Appendix 3

# SPECIFICATIONS: CAPITAL MARKET TRADING DATA

## A3.1  Introduction

Put simply, Exchange Trade Data it is any time-ordered data generated by a trading exchange as a record of the trading activity on that exchange and as such it falls into the broad category of multivariate time series. Typically, the complete set of all exchange data is quite voluminous. For example, the ASX, about 1% of the global economy, generates over 100 Mb of ASCII data daily. The structure of trading data from other trading engines is similar, though not identical, with that described here.

A TRADE is the exchange registration of the movement of a specified number of a security from one member of the exchange to another, including possibly from a member to themselves. Members are typically brokers, and all TRADEs are novated[1] by the exchange. All off-market transactions between brokers have to be reported to the exchange within a specified time period, as determined by the rules of the exchange, usually on the same day of trade and usually after normal trading times.

A TRADE is an amalgam of individuals' trading activity. TRADEs in an exchange's Orderbook contains <brokerID> but not <traderID> tags. Under some circumstances <brokerID> is undisclosed, either for a period of time (say 3 days) or until the outstanding monetary value of the order is less than a fixed amount, as set by the exchange.

Trading engine data can be divided into a few broad categories:

- Market Depth Data. Records of Buy and Sell Bids [<Price>, <Volume>, <TraderID>] generated by potential traders "positioning" themselves around a price, frequently that of the last trade, with instructions which are executed under certain circumstances, such as when the last trade is at the same price as that of another order in the orderbook.

- Trade Data. For all instruments on a exchange, each Trade Data packet (called a Tick) contains information about the date-of-trade (Date). time-of-

---

[1] Novation is a process in law in which, with the consent of all parties, one contract is entered into in substitution for and supersession of another contract.

trade (Time), trading session (Day, Night), number of units traded (called the Volume), and the Price at which they are traded. In addition, the Broker ID code and for large (institutional) traders, the Trader-ID-Code is sometimes revealed.

- Open Interest. For derivative instruments such as Futures contracts and Options, Open Interest reflects the number of currently open contracts. Data for each trading session is available sometime after market close and before the next trading session commences.

- Pre– and End–Of–Trading Events that contains information about the Time (Duration), Price Statistics (Open, High, Low and Close) Period Volume.

## Intraday trading data: Cleaning procedure and metadata structure

The dataset used for the intraday trading data sonifications is of one financial quarter (sixty-five trading days) of a generic market's order-book day files. The data contains the multiplexed orderbook activity of approximately 3000 thousand trading instruments, listed firstly in action-time sequence at the resolution of one second and then by unique number according to their entry into the order-book. This data had to undergo extensive cleaning before being usable. The magnitude of the dataset (each day file was approximately 100 MB of ASCII data) demanded that the data–cleaning be undertaken algorithmically. The Metadata Specification here is from the cleaned dataset. Like the dataset itself it represents a generic structure as all references to a specific exchange, specific dates and specific securities have been omitted in–keeping with the confidentiality agreement between the parties signatory to the research proposal.

# Metadata: Description of order types

`<date code>.csv` files contain one message per line (separated by a newline character (`<\n>`)). Lines in each file are sorted in the order that the messages were sent from the trading engine. This should be in the same order as the <timestamp> attached to each message, but this has not been verified.

**Descriptions of the fields used in the ORDER TYPES follow them.**

**ENTER:**
`<timeStamp>,ENTER,<securityID>,Bid|Ask,<orderID>,<price>,`
`<disclosedVolume>,<disclosedValue>,<flags>,U,<brokerTrade`
`rRefs>,<\n>`
Description: A new order, or the untraded volume & undisclosed volume of an order carried forward from the last trading day (will have 00:00:00 timeStamp). Ignore the "U" field since this is for undisclosed volume (which is not provided in this dataset).

**DELET:**
`<timeStamp>,DELET,<securityID>,<orderID>,B|A,<\n>`
Description: Deletion of the untraded volume and untraded undisclosed volume of an order. i.e. an order cancellation. B for a bid order, A for an ask order.

**AMEND:**
`<timeStamp>,AMEND,<securityID>,Bid|Ask,<oldOrderID>,<newO`
`rderID>,X,<newPrice>,<newDisclosedVolume>,<newDisclosedVa`
`lue>,<flags>, <traderID>,<\n>`
Description: An amendment to an order. After an order is amended, it is referred to by its new ID, not its old ID. The traderID is that of the trader who amended the order.

**TRADE:**
<timeStamp>,TRADE,<securityID>,<tradeID>,<price>,<disclosedVolume>, <disclosedValue>,<flags>,<bidorderID>,<askorderID>,<\n>
Description: An on-market trade.

**OFFTR & CANTR:**
`<timeStamp>,OFFTR|CANTR,<securityID>,<offtrID>,<executeTi`
`meStamp>,`
`<price>,<disclosedVolume>,<disclosedValue>,<flags>,`
`<bidBrokerTraderRefs>,<askBrokerTraderRefs>,<\n>`
Description: An off-market trade, or a cancelled trade (the orders involved are not re-inserted into the order-book).

**FIELD:**
`<timeStamp>,FIELD,<securityID>,<Security><[1-5]>`
Description: An update to the group number of a security. The last field contains the group number.

**CONTL:**
`<timeStamp>,CONTL,<securityID><statusText>,<\n>`
Description: Pre-open/opening/open/suspend/close/adjust period/trading halt for a particular security or a particular group or all securities.

**FINIS:**
`<timeStamp>,FINIS,<\n>`
Description: Signifies that there are no more messages for this day (can probably be ignored – only useful in real-time data). `<timeStamp>` is usually `24:00:00.`

| Field | Description |
|---|---|
| `<timeStamp>` | Message time stamp reported by the trading engine. `HH:MM:SS` |
| `<securityID>` | Same as `<securityID>` in securityinfo.txt file. If this is "-" then there is a mapping problem in the data. In this case, we suggest you ignore these messages. |
| `<orderID>` | Order id. A natural number. Unique for the current day. |
| `<price>` | Price to exactly 3 decimal places, prefixed by '$'. |
| `<disclosedVolume>` | Disclosed volume. A natural number. |
| `<disclosedValue>` | Disclosed value, rounded to exactly 2 decimal places. |
| `<flags>` | Containing a list of flag codes. If there are no flags a double-quoted blank string is emplaced. Refer to market.cfg for descriptions of flags. |
| `<brokerTraderRefs>` | Broker reference and trader reference of the order (if they exist). ()\|(@<brokerID><" ">&<traderID>)\| (@<brokerID>)\|(&<traderID>) |
| `<newOrderID>` | New order id of an amended order. After an order is amended, it is referred to by its new id (not its old id). |
| `<newPrice>` | New price of an amended order, to exactly 3 decimal places, prefixed by '$'. |
| `<newDisclosedVolume>` | New disclosed volume of an amended order. A natural number. |
| `<newDisclosedValue>` | New disclosed value of an amended order, rounded to exactly 2 decimal places. |
| `<brokerID>` | Broker id. A natural number prefixed by '#'. |
| `<traderID>` | Trader id. A natural number prefixed by '#'. |
| `<tradeID>` | Trade id for an on-market trade. A natural number. |
| `<bidorderID>` | `<orderID>` of the bid order. |
| `<askorderID>` | `<orderID>` of the ask order. |
| `<offTrID>` | Trade id for an off-market trade. A natural number. |
| `<executeTimeStamp>` | TimeStamp for the time the trading engine reports an off-market trade as happening (will probably be different to `<timeStamp>`). `HH:MM:SS\|H:MM:SS` |
| `<bidBrokerTraderRefs>` | Broker reference and trader reference of the bid order (if they exist). B()\|B(@<brokerID><" ">&<traderID>) \|B(@<brokerID>)\|B(&<traderID>) |
| `<askBrokerTraderRefs>` | Broker reference and trader reference of the ask order (if they exist). A()\|A(@<brokerID><" ">&<traderID>) \|A(@<brokerID>) \|A(&<traderID>) |
| `<statusText>` | Control message state for this security, and whether it was initiated for a particular security (in this case no control type is stated), a security group or all securities. <status>[<" "><controlType>] |
| `<status>` | The status of the security., Can be either: Open, Closed or Suspended Opening: Not open for trading nor order entry, because opening-initiated trading is occurring. PreOpen: Open for order entry but not trading. Adjust-Period Trading-Halt |
| `<controlType>` | Whether the control was initiated for a particular security group ("`group`"), or for all securities ("`system`"). |