THE BENEFITS OF IMPLEMENTING A ‘ONE SOURCE OF TRUTH’ ASSET MANAGEMENT DATASET

Martyn Simpson, Wellington City Council
Louise Lindsay, Jeff Booth Consulting Ltd

ABSTRACT

Data integrity is extremely important in all facets of life. If data is incomplete, inaccurate or unreliable, its value becomes negligible. Having timely and accurate infrastructure data is crucial to the success of any network planning or development project.

Many local authorities store asset attribute data in a database and spatial data in a GIS system with a link between the two. They also have many other datasets such as customer complaints, survey and repair information, all in other formats. Having one central master database holding both attribute and spatial data enables efficient management of network information.

This paper discusses the implementation project of a pipe information management system that was undertaken for Wellington City Council (WCC) “three waters” reticulation. It looks at the main motivators for undertaking such a project, identifying the main data issues central to the project and how these could be best resolved.

A number of sources of data inaccuracies that were previously unknown were discovered and able to be quantified throughout the project. A process was developed to correct and validate the existing data and all future recorded data. The benefits to WCC of having a ‘single source of truth’ master dataset containing both attribute and spatial data is that the sophisticated analytics required to manage our networks are underpinned by a complete and accurate data set.

KEYWORDS

Asset Management, Underground Assets, Pipe Networks, Network Data

1 INTRODUCTION

The challenges of a renewals program that WCC now faces has meant historical asset systems and business processes at WCC were not sophisticated enough for the Planning and Assets team to manage Wellington’s underground infrastructure networks (wastewater, stormwater and water supply) with the challenges ahead.

The WCC underground facilities include an inventory of 2,800 kilometres of pipe and 250,000 associated fittings and facilities – described as ‘nodes’ in information systems. This asset inventory is valued at $2B and forms a significant proportion of WCC’s asset inventory.

Historically the Council stored asset attribute data in an enterprise database and the related spatial records in a geographic information system (GIS) with a scripted link between the two systems. Updating asset data to two disparate systems from as-built plans was overly time consuming, labour intensive and had a high potential for error. The Council’s closed circuit television (CCTV) survey, maintenance and repair data was being held in a form that constrained its potential use to improve asset data information and inform asset replacement analytics. The Council recognised that it required another level of sophistication to enable robust decision making. Put another way, it needed a single source of the truth.

For WCC this translated into a dedicated pipe asset management tool – a Pipe Information Management System (PIMS). It needed to be a best of breed system that would be capable of storing, interfacing, updating and inter-relating multiple forms of three waters network data from numerous sources to provide a repository of all information. WCC required a system that would be efficient and effective, and would enable the right information to be available at the right time for network management decisions.

InfoNet was chosen to meet the PIMS requirements to bring significant benefits to the Council by providing a consistent data foundation and the tools for cost effective measurement, monitoring, maintenance, and improvement programmes for the three waters pipe networks.

2 THE AMS ENVIRONMENT - A HISTORICAL VIEW

2.1 DISPARATE DATA SETS

WCC stored its underground infrastructure asset data in two locations. The primary asset record was stored in an enterprise database while the spatial record was held in the corporate GIS. A proprietary link between the two systems provided the means to keep the two systems synchronised. This is not an unusual approach within the local authorities.

2.1.1 ENTERPRISE ASSET PLATFORM

The enterprise database used by WCC is also similar to many other enterprise systems used by local authorities in New Zealand. It is a generic relational database with data stored on a SQL server. It is designed to meet the requirements of all aspects of a business's needs. In reality, enterprise systems must compromise in their ability to meet all needs by simplifying the more technical aspects of the business requirements. At WCC, the enterprise system was a manually entered data repository for all three water networks' asset attribute data.

Database configuration was done on an as needed basis by users and database managers with many changes over an extended period of time. Attributes were added to the schema of each asset type in an ad hoc manner. Clear business rules had not been adopted. This resulted in attributes being added to asset types that were not required or attributes not being collected which were. The inconsistence approach to the management of infrastructure data was a significant problem.

Historically, like assets were split according to attributes. For example, for each drainage network there were three types of manhole asset – brick manhole, pre-cast concrete manhole and concrete manhole cast in-situ. Similarly, there were two types of cleaning eye available in each network. Chambers existed as a separate asset type to manholes, but were also split by attribute. .

Significant challenges were also faced with the coordinate attributes between the enterprise system and those found in GIS. Updating these coordinates to the enterprise system were extremely problematic. Whilst the batch tools provided to update data in the enterprise system would allow for the addition of new data, existing coordinates could only be changed manually. This posed several problems, it was simply not feasible to update the coordinates of all assets in the three water networks manually.

Features within the AMS possessed three identifiers. These identifiers were the *central_asset_id*, *feature_id* and a combination of *site_code* with *plot_number*. The problems encountered with these identification formats are outlined below.

a) The *site_code* with *plot_number* is a numeric sequence of all assets numbered within a particular site resulting in a unique combination. Sites were created and managed according to the needs of business units that used surface features and assets. Consequently, the *site_codes* that were being used to identify and manage the underground assets networks changed independently to those networks. Therefore, it was not possible to find all of the assets within a particular site by listing the assets with that particular *site_code*. The *site_code* with *plot_number* is a poor reference to use for assets.

b) The *central_asset_id* was initially designed as a simple sequential unique identifier for all assets. At WCC the *central_asset_id* was over-complicated at the database design stage of the enterprise database implementation. The configuration of the identifier was changed to be a manual entry of the *site_code* with *plot_number* with a zero between the strings. This resulted in a complicated, manually entered 13 digit string. There was high potential for error and mismatches between the intended *site_code* and *plot_number* combination and the *central_asset_id* were common.

c) The *feature_id* is a string used to identify any feature. It is the primary method of identifying all features – nodes and pipes. At WCC the *feature_id* of assets was based on a drainage grid or a water grid as appropriate. The node *feature_id* has a four digit prefix from the name of the grid square and a sequential number based on its order within the grid square. It was not used as a primary key by the AMS so was never intended to be unique. Consequently there is no validation to ensure that duplicate *feature_id's* are not created. The problem of duplication extended when the drainage grid was corrupted and offset. This resulted in a large number of *feature_id's* being replicated for nodes and put into use.

   The pipe *feature_id* was based on the *feature_id's* of the nodes at each end of the pipe. There were two systems used to name the pipes. These were from low node number to high node number and from upstream node to downstream node. This mix of systems caused confusion for data users and added difficulty when updating, locating and reporting on pipes.

The enterprise AMS was not able to be configured to meet the business needs to support the effective management of its underground assets. The data model, asset naming conventions and business rules all further contributed to the problems users were experiencing in the normal operating environments of an enterprise system. The complexity that had ensued quite simply required a complete rethink

This and the requirement to significantly elevate the level of sophistication in information to support the understanding of our networks led WCC to consider its future options.

2.1.2 SPATIAL DATA PLATFORM

Historically, data held in the GIS was limited to only that information necessary to map the three water networks. The three identifiers used by the enterprise system and the feature type were held for all assets in the GIS. The operational status of pipes in the network were also held so that operational features could be differentiated from abandoned features. Complexities like manually generated pipe lengths from the enterprise database were held despite the length field of linear attributes being calculated by the GIS has also arisen over time.

The data management was not always fit for purpose. Level data for example was not held as part of the GIS dataset. Any analysis using levels needed to have the level data combined from the enterprise dataset. There were no rules governing topology included in the dataset or on the enterprise platform. Directionality was held in the enterprise database as a 'connectivity' attribute, stating which pipes fed to which nodes and vice versa. Because there were no defined drawing rules, the in-built tracing tools within the GIS could not be used analytically.

Historically, processes were developed in isolation with respect of one system to another and this has highlighted another range of issues. For example, spatial data with regard to any stormwater, wastewater or water network is usually used to represent what is currently in the ground. That is, GIS is used to display the assets that are in service. Asset managers and GIS technicians alike will extend this to include abandoned assets. Contractors, network operators and anyone else that may be digging near assets must know what they are likely to encounter and what its current service state is. When a node or pipe was removed from the ground it was deleted from the GIS. This created an imbalance in the number of features held in the GIS when compared to the features held in the enterprise database. Furthermore, through a lack of understanding of the value and need for a complete asset lifecycle, many features were historically deleted from the enterprise database at the same time as they were removed from the GIS.

Whilst the data held in the GIS and the functionality of the GIS platform allowed for the production of good quality maps (without network directionality) to be prepared, the data quality was not of a high enough standard to be reasonably used for analysis of the three waters networks without fundamental improvement in the data which underpinned these maps, notwithstanding a significant change in improving the processes which maintained data quality.

This compounded WCC's desire to find a solution which ameliorated, if possible, all of these inefficiencies.

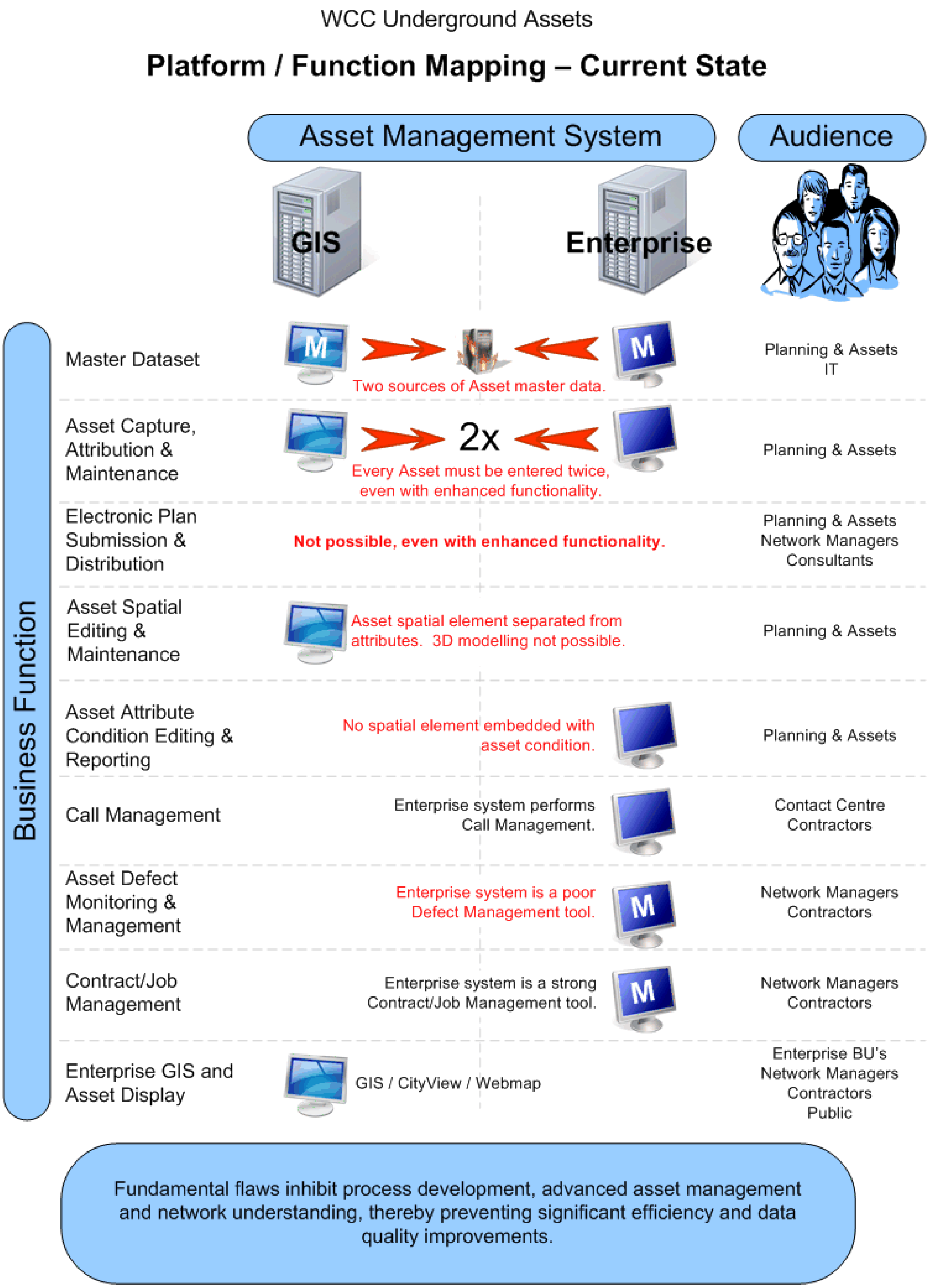

Figure 1: The Current State

2.1.3 THE LINK

Historically, data integrity was maintained between the two components of the AMS through the use of a proprietary link. The link is able to select a range of pre-registered asset types based on their *central_asset_id*. Assets selected in the GIS can be made available for editing in the enterprise system using the link provided. Selections in the enterprise system can be pushed to the spatial system. There were a number of challenges with this link and the data accuracy, particularly with the dependent data of the enterprise system.

Historically, a number of errors occurred where the link was unable to match to the corresponding asset data in the alternate system. This was due to assets not being presented correctly in one of the systems to the other; the *central_asset_id* is one good example of this.

The link was slow, and selecting large amounts of data from the spatial database in the enterprise system was extremely time-consuming and inefficient. This had the non-sustainable effect of creating an environment where WCC staff could not validate whether the data held in the enterprise system matched the data held in the GIS.

It was clear through our analysis that 'two versions of the truth' was simply not sustainable.

2.2 ADDING NEW ASSET DATA

New asset related data could be added to the AMS in a number of different ways. The methods of getting data into the network were dictated by the type of data involved, the skill set of the staff doing the data entry and the tools available. There was no single method of data entry. The implications of this approach are well understood.

2.2.1 CAPITAL WORKS AND VESTED ASSET DEVELOPMENTS

Historically, data entry of new or upgrade capital works and developments that resulted in vested assets was a manual process. All spatial information about the assets needed to be entered into the GIS and the attribute information needed to be entered into the enterprise system. The sources for the majority of the data were as-built plans. These were received as a mix of paper plans and *.pdf files. The *.pdf plans were always printed so that the detail contained on the plan was clear.

Features were created in the GIS from the as-built plans as received. Node features were usually created by manually entering a point feature at the coordinates provided in a coordinate table on the as-built plan. Where coordinates were not available, the features needed to be created using measurements from the as-built plan or measured using a scale ruler. Geo-referencing was not often used as a data entry method due to the large amount of superfluous data contained on the as-built plans. Frequently plans contained trees, gas mains, power cables and other unnecessary detail that prevented clarity when attempting to read the as-built plans.

Once the nodal features were created in the GIS, the pipe network was drawn in. It was not historically understood that the directionality of the drawn linear feature was important to the state of the network, functionality of analytic tools or future network modelling. All pipes recently entered have been drawn from upstream node to downstream. However, there are still a large number of legacy directionality errors to be corrected.

After all features were entered they needed to have the limited number of attributes held in GIS populated. This could either be done by manually entering the data into the GIS feature record or by updating the data from the enterprise system using the linking tool once the enterprise data had been entered.

The enterprise data records could only be entered from the as-built plan manually. This was a slow process due to the design of the software. Each attribute also needed to be saved before moving to the next. Hitting the enter key inadvertently resulted in the record being prematurely saved with a sequential *central_asset_id* rather than the *site_code + 0 + plot_number* format. The only way to fix this error was to delete the record being entered and start it again.

2.2.2 CCTV DATA

Historically, CCTV data was entered into the enterprise database using a proprietary batch data import process. There was no spatial representation of CCTV data.

The format of the data import file was time consuming and difficult to produce from raw CCTV information. A Microsoft Access database was introduced as an interim import tool and to simplify the data entry process. This allowed CCTV data to be easily entered while the output format of the interim database was written to match the import format of the enterprise system. However, it added another step to the process and another point at which errors could be introduced.

CCTV data can only be entered against existing pipes from node to node. If a CCTV survey finds an existing node, the survey is modified to reflect that in the field. This changes the survey parameters and the data that is to be entered into the enterprise system. Because a node has been found it will not exist in either the GIS or enterprise systems. Before the CCTV data can be entered, the new feature, including the original pipe split into two new pipes, must exist in the database. This forms a chicken and egg style problem where there is a need to update assets based on a survey but the survey can't be entered until the assets are updated. The solution was to enter the new assets manually based on a description and schematic of their situation. The CCTV data could then be loaded.

Data loaded against a pipe in the enterprise system effectively locks that record. The CCTV survey does not exist in its own right within the database. It is tied fundamentally to the pipe that it is loaded against. If a pipe is mistakenly entered into the AMS and a CCTV survey is mistakenly 'performed' on the non-existent pipe, it is not possible to move either the survey to the correct location or to remove the erroneous pipe. The records are locked until the CCTV survey is deleted from the erroneous pipe. Unfortunately mistakes such as this have occurred where pipes have been entered in error very close to existing assets.

It is becoming clear how difficult it is to maintain more than 'one version of the truth'. System compatibility and data integrity immediately and significantly impact on the ability to optimise the information for analytical interrogation.

2.2.3 MAINTENANCE AND REPAIR DATA

Historically, maintenance and repair was entered into the AMS in two ways. It was entered manually (extremely time consuming) or it was loaded using a proprietary batch data import process with limited data fields reducing its usefulness to be used for analytics. The majority of maintenance and repair records were entered manually.

Significant benefits were able to be gained in this process with the upgrade of our PIMS solution.

2.2.4 DATA VALIDATION

Data validation was a significant challenge for WCC.

There was limited data validation available within the enterprise database when records were entered. This validation was in the form of drop-down lists for attributes with required fields and some formats for asset numbering. Where an attribute did not exist in a drop down list historically it was simply added. This subverted any validation benefit that the drop-down lists could provide. CCTV surveys and maintenance data needed the target asset to exist, but there was no validation on the asset type.

No other form of data validation was in place.

2.3 HISTORICAL SUMMARY

WCC knew that there were data issues, and required a significant shift in systems capability and sophistication. With a tenuous link between the spatial dataset and the attribute database there was a significant data quality issue which required rectification. The quantum of the challenge was significant, so WCC made a plan.

3 THE BUSINESS NEED

Water managers rely on data held in their asset management systems for many reasons. It is this data that is used for many important decisions across capital investment planning, city development and growth planning, maintenance planning and monitoring. It also provides important information to developers, other utility companies / agencies and regulators for their own organisations. Having an inaccurate representation of assets can have far-reaching consequences both within and outside an organisation.

Water networks are living entities, and they change on a day by day basis. Accordingly, vast quantities of data about these networks are held by councils in varying ways. As-built information is provided when an asset is initially put in the ground. Maintenance information in the form of repair data or pipe cleans and customer complaint information is captured. Surveys such as smoke tests, dye tests and CCTV are also carried out. Monitoring information from data loggers or telemetry is routinely collected. Incidents such as flooding, bursts and pollution are recorded.

This data is all collected by different teams of people, and often in many different formats. This disparity of data sets makes it difficult for Council to validate the data accuracy or to use the information collectively and effectively in their decision making processes.

It was decided a number of key considerations would be had in the selection of a new PIMS. These included;

3.1 CONSOLIDATION OF DATA SETS

Having disparate data sets for spatial information and asset attribute data is not an effective or efficient method to enable the efficient managing of assets. WCC had come to the conclusion, through its self assessment, a single source of truth was needed. This called for the development of one database, a database which would hold all wastewater, stormwater and water supply network information. It would be the one point of data entry, data management and reporting. It would hold all identification, spatial and attribute data for each asset.

In this decision, it also became apparent other business units within council, consultants and the public, all obtain large amounts of data from these systems. While it was felt the PIMS solution would be the master data source, the solution also needed to ensure it pushed data to other enterprise systems to ensure the wider business needs were met.

3.2 DATA ENTRY

Data entry should not be something that is difficult for users. Information should follow a process that is consistent and as common as possible for all data streams arriving. Data entry methods should move away from being labour intensive and utilise improvements in technology. Electronic processing should be the norm for all as-built plans.

WCC can see that through the use of technology, it is possible to change the role of data entry into process monitoring. Staff can step back from the detail in the data entry process by allowing systems to do what they should do. Staff involvement then concentrates on improving data quality at the front end by ensuring that the processes are working correctly and efficiently. There is also more time for staff to perform analysis and problem solving. Value can be added to the network by redirecting effort and focusing on what matters.

3.3 DATA INTEGRITY

Data integrity is fundamental to all data that the Council collects, manages and stores. The data must be checked on its way in, held without degradation and made available while meeting uncompromised data quality standards.

WCC must ensure that all data collected contains the attributes required to manage the three water networks. Additionally, Council must ensure that any exception to the published codes and guidelines for network construction or development are identified. This would be done through the use of extensive validation tools.

WCC Underground Assets

Platform / Function Mapping – Desired Position

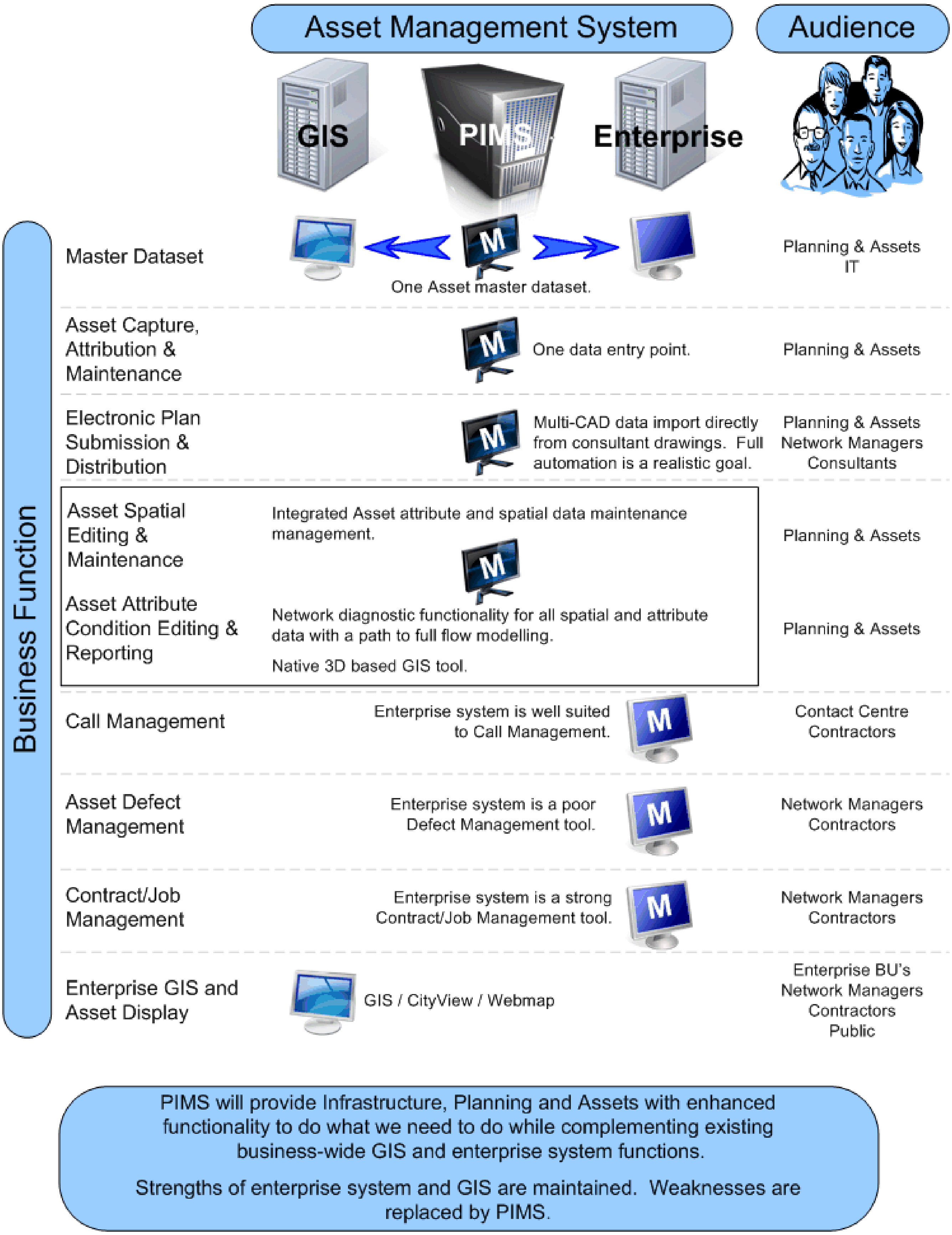

Figure 2: The Business Need

An extensive audit trail is needed so that we can determine the source of errors that may be found in the network. Systemic errors can be easily corrected provided the nature of the fault known and the source can be identified. The type of fault is generally easy, but without a comprehensive audit trail, the source can rarely be identified.

The quality of existing data needs to be improved as better information becomes available. With each inspection, survey and repair performed on a network the available knowledge base is increased. Information gained during inspections and surveys such as pipe diameter or material needs to be easily updateable to the asset dataset. In this way, data quality can be continually improved using maintenance, inspection and survey activity in the networks as the data source.

All inspection, survey and maintenance data needs to exist independently of the asset to which it relates. If a problem is found with the asset, the asset must be able to be corrected without constraint. If any activity is erroneously allocated to an asset there should be nothing hindering that record from being moved to the correct asset.

3.4 THE ASSET LIFE CYCLE

The full asset life cycle of every asset must be captured and retained for future analysis. Assets stay within the database regardless of their status from proposed through operational to abandoned or removed.

New assets added to a network and AMS are almost always added with an operational status of 'in use', 'operational' or something similar. They progress through their useful life until they are either physically abandoned or removed. There is a clear understanding that assets remaining in the ground should be marked as 'abandoned' in PIMS which marks their effective end of life as part of the three water networks. Asset managers will also recognise that assets physically removed from the network need to be attributed as 'removed' but their record must remain intact. It is the combination of the abandoned and removed assets that set the life expectancy of similarly attributed assets within the network. Additionally, PIMS must be able to analyse, display and report on assets based on their life cycle stage and across the full life cycle.

Historically, a large number of valid asset records were deleted from Wellington City Council's AMS. Through education, greater understanding and systems improvement, the Council can ensure that this will not happen again.

3.5 ANALYSIS

WCC needed analysis tools for use with the data of the three waters underground networks. Council realised that the existing enterprise system did not have a suitable analysis capability for many of the activities undertaken in the underground assets networks. This had been a significant functionality gap since the implementation of the enterprise system. Consequently, the Council wanted PIMS to have the ability to natively analyse the asset data for condition and criticality, assess repair and maintenance data and be a suitable platform for the review of CCTV data.

The Council needed to enhance its capability to manage the financial components of capital expenditure projects, repair activities, its maintenance contracts for the three waters networks and asset valuations. These activities would be better performed inside PIMS, a dedicated asset management system, than the existing process of exporting the asset data to another financial package or spreadsheets and then applying dollar values to each component.

3.6 REPORTING

PIMS needs to have a comprehensive reporting suite. As a minimum, it must be able to report spatially to the level of the existing GIS and on attribute data, to the level of the existing enterprise system. The reporting component must output to common business tools through a user-friendly interface. It is imperative that anyone that needs to get access to the information held in PIMS can do so following basic training on the system. Advanced database and report writing skills should not be necessary to get the data that keeps the business running on a daily basis.

3.7 RELATIONSHIP TO EXISTING SYSTEMS

PIMS needs to seamlessly transfer data to the existing GIS and the enterprise AMS system. The intent of PIMS was to give the asset management capability at the Council a significant boost. It was not to globally replace the functions and processes currently being performed well by other applications and platforms. It was important that these processes were not disrupted. Automated data transfers from the GIS to the Council's business partners and to a third party data provider needed to continue without significant change. Similarly, asset data that was currently available to other business units from the enterprise system needed to remain available so that capital project management activities and the job management of maintenance and repair activities continued as business as usual.

4 THE PIMS RFP

Wellington City Council issued a request for proposal (RFP) for a Pipe Information Management System in July 2011. Following the evaluation of responses, the shortlisted vendors were invited to present to the RFP evaluation panel and demonstrate their product. The evaluation criteria for responses and presentations as stated in the RFP documents were based on the Council's business needs. The RFP process was comprehensive and covered all aspects of the evaluation criteria.

In December 2011 InfoNet was selected by the evaluation panel and approved by the project steering group to become Wellington City Council's Pipe Information Management System.

An implementation project was scheduled to commence in January 2012. The wastewater and stormwater networks were combined to be done in Phase 1 of the implementation project. The water supply network was allocated to Phase 2 with a commencement date of August 2012.

5 THE SINGLE SOURCE OF TRUTH

InfoNet is the platform for all Wellington City Council three water networks information (hereafter referred to as WCC PIMS – Pipe Information Management System). All data relating to the networks will be stored in the PIMS master database. This includes all asset spatial and attribute data, all survey data including CCTV, smoke testing, pressure monitoring and manhole inspections, along with all maintenance, defect and repair data. All asset values, activity resource allocations and maintenance schedule of rates (SOR) items will also be held in PIMS.

PIMS will be the database where all functions relating to data management are performed. Validated raw data will be imported into PIMS according to the Council's business rules, but no processed data will be passed to PIMS from any other platform.

PIMS is our single source of truth.

5.1 CURRENT STATE – QUANTIFYING THE PROBLEM: PART 2

The first task following the purchase of PIMS was to determine the size of the data quality problem faced by the Council. This was done through a series of data imports of our existing data from both the GIS and enterprise system and evaluating raw numbers as PIMS saw them.

PIMS is an intelligent application. It understands that if it is holding a pipe asset, then there must be a node at each end of the pipe. If a pipe is imported to a position such that it does not terminate at a node, then PIMS will add a node at the end point. If a node does not have an identifier, then PIMS has the ability to create one for it.

The wastewater network was selected as the first network to import and staff would use the lessons learned in subsequent networks. All wastewater nodes were imported as a starting point and 38,324 nodes were successfully imported into PIMS. When wastewater pipes were imported an additional 2,488 nodes were created by PIMS. There were two reasons for this. The first was that there were a large number of nodes physically missing from the AMS. The second reason was that there a large number of snapping errors. Snapping errors are where features are not correctly connected. In this case WCC had 1,300 instances where pipes were not correctly connected to nodes.

The stormwater network was found to be in a similar state. Of all the nodes imported into PIMS as the primary data import, 1,124 were auto-generated by PIMS.

In most cases it will be possible to investigate historic data for both the wastewater and stormwater networks. The auto-generated assets can be identified from orphaned records in the AMS and from as-built plans. There are also nodes that were never entered into the AMS.

It was also possible to quickly understand the state of the asset attribute data using PIMS. Some attributes such as *Operational Status* for nodes had a 95% population rate. Others, such as pipe *Install Date* are only populated in 19% of cases. This clarity of data quality had not been previously available.

5.2 IMPLEMENTATION

It was decided to use the GIS data as a starting point for the implementation. Data quality was higher in GIS than it was in the enterprise system primarily because it was visible. If the asset physically exists then generally it is displayed in GIS, with the inverse also true. The attribute data was also imported from the enterprise database. At the point of data import, the mismatch between the data held in GIS and the data held in the enterprise could be quantified. It was clear that some assets existed in both components of the AMS, some existed in only GIS and some existed only in the enterprise system. It must also be assumed that there are assets in the networks that were not held in either database. The situation detailed in Figure 3 is representative of the state of nodes and pipes in both the wastewater and stormwater networks.

When the initial state of the imported data was combined with the attributes that were clearly missing, a sound foundation was set for data cleansing - using PIMS to return the data to the desired quality.

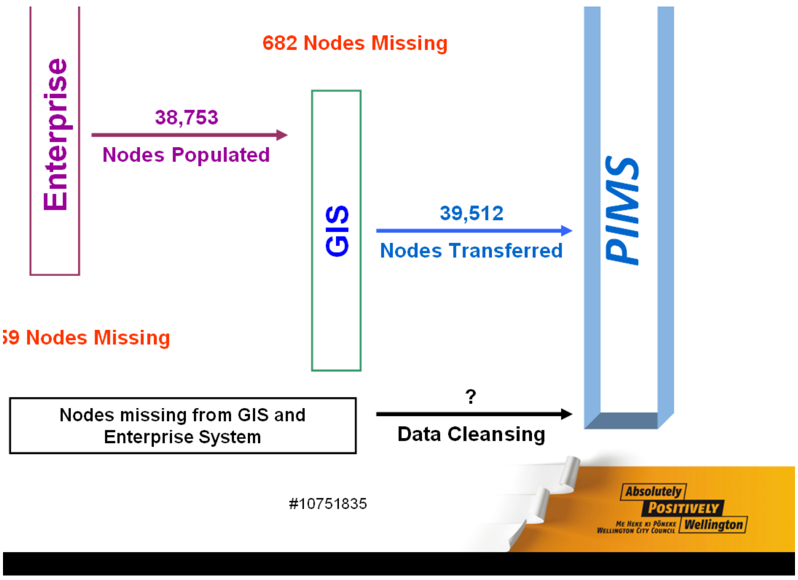

Figure 3: Wastewater Nodes as Imported to PIMS

5.2.1 DATA CLEANSING

The enhanced analytic tools in PIMS were used to assess the asset data as it was imported. As a starting point, a desktop study was performed of data completeness and quality. This has shown that there are distinct batches of work that can be done through desktop analysis and correction, non-field investigations of plans and related records, minor site visits and detailed field investigations. A comprehensive programme of works has been developed that will result in robust asset management data in PIMS of a known high quality and completeness.

The desktop analysis and correction was started as soon as the asset data had been migrated into PIMS. As a data cleansing starting point, all asset identifiers were renumbered with a new generic asset id. This ensures that the duplication of the legacy asset numbering system is removed when dealing with other AMS's. While PIMS uses the dynamic upstream_node.downstream_node.# format for identifying a pipe, most other AMS's that PIMS will interface to, both internally and externally, will need a static *asset_id*. The new generic number will meet their need.

Pipe diameters have been consolidated from three fields (*internal_diameter*, *external_diameter*, *nominal_diameter*) into a single *nominal_diameter* field. This was done using PIMS' SQL tools with care not to lose valid data and to ensure that the correct form of the nominal diameter was allocated based on pipe type.

The topology issues that were present in the pre-PIMS GIS data were identified and corrected using PIMS tracing tools in conjunction with its longitudinal section views. The CCTV data import subsequent to the asset import could not be done cleanly until all of the pipes "pointed downstream". This then allowed the CCTV data to be imported and the survey information used to populate missing attributes.

5.2.2 LINKING DATA SOURCES

As the single source of truth for Wellington City Council's three water networks, PIMS is able to receive all asset related data. The types of data currently being received, imported, validated, stored and analysed include as-built plans, defects / repairs and CCTV. These are standard data sources for the three waters networks. They are quick to establish and provide immediate benefit by improving the asset related data knowledgebase. This is not the limit of the data types that will be imported, but these three data types provide a quick return on investment.

As-built plans received from our partner organisations are currently being entered on a plan by plan basis. There are a number of ways that this can be done based on the format of the as-built plan. Plans received in CAD or shapefile format can be imported into PIMS using a tool called the Open Data Import Centre. This tool reads the spatial data available and has a field-mapping capability for any attribute fields that are included in the data. Attributed shapefiles are particularly suitable for import and are likely to become the preferred format of as-built plans for capital expenditure projects and new developments. If there are no attributes embedded in the data, then a separate attribute file can be imported to add the asset attributes. Naturally, attribute data can always be added manually, even when dealing with a spatial data import.

When paper plans or *.pdf's are received for assets, the data can easily be plotted manually and the attributes added. It is also possible to georeference nodes that should be added to the network in any standard GIS package and then import the resultant shapefile. It is then very simple to simply add the pipes in PIMS.

WCC is moving toward only accepting electronic as-built plans in an import ready format for the three water networks. This is being done in conjunction with the D-spec Consortium and will result in standard formats for all as-built data received by the Council. This development work is still on-going and is expected to be ready for consultation with stakeholders by the end of the year. When D-spec is implemented, as-built plans for all capital works and new developments will be submitted to Council electronically in the required format. These submitted plans will be imported into PIMS using the Open Data Import Centre and validated using PIMS's validation tools prior to acceptance. When the plans being received are in a standard format, a configuration file can be saved that maps attributes to their equivalent fields in PIMS. This reduces the risk of error and increases efficiency within the Planning and Assets team.

In the as-built plan future state, PIMS will use a proprietary tool called InfoNet Exchange (IE) to automatically import as-built plans to a holding area for validation. Submitted plans received will be moved to a generic import folder to await a nightly import to PIMS. IE uses Ruby scripting to automate its actions and set its configuration. This will save staff time and free them up to focus other key tasks.

CCTV data is stored in PIMS as stand-alone surveys, associated with a pipe. They are stored in a separate table to other survey types, repairs and asset data and have their own attribute types. CCTV data is being imported using the Open Data Import Centre with a configuration file as all CCTV data received by Council is in a standard pre-defined format. Once the data passes validation it is included in the CCTV survey dataset and linked to the appropriate pipe asset. The survey is then used to update any details of the existing assets such as attributes (pipe material, diameter etc.), spatial information or create any assets that may have been found during the survey. If an asset is created the survey is automatically split to ensure the relational integrity of the survey to the asset.

Repair and defect information will be imported in a similar manner to CCTV data. The output from Council's partner organisations will be submitted in a standard format. This allows the use of the Open Data Import Centre with a standard configuration file. Once imported the data is standalone in nature, but is also linked to an asset.

As with as-built plans, the CCTV and defect/repair data imports can be automated using Infonet Exchange and a small amount of Ruby scripting. By using a simple batch file, all of the data files to be imported can be placed in a collective bucket and scheduled for import to a working file. Validations can then be performed prior to the data being included in a certified dataset.

All of this data is available spatially once imported. Other data types will be included in the datasets being imported as the PIMS project progresses.

5.2.3 DATA EXPORTS

PIMS will export data on a nightly basis to the Council GIS and to the enterprise system – the two parts of the pre-existing Council AMS. This is being done so that existing Council systems and processes that use these as data sources will not be compromised and business can continue seamlessly. There is no noticeable change to the nature or levels of service provided to other business units, partner organisations or the public.

The PIMS data export to the Council AMS is done through an additional automated export feature (IE) available with the PIMS Software. This looks at the current version of the PIMS database and compares it to the version of the database at the time of the previous export. This comparison allows IE to select only assets with changed data to be exported. New assets, assets whose position has been 'clarified' or assets that have had an attribute updated are examples of what are selected. This 'changes only' export is effectively a one-way synchronisation and results in much faster data transfers than would be achieved by attempting to update or overwrite the entire dataset.

One of the key drivers for the PIMS project was to ensure that the full life-cycle and history of assets could be managed. Assets remain in PIMS even after they are physically removed from a network. It is important, though, that these assets are not displayed in the GIS. When a consultant, designer or contractor is viewing data in a GIS environment, particularly if they are in the field, they are only interested in what is actually in the ground. The fact that a pipe was once buried there is not of interest. Furthermore, any removed assets will clog plans and maps if they are included in the GIS. The nightly update of data from PIMS to GIS will not include assets that have been removed. In this way Wellington City Council is able to fully record and manage the history and life cycle of assets without assets that have been removed clogging the spatial data platform.

6 THE BENEFITS OF PIMS

6.1 SINGLE SOURCE OF TRUTH

Underpinning any asset management system is data. The first step in any asset management process is to focus on the consolidation of existing information and data held in different silos by different departments in an organisation. The next step will be to report on the quality, completeness and reliability of this data. Lessons can be learned at this stage from the success and failures of previous data collection and management programs. This is the stage when an organisation can start thinking about improving their processes.

All the information WCC owns will be consolidated, and integrated with their existing systems. The PIMS is designed to be a single source data set for all Water, Wastewater and Stormwater Assets. It allows asset attribute, incident, survey, spatial, work order, repair/intervention and resource data to be entered/imported, dispatched, stored, improved, validated, interrogated and reported. Once the asset registry has started to be organised, Validation and Quality Control tools are available to check the completeness, accuracy, consistency and connectivity of the plethora of data.

It is crucial for organisations starting out on an asset management program to ensure that data consolidation, quality control, and data collection protocols are high priorities. With data that is verified and audited according to an agreed company standard, it is possible for organisations to be confident in the operational, planning and financial decisions they will have to make when undertaking proactive, preventative and responsible asset management.

6.2 DETERMINE CRITICALITY – ASSESS RISK

Asset management seeks to understand how an asset performs (and degrades) over time. From this understanding we can derive optimal times for rehabilitation, repair, replacement, etc. From the aforementioned examples we can conclude that asset data must form the basis for any type of asset management program that seeks to quantify risk. Differing asset management approaches might require different sets of data, but data and data management are generically required for all asset management initiatives. In particular, management of historical data for each asset is required to complete the risk prediction.

Effective decision making about assets is all about managing risk. Risk management has in the past been driven by decisions that have been based on perceptions rather than facts. Tools to analyse risk have been basic or non-existent. Information and data that is required in order to understand risk in the first place were decentralized, located in different silos and therefore difficult to get access to, unverified and of poor quality.

Taking a strategic and knowledge based approach allows WCC to begin the process of risk management. Although there are other factors, this requires reliable and quality controlled data, easy and flexible access to this data, and standardised, repeatable analytical methods. PIMS has the ability to provide all of the above.

6.3 IMPROVED FINANCIAL MONITORING

One of the keys to successful monitoring and financial planning in an asset management system is the ability to hold the correct level of detail relating to costs. Too much information and you become swamped and the retrieval of this information becomes a difficult task. On the other hand each activity relating to asset management requires costing information in order to determine Capital Investment Planning and Operational and Maintenance costs over time.

Collecting direct and indirect cost information (such as pipe age) is a difficult task. Understanding the provenance and accuracy of the cost information is crucial when making financial statements. Incorrect data or gaps can have a profound impact on the accuracy and efficacy of decisions.

PIMS provides the necessary tools to validate the accuracy of the costs and associated data, that can be configured by WCC to perform the required standardised calculations, and flexible reporting tools.

6.4 LONGITUDINAL SECTIONS AND 3D REPRESENTATION

Through the use of network topology rules, tracing tools and connectivity analysis, visualization of connected features in long-section profiles and 3D views enables users to visually understand the operation of networks, and quickly and easily identify the data inaccuracies.

6.5 IMPROVED ANALYSIS TOOLS

The key query/analysis tool utilised in PIMS is Structured Query Language (SQL). The SQL tool incorporates both attribute and geospatial querying capabilities. It takes full advantage of the relationships between assets and "events" (CCTV, smoke tests etc) in order to carry out nested queries (e.g. query pipe information from a CCTV inspection) and through spatial querying can interact with GIS layers such as soils and traffic loading. In doing so, users are able to leverage current and historical information and data pertaining to an asset (e.g. showing deterioration of condition over time).

The SQL Tool provides a mechanism for opening up the data in order to:

- Carry out analysis for decision support
- Organise information and data
- Produce tabular and graphical reports
- Quality control and validation of existing data
- Carry out calculations that can add value
- Create new information based on pre-defined standards and methods.

6.6 VERSION CONTROL WITH AUDITING HISTORY

The version control system maintains a complete and permanent record of all changes to asset and survey data. Version control supports multiple users editing the same network at the same time, and includes tools to resolve any conflicting changes that may occur. Should errors be found in the data, WCC staff can track when the erroneous changes were made, and who made them. Version comparison is also enabled for single objects to enable users to view and query how an asset has changed over time.

6.7 DATA FLAGS

There is comprehensive auditing available through the use of data flags. Every data field on all network objects have a designated flag field. For example, WCC has implemented flags on data identifying the source of the data such as as-built, survey inspection, consultant and also the accuracy of the data such as identifying inferred data.

6.8 VALIDATION

PIMS has built in data validation rules for collection systems and water distribution systems. Examples of these are: "invert is above cover level of manhole", or "this pipe is crossed or overlain by another pipe or pipes, only fittings are allowed on lateral pipes". These tools tell WCC where errors in the data exist. Additionally users are allowed to set required fields for every object and data field. If these fields are blank this information is reported to the user. Entries that are not consistent with the choice list defined for a particular field are also reported. Numeric range control allows the user to set minimum, maximum and unit constraints on each field. Users can also set an unlimited number of user defined validation rules using SQL. For example, if a work order was issued to repair a 6 inch pipe and the pipe repair record indicates a 10 inch pipe was used in the repair, this inconsistency can be flagged. Each validation rule is prioritised into 3 categories: 1 is error, 2 is warning, 3 is information. Validation rules are saved within the master database and can be used across all networks allowing for standardisation. There are potentially over 5,000 separate validation rules that can be used to check data.

6.9 INFERENCE

Inference tools are available to fix errors in the underlying data or populate where data is missing based on rules and sound engineering judgement. The data auditing flag system is used to show the user that the data is inferred and cannot necessarily be relied upon.

6.10 REPORTING

Extensive in-built reporting functionality means that reports can be easily generated and replicated on any required dataset such as incidents, survey and inspections, and operations and maintenance. In addition, quality control validation reports can be generated as well as depreciation and rehabilitation cost reports.

6.11 SEAMLESS LINKS TO OTHER PLATFORMS

PIMS will support the exchange of data in a wide variety of GIS, text based and database formats such as the ESRI personal, file and enterprise Geodatabases, shapefiles, and coverages; AutoCAD drawing files (dwg, dxf); comma separated variable (CSV) and Tab separated formats, Microsoft Access, Oracle and Oracle Spatial and SQL Server database tables. This flexibility is important to WCC as data is being imported and exported from a wide variety of sources.

6.12 IMPROVED DECISION MAKING

PIMS allows the opportunity for cost-effective decision-making for both day-to-day operational management and long term network planning. It is now possible for WCC to standardize on criticality and risk methods, configure and develop their own querying and analysis tools, since the 'single source of truth database' comprises all the required asset data and information, coupled with the various inspections, surveys, incidents, customer complaints, and work orders etc. Data, tools and methods are now easily available to engineers to undertake decision support.

With one single data source, calculation of the number of bursts per mile per year, or the number of times a customer has had low water pressure can be carried out, for example. With appropriate data it is possible to ask virtually any question of the data. Analyses can take the form of decision matrix, step-wise decision tree, what if scenarios.

6.12.1 CAPITAL PLANNING

The development of a Capital Improvement Plan (CIP) is a data driven process. High level planning tools require a minimum amount of asset data, but deliver only rudimentary assumptions as to how an asset will perform over time. Improving the accuracy of capital forecasting requires a corresponding increase in the amount and types of data associated with the assets in question. At the most fundamental level, an asset management system that supports capital planning processes must answer basic questions. Where is the asset? How old is it? What are the material properties? What condition is it in? Answering these questions enable a high level understanding of our infrastructure, facilitating high level capital planning decisions.

The potential exists to use asset management techniques for more detailed capital improvement planning. If the required data is managed properly, we can understand our assets from a "bottom up" perspective. We can then answer questions about individual assets, including their interrelation to each other. How is this asset performing? At what rate is it degrading? How much would it cost to repair or replace it? What are the ramifications of differing maintenance/replacement scenarios? With a fixed level of investment, which assets do we repair or replace first? What are the criteria for prioritisation (condition, criticality, expected life, etc.)?

Capital plans that seek to answer these questions not only require detailed asset information, they demand to know how that asset has performed over time. How many times has an asset needed repair? What type of repair was required? How much has maintenance cost over the last 10 years? Hence, a decision support platform for capital planning needs both detailed asset data and a record of how that asset has performed over time.

Complicating this equation is the fact that the assets of water/wastewater operators are often buried, so historical performance and condition data must often be derived or inferred. Hence, any asset management platform must offer a set of flexible data management capabilities that support everything from high level estimation to detailed asset analysis. Regardless of the level of complexity, a decision support system for capital improvement planning must offer the ability maximize the value that is inherent in the asset data.

PIMS supports a significant array of capital planning decision support tools. It provides the three components that are critical to unlocking the value of asset data for capital planning: Water/Wastewater Data Model; Enterprise Data Management of Multiple Technologies; and Management of Asset History Data

7 THE FUTURE FOR WELLINGTON CITY COUNCIL

There are multiple work streams that the WCC Infrastructure asset data management team are considering undertaking. These work streams include:

- Data business process mapping projects
- Data improvement projects
- Data digitisation projects
- Data collection projects
- Data validation projects
- Data interfacing / access projects

WCC plans to leverage off PIMS analytical tools and data management capabilities to get to the future state as illustrated in Figure 4 below.

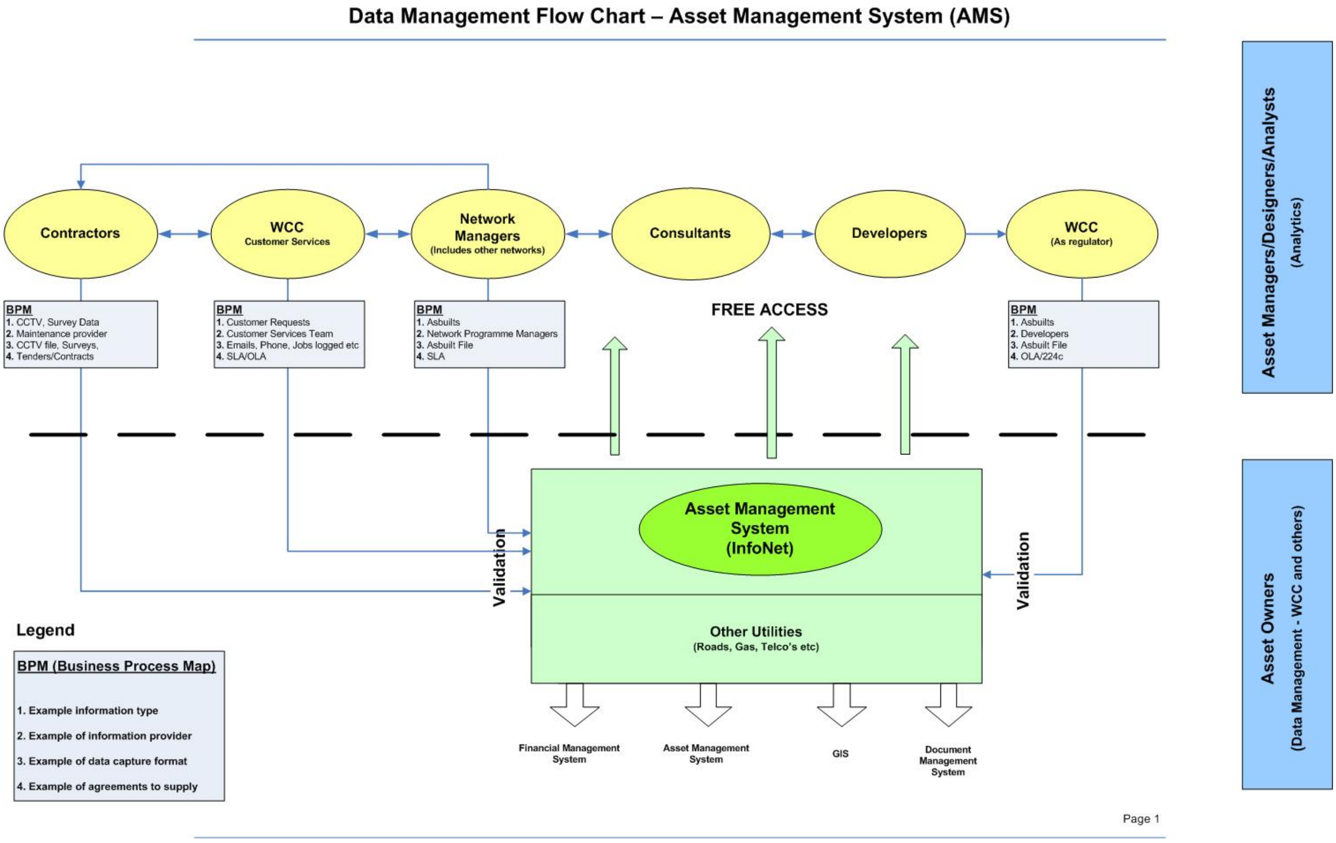

Figure 4: Data Management Flow Chart

8 CONCLUSIONS

It has become apparent to the WCC data management team, through the implementation and migration of data into PIMS from a number of different systems, that there are a number of key gaps within the data management processes which impact negatively on data integrity. Parallel datasets, incomplete records, inconsistent data capture and recording techniques have all added to inaccurate and disparate sets of data and information in the three water networks.

The project benefits were in both effectiveness (joined up, comprehensive insights into infrastructure assets) and efficiency (through technology solutions and reduced processes). There are significant benefits to be gained from a "single source of the truth" which offered real transparency and accountability around asset performance.

Wellington City Council sees InfoNet as a tool which will provide not only a solution to the information challenges currently at hand, but will also provide the foundations for another level of sophistication to address the specific future infrastructural issues of the City.

ACKNOWLEDGEMENTS

Thanks to the following for their contributions to this project. Their hard work in developing a long term system for the stewardship of Wellington City Council's wastewater, stormwater and water supply assets is greatly appreciated. The Council and people of Wellington will benefit from their efforts:

Wellington City Council
WCC Infrastructure Directorate
WCC Spatial Business Intelligence
WCC IT

Capacity Infrastructure Services

D-spec Consortium

Jeff Booth