEANSING OF ACCIDENT DATA / STRATEGY FOR ACCIDENT DATA CLEANSING

### Background

The purpose of this document is to gather a list of problems relating to the task of cleansing the accident data stored in the Accibase system used by the BCM Traffic Department. These problems/requirements were identified through discussions with the staff from the BCM Transportation Planning Division and EL and KWT Traffic Departments and analysis of the electronic data.

The strategy for accident data cleansing is presented as proposed solutions to the various problems identified. A detailed process to resolve the problem is presented along with expected completion dates for each of the solutions.

### TS1.11 : PROBLEM IDENTIFICATION

To cleanse the accident data in the Accibase system, problems with the data must be identified and documented. These findings will then determine the required tasks to ensure the availability of a reasonably reliable accident data set.

### PROBLEMS / REQUIREMENTS

NO	DESCRIPTION	DETAILS	SOLUTION	STATUS
PROB1	EL Traffic department accident report backlog	As of 2005-01-14 the backlog of uncaptured calls at EL Traffic department was ±1600	<ul style="list-style-type: none"> <li>• Set up 3 PCs at Transportation Planning Division offices and make 3 capturers available to deal with backlog</li> </ul> <p><b>Estimated completion dates:</b>  <b>[2005-01-19]</b> 2003, 2004 data  <b>[2005-01-26]</b> 2001, 2002, 2005 data</p>	<p><b>[2005-01-14]</b> Capturing is proceeding smoothly</p> <p><b>[2005-01-20]</b> Caught up 2003, 2004 backlog</p> <p><b>[2005-01-26]</b> All backlog captured</p>

**PROBLEMS / REQUIREMENTS**

NO	DESCRIPTION	DETAILS	SOLUTION	STATUS
PROB2	Poor location coding on Accident Report (AR) forms received from SAPS in EL	Many of the accident report forms received from the SAPS in the EL area have little or no location information filled-in, making it impossible to assign accidents to the correct locations in Accibase.	<ul style="list-style-type: none"> <li>• Hire casual staff to correctly code locations on AR forms which have little or no location information available.</li> <li>• This can be achieved through a number of means such as:                             <ul style="list-style-type: none"> <li>○ Contacting the people involved in the accident telephonically</li> <li>○ Checking maps to get additional information with regard to the accident</li> <li>○ Personal experience of the EL area</li> </ul> </li> </ul> <p><b>Estimated completion dates:</b>  <b>[2005-01-10]</b> Locations coded on AR reports</p>	<b>[2005-01-14]</b> Completed

**PROBLEMS / REQUIREMENTS**

NO	DESCRIPTION	DETAILS	SOLUTION	STATUS
PROB3	Inaccurate GIS nodes and centerlines for EL locations	The current GIS spatial data (nodes and intersections) comes from a previous inaccurate set. This spatial data is currently being corrected. Once these corrections are complete, Accibase needs to be updated with the new spatial data.	<ul style="list-style-type: none"> <li>• Identify locations associated with accidents logged in Accibase</li> <li>• Run a query to match-up the locations in Accibase to the locations in the corrected GIS data-set using the "Location Name" field as a reference</li> <li>• Replace the old Accibase location data with the corrected GIS data-set locations on all accidents where the fields match</li> <li>• Extract the Accibase locations which do not match-up to any locations in the new GIS data-set</li> <li>• Generate a map of these non-matching locations overlaid on the new GIS data-set locations to allow for visual matching of these locations</li> <li>• Prepare a mapping of the previously unmatched Accibase locations to new GIS data-set locations in electronic format (Excel or CSV)</li> <li>• Use the above mapped data to update accident locations on Accibase</li> <li>• Replace the old, inaccurate GIS dataset with the new, corrected data-set on the SDE database</li> </ul> <p><b>Estimated completion dates</b>  <b>[2005-01-26]</b> GIS spatial data corrected  <b>[2005-01-31]</b> Accibase update</p>	<p><b>[2005-01-14]</b> Awaiting corrections</p> <p><b>[2005-02-04]</b> Spatial data corrected. Accibase still to be updated.</p> <p><b>[2005-08-15]</b> Accibase Updated</p>

**PROBLEMS / REQUIREMENTS**

NO	DESCRIPTION	DETAILS	SOLUTION	STATUS
PROB4	Incorrect locating of accidents in EL	<p>Because there are no street names in Mdantsane, many accidents and their matching locations are incorrectly allocated.</p> <p>Other locations such as the Vincent Park Shopping Center parking lot appear to have unreasonably high accidents associated with them.</p> <p>In a meeting held with the BCM Transportation Planning Division in December, it was decided that because of the difficulties of identifying street names, as a temporary measure, accidents in the Mdantsane area would be coded by zone rather than by node/centerline</p>	<ul style="list-style-type: none"> <li>• Add zone information to SDE database</li> <li>• Update Accibase to allow zones to be selected. Zones will be added to the intersections table and all labels previously referring to "intersections" will now refer to "Zones/Intersections". This allows the accommodation of zones in Accibase without any major overhauling of the way locations are coded in Accibase.</li> <li>• Perform a GIS shape query to map existing nodes and centerlines to zones in a table.</li> <li>• Use the table above to convert Accibase nodes and intersections to zones for the Mdantsane area.</li> </ul> <p><b>Estimated completion dates</b>  <b>[2005-01-31]</b> Add zones to GIS  <b>[2005-02-04]</b> Convert Accibase data</p>	<p><b>[2005-01-14]</b> Not started</p> <p><b>[2005-02-14]</b> Zones added. Data not yet converted.</p> <p><b>[2005-08-15]</b> It has been decided to leave data as is because mechanisms exist within the Accident Reporting to summarise data by ward and by area. The accidents in these locations will thus be correctly allocated in these reports.</p>

**PROBLEMS / REQUIREMENTS**

NO	DESCRIPTION	DETAILS	SOLUTION	STATUS
PROB5	No GIS centerline and node data is available for Kidd's Beach	<p>Currently, accidents which have taken place in the Kidd's Beach area have not been captured on the Accibase system due to the unavailability of centerline and node data on the SDE system.</p> <p>As of 2005-01-20 the backlog of uncaptured calls at Kidd's Beach was ±300.</p>	<ul style="list-style-type: none"> <li>The GIS centerline/node data for Kidd's beach is included in the GIS centerline/node conversion project discussed in PROB3</li> <li>Relocate accidents captured for Kidd's Beach to correct nodes/intersections</li> </ul> <p><b>Estimated completion dates</b>  <b>[2005-01-31]</b> Update Accibase locations  <b>[2005-02-04]</b> Kidd's Beach data capture  <b>[2005-03-31]</b> Kidd's Beach data correctly located</p>	<p><b>[2005-01-14]</b> Awaiting corrections</p> <p><b>[2005-01-24]</b> Kidd's Beach GIS information is NOT available on conversion data set – this needs to be imported</p> <p><b>[2005-01-26]</b> Kidd's beach ARs being captured to single location – to be corrected when locations are available</p> <p><b>[2005-01-31]</b> Kidd's Beach ARs capture completed, still to be converted. Kidd's Beach GIS data added to dataset</p> <p><b>[2005-02-14]</b> All new elements added. Accidents in these areas still need to be relocated to the new centerlines/nodes</p> <p><b>[2005-08-15]</b> Completed</p>

**PROBLEMS / REQUIREMENTS**

NO	DESCRIPTION	DETAILS	SOLUTION	STATUS
PROB6	There is no historical data available on Accibase for the KWT area for dates prior to 2004-05-20	The KWT traffic department has only had the Accibase system installed since 20 May 2005. Previous accident history is stored in the TCS system.	<ul style="list-style-type: none"> <li>• Export historical accident data out of the TCS system</li> <li>• Import TCS data into Accibase system</li> <li>• Convert TCS descriptive location codes to GIS compatible locations (either nodes or centerlines) by editing imported accident reports manually in Accibase</li> </ul> <p><b>Estimated completion dates</b>  <b>[2005-01-10]</b> Import TCS data into Accibase  <b>[2005-01-28]</b> Convert location codes</p>	<p><b>[2005-01-10]</b> Data import completed</p> <p><b>[2005-01-10]</b> Manual conversion 70% complete</p> <p><b>[2005-02-14]</b> Completed</p>

**PROBLEMS / REQUIREMENTS**

NO	DESCRIPTION	DETAILS	SOLUTION	STATUS
PROB7	<p>A backlog of duplicate entries for the same accident exist within Accibase</p>	<p>The same accident may be reported (and captured) more than once. This is not uncommon in accidents where more than one person is involved.</p> <p>In the standard operation of Accibase, the system identifies potential duplicates (by fields such as vehicle registration numbers, accident date and location and details of people involved) and presents this data to the capturer. The capturer then, on a regular basis, analyses the potential duplicates and makes corrections where applicable.</p> <p>At BCM, this process has not been rigorously followed since the inception of the system at the beginning of 2003.</p>	<ul style="list-style-type: none"> <li>• Select potential duplicate data based on multiple accident reports linked to the same vehicle registration number</li> <li>• In cases where no vehicle registration details and no driver details are available at all and where pedestrian data is available, the accidents are assumed to be hit-and-run and cannot be analyzed for duplicates. It is anticipated that this set of data will be too small to impact significantly on accuracy.</li> <li>• Generate a printout of all other potential duplicates</li> <li>• Manually scan the printout to identify potential accidents based on the vehicle registration numbers, accident date and location and details of people involved.</li> </ul> <p><b>Estimated completion dates</b>  <b>[2005-01-25]</b> Generate and extract reports  <b>[2005-02-04]</b> Identify and correct duplicates</p>	<p><b>[2005-01-26]</b> Report generated (96 pages). Assumed hit-and-run dataset is 1496 records of 44 215 (3.3%)</p> <p><b>[2005-01-31]</b> 10% of duplicates corrected – estimated 2 days to complete task</p> <p><b>[2005-02-14]</b> Completed</p>

**PROBLEMS / REQUIREMENTS**

NO	DESCRIPTION	DETAILS	SOLUTION	STATUS
PROB8	<p>Injury totals have not been completed on the new AR form resulting in skewed accident statistics for 2004</p>	<p>The fields for the number of people killed, with serious injuries and with minor injuries have not been completed on the new AR forms which were introduced at the beginning of 2004.</p> <p>These fields were regularly filled in on the old OAR forms because they appeared at the beginning of the form. However these fields were moved to near the end of the new AR form and consequently were very poorly completed.</p> <p>This problem is especially prevalent on forms captured by the SAPS</p>	<ul style="list-style-type: none"> <li>Assign and train an officer at each of the police-stations at BCM. This officer will be held responsible for the sign-off on the quality of the AR forms</li> <li>Use the data available in DriversTBL and CasualtiesTBL in Accibase to reconstruct the total injuries for each form.</li> </ul> <p><b>Estimated completion dates</b>  <b>[2005-03-31]</b> Training  <b>[2005-03-31]</b> Reconstruct totals</p>	<p><b>[2005-02-14]</b> Totals reconstructed  <b>[2005-08-15]</b> Training Outstanding</p>