Toitu te Land whenua Information New Zealand

Landonline Bulk Data Extract

Overview

June 14

Contents

Section Page

1. Bulk Data Extract 1

1.1 Introduction 1
1.2 Data Content 1
1.3 Purpose 3
1.4 Amendments 3

2. Bulk Data Extract (BDE) 5

2.1 Data Supply Frequency and Updates 5
2.2 Potential issues related to incremental extracts 6
2.2.1 Usage of Static and Unique Entity identifiers (SUE's) 6
2.2.2 Audit Table 6
2.2.3 Currency 7

3. Data definitions 8

3.1 Bulk Data Extract Data Models 8
3.1.1 High Level Data Models 8
3.1.2 Roads & Addresses Dataset 10
3.1.3 Parcels Dataset 11
3.1.4 Parcel Topology Dataset 12
3.1.5 Survey Dataset 13
3.1.6 Marks, Nodes & Coordinates Dataset 14
3.1.7 Survey Observations Dataset 15
3.1.8 Adjustments Dataset 16
3.1.9 System / Shared Dataset 17
3.1.10 Title Dataset 19
3.1.11 Title Memorial Dataset 20
3.2 Landonline BDE - Data Dictionary 21
3.3 Spatial Data 21
3.3.1 Spatial Reference ID (SRID) 21
3.3.2 Geometry Shift 22

4. Data supply formats 23

4.1 File Format 23
4.1.1 Metadata 23
4.1.2 File Header Format 23
4.1.3 File Body Format 24
4.1.4 Sample File Layout 24
4.2 Data Formats 25
4.2.1 Data Types 25
4.2.2 Graphic Entity Representation 26
4.3 Media Supply Format 27
4.4 Schema Changes 27

5. Estimated extract volumes 28

6. Pricing and Licensing 29

1. BULK DATA EXTRACT

1.1 Introduction

This document explains how Bulk Data can be obtained from Landonline via Bulk Data Extract (BDE). BDE includes survey, geodetic, title, road and street address that is maintained by LINZ in undertaking its statutory functions.

This data is typically used to provide a land parcel framework for agencies that operate property-related geographic information systems. e.g. local and regional authorities and utility companies.

1.2 Data Content

The data that is available consists of the following broad datasets:

- Roads & Addresses
- Parcels
- Parcel Topology
- Survey
- Marks, Nodes & Coordinates
- Survey Observations
- Adjustments
- System / Shared
- Title
- Title Memorials

Further details of these datasets are contained in sections 3 and 5.

BDE does not include image data, e.g. scanned documents and survey plans.

The following BDE Related Datasets may be useful:

Electoral Streets & Places

For those users receiving the Roads & Addresses Dataset, additional road name data can be obtained from the Electoral Streets & Places. Information can viewed or downloaded from the LINZ Internet site. The data was posted weekly and needed to be obtained to coincide with the monthly BDE unloads. This was previously included as an additional dataset with monthly supplies. From June 2014, the Electoral Streets & Places maintenance stopped and the street

location data has replaced the SUFI previously stored in the RNA.location field. The asp.zip is no longer supplied with BDE.

Digital Meshblocks

Digital Meshblock boundaries are maintained in Landonline on behalf of Statistics New Zealand. This data is not available as part of BDE as Statistics New Zealand has a requirement that the boundaries released to the public are 'as at' specific dates ie to coincide with the release of their annual and census datasets. For access to these official releases refer to "Download Digital Boundaries" on the Statistics New Zealand Internet site (www.stats.govt.nz).

Geodetic Database

Geodetic data is available as part of BDE, but users of geodetic data will find it more convenient to download geodetic mark data directly from the Geodetic Database on the LINZ Internet site where it is freely available.

Digital Cadastral Database (DCDB) related datasets

For those users that previously obtained cadastral data from the DCDB, the DCDB SUFI-SUE & Landonline Final DCDB Coordinates datasets may be useful. Further information on these datasets can be found on the LINZ Internet site and in the "Mapping of DCDB features into Landonline" document. These datasets can be added to a BDE supply as a one-off exercise.

1.3 Purpose

The interface used to extract Bulk Data is specifically intended to cater for those users who have a requirement to access large volumes of data on a regular basis. The extract allows the recipient to recreate the database held in Landonline. It does not enhance the data in any way, other than to present it as a series of standardised text files.

Entire table views are exported. Consequently, the extracted data covers the entire country, rather then individual areas or pre-defined geographic extents. This approach allows the data to be made available with minimum intervention, minimum cost, and with minimum overhead to Landonline itself.

The data is only provided on DVD, and will be up to 8.21 Gb in its compressed format, depending on the datasets selected (see Section 5 for dataset sizes)

Prior consultation with a Landonline BDE reseller is strongly encouraged, as they may already hold the data in the format you require. A list of resellers is available on the LINZ Internet site.

1.4 Amendments

Version Number	Amendment(s)
1.0	First draft for internal comment. Wholesale changes.
2.0	First draft for external comment. Wholesale changes.
3.0	First official release, 1 May 2000. Changes made to format and removal of "draft" status.
3.1	Corrections due to Informix Spatial 8 in Landonline 1.6 release Shape fields change from geometry to st_geometry. Format for coordinates no longer has a space following the bracket.
3.2	2 Gigabyte file size limit – LIN, PAR, VCT, AOC, OBA and OBN affected.
3.3	DCDB to CRS mapping details added
4.0	Addition of certain title related tables, as available from the July 2002 BDE.
5.0	Change to Marks, Nodes & Coordinates Dataset file names.
6.0	Remove VTS table, WRK.annotations field and note about SFI table supply
7.0	Remove AOC table
8.0	OBN table split into 3 files due to 2GB file size limit & file sizes

	Remove reference to stt1 shape file
9.0	DVD Supply & update to file sizes
10.0	No longer offer DLT supply
11.0	EEZ in SRID diagram corrected and dataset sizes updated
12.0	RNA.location description to include where to find ASP
13.0	TIN.audit column & addition of ALS unload
14.0	RSU.bearing_corr and dataset sizes updated
15.0	DVD Supply on DL 8.5GB
16.0	PRP.name_suffix column added
17.0	ETS.share_memorial field size for Landonline 3.0 release, all lvarchar to lvarchar(2048) and major update of whole document (including name change from "Introduction to the Provision of Bulk Data from Landonline" to "Overview")
18.0	Remove DCDB Mapping Dataset (now in "Mapping of DCDB features into Landonline" document), ASP to Electoral Streets & Places, epa1.crs removed from supply and minor edits.
19.0	Changes to Titles Dataset unloads to remove TTL.audit_id link and addition of SAV2 & STT2 unloads to Survey Dataset
20.0	STE unload added
21.0	Change to dissemination charge, effective 11th September 2009
22.0	Change for file compression to GNU zip (Unix gzip)
23.0	Addition of Titles Memorial Dataset
24.0	Additional fields added to RSU, WRK1, MPS1, SIT1, TTE, ENC, ENS, ETS, ETT, PRP, TDR, TIN, TLH, TTL Renamed unloads MPS1 to MPS2, NPO1 to NPO2 & MNT1 to MNT2 New unloads CMP, DVL
25.0	Change to dissemination charge, effective 1st October 2010
26.0	Additional field added to LDT & TTL New unload CPTA
27.0	STE.name_and_date field size changed from varchar(100) to varchar(200).
28.0	ASP no longer maintained and changes to road name & location in RNA unload.

2. BULK DATA EXTRACT (BDE)

Bulk Data Extract (BDE) is the Landonline interface that has been designed to provide access to Bulk Survey and Title Data. It has been designed as a batch process, which incrementally extracts data on a daily basis, with full extracts generated once per month.

The objective of BDE is to enable access to all survey related textual and spatial data except where this data is restricted due to confidentiality or commercial sensitivity. For titles data, only certain Landonline tables are extracted. Limited data in the Titles and Titles Memorial datasets, protected under section 107 Domestic Violence Act 1995, has been excluded from BDE supply.

Entire table views are exported. Consequently, the extracted data covers the entire country, rather then individual areas or pre-defined geographic extents.

BDE uses a "view" of each Landonline table, allowing some fields to be omitted, and in some cases combining two or more Landonline tables into one. Therefore, table views accessed by BDE are not always identical to the tables that exist in Landonline.

The files supplied are 8 bit text files with ANSI (Windows) 1252 encoding and Unix linefeeds (ASCII code 10). Each file consists of a header section and then a pipe delimited data section, with any spatial data represented in Modified Well Known Text format (see section 4.2.2). The Unix gzip command is used to compress these files, which may be extracted using most unzip utilities.

2.1 Data Supply Frequency and Updates

At present only monthly (full) extracts are available. These extracts will normally be run on the first Saturday of every month.

Daily extracts will also be run to provide a record of change information. Incremental extracts are not available for supply.

Once a month; daily and monthly extracts will be performed. Daily extracts are also run, where daily changes apply since the last daily increment. The extract process runs either over-night or on the weekend when business transactions do not occur.

The table views used by the extract process generally limit extracted data to "current" records. Data that has yet to be Approved or Authorised (status of "Provisional" or "Pending"), and data that has been retired (status of "Historic") are generally not included in the extract.

Data extracts from Landonline originally had a file size limit of 2 Gigabytes. Some tables have been split into multiple files in the monthly (full) extract. An incremental extract will be to one file. At present, the tables affected are the Line, Parcel, Vector, Obs Accuracy and Observation. In all cases, the following logic has been used :

- Incremental files – lin1.crs, par1crs, vct1.crs, oba1.crs and obn1.crs
- Monthly files – lin1.crs & lin9.crs, par1crs & par9.crs, vct1.crs & vct9.crs, oba1.crs & oba9.crs and obn1.crs, obn3.crs & obn9.crs

As the 2 Gigabyte file size limitation no longer applies, tables in the Title Memorial Dataset have not been split into multiple files. At present, the Title Memorial and Title Mem Text tables exceed 2 Gigabytes.

2.2 Potential issues related to incremental extracts

2.2.1 Usage of Static and Unique Entity identifiers (SUE's)

A Static and Unique Entity (SUE) identifier is a unique reference made to a Landonline record by referring to both the table name and the record identifier for that table. i.e. there may be many records in Landonline with the same record identifier, but the combination of the table name with the table record ID will always be unique.

A SUE is considered to be static as it will not change for the life of that record, and will not be used by any other record.

A SUE should not be confused with a Persistent Identifier (PID) or a Static and Unique Feature Identifier (UFI or SUFI), both of which remain static for the life of a "real world" feature, e.g. a survey mark or street address.

BDE clients should not automatically assume that a SUE can be used in lieu of a PID or a SUFI, as there are many situations in Landonline where changes to real world features are recorded by retiring the existing record of that feature and replacing it with a new record with a new record identifier (SUE).

2.2.2 Audit Table

Row level triggers placed on all tables accessed by BDE allow changes (insert, update and delete) to be automatically recorded in an audit table as soon as they occur. This audit table is required by BDE's incremental extract process to identify records that have changed since the last extract was performed. The records accessed in the audit table are included in each incremental extract so that clients can identify the type of change that has occurred. The audit table is not relevant for full extracts, and hence will appear as an empty file for this type of extract.

Clients should be aware that some records flagged as changed will appear to be identical to those previously supplied. This is most likely to occur where attributes have been edited in fields that are not included in the BDE table view.

There will also be some instances where Landonline is unable to correlate an original record set with an edited record set. Landonline manages these situations by retiring the original set and inserting the edited set as new records, and as a result there will be situations where existing records are retired and replaced with identical new records which have new record ids. This has yet to be fully documented but is most likely to occur to road centrelines, street addresses, and place names, all of which are edited via an extract/merge process.

For the Titles and Titles Memorial datasets, only the Title (ttl.crs), Ttl_inst (tin.crs) and Transact_type (trt.crs) tables have an AUDIT_ID associated to them. The remaining tables are not currently able to be delivered incrementally.

2.2.3 Currency

The omission of unapproved and retired records from Bulk Data Extract has the following implications for incremental extract clients;

- Most records in Landonline are "retired" by either setting their status to historic or decommissioned, or by population of their end date. Very few records are deleted. Retired records will appear in the audit table as an 'Update' rather than a 'Delete', and the retired record will not be extracted as it is no longer current. For example, if the status of a parcel is changed from current to historic, the audit table will record this as 'Update', but the historic record will not be included in the extract. A list of records which have been "retired" may be generated from an incremental extract by identifying SUE's present in the audit table (where their status is "Update"), that have no corresponding record in the balance of the extract. SUE's for tables that have not been requested in the extract will need to be ignored.

- Records that have a status of "Update" that do not exist within the client's database will need to be treated as an "Insert". Explanation; In the above example, if this parcel was then manually changed from historic to current via a back end process (e.g. an error was discovered) it would appear as an "Update" in the next incremental extract. This would be valid to a client who had received a prior incremental extract that set the record to historic, provided they had set the record to historic in their own database. It does however present a problem to a new client who would not have a record to update, similarly it would be a problem to an existing client if that same client deleted the original record.

- When the same SUE's appear more than once in an incremental extract's audit table, only the most recent entry should be used. Still using the parcel example, a manual over-ride that changed the parcel status from current to historic then back to current would cause the parcel to re-appear with the same SUE. In this example the parcel would appear as an update, but a comparison with the client's data would indicate no change.

3. DATA DEFINITIONS

The data available through BDE is described in this document and a Data Dictionary. At a later date some documentation on data characteristics may also be prepared.

3.1 Bulk Data Extract Data Models

Two high level data models are shown here to give an overview of Bulk Data Extract. This is followed by 11 detailed models, each representing a Bulk Data Extract dataset that may be requested. There is no overlap between datasets.

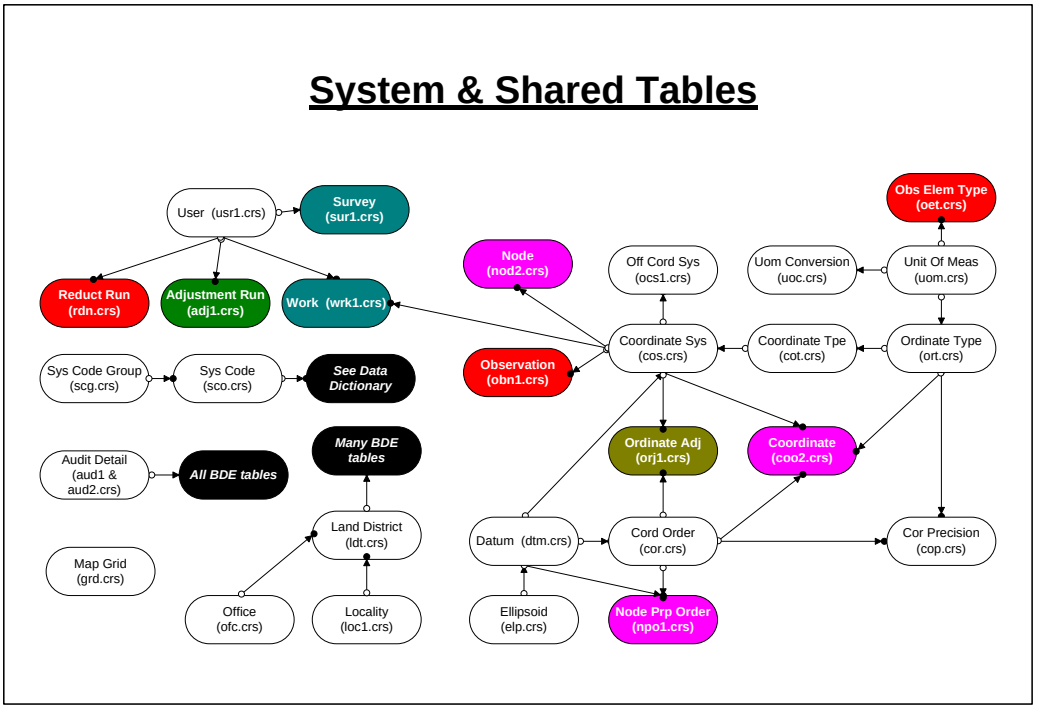
All of the following models have been presented in a diagrammatic format. No attempt has been made to indicate cardinality, however the arrows drawn between entities point from the entity containing the unique record, to the table that may contain one or more references to that record (i.e. primary key -> foreign key). In the detailed models the relationship between tables is indicated in the same direction as the arrow (i.e. primary key = foreign key).

3.1.1 High Level Data Models

Note: This sub-section is legible in black & white but best viewed in colour.

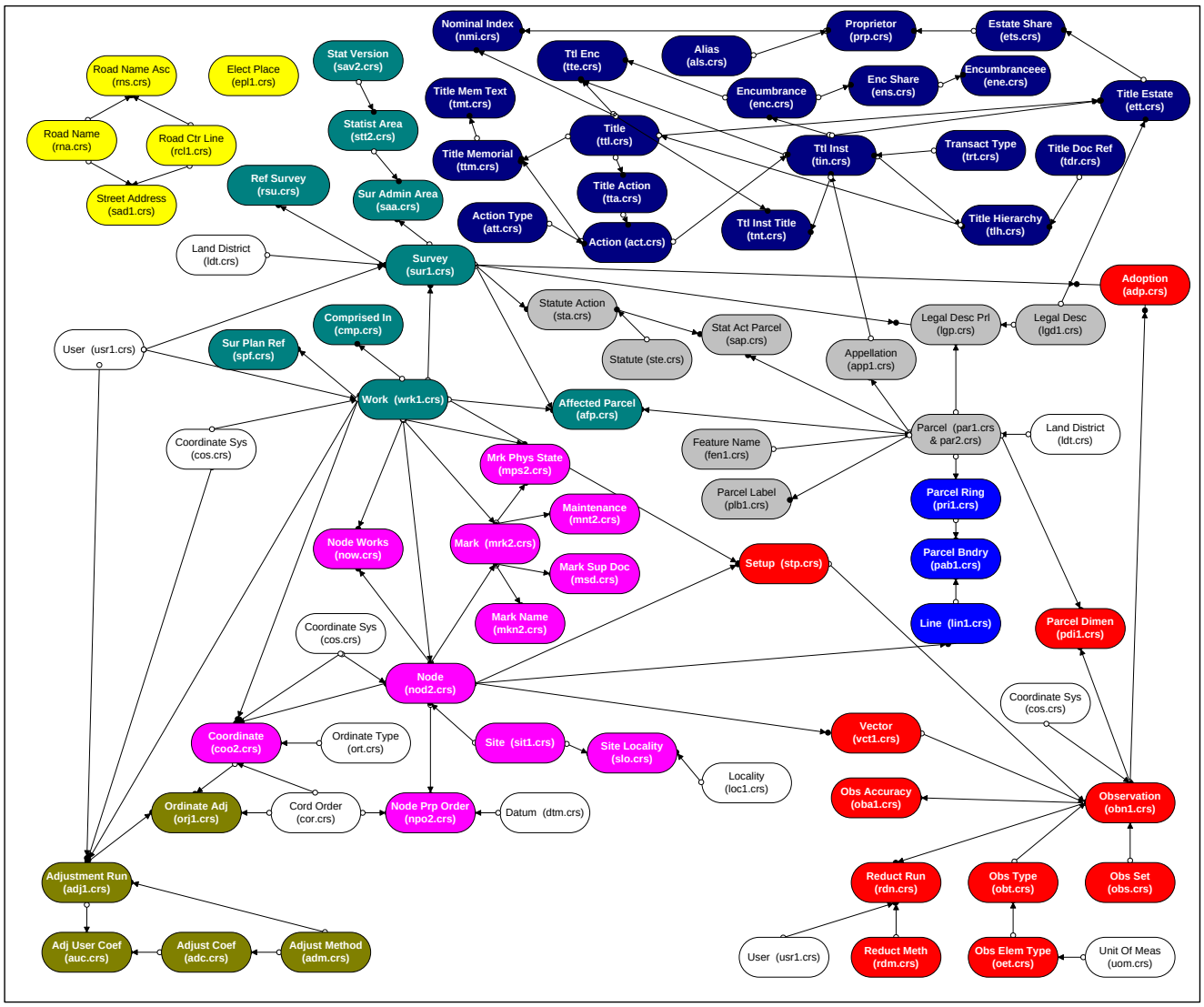
Colour coding;

Yellow = Roads & Addresses	Magenta = Marks, Nodes & Coordinates
Grey = Parcels	Red = Survey Observations
Blue = Parcel Topology	Brown = Adjustments
Aquamarine = Survey	White = System & Shared
Dark Blue = Title	



Main Tables

(excludes some system / shared tables)

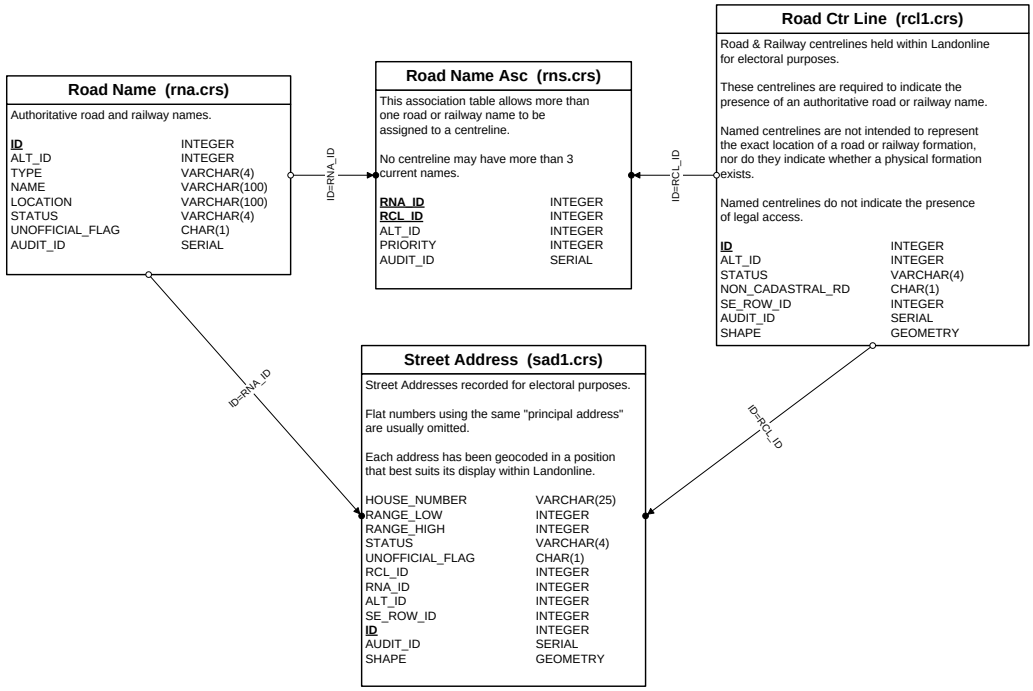


3.1.2 Roads & Addresses Dataset

This dataset covers the road, address, and place names stored in Landonline. This data is stored and maintained for electoral purposes, however is also used in Landonline for spatial searching and location.

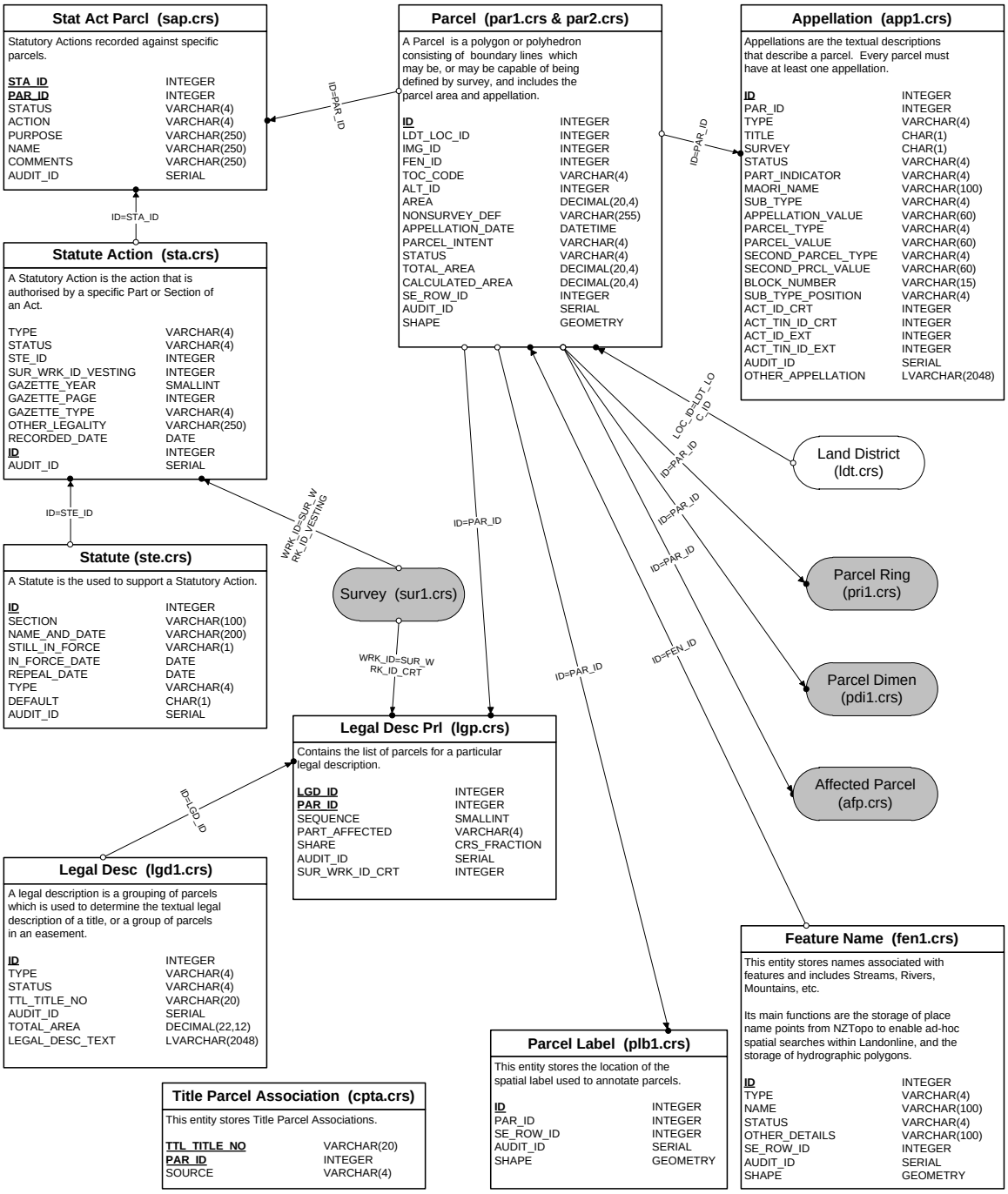
Road centreline and street address data is expected to be popular with a broad range of clients, however both road centrelines and street address should be used with a degree of caution. Road centrelines in Landonline are currently used as a mechanism to display authoritative road names along legal road corridoors and do not purport to represent what exists on the ground. Similarly Landonline's street addresses represent the principal addresses assigned by Territorial Authorities, and seldom include addresses assigned to individual flats, or addresses that differ from those which have been officially assigned.

For all named roads the RNA.location field contains the name or description, as previously stored for the equivalent Electoral Street & Places street.



3.1.3 Parcels Dataset

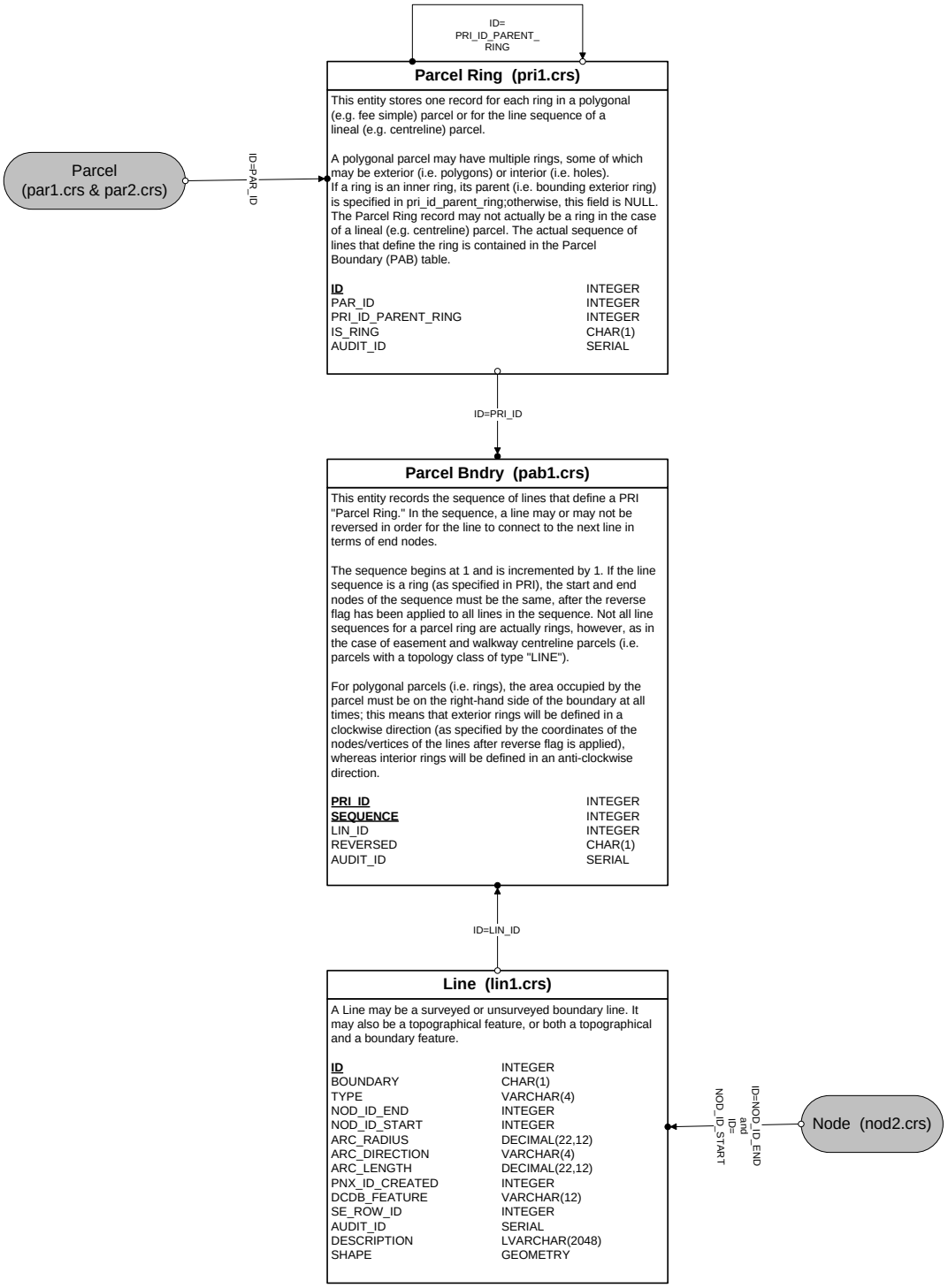
This dataset covers the parcel polygon and information directly related to it, e.g. appellation and legalisation. It is anticipated that this dataset, the system dataset, the road & address dataset, and possibly the Survey dataset will satisfy the majority of BDE clients.



Note: The rounded rectangles represent tables that exist in other BDE datasets. (those with no shading are in the System dataset)

3.1.4 Parcel Topology Dataset

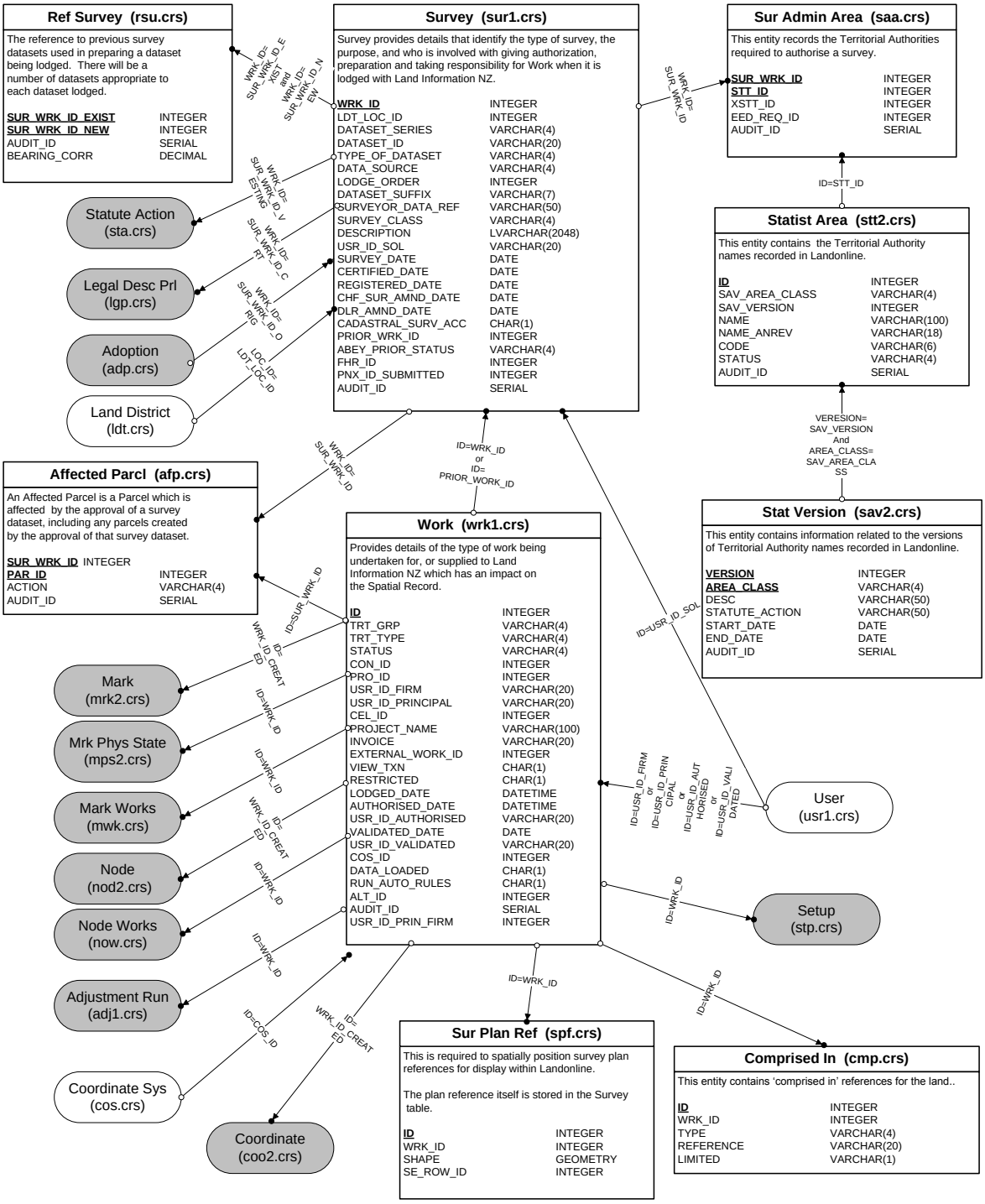
This table contains the intermediate "building blocks" for parcel polygons. Some Geographic Information System users may require this linear spatial data rather than the SHAPE's stored in the parcel table.



Note: The rounded rectangles represent tables that exist in other BDE datasets.

3.1.5 Survey Dataset

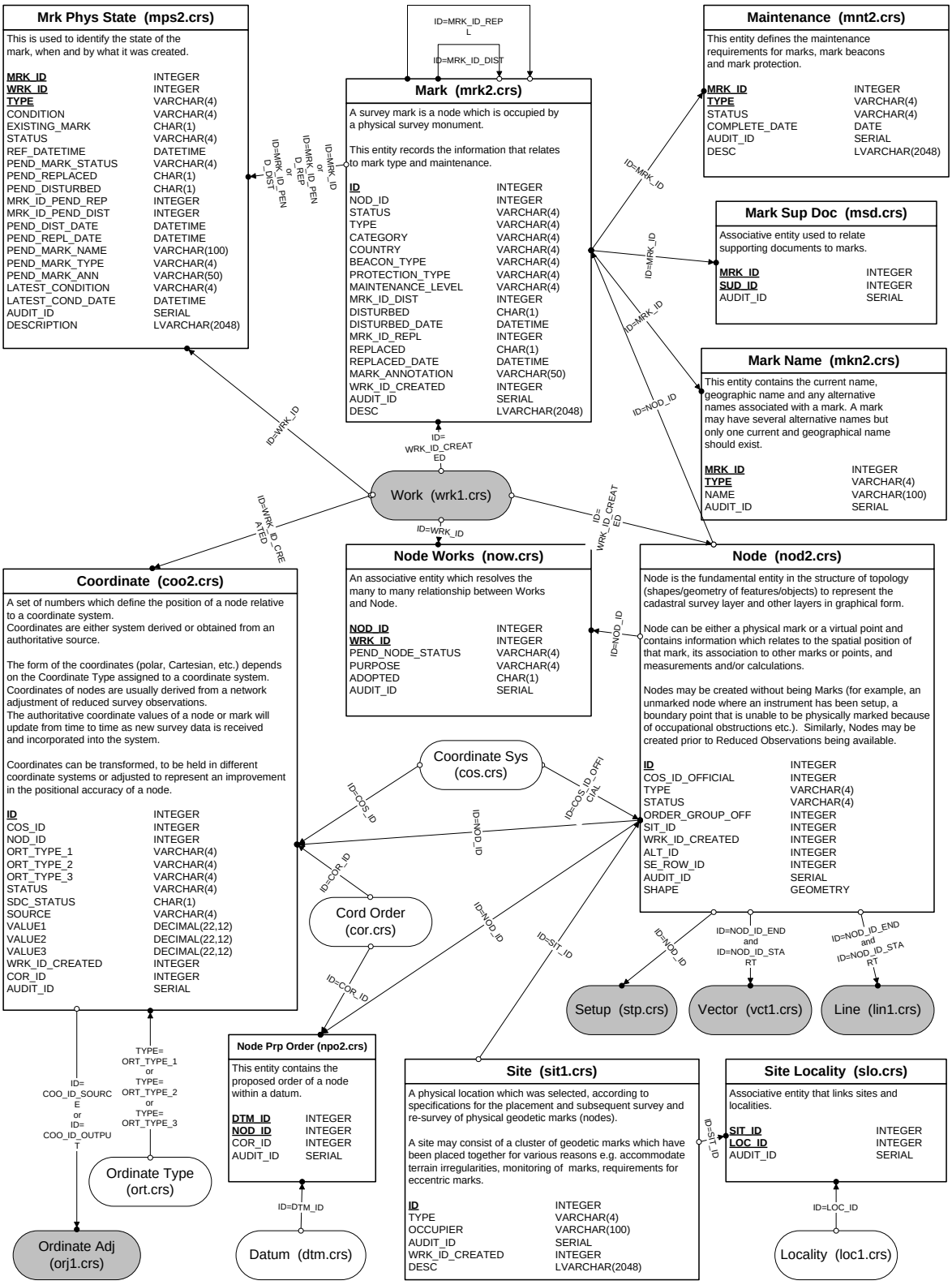
This dataset contains those tables related to survey documentation, including survey plan references.



Note: The rounded rectangles represent tables that exist in other BDE datasets. (those with no shading are in the System dataset)

3.1.6 Marks, Nodes & Coordinates Dataset

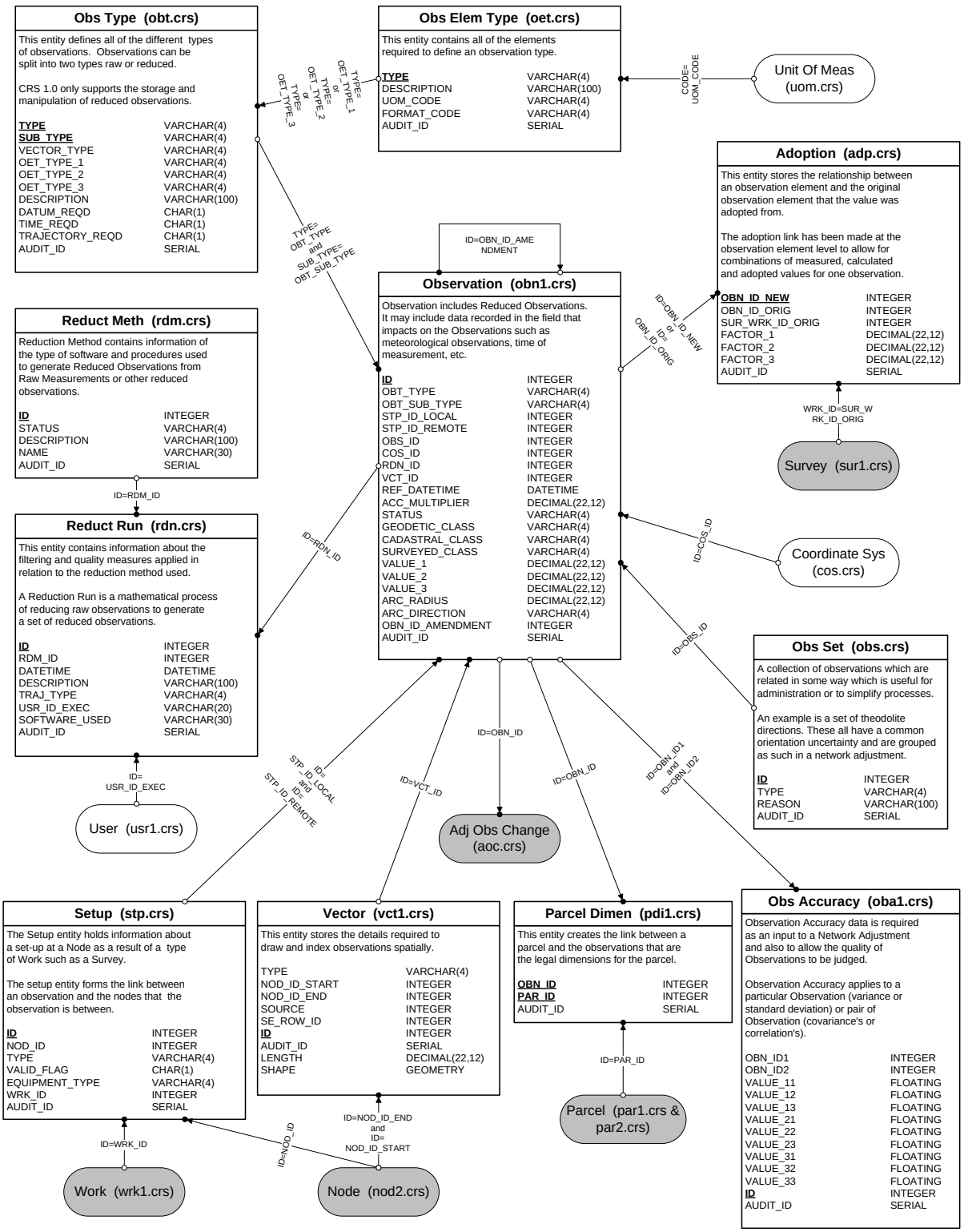
This dataset focuses on the tables directly related to nodes and survey monuments (marks).



Note: The rounded rectangles represent tables that exist in other BDE datasets. (those with no shading are in the System dataset)

3.1.7 Survey Observations Dataset

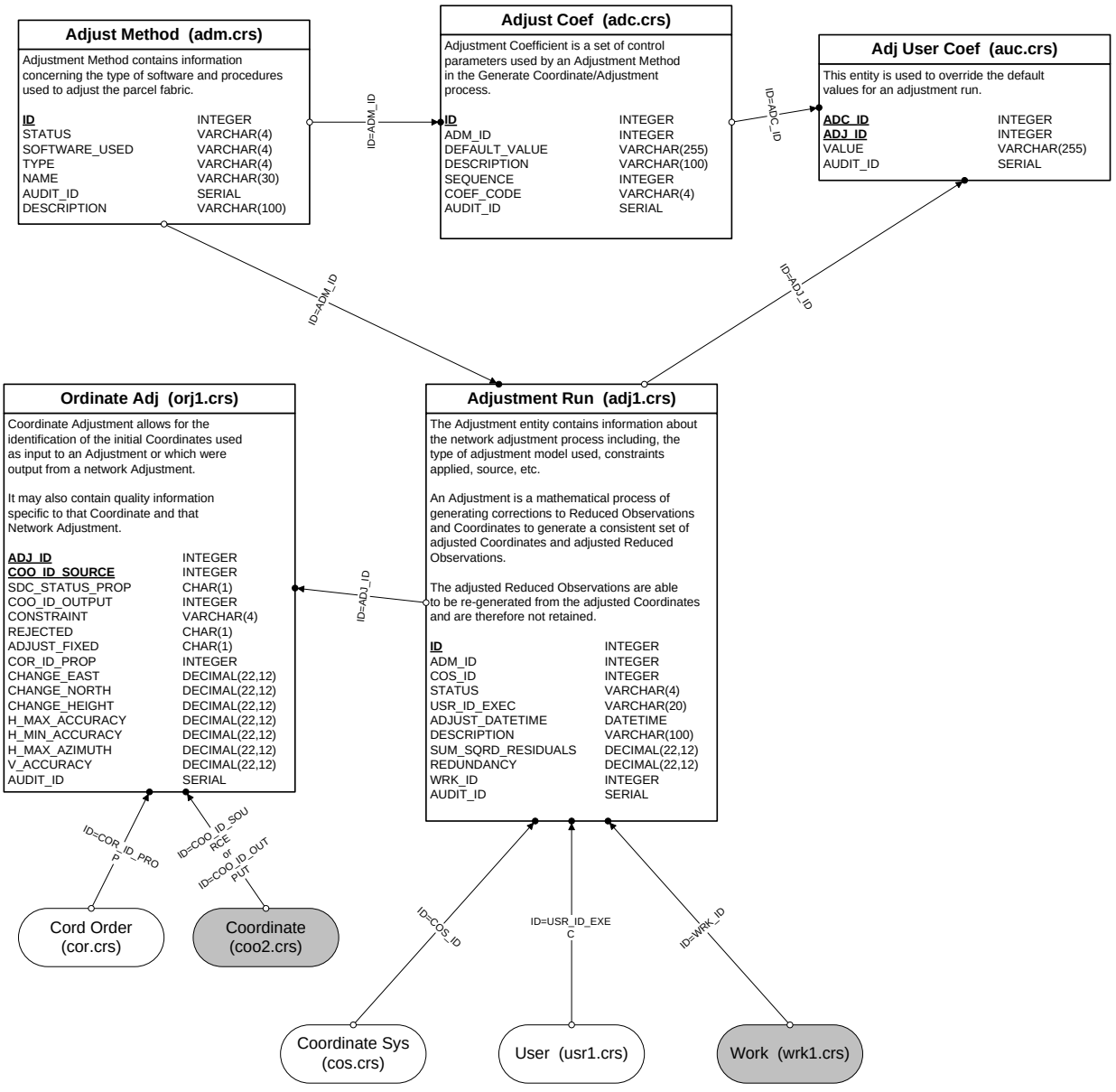
This dataset relates to the survey observations stored in Landonline.



Note: The rounded rectangles represent tables that exist in other BDE datasets. (those with no shading are in the System dataset)

3.1.8 Adjustments Dataset

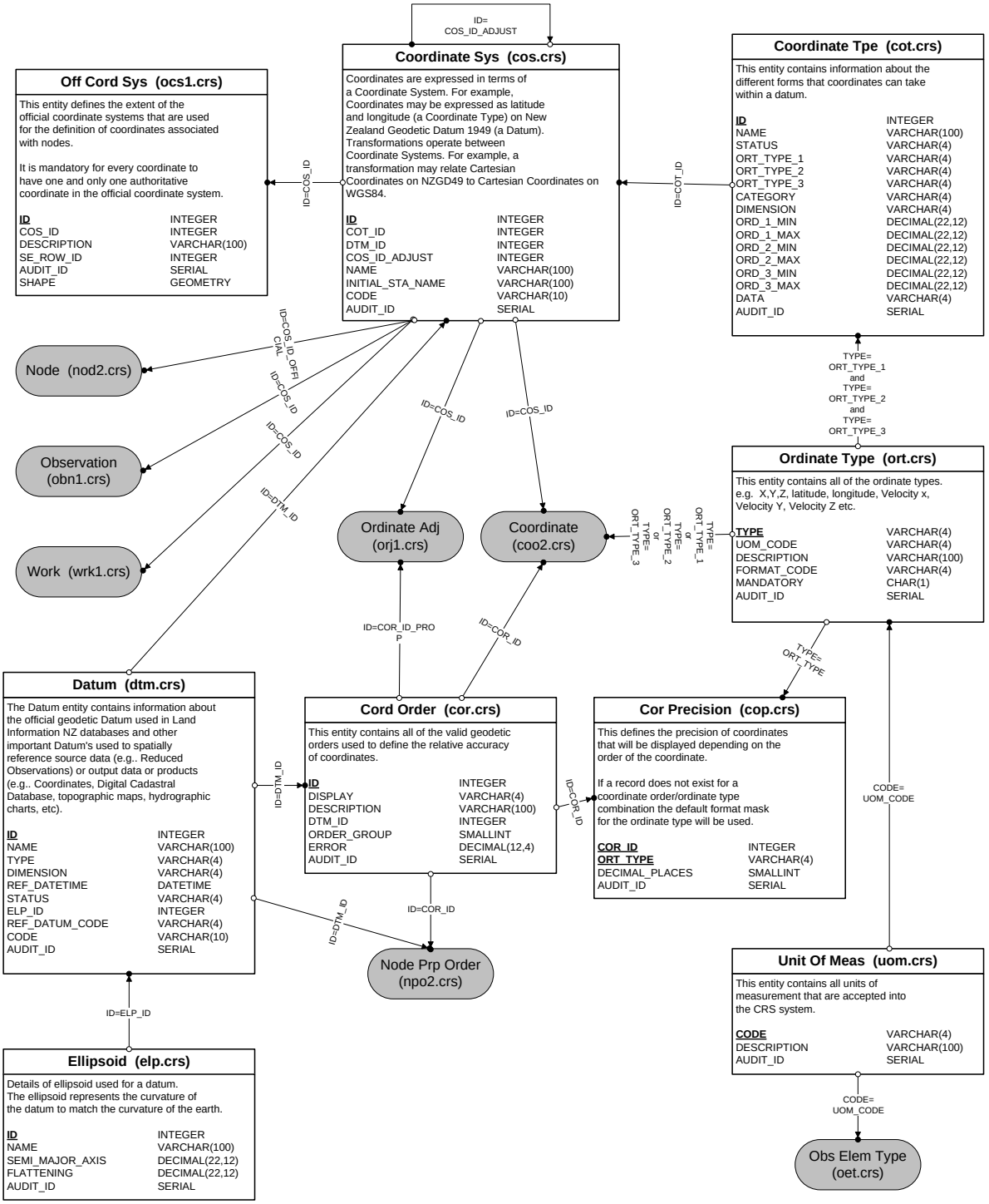
This dataset relates to survey & geodetic adjustments that have been carried out within Landonline.

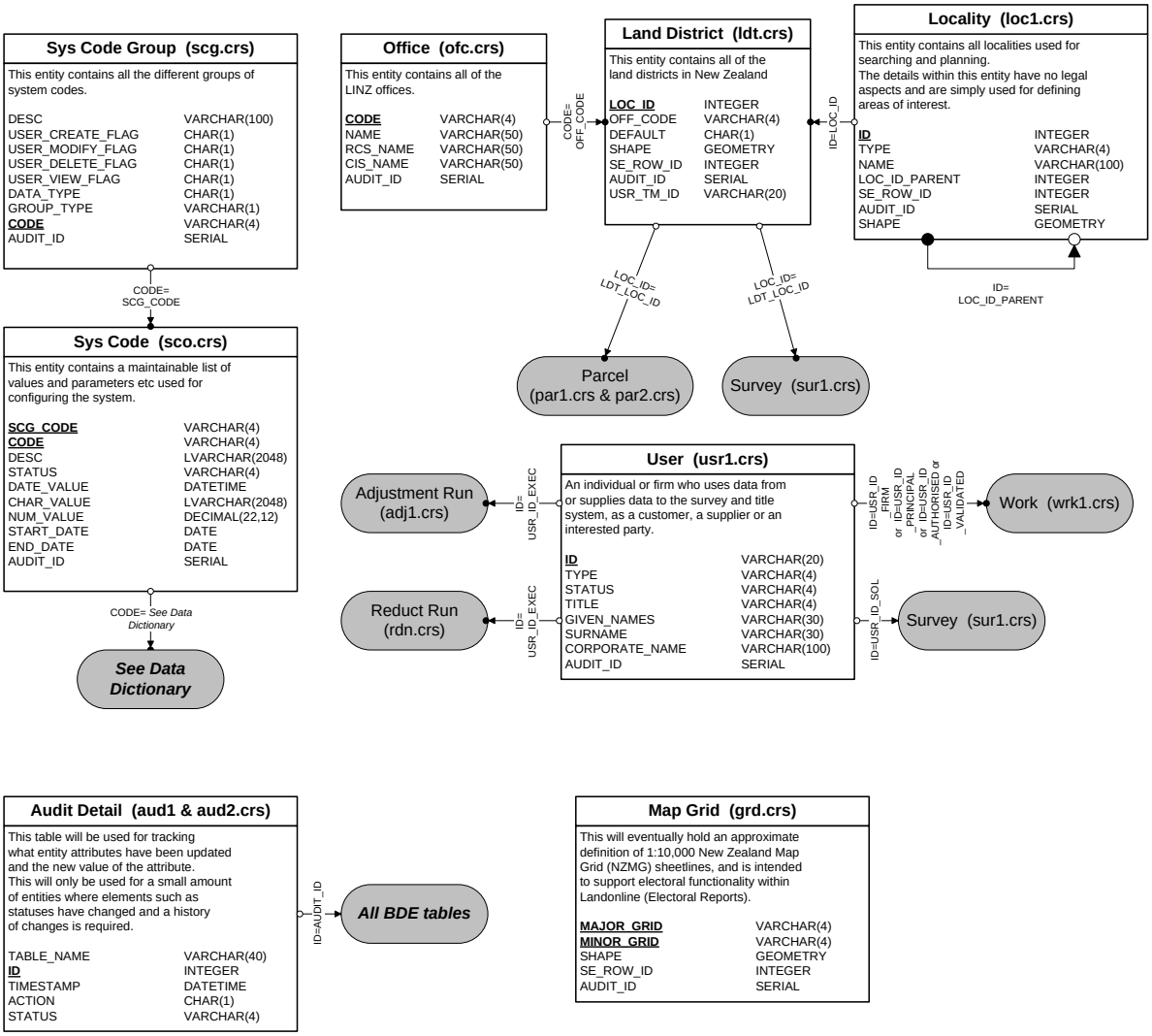


Note: The rounded rectangles represent tables that exist in other BDE datasets. (those with no shading are in the System dataset)

3.1.9 System / Shared Dataset

The tables included in this dataset are generally small in size and are typically those which are required by more than one dataset.

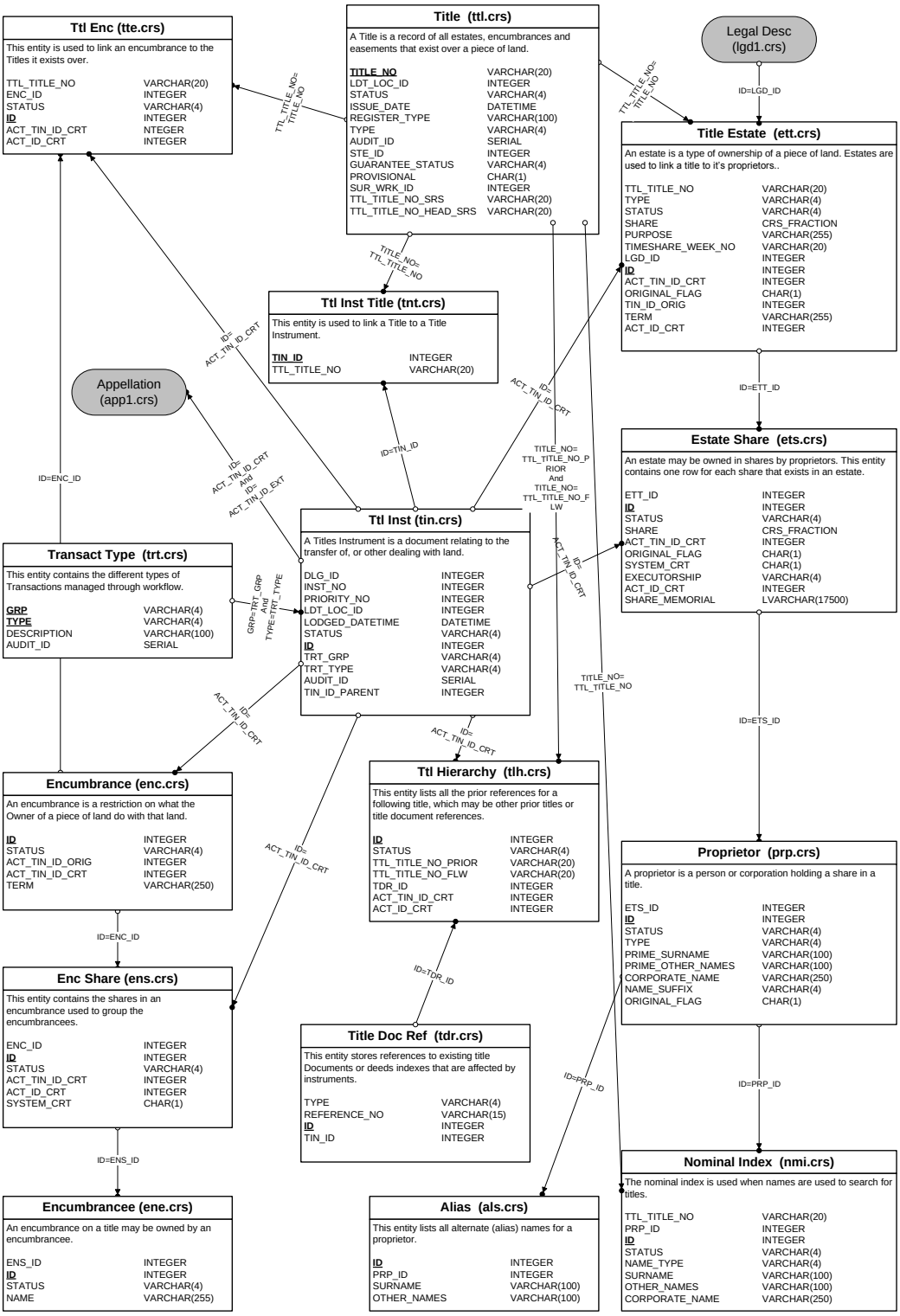




Note: The rounded rectangles represent tables that exist in other BDE datasets.

3.1.10 Title Dataset

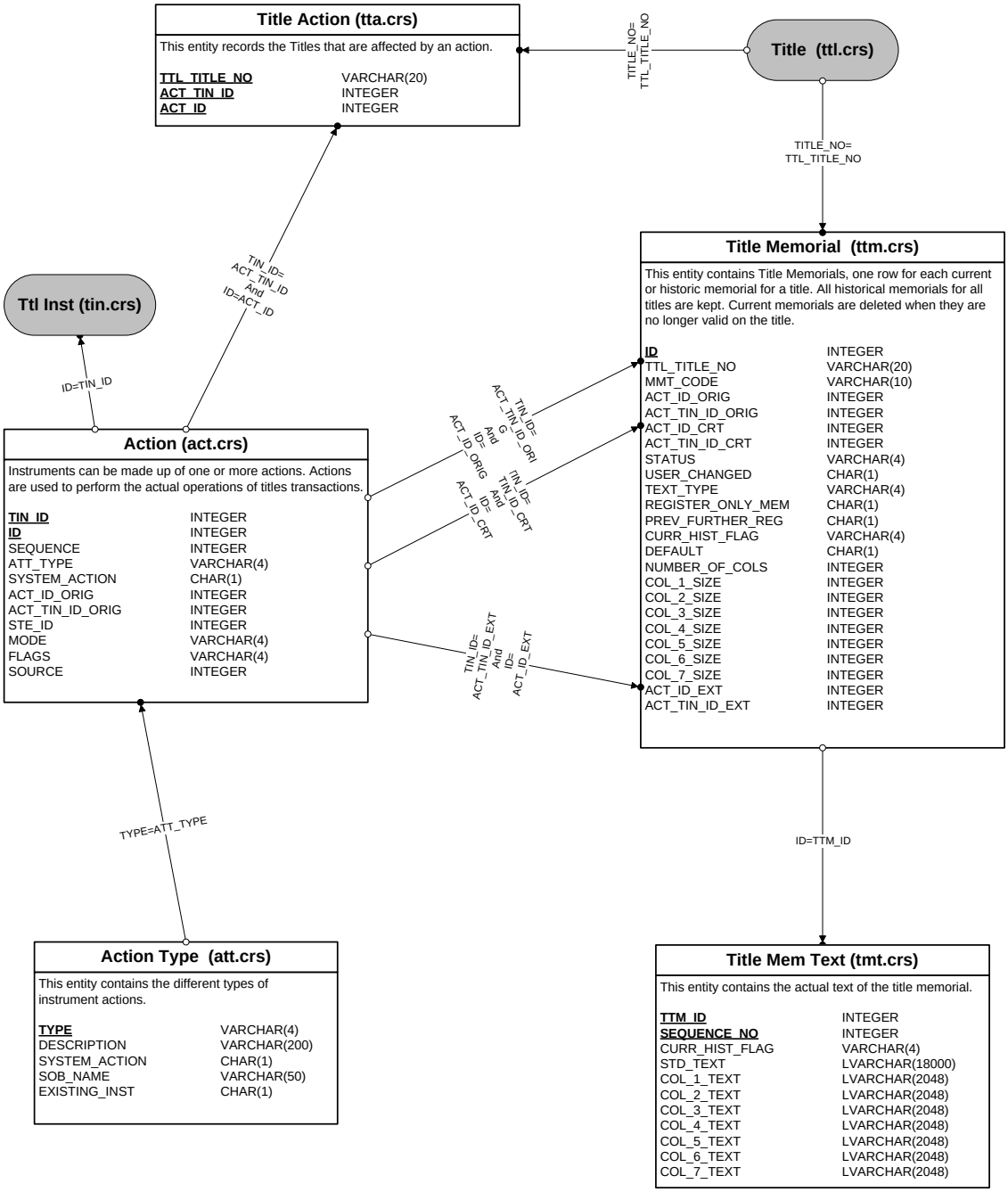
This dataset relates to the title information stored in Landonline. The dvl.crs unload lists those titles that are protected by a non-disclosure direction. It is recommended that the following wording be displayed for a title where the ownership would have been displayed : “Ownership details for this title are protected by a non-disclosure direction”.



Note: The rounded rectangles represent tables that exist in other BDE datasets.

3.1.11 Title Memorial Dataset

This dataset relates to the title memorial information stored in Landonline.



Note: The rounded rectangles represent tables that exist in other BDE datasets.

3.2 Landonline BDE - Data Dictionary

The "Landonline BDE - Data Dictionary" is a separate document that details all table views that are available for extract by BDE. Refer to the LINZ Internet site where it can be viewed or downloaded.

3.3 Spatial Data

All spatial geometries are supplied in terms of latitude and longitude on NZGD2000 (New Zealand Geodetic Datum 2000). No other coordinate system (eg. NZMG) will be supplied.

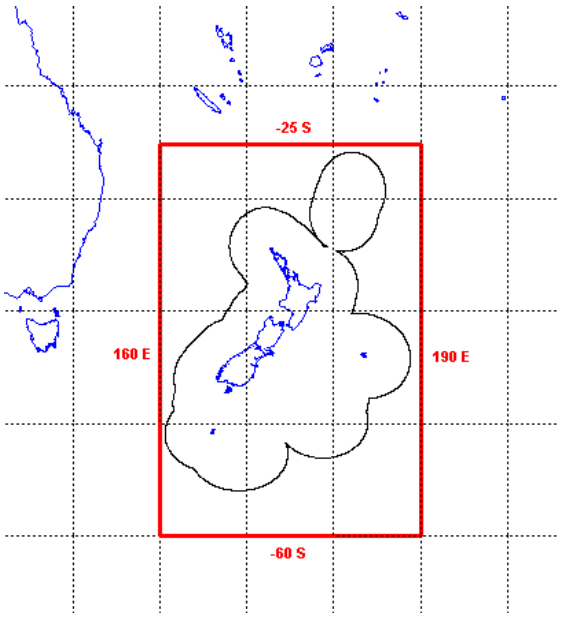
Most geometries have been built by Landonline from aspatial definitions of the data. This is based on a defined node coordinate and subsequent relationships (eg. through boundary lines, to parcels, etc.).

3.3.1 Spatial Reference ID (SRID)

Two spatial reference areas have been defined, one covering the world's entire extent (SRID=2), and a second covering New Zealand (SRID=1). This allows geometries (i.e. shapes) stored within SRID 1 to be stored at a higher resolution of 2mm.

Note however that this only applies to the **geometries** stored. All **values** stored in the aspatial area of Landonline will be stored to the same precision as captured from survey (or DCDB conversion).

The following diagram shows the spatial extents of SRID 1. The table shows the extent and graphical resolution for each spatial reference area;



SRID	Min X	Max X	Min Y	Max Y	Resolution
1	160	190	-60	-25	0.002 metres
2	-170	190	-100	+100	0.02 metres

In BDE files, the SRID of extracted geometries can be identified by the number which appears immediately before the geometry type

e.g. the following geometry is in terms of SRID 1;

```
1 POLYGON ((10.08986322 -45.91888600, 10.08981167 -45.92094295, 10.09507035 -45.91686293,
10.09365317 -45.91595345, 10.09309980 -45.91643815, 10.09063387 -45.91801058, 10.08986322 -
45.91888600))
```

Refer to Section 4.2.2 for further information concerning geometry representation.

3.3.2 Geometry Shift

Landonline has encountered technical problems with both Map Objects and ArcInfo when trying to project geometries either side of 180 degree longitude where they exist in the same SRID. Projecting around the 180 degree line is a recognised problem for most GIS platforms.

All spatial geometries stored in Landonline are stored with a 160 degree shift applied using the following algorithm;

```
set longitude = longitude - 160
if longitude < -180 then set longitude = longitude + 360
end if
```

e.g. 160°E becomes 0°E (160-160 = 0),
180°E becomes 20°E (190-160 = 20), and
190°E (or -170°E) becomes 30°E
(190-160 = 30) or (-170-160 = -330, -330+360 = 30)

In Landonline, this has the added advantage of displaying both New Zealand and the Chatham Islands in the centre of the world view. This also allows polygons that fall on or across the 180 degree line to be stored and manipulated by Landonline.

Clients can use a simple script to reverse the 160 degree shift. The conversion will get a little more complex when dealing with the geometry of features that cross the true 180 degree line. Most applications dealing with a -180 to +180 coordinate range will probably require polygons spanning the 180 line to be split. At this point no polygons in Landonline span the +180 degree line, but it is feasible that they may be created in the future, e.g. for meshblocks that encompass the New Zealand Exclusive Economic Zone (shown in the above diagram)

Note: Only spatial geometries have had this shift applied. All textual (non-geometry) coordinates stored in Landonline will have their correct real world values.

4. DATA SUPPLY FORMATS

4.1 File Format

Table views will be supplied as separate Unix ASCII files and compressed using the Unix gzip command.

4.1.1 Metadata

Metadata related to the exported table is contained in the header section of each extract file.

4.1.2 File Header Format

Each ASCII file representing a BDE table view will include a schema definition, along with other descriptive data as described in the following table:

Field	Format	Description/Rules	Example
File format version	Character (16)	Header file version information in the form HEDR xx.xx.xx	`HEDR 1.0.0`
Software	Character (100)	Software name and version info that produced the file	`SOFTWARE cbe_b30 V1.0.0`
Table Version	Character (20)	Table schema version	`SCHEMA V1.0`
User Name	Character (100)	Name of the individual that created the file.	`USER crsprd`
Time Span	DateTime to DateTime	START = datetime of last extract, END = datetime of this extract. For incremental extracts these dates reflect the applicable timespan, i.e. the file will contain new and changed data between the specified dates only.	`START 2001-01-06 12:42:20` `END 2001-02-17 11:49:52`
Table	Character (18)	"TABLE", Landonline source table name	`TABLE crs_adjust_coef`

Column(s)	Character (100)	"COLUMN", column name, type, and null fields.
		See data types below for a full description of available types.
			`COLUMN id integer NOT NULL`
			`COLUMN adm_id integer NOT NULL`
			`COLUMN default_value varchar NULL`
			etc.

Description	Character (256)	"DESC", General table description field	`DESC Non simplified appellation`
File size	Character (50)	"SIZE", size of this file in bytes.	`SIZE 4235753`
Beginning of data	Character(10)	Text string "{CRS-DATA}" indicating the end of the header, and the beginning of the data section.	`{CRS-DATA}`

Clients can confirm the integrity of each file supplied by comparing the uncompressed file size to the size recorded within the file header.

Note that where table views have been requested that contain no current data, the files supplied will consist only of the header information.

4.1.3 File Body Format

The body of the file follows immediately after the file header, and contains all the data of the specified Landonline table in an ASCII format. Data for each field will be terminated by a '|' (pipe) and there will generally be one record per line. If the '|' symbol is part of the Landonline data for extract, then it will be preceded by the '\' character. The symbol '|' has been chosen as being the least likely to exist within Landonline.

The intent is that each record is terminated by a Unix linefeed (ASCII code 10), however carriage returns and linefeeds sometimes occur within Landonline fields and are not filtered by the extract process. When these exceptions exist (eg in the lgd1.crs & app1.crs tables), these files may require pre-processing to load successfully.

If the file header is removed, the format will allow most database load facilities such as Informix dbload, Oracle SQL Loader, or Microsoft SQL Server to import the data.

4.1.4 Sample File Layout

```
HEDR    1.0.0
SOFTWARE cbe_b30 V1.0.0
SCHEMA          V1.0
USER    crsprd
START   2001-01-06 12:37:22
END     2001-02-17 11:49:52
SQL     SELECT * FROM viw_cbe_par1
TABLE   crs_parcel
COLUMN          id              integer NOT NULL
COLUMN          ldt_loc_id      integer NOT NULL
COLUMN          img_id          integer NULL
COLUMN          fen_id          integer NULL
COLUMN          toc_code        varchar NOT NULL
COLUMN          alt_id          integer NULL
COLUMN          area            decimal NULL
COLUMN          nonsurvey_def   varchar NULL
COLUMN          appellation_date datetime NULL
COLUMN          parcel_intent   varchar NOT NULL
COLUMN          status          varchar NOT NULL
COLUMN          total_area      decimal NULL
COLUMN          calculated_area decimal NULL
COLUMN          se_row_id       integer NULL
COLUMN          audit_id        serial NOT NULL
COLUMN          shape           st_geometry     NULL
DESC
SIZE    578566444
{CRS-DATA}
3000004|1011|||PRIM|||DCDB Document Id: CT 18C/120||DCDB|CURR|326110.0||3000004|1|1 POLYGON ((10.08986322 -45.91888600,
10.08981167 -45.92094295, 10.09507035 -45.91686293, 10.09365317 -45.91595345, 10.09309980 -45.91643815, 10.09063387 -
45.91801058, 10.08986322 -45.91888600))|
3000008|1011|||PRIM|||||DCDB|CURR|17320545.0||3000008|2|1 POLYGON ((9.58056935 -45.79915613, 9.58036890 -45.79926868,
9.58023758 -45.79936518, 9.58014245 -45.79946562, 9.57935250 -45.80014947, 9.57851672 -45.80106793, 9.57845165 -45.80122287,
9.57816800 -45.80448555, 9.57811385 -45.80461163, 9.57816093 -45.80458673, 9.57922292 -45.80432467, 9.57947893 -45.80421117,
9.57958870 -45.80412463, 9.57967223 -45.80392597, 9.57973187 -45.80371452, 9.57973290 -45.80371142, 9.57979013 -45.80354043,
9.57988795 -45.80299448, 9.57988985 -45.80298790, 9.57991798 -45.80289018, 9.57991758 -45.80180468, 9.57991970 -45.80179318,
9.58010118 -45.80080365, 9.58010732 -45.80078938, 9.58030822 -45.80032168, 9.58031717 -45.80030873, 9.58038487 -45.80021078,
9.58045578 -45.80006558, 9.58056935 -45.79915613))|
3000012|1011|||PRIM||1260.0|||DCDB|CURR|||3000012|3|1 MULTIPOLYGON (((10.97852097 -45.08208233, 10.97887587 -45.08224682,
10.97913818 -45.08196282, 10.97878328 -45.08179832, 10.97852097 -45.08208233)))|
3000016|1011|||PRIM||52199.0|||DCDB|CURR|||3000016|4|1 POLYGON ((10.09505287 -45.11478175, 10.09518427 -45.11737200,
10.09764092 -45.11736150, 10.09766538 -45.11521532, 10.09505287 -45.11478175))|
```

4.2 Data Formats

4.2.1 Data Types

The table below details what data types will be described in the header column record, and how that data will appear in the body of the file.

Type	Description	Example
CHAR(x)	The CHAR data type stores any sequence of letters, numbers and symbols. It can store single-byte and multi-byte characters, based on what the chosen locale supports. A character column has a maximum length of n bytes, where 1<= n <= 32,767. Because the length of the column is fixed, if the character string is shorter than n bytes, the string is padded with spaces to make up its length. If the character string is longer than n bytes, the string is truncated.	N
CRS_FRACTION	Data type extension developed specifically for Landonline to support fraction mathematics. The most applicable storage format in external systems would be as a string, ie: CHAR or VARCHAR equivalent.	1/1
DATE	Calendar date in the form yyyy-mm-dd	2000-02-23
DATETIME	Calendar date combined with time of date in the form yyyy-mm-dd hh:mm:ss where yyyy = 4 digit year, mm = 2 digit month, dd = 2 digit day, hh = 2 digit hours since midnight (24 hour clock), mm = 2 digit minutes, ss = 2 digit seconds	2000-02-23 12:03:32
DECIMAL(x,y)	Numbers with a pre-defined scale and precision.	0.019254934026
FLOAT	Double precision floating point number corresponding to the "double" data type used in the C programming language.	1.3130999999999999
ST_GEOMETRY	See Graphic Entity representation below	1 POLYGON ((10.08986322 -45.91888600, 10.08981167 -45.92094295, 10.09507035 -45.91686293, 10.09365317 -45.91595345, 10.09309980 -45.91643815, 10.09063387 -45.91801058, 10.08986322 -45.91888600))
INTEGER	Whole number from -2,147,483,647 to +2,147,483,647	36725001
LVARCHAR(x)	Long Varchar. Stores multibyte strings of letters, numbers, and symbols of varying length to a maximum of 32,739 bytes.	2 inch galv iron tube situated in mound of rocks on highest point of Mt Rakeahua. Quadripod in bad repair.
SERIAL	Sequential integer	123
SMALLINT	Number from -32,767 to +32,767	123
TEXT	Stores <u>any</u> kind of text data up to a maximum size of 2 Gb as a simple large object type (SLOB).	12-Feb-2001 Kerr, Peter \ a) The plan title should read Pts. Section 6S, not Pt. b) I have created links from this survey to the underlying DCDB at locations...*etc etc etc*
VARCHAR(x)	Stores multibyte strings of letters, numbers, and symbols of varying length to a maximum of 255 bytes.	FNAD Free net adjustment

4.2.2 Graphic Entity Representation

Geometry data from Landonline is supplied in Modified Well Known Text (MWKT) format. This is a variation of the Open GIS (OGIS) standard definition for Well Known Text (WKT) format.

The well known text representation for OGIS Geometry is made up of three basic components; *<geometry type> [coordinate type] [coordinate list]*

The *geometry type* is defined as either: point, linestring, polygon, multipoint, multilinestring, or multipolygon.

The *coordinate type* specifies whether or not the geometry has Z coordinates and/or measures. This argument is left blank if the geometry has neither. Landonline does not include any Z coordinates or measures.

The *coordinate list* defines the double precision vertices of the geometry. Coordinate lists are comma delimited (with each coordinate for a vertex being space delimited) and enclosed by parentheses. Geometries having multiple components require sets of parentheses to enclose each component part. If the geometry is empty, the EMPTY keyword replaces the coordinates.

The Modified Well Known Text Representation is created by simply removing the quotes and preceding the text description with Landonline's Spatial Reference ID (SRID) separated by a space, i.e.

Well Known Text Representation = **'LINESTRING (175.668126 -40.944098, 175.668242 -40.944138)'**

Modified Well Known Text Representation **= 1 LINESTRING (175.668126 -40.944098, 175.668242 -40.944138)**

The following are examples of Geometry types that may be found in Bulk Data Extract files;

Geometry type	Modified Well Known Text Representation
Point	1 POINT (10.09243073 -45.91778598)
Linestring	1 LINESTRING 8.04836990 -44.39847608, 8.04929795 -44.39867643, 8.05029350 -44.40027603, 8.05009282 -44.40093845)
Simple polygon Polygon with internal ring	1 POLYGON ((10.49079508 -45.90559523, 10.49118440 -45.90543437, 10.49109617 -45.90533160, 10.49102887 –45.90535932, 10.49070815 -45.90549138, 10.49079508 -45.90559523)) 1 POLYGON ((9.31149628 -45.48237713, 9.31095698 -45.48238873, 9.31106112 -45.48287398, 9.31083202 –45.48336223, 9.31170485 -45.48401058, 9.31371167 -45.48355873, 9.31211082 -45.48257198, 9.31162413 -45.48227198, 9.31149628 -45.48237713),(9.31162388 -45.48339993, 9.31215902 –45.48297213, 9.31276920 -45.48334558, 9.31223180 -45.48376873, 9.31162388 -45.48339993))
Multipolygon (1 polygon) Multipolygon (2 polygons)	1 MULTIPOLYGON (((10.97852097 -45.08208233, 10.97887587 -45.08224682, 10.97913818 -45.08196282, 10.97878328 -45.08179832, 10.97852097 -45.08208233))) 1 MULTIPOLYGON (((9.31891105 -44.76125265, 9.32010992 -44.76124942, 9.31997402 -44.76042467, 9.32001870 -44.75983933, 9.31997302 -44.75981192, 9.31900730 -44.75923200, 9.31757682 -44.75837295, 9.31756092 -44.75850670, 9.31827625 -44.75955302, 9.31834038 -44.76118975, 9.31891105 -44.76125265)),((9.32001870 -44.75983933, 9.32040742 -44.75934172, 9.32024085 -44.75927673, 9.32001870 -44.75983933)))

4.3 Media Supply Format

Files will be written to DVD, with one folder for each dataset being written. The DVD's used will be either 4.7GB DVD-R or 8.5GB DVD+R DL (Double Layer). All datasets, in a compressed form, is approximately 8.2 GB.

Notes;

- The files are compressed by the BDE batch process. They can be uncompressed using the Unix command "gunzip *.gz", or via a Windows compression utility such as Winzip.

- Where there is no actual data for the table view, only the header portion of the BDE file is generated. Files that actually contain data can be found by doing a "Find" (Windows) or a "grep" (Unix) for all files containing the pipe delimiter.

4.4 Schema Changes

Occasionally LINZ will need to change the Landonline schema or the BDE extract schema. In the latter case this may be to improve the data delivery to clients. These changes will not occur frequently and will be communicated to existing BDE clients as far in advance as is possible.

As the file format is self describing, new schema information will be automatically documented in the file header of the affected BDE file.

5. ESTIMATED EXTRACT VOLUMES

All files created by BDE are compressed to approximately 20% of their original size. Compression rates will vary depending on content.

Uncompressed data set sizes for all of NZ are (as at April 2014 unload);

Data Set	Compressed	Approx. Size
Roads & Addresses	68 Mb	267 Mb
Parcels	1.01 Gb	4.08 Gb
Parcel Topology	957 Mb	3.76 Gb
Survey Dataset	91 Mb	406 Mb
Marks, Nodes & Coordinates	921 Mb	4.12 Gb
Survey Observations	3.15 Gb	15.08 Gb
Adjustments	10 Mb	46 Mb
System / Shared	9 Mb	60 Mb
Title	866 Mb	3.95 Gb
Title Memorial	1.13 Mb	7.51 GB
All tables available via BDE	**8.21 Gb**	**39.28 Gb**
Daily incremental extracts for all BDE table views		Not available

Warning: Some tables may grow significantly as new data is captured into Landonline.

Due to files sizes, only the latest extract is likely to be available for supply.

6. PRICING AND LICENSING

The LINZ price for Bulk Data supplied from Landonline has been set at $54.18 inclusive of GST, per supply, irrespective of whether it contains a few megabytes of data or tens of gigabytes of data. This reflects LINZ's estimated cost of dissemination and is in line with Government policy of making Government-held information more readily available.

Costs and distribution mechanisms for incremental extracts have not yet been confirmed.

LINZ does not impose any copyright fees for the use of Landonline bulk data. Those purchasing this data from LINZ can on-sell it or share it with other parties without the requirement to pay copyright fees. However, while the Crown does not impose any charges for the use of its intellectual property, the Crown maintains its copyright and requires its acknowledgement. This is necessary to protect the authoritative nature of the Crown's data and to reduce the risk of any liability arising from its misuse.

Those purchasing Bulk Data from LINZ will be required to sign a licence agreement.