Irving Fisher Committee on
Central Bank Statistics

BIS

12th biennial IFC Conference: "Statistics and beyond: new data for decision making in central banks"

22-23 August 2024

How to use Trade Repository data on OTC derivatives for analysis – a practical framework[1]

Silvia Pezzini, Henry Chan and Wenjing Shi,
Hong Kong Monetary Authority

[1] This contribution was prepared for the conference. The views expressed in this publication are those of the authors and do not necessarily represent the official views of the Committee, its members, or the BIS.

How to use Trade Repository data on OTC derivatives for analysis – a practical framework

Silvia Pezzini, Henry Chan and Wenjing Shi [+]

31 October 2024

Abstract

With the increased speed of market movements, granular data on market activities collected by Trade Repositories (TR) have become a valuable source of information for market monitoring. This paper describes how the Hong Kong Monetary Authority (HKMA) developed a framework for using the TR data on over-the-counter (OTC) derivatives. The "SESA" framework follows four steps: (i) Standardisation of entity and stock names by building on Legal Entity Identifiers and other market identifiers; (ii) Enrichment of entity and trade information with market data; (iii) Scanning for anomalies; and (iv) Aggregation of exposures. The framework is shared as a reference for other central banks and regulators exploring the use of TR data for analysis.

Keywords: Trade Repository, OTC derivatives, data analysis

JEL classification: C55, G01, G15, G28

[+] This paper has been prepared by Silvia Pezzini, Henry Chan and Wenjing Shi from the HKMA.

Acknowledgements: We wish to thank Darryl Chan, Howard Lee, Clara Chan, Daryl Ho and Kevin Cheng for guidance in developing the framework and a number of HKMA colleagues who contributed over the years: Pansy Pang, Ceara Hui, Tak-Huen Chau, Michael Wong, Zhang Wu, Ben Lo, Iris Tang, Vincent Leung and several interns. This paper represents the views of the authors and not necessarily those of the HKMA.

Contents

1. Executive summary

The global financial crisis (GFC) and the collapse of Lehman Brothers triggered a global movement to improve transparency and reduce counterparty risk in the over-the-counter (OTC) derivative markets. In 2009, the G20 leaders agreed that all OTC derivatives should be reported to Trade Repositories (TR), aiming to improve transparency in the OTC derivative markets.

In Hong Kong, the Hong Kong Monetary Authority (HKMA) and the Securities and Futures Commission (SFC) have established a regulatory regime for the mandatory reporting of OTC derivative trades to the Hong Kong Trade Repository (HKTR). At the HKMA, we have developed a practical framework called "SESA" for using HKTR data for analysis.

The "SESA" framework consists of four steps, namely (i) Standardisation, (ii) Enrichment, (iii) Scanning for anomalies and (iv) Aggregation of exposures. With the "SESA" framework, we start with standardisation of the names of trade parties and equity derivative underlying stocks or indices, and enrich the standardised data with additional information via mapping and computation. We also conduct data scanning to identify anomalies for clarification with reporting entities and make necessary adjustments to facilitate timely analysis. Finally, we deal with challenges from the double-sided nature of TR reporting when aggregating the clean HKTR data to compute aggregate and entity-level exposures.

The application of the "SESA" framework ensures a high standard of data quality for OTC derivative data being used in financial stability monitoring and systemic risk assessment at the HKMA. Its modularity and scalability helps staff cope with the growing size and complexity of HKTR data. The "SESA" framework can serve as a reference for other central banks and regulators, and help promote the use of TR data in different jurisdictions.

2. Introduction

With the increased speed of market movements, trade-level data on OTC derivatives collected by TR have become a valuable source of information for regulators to monitor the build-up of systemic risk in financial market. A trade repository is a centralised database that maintains electronic records of OTC derivative trades across multiple asset classes. In response to the GFC, in 2009 the G20 leaders agreed that all OTC derivatives should be reported to TR, aiming to improve transparency in the OTC derivative markets.

Since then, the TR reporting obligation has been implemented across different jurisdictions. For example, in the United States, TR reporting is required by the Dodd-Frank Act and has been implemented by the Commodity Futures Trading Commission (CFTC) and the Securities and Exchange Commission (SEC). In the European Union, the TR reporting obligation is implemented through the European Market Infrastructure Regulation (EMIR). In Hong Kong, the HKMA and the SFC have established a regulatory regime for the mandatory reporting of OTC derivative trades to the HKTR.

According to the 2018 survey conducted by the Irving Fisher Committee on Central Bank Statistics (IFC) on central banks' access to and use of TR data [1], central banks have a strong interest in the use of TR data for market monitoring and policy purposes. At the HKMA, TR data have been used intensively for financial stability surveillance for their high frequency and level of detail which facilitate the timely monitoring of market activities [2, 3]. In addition, when TR data are combined with other datasets such as granular banking exposures, they enable a more comprehensive analysis of exposures and risks than previously possible.

Despite the benefits of using TR data, the size and complexity of TR data often pose challenges. Hence, we developed a practical framework called "SESA" for using TR data for analysis. The framework could serve as a reference for other central banks and regulators, and promote the use of TR data in different jurisdictions.

The "SESA" framework consists of four steps, namely (i) Standardisation, (ii) Enrichment, (iii) Scanning for anomalies and (iv) Aggregation of exposures. Through these four steps, the raw data are transformed into a dataset that is ready for analysis.

One use case of TR data is for non-bank financial institution (NBFI) surveillance. Following the Archegos' incident in 2021, the HKMA used HKTR data to develop an enhanced surveillance framework to identify potentially problematic NBFIs, in particular among hedge funds and family offices, for further monitoring [4]. To make this possible, we apply the "SESA" framework to transform the raw HKTR data into an enriched dataset ready for analysis.

The rest of the paper is structured as follows. Section 3 covers the development and scope of the HKTR. Section 4 highlights the benefits and common pain points of using TR data. Section 5 explains the "SESA" framework in detail. Section 6 presents the key utilities that transform TR data into data available for analysis. Section 7 covers the development of use cases. Section 8 discusses how Critical Data Element (CDE) implementation will help with TR data analysis. Section 9 concludes the paper.

3. Development and scope of the Hong Kong Trade Repository

Development of the HKTR

The GFC and the collapse of Lehman Brothers triggered a global movement to improve transparency and reduce counterparty risk in OTC derivative markets. In 2009, the G20 leaders agreed on comprehensive reform measures to mitigate systemic risks, improve transparency and prevent market abuse, including requiring that all OTC derivative trades be reported to TRs and all standardised OTC derivative trades be cleared at central counterparty clearing facilities [5]. The mandate for OTC derivative reporting has several benefits: it has promoted standardisation and consistency in the quality and availability of transactional data; it has provided financial authorities with rich data for market surveillance with the aim of maintaining financial system stability; and where statistics and analyses are shared with the public, they increase transparency in financial markets.

In August 2013, the HKMA introduced interim reporting requirements that required licensed banks (LBs) to report OTC derivative trades with other LBs to the HKTR. On 26 March 2014, the Legislative Council in Hong Kong enacted the Securities

and Futures (Amendment) Ordinance 2014 (Amendment Ordinance) [6]. The Amendment Ordinance is a regulatory framework for the OTC derivative market in Hong Kong. The regulatory framework was made operational with the Securities and Futures (OTC Derivatives Transactions – Reporting and Record-keeping Obligations) Rules (the "Reporting Rules") which were implemented in two phases.

The first phase commenced on 10 July 2015, introducing mandatory reporting of certain interest rate swaps and non-deliverable forwards. The interim reporting requirements for LBs were terminated when the Reporting Rules came into effect.

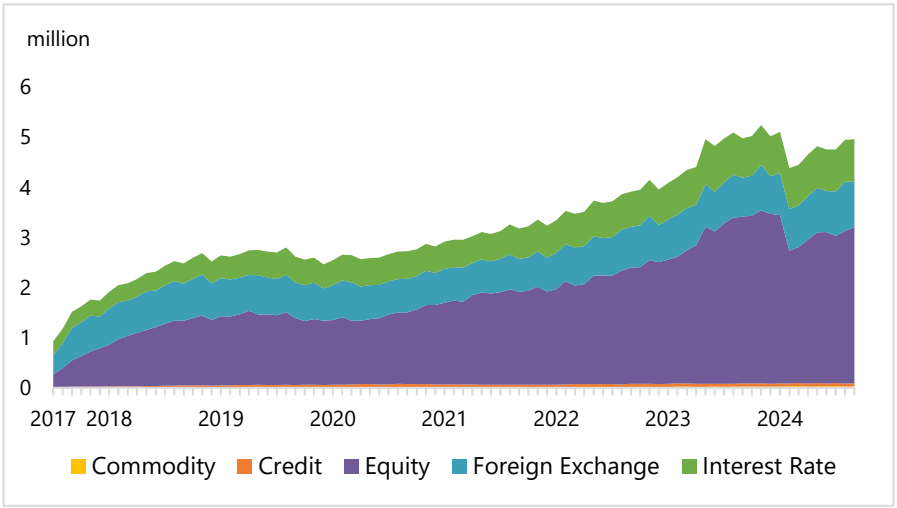
The expanded mandatory reporting, or second phase reporting, was implemented on 1 July 2017 and covers all OTC derivatives. Reporting is delivered with templates reflecting five asset classes (interest rate, foreign exchange, equity, credit and commodity), with space for exotic products that may not fit squarely in any template.

Under the new regulatory regime, the following entities are required to report their OTC derivative trades to the HKTR: Authorised Institutions (AIs) and Approved Money Brokers (AMBs) licensed and regulated by the HKMA under the Banking Ordinance, Licensed Corporations (LCs), recognised clearing houses (RCHs) and automated trading services – central counterparties (ATS-CCP) licensed and regulated by the SFC under the SFO [7].

Growing volume of trades

The volume of TR data has increased over time. In Graph 1, the number of outstanding OTC derivative trades increased from 0.9 million since the launch of second phase reporting in July 2017 to 5.0 million in September 2024. Across all five asset classes, as of September 2024, OTC derivatives in equity accounted for the largest share of number of trades (63%), followed by trades in foreign exchange (18%) and interest rate (17%). OTC derivative trades in commodity and credit had a relatively low share of trade count (2%).

Trend of Trade Count in the HKTR by Asset Class — Graph 1



Scope

HKTR data provide comprehensive information about OTC derivative trades, covering more than 250 key data attributes, including trade party information, key economic terms (e.g. gross notional amount, product, maturity, currency, strike price), booking location, trading desk, daily mark-to-market valuation and other data elements. Table 1 shows the list of products supported by the HKTR across the five asset classes. For more details on reporting template layouts, please refer to the appendixes of the HKTR Administration and Interface Development Guide (AIDG) [8].

Summary of Products supported by the HKTR Table 1

	Interest Rate	**Foreign Exchange**	**Equity**	**Credit**	**Commodity**
Base Product	Interest Rate Swap	Forward	Option	Single Name	Swap
	Cross Currency Swap	Non-deliverable Forward	Swap	Index Tranche	Option
	Cap Floor	Vanilla Option	Other	Index	Forward
	Option	Non-deliverable Option		Other	Other
	Forward Rate Agreement	Swap			
	Other	Other			

In the HKTR, a reporting entity is required to report an OTC derivative trade to the HKTR if the transaction is "booked in Hong Kong" (e.g. recorded on a bank's balance sheet in Hong Kong). Besides the trades booked in Hong Kong, trades are also reportable if "conducted" in Hong Kong on behalf of the head office, an affiliate or a branch/office outside Hong Kong (e.g. traded by the Hong Kong trading desk but not booked on the Hong Kong's balance sheet).

Looking ahead: CDE implementation

The HKMA and the SFC intend to implement the new derivative reporting regime for CDE in September 2025 in accordance with the International Organisation for Standardisation (ISO) standard. CDE reporting adopts a scope that maximises the data elements common to other jurisdictions and minimises those unique to Hong Kong. Hence, some HKTR data fields may change due to the CDE implementation. The implementation of CDE will enable Hong Kong's OTC derivative reporting regime to better integrate with international reporting frameworks. More details can be found in section 8 of this paper.

4. Benefits and pain points of using TR data

Benefits

TR data have been used intensively by the HKMA for financial stability surveillance because of its three features and corresponding benefits: timeliness, detail and standardisation.

First, TR data are reported at high-frequency, thus facilitating timely analysis. Transaction-level TR data are collected on a daily basis. With the use of automation and digitalised workflows, TR data are made available to users shortly after day close. While reporting entities are allowed to report the trades within T+2 days, over 90% of the trades in the HKTR are reported within T or T+1 day.

Second, the over 250 key data attributes reported for each trade allow regulators to slice and dice the data according to different dimensions, including by counterparties, types of activities or timing of actions. Thanks to the granularity of TR data, regulators can deploy the same data for different use cases, from mapping interconnectedness between market participants, to outlining the market dynamics surrounding stress events, to identifying channels of transmission of liquidity shocks in the financial system. In addition, counterparty information helps regulators identify specific group of entities of interest, understand their trading behaviour and monitor their activities and exposure concentration.

Third, the standardisation of TR data enables combining information on the OTC derivative market with information from other markets, leading to improvements in developing a comprehensive analysis of risks. The implementation of the mandatory use of Legal Entity Identifiers (LEI) in April 2019 for all reporting entities allows precise matching of entities observed across different datasets, enabling a more comprehensive analysis of exposures across different markets.

Pain points

Despite the benefits of using TR data, the size and complexity of TR data often pose challenges.

First, while standardisation is advanced, and will be further progressed with the move to the CDE, it is not complete, which creates pain points when aggregating exposures related to specific dimensions. For entities, the use of LEIs as universally unique identifiers is not complete for counterparties in TR data. Since April 2019, reporting entities are expected to establish a process to request the LEIs of their counterparties, but when counterparties do not have an LEI, reporting entities are allowed to identify counterparties using various third-party assigned identifiers supported by the HKTR with a specified order of priority (see Table 2). The outcome is that a counterparty can be represented by different types of identifiers by different reporting entities.

Priority Order of Identifiers for Counterparties not on the Reporting Side Table 2

Priority Order	Third-Party Assigned Identifiers
1st	Global LEI
2nd	SWIFT BIC (Business Identifier Code) issued by SWIFT under ISO 9362
3rd	Number of the Certificate of Incorporation ("CI") (for locally incorporated companies) or Certificate of Registration ("CR") (for companies incorporated overseas) issued by the Companies Registry of Hong Kong
4th	Business Registration Number ("BRN") issued by the Inland Revenue Department of Hong Kong

Another similar pain point is that, for equity derivatives, a range of identifiers is allowed for underlying stocks or indices. Allowed identifiers are the International Securities Identification Number (ISIN), or if ISIN is not available, any third-party assigned identifiers supported by the HKTR (see Table 3). The variety of identifiers creates difficulty when aggregating exposures related to specific stocks or indices.

List of Identifiers for Equity Derivative Underlying Table 3

Third-party assigned identifiers	Description
ISIN	International Securities Identification Number (ISIN) is a 12-character alphanumeric code that uniquely identifies a security globally for the purposes of facilitating clearing, reporting and settling of trades. Its structure is defined in ISO 6166.
FIGI	Financial Instrument Global Identifier (FIGI) is a 12-character alphanumeric code serving as a unique identifier of financial instruments.
RIC	Refinitiv Instrument Code (RIC) is a ticker-like code used by Refinitiv to identify financial instruments and indices.
SEDOL	Stock Exchange Daily Official List (SEDOL) code is a unique 7-character alphanumeric identifier assigned to securities traded on the London Stock Exchange and other smaller exchanges in the UK.
CUSIP	Committee on Uniform Security Identification Procedures (CUSIP) number is a unique 9-digit identification number assigned to financial securities in the US and Canada.
Valoren	A valoren number is an identification number of 6 to 9 digits assigned to securities in Switzerland.
SICC	A Securities Identification Code Committee (SICC) code is a 4-character alphanumeric code for listed shares and securities in Japan.
Local stock exchange ticker	A ticker is a string of characters or numbers to uniquely identify financial securities (e.g. stocks, indices) in local stock exchange.

Second, the limited standardised information on how entities with different LEIs belong to the same financial group introduces a pain point when computing the consolidated exposure of a parent group where different subsidiaries are separate

legal entities and are reported to the HKTR using different LEIs. This is the case for banking groups, investment funds and financial groups more broadly. Level 2 data under the LEI framework which show the direct and ultimate parents of legal entities are meant to tackle this shortcoming and are in progress [9].

Third, reporting errors may occasionally create false alarms of impending risk when trade reports contain exceptionally large notional amounts and out-of-the-ordinary strike prices.

Fourth, double-sided reporting needs to be handled carefully to avoid double-counting of trades. Double-sided reporting takes place when both parties of a trade need to report the trade to the TR if they fall under the reporting requirement. In this case, there are two reporting records for one single trade, which may lead to double-counting when aggregating either total positions or the positions of individual entities which appear across the TR data both as reporting entities and as counterparties.

5. HKMA's "SESA" framework

To tackle the mentioned pain points systematically, we developed a four-step "SESA" framework to clean and enrich the TR data for analysis (see Graph 2):

Step 1: We standardise the entity names and the names of equity derivative underlying stocks or indices, which are both reported using different types of identifiers in the TR data.

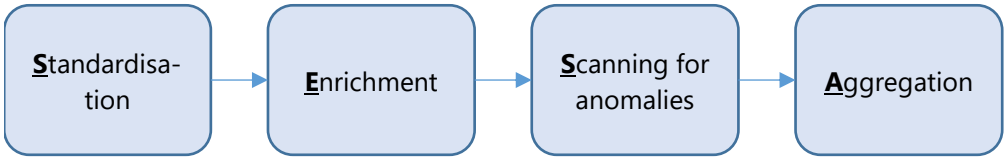
Step 2: We enrich the TR data with additional information via mapping tables or by computing useful statistics. The enrichment takes place for some common fields applicable to all trades (e.g. counterparty's sector) as well as for fields in specific products or asset classes (e.g. currency pair for FX derivatives).

Step 3: We conduct scanning to identify anomalies and potential data quality issues (e.g. exceptionally large notional amounts and out-of-the-ordinary strike prices) and make necessary adjustments for analysis (e.g. outlier removal pending rectification by reporting entities).

Step 4: With the clean and enriched TR data, we perform data aggregation while taking into account double-sided reporting when an entity can act as a reporting entity or a counterparty in different trades.

We explain each step in detail in the following sub-sections.

Workflow of the "SESA" framework — Graph 2



5.1. Step 1: Standardisation

In the first step of the SESA framework (Standardisation), we standardise the entity names and the names of equity derivative underlying stocks or indices using in-house databases.

Entity names

Standardising entity names provides two key benefits. First, it enables the computation of consolidated exposure of an entity, its parent or a fund manager by aggregating individual trades under the same entity. In addition, standardisation enables the combination of TR data with other data sources (e.g. granular banking data on loans, market data, news data) to analyse different types of exposures of an entity from different perspectives.

We developed an in-house entity database with the mapping of around 29,000 entity identifiers by leveraging (i) public data sources such as Global Legal Entity Identifier Foundation (GLEIF) database and SEC filings (Form D and Form ADV) (ii) commercial databases such as Eurekahedge, S&P Capital IQ, and (iii) company websites. The entity database is intended to tackle two key challenges in standardisation.

One challenge is about an entity being represented by different types of identifiers. Table 4 provides a sample of the entity database. It shows that ABC Capital has different identifiers e.g. LEI, CICR, BRN, which can be reported to the HKTR. With the entity database, all these identifiers will be mapped to the same entity name and such standardisation enables aggregation of trades related to the same entity.

Another challenge is about a fund manager conducting trades via multiple funds or entities and each of them has its own LEI and name, which makes it difficult to aggregate the trades related to the same fund manager. In Table 4, XYZ Capital, as a parent group, manages three individual funds with their own LEIs. With the entity database, all the trades reported under these three LEIs can be grouped under the same parent for analysis. The Bank of England (BoE) and the De Nederlandsche Bank (DNB) highlighted this limitation, and the BoE used alternative solutions to tackle this challenge by developing an algorithm which uses information on intra-group trades reported in the data and group firms by their names [10, 11].

In addition to standardising entity names by mapping different identifiers, the entity database helps us enrich the data with additional entity-level and parent-level information to facilitate the grouping of trades under the same fund manager or parent group. Being able to consolidate the positions under the parent group is important when computing the total gross OTC derivative exposures of the parent group in order to assess the size of its portfolio and potential impact on the market. One may also note that different subsidiaries under the parent group can have offsetting (long and short) exposures. Consolidating positions under the parent group also allows computation of net exposures.

Sample of the HKMA's In-house Entity Database Table 4

ID Type	LEI	Other Identifier	Entity Name	Entity Sector	Entity SubSector	Parent Name	Parent Sector	Parent Country	Entity Country
LEI	EHWNDI2G5WXXXXXXXXXX	NA	ABC CAPITAL MASTER FUND LP	Non-bank financial	Hedge Fund	ABC CAPITAL MANAGEMENT LP	Non-bank financial	USA	USA
CICR	NA	C12345	ABC CAPITAL MASTER FUND LP	Non-bank financial	Hedge Fund	ABC CAPITAL MANAGEMENT LP	Non-bank financial	USA	USA
BRN	NA	B45678	ABC CAPITAL MASTER FUND LP	Non-bank financial	Hedge Fund	ABC CAPITAL MANAGEMENT LP	Non-bank financial	USA	USA
LEI	5493000ALK1111111111	NA	XYZ CHINA OPPORTUNITIES FUND LP	Non-bank financial	Hedge Fund	XYZ CAPITAL MANAGEMENT LP	Non-bank financial	USA	USA
LEI	54930058A22222222222	NA	XYZ HK OPPORTUNITIES FUND LP	Non-bank financial	Hedge Fund	XYZ CAPITAL MANAGEMENT LP	Non-bank financial	USA	CYM
LEI	I4BOGE0V4Y3333333333	NA	XYZ JAPAN OPPORTUNITIES FUND LP	Non-bank financial	Hedge Fund	XYZ CAPITAL MANAGEMENT LP	Non-bank financial	USA	CYM

Note: Only selected fields are shown.

Equity derivative underlying names

The standardisation of equity derivative underlying names is crucial for analysing equity derivative portfolios of market participants. Similar to the standardisation of entity names, we developed an equity security database using APIs and web scraping for the purpose of standardising equity derivative underlying names and identifying country of exchange by mapping relevant identifiers.

The in-house database supports the mapping of different types of equity identifiers to standardised equity security names. Common identifiers include International Securities Identification Number (ISIN), Financial Instrument Global Identifier (FIGI), Stock Exchange Daily Official List (SEDOL) code, Committee on Uniform Security Identification Procedures (CUSIP) number, Refinitiv Instrument Code (RIC), Valoren number, Securities Identification Code Committee (SICC) code and local stock exchange ticker. The OpenFIGI API is used to map third-party identifiers to FIGIs.

Section 6 provides more detail on the in-house entity database and equity database.

5.2. Step 2: Enrichment

After standardisation, we proceed to the second step of the framework, Enrichment, through which we supplement the HKTR data with more detailed information that is useful for analysis. In this step, we make use of the in-house entity database and some external reference data such as FX rates and country codes to conduct two levels of enrichment.

General enrichment

The general enrichment is applicable for all asset classes.

First, based on the in-house entity database, HKTR data are enriched with entity-level information (e.g. sector and country of incorporation) as well as parent-level information (e.g. name of parent group, parent sector, parent country of incorporation) to support further slicing and dicing. Take entity sector as an example. This allows us to analyse the positions of a particular sector of interest.

Second, we also derive new data fields based on existing ones to facilitate analysis. For example,

- USD equivalent of the notional amount in a different currency by applying corresponding foreign exchange rates for comparison;
- Original and remaining maturity in days and by buckets (e.g. less than one month, one to three months) using the fields on start, end and position dates of the trade;
- Intra-group indicator if two sides of the trade belong to the same parent group;
- Full name of the location of the branch where the trade is booked and that of the trading desk entering into the trade using ISO country codes;
- Role of the reporting entity and the counterparty across different products as to whether they are option buyers/sellers, base currency payers/receivers, equity return payers/receivers etc.

Third, by looking at multiple data fields, we determine whether an entity's trade is long or short, and hence its gross and net positions. This is particularly useful if we need to assess whether a particular sector is building up substantial short position in certain currencies or concentrated positions in certain underlying stocks or indices. For example, in the case of USD/HKD options, we can define the position of an entity as a long USD/short HKD position if the call currency of the option is USD and the entity's role is option buyer.

Asset-class specific enrichment

On top of the enrichment applicable for all asset classes, there is some further enrichment specific to a particular asset class.

In the foreign exchange asset class, we clean and extract currency pairs from trades using the currency standard of the International Swaps and Derivatives Association (ISDA), making it feasible to filter out particular currency pairs (e.g. USD/HKD, USD/CNY) for market monitoring and topical research analysis.

In the equity asset class, in order to analyse the positions related to the underlying securities in a particular jurisdiction, we derive the underlying country of exchange based on the identifier of exchange and ISO Market Identifier Codes (MIC). Also, for some exotic product trades without an explicit notional amount, we derive the notional amount using a number of fields such as strike price, number of options and number of units of shares per option comprised in the transaction.

In the commodity asset class, HKTR data record the unit price and the number of units instead of the notional amount. Therefore, we enrich the commodity data by computing the notional amount based on product features (e.g. by multiplying the unit price by the number of tons).

5.3. Step 3: Scanning for anomalies

With standardised and enriched HKTR data, we carry out scanning for anomalies. To reduce the occurrence of false alarms and ensure the quality of analysis, we have developed scripts to scan trades, spot anomalies and exclude relevant trades during initial analysis while clarifying with reporting entities regarding potential reporting errors.

In cleansing HKTR data, we have two main criteria for identifying anomalies.

First, we identify trades with exceptionally large notional amounts using a range of cut-offs depending on asset classes. This sometimes helps spot reporting errors in the currency fields.

Second, we check whether some out-of-the-ordinary strike prices are due to reporting errors. For example, in the foreign exchange asset class, some reporting entities may report 780 as the strike price for USD/HKD options, which looks abnormal given the range of USD/HKD FX rate being 7.75 to 7.85.

To deal with data quality issues, we have a set of well-established follow-up actions. A designated operational team at the HKMA reaches out to the reporting entities for clarification such that they can rectify any reporting errors if necessary. Simultaneously, data users at the HKMA may make necessary adjustments to enable timely analysis. For example, we may exclude trades with exceptionally large notional amounts or abnormal strike prices while pending rectification by the reporting entity.

5.4. Step 4: Aggregation of exposures

The HKTR reporting framework imposes a double-sided reporting regime, i.e. both counterparties need to report the trade to the HKTR if they fall under the reporting requirement. While double-sided reporting allows the HKTR to perform cross-checks which improve data quality, it also results in duplication of trade records where an entity can be the reporting entity in one trade record and the counterparty in another trade record. Therefore, handling such duplication is a crucial final step when aggregating exposures.

Our aggregation approach is based on the Unique Transaction Identifier (UTI), a unique and paired code to identify a transaction, which links the two sides of a trade reported by the two transacting TR members.[1] With that, we can identify duplicate trades and adjust the notional amount accordingly to avoid overstating total notional amounts. If two records refer to the same trade (i.e. same UTI), a linking factor of 0.5 is applied to the notional amount of each trade to arrive at an "adjusted notional amount" that is used when computing total notional amounts to avoid double counting (see Table 5).

[1] The HKTR recognises the Unique Transaction Identifier–Unique Swap Identifier (UTI-USI) required under the US CFTC reporting regulations and the Unique Transaction Identifier–Unique Trade ID (UTI-TID) required under the European regulations of the EMIR as reportable UTIs.

Example of Double-sided Reporting Adjustment Table 5

Scenario:
Entity A bought a USD/HKD call option of USD 2 bn from Entity B. Both entities fall under the reporting requirement and are required to report the trade to the HKTR. So there are two trade records (T1 and T2). Since both records refer to the same trade (i.e. same UTI), a linking factor of 0.5 is applied to the notional amount of each trade to arrive at an "adjusted notional amount".

Entity C bought a USD/HKD call option of USD 1 bn from Entity A but Entity C does not fall under the reporting requirement. So only Entity A reports a trade (T3) to the HKTR. Since there is only one trade record under this UTI, no adjustment is required (i.e. linking factor of 1) and hence the "adjusted notional amount" is equal to the original notional amount.

By summing up the "Adjusted Notional Amount", we can arrive at total outstanding option positions of USD 3 bn without double counting.

Trade Reference	Unique Transaction Identifier	Reporting Entity	Counter-party	Base Product	Notional Amount (USD bn) *(i)*	Linking Factor *(j)*	Adjusted Notional Amount (USD bn) *(=i x j)*
T1	X12345	A	B	Option	2	0.5	1
T2	X12345	B	A	Option	2	0.5	1
T3	X54321	A	C	Option	1	1	1

Note: The value of UTI is simplified for illustration purpose.

Another complication is to compute the entity-level exposures using the TR data in its original form because an entity can sometimes be on reporting side and at other times on counterparty side. We need to restructure the data in a way that facilitates the computation of an entity's total position, be it the reporting entity or the counterparty in different trades.

To do so, we leverage the idea of pivoting the data. In essence, the process rearranges the columns and rows of the dataset related to the role of the entities. Continuing with the scenario in Table 5, we pivot the data in two fields i.e. "Option Buyer" and "Option Seller" (see Table 6). After pivoting, the option buyer and option seller roles become values in the "Role" field while the entity names become values in the "Entity" field. Leveraging the Role and Entity fields, we can compute the entity-level exposures easily.

Pivoting of the Roles of Reporting Entity and Counterparty Table 6

Before pivoting, the data are in the original form.

Trade Reference	Unique Transaction Identifier	Reporting Entity	Counterparty	Adjusted Notional Amount (USD bn)	Option Buyer	Option Seller
T1	X12345	A	B	1	A	B
T2	X12345	B	A	1	A	B
T3	X54321	A	C	1	C	A

After pivoting the data in two fields i.e. Option Buyer and Option Seller, the roles of option buyer and option seller become values in the "Role" field while the entity names become values in the "Entity" field. Following on the adjustment of notional amount into “Adjusted Notional Amount” in Table 5, by summing up “Adjusted Notional Amount” by “Entity”, we can compute the gross positions of USD 3 bn for Entity A, USD 2 bn for Entity B and USD 1 bn for Entity C.

Trade Reference	Unique Transaction Identifier	Reporting Entity	Counterparty	Adjusted Notional Amount (USD bn)	Role	Entity
T1	X12345	A	B	1	Option buyer	A
T1	X12345	A	B	1	Option seller	B
T2	X12345	B	A	1	Option buyer	A
T2	X12345	B	A	1	Option seller	B
T3	X54321	A	C	1	Option buyer	C
T3	X54321	A	C	1	Option seller	A

When it comes to computation of entity-level net exposures, we determine the long vs short positions based on the role of the entities and apply the +/- sign to the corresponding notional amount (see Table 7).

Computation of Net Exposures Table 7

Continuing with above table after pivoting, if we focus on Entity A's net positions, when it is buying a USD/HKD call option, its position is considered as short HKD and hence we apply a negative sign to its position (USD -2 bn); when it is selling a USD/HKD call option, its position is considered as long HKD, resulting in a positive sign to its position (USD +1 bn). Summing up its "Long/Short HKD Position" gives us Entity A's net short HKD position of USD 1 bn.

Base product	Entity	Role	Adjusted Notional Amount (USD bn)	Long/Short HKD Position (USD bn)
Option	A	Option buyer	2	-2
Option	A	Option seller	1	+1
			Total	**-1**

Note: Negative value refers to a short HKD position.

The same treatment is applied to the HKTR data in all asset classes in order to compute the entity-level gross and net exposures.

6. Key utilities for transforming TR data for analysis

To deal with the size and complexity of HKTR data, we apply the "SESA" framework to transform TR data into data available for analysis and distil insights to support financial stability surveillance. Several utilities are essential and used along with the framework.

- In-house entity database: Our in-house entity database with around 29,000 entity identifiers helps us standardise entity names and enrich the data with additional entity-level information and parent-level information (e.g. sector, country of incorporation for the entity and its parent). We use automated scripts to marry the entity database with HKTR data to perform entity-level analysis.
  - To keep the mapping process effective and accurate, we regularly update the entity database by adding new entries as well as further standardising and enriching existing entities when additional information becomes available.
- In-house equity security database: We use our equity security database to standardise equity derivative underlying names and to enrich the data with more detailed information such as country of exchange.
  - From time to time, we update the equity security database via web scraping to take in account new equity securities for comprehensive analysis. Given the large amount of data being covered, the process takes some hours to complete.
- Process for timely follow-up with reporting entities: Once we spot potential data quality issues, a dedicated team at the HKMA reaches out to the reporting entities for clarification such that the reporting entities can rectify their reporting errors as soon as possible. This timely follow-up improves the quality of HKTR data to ensure accurate analysis.

- Over the years, this process has been fine-tuned and streamlined such that the follow-up with reporting entities is smooth and effective.

7. Development of use cases

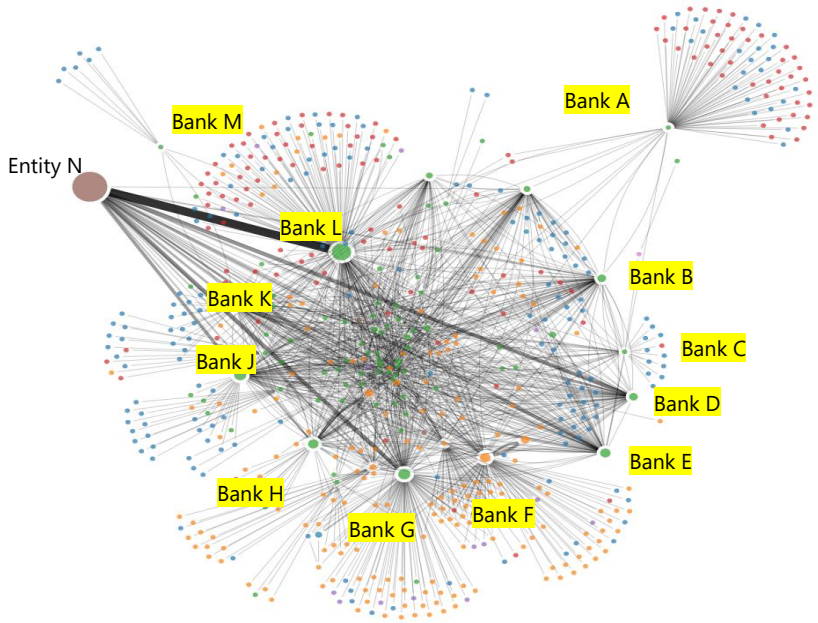
At the HKMA, we have developed several use cases of HKTR data to support financial stability surveillance.

– Monitoring of monetary stability as part of the Linked Exchange Rate System (LERS): On a daily basis, we conduct screening of HKD derivatives positions in option, forward and swap markets in the HKTR to understand market trends.

– Gauging exposures and interconnectedness of targeted entities during stress events: For example, when major international market entities are subject to market stress, with HKTR data we can quickly gauge their exposures in OTC derivatives within the scope of our regulatory framework and to identify their connections with the local financial system. This is particularly useful for regulators to assess the potential impact of the stress events.

  - Network analysis is a useful tool to visualise the interconnectedness of different market participants in OTC derivative transactions and to spot patterns and concentrations in trading relationships. For example, Graph 3 maps the network structure of the HKD interest rate swap market in the HKTR. Financial institutions are the nodes of the network and the links between nodes are weighted by the gross notional amount. The links between nodes could also be directional in other cases (e.g. HKD/USD cross currency swaps), with the arrow pointing from the HKD lender to the HKD borrower. In addition, Graph 3 indicates a core-periphery structure, with the core being a small set of banks conducting the majority of trades while the periphery being less active in the market. At times of market stress (e.g. if any risk event occurs to entity N), network analysis can support regulators to understand the market interconnectedness and assess potential channels of financial contagion.

Structure of the HKD interest rate swap market Graph 3



– NBFI surveillance framework: The HKMA adopted its NBFI surveillance framework in Q3 2021 as part of its financial stability surveillance toolkit. The NBFI framework combines a number of regulatory and non-regulatory datasets, from the HKTR data to granular banking data to news, in order to identify the risk exposures and risk characteristics of NBFIs. As shown in Table 8, the NBFI framework computes a set of risk indicators to quantify the potential impact of an NBFI default, measured by the size of the NBFI positions, and its vulnerability, measured by the potential illiquidity and concentration of its portfolio, its interconnectedness, leverage, and negative market news about the NBFI, with an additional role for volatility in the macro environment. Risk indicators are standardised and aggregated to produce impact and vulnerability scores for each NBFI as well as an overall score, simplifying metrics to Red-Amber-Green buckets. Generally speaking, NBFIs with Amber risk require close monitoring and NBFIs with Red risk are considered for possible follow-up actions.

Risk Indicators in NBFI surveillance framework Table 8

Category	Risk indicator
Impact	**Size of NBFI's positions** – This indicator measures the gross positions taken by the NBFI using bank loans and OTC derivatives.
Vulnerability	**Portfolio characteristics** – The indicators capturing portfolio dynamics include the share in high-volatility assets, as well as the share of the top five underlying holdings, small-cap stocks and substantial interest stocks in equity derivatives.
	Interconnectedness – The degree of interconnectedness measures how an NBFI amplifies its systemic vulnerability in relation to the banking

	sector. Indicators include the number of counterparties, prime broker concentration and crowded trades.
	Leverage – Two proxy indicators are the ratio of total gross positions to asset under management and the portfolio growth rate.
	Market news – The entity-specific market sentiment score is included to factor in negative news about specific NBFIs by using textual analysis.
	Macro environment – This indicator measures the degree of overall financial market uncertainty or stress based on stock market volatility indices. The indicator is market-wide and not specific to individual NBFIs.

8. Looking ahead: how CDE implementation will help with TR data analysis

At the international level, to facilitate the standardisation and harmonisation of OTC derivative data, the Committee on Payments and Market Infrastructures (CPMI) and the International Organisation of Securities Commissions (IOSCO) have published several technical guides on the harmonisation of UTIs, unique product identifiers (UPIs) and other CDEs [12, 13, 14]. The CDE initiative aims to standardise and harmonise OTC derivative data elements across different regulators' TR datasets, with the aim of promoting transparency in OTC derivative trades and facilitating data aggregation for analysis. The HKMA and the SFC intend to implement the mandatory use of UTI, UPI and CDE reporting in September 2025.

The adoption of CDE for TR data reporting will provide several key benefits for HKTR data analysis:

- Standardising the definitions, formats and allowable values of data elements, to streamline the reporting process, reduce the reporting burden for reporting entities and enhance data quality and comparability.
- Aligning Hong Kong's OTC derivative reporting regime with international reporting frameworks and extending it to important data elements that are not under Hong Kong's current reporting requirements, to facilitate cross-border data analysis for financial stability monitoring and systemic risk assessment.
- Harmonising Hong Kong's reporting regime with international reporting standards, to strengthen Hong Kong's regulatory framework and consolidate Hong Kong as an international financial centre.

9. Conclusion

Since the initial collection of TR data in Hong Kong in 2015, three developments have turned TR data from a fuzzy new data source available to central banks and regulators with no textbook on how to use them to a source of neat insights on the Hong Kong financial system. The three key developments were:

- The development of the SESA framework: (i) Standardisation of entity and securities names by building on LEIs and other market identifiers; (ii) Enrichment of entity and trade information with market data; (iii) Scanning for anomalies; and (iv) Aggregation of exposures;
- Timely follow-up with reporting entities for clarification and rectification if needed once we spot potential data quality issues; and
- Developments of use cases with useful insights for financial stability surveillance.

The TR data provide central banks and regulators with the opportunity to conduct timely analyses for financial stability monitoring. There are pain points to using the TR data which can be overcome. Hopefully this paper will contribute to increasing the adoption of TR data for financial stability analysis by central banks and regulators.

References

[1] 2018 survey conducted by the Irving Fisher Committee on Central Bank Statistics (IFC), "Central banks and trade repositories derivatives data", BIS IFC Report, no 7.

[2] "A first analysis of derivatives data in the Hong Kong Trade Repository", Hong Kong Monetary Authority, Quarterly Bulletin, June 2015.

[3] "Understanding Foreign Exchange Derivatives Using Trade Repository Data: The Non-deliverable Forward Market", Hong Kong Monetary Authority, Quarterly Bulletin, March 2018.

[4] K Cheng, Z Liu, S Pezzini and L Yu (2023), "Building an integrated surveillance framework for highly leveraged NBFIs – lessons from the HKMA", BIS Working Papers, No 137.

[5] U.S. Department of State, "The Pittsburgh summit: key accomplishments", Available at: https://2009-2017.state.gov/e/eb/ecosum/pittsburgh2009/resources/165061.htm

[6] Securities and Futures (Amendment) Ordinance (2014). Available at: https://www.gld.gov.hk/egazette/pdf/20141814/es1201418146.pdf

[7] HKMA HKTR Info Page. Available at: https://hktr.hkma.gov.hk/

[8] Administration and Interface Development Guide (AIDG). Available at: https://hktr.hkma.gov.hk/ContentDetail.aspx?pageName=HKTR-RPT-Administration-and-Interface-Development-Guide

[9] Global Legal Entity Identifier Foundation (GLEIF), "Level 2 Data: Who Owns Whom". Available at: https://www.gleif.org/en/lei-data/access-and-use-lei-data/level-2-data-who-owns-whom

[10] Cielinska O, A Joseph, U Shreyas, J Tanner and M Vasios (2017), "Gauging market dynamics using trade repository data: the case of the Swiss franc de-pegging", Bank of England, Financial Stability Paper, no 41.

[11] Iman van Lelyveld (2017), "The use of derivatives trade repository data: possibilities and challenges", Netherlands Bank, BIS IFC National Bank of Belgium Workshop on "Data needs and statistics compilation for macroprudential analysis".

[12] CPMI-IOSCO (2017), "Technical guidance – harmonisation of the unique transaction identifier".

[13] CPMI-IOSCO (2017), "Technical guidance – harmonisation of the unique product identifier".

[14] CPMI-IOSCO (2018), "Technical guidance – harmonisation of critical OTC derivatives data (other than UTI and UPI)".

HONG KONG MONETARY AUTHORITY
香港金融管理局

12th Biennial IFC Conference 22-23 Aug 2024
"Statistics and Beyond: New Data for Decision Making in Central Banks"

How to use Trade Repository data on OTC derivatives for analysis – A practical framework

Silvia Pezzini and Henry Chan

Development and scope of Hong Kong Trade Repository (HKTR) data

- **HKTR data**: Trade-level data on OTC derivatives traded in HK reported to HKTR, **~4.8 mn trades outstanding.**
- **Reportable by banks and CCPs** if the trade is: (i) **Booked on HK balance sheet**, or (ii) **Conducted in HK** (but not booked on HK b/s), or (iii) **Cleared** by CCPs with HK-incorporated counterparties.
- **Timely**: **Next-day availability for most trades.** 90%+ within T+1, latest T+2.
- **Comprehensive coverage**: Five asset classes, typical products, 250+ key data attributes (incl. valuation).
- **Timeline of HKMA use of TR data:**

2008-2009	2010-2017	2018	2019-2020	2021 - now
2008 (GFC): **Lehman** default, **interconnectedness** from OTC derivatives **2009: G20 reform OTC derivatives markets** Aim to improve **transparency**, prevent market abuse, **reduce systemic risks**	**HKMA builds own TR (HKTR)** 2013-2017: Reporting entities start reporting in **phased** manner Initial analysis: Interest Rate Swaps, Non-Deliverable Forwards	**Good quality HKTR data available** **5 asset classes**: Interest Rate FX Equity Credit Commodity Typical products: Swaps, Forwards, Options, Exotics	**HKTR analysis** becomes **regular input in financial stability monitoring** Focus on **currency stability** and interaction with **equity derivatives**	**Equity derivatives** regular monitoring Development of **NBFI surveillance framework** **Link** with **banking data** and **news** for **textual analysis**

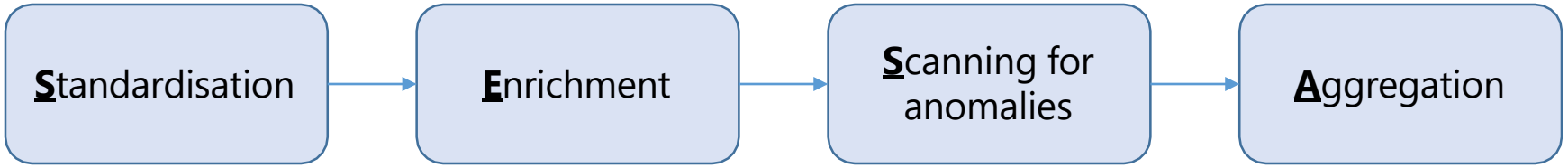
Benefits and pain points of using HKTR data

HKTR data has been used intensively by the HKMA for financial stability surveillance for its benefits in **timeliness**, **detail** and **standardisation**.

However, the size and complexity of HKTR data often pose **challenges**.

- **Identifier standardisation** is **not complete** (e.g. use of LEIs for counterparties, use of ISIN for equity derivatives underlying stocks or indices).
- **Limited standardised information on how entities with different LEIs belong to the same financial group** (Level 2 data under LEI framework which show the direct and ultimate parents of legal entities are in progress).
- **Reporting errors** may occasionally **create false alarms** (e.g. exceptionally large notional amounts or strike prices).
- **Double-sided reporting** needs to be handled carefully to avoid double-counting of trades.

At the HKMA, we have developed a practical framework called "SESA" for using HKTR data for analysis.



"SESA" framework - Standardisation

To enable **computation of consolidated exposure** of an entity or its parent and allow **combining** TR data with **other data sources** for further analysis.

1. **Standardise entity names** by mapping different identifiers (e.g. LEIs, CICR*), using an **entity database** with ~29,000 entity identifiers.
   - Regular update by adding new entries and further enriching existing entries.

2. **Standardise equity derivative underlying names** by mapping relevant identifiers (e.g. ISIN, FIGI), using an **equity security database.**
   - Update using APIs and web scraping to take into account new equity securities.

*Certificate of Incorporation, Certificate of Registration number

Sample of the HKMA's In-house Entity Database

ID Type	LEI	Other Identifier	Entity Name	Entity Sector	Entity SubSector	Parent Name	Parent Sector	Parent Country	Entity Country
LEI	EHWNDI2G5WXXXXXXXXX	NA	ABC CAPITAL MASTER FUND LP	Non-bank financial	Hedge Fund	ABC CAPITAL MANAGEMENT LP	Non-bank financial	USA	USA
CICR	NA	C12345	ABC CAPITAL MASTER FUND LP	Non-bank financial	Hedge Fund	ABC CAPITAL MANAGEMENT LP	Non-bank financial	USA	CYM
BRN	NA	B45678	ABC CAPITAL MASTER FUND LP	Non-bank financial	Hedge Fund	ABC CAPITAL MANAGEMENT LP	Non-bank financial	USA	CYM
LEI	5493000ALK1111111111	NA	XYZ CHINA OPPORTUNITIES FUND LP	Non-bank financial	Hedge Fund	XYZ CAPITAL MANAGEMENT LP	Non-bank financial	USA	USA
LEI	54930058A22222222222	NA	XYZ HK OPPORTUNITIES FUND LP	Non-bank financial	Hedge Fund	XYZ CAPITAL MANAGEMENT LP	Non-bank financial	USA	CYM
LEI	I4BOGE0V4Y3333333333	NA	XYZ JAPAN OPPORTUNITIES FUND LP	Non-bank financial	Hedge Fund	XYZ CAPITAL MANAGEMENT LP	Non-bank financial	USA	CYM

Different types of identifiers of ABC Capital will be mapped to the same entity name to enable aggregation of trades related to the same entity.

XYZ Capital manages three individual funds with their own LEIs. All the trades reported under these LEIs can be grouped under the same parent for analysis.

"SESA" framework - Enrichment

To enrich HKTR data with more detailed **dimensions for meaningful analysis**.

1. General enrichment:

- **Additional entity-level** (e.g. sector and country of incorporation) and **parent-level information** (e.g. name of parent group, parent sector and parent country of incorporation) based on the in-house entity database.
- **Derive new data fields** for analysis (e.g. USD equivalent notional amount, original and remaining maturity, intra-group indicator) based on existing data fields.
- **Define** the position of an entity as **long or short** for computation of **gross** and **net positions**.

2. Asset-class specific enrichment:

- FX: Clean and extract **currency pairs** for market monitoring and topical research analysis.
- Equity: Derive the underlying country of exchange to analyse the **equity positions in a particular jurisdiction.**
- Commodity: Compute the **notional amount based on product features** (e.g. unit price *X* number of tons).

"SESA" framework – Scanning for anomalies

To **identify anomalies** and potential **data quality issues.**

- Exceptionally large notional amounts and exceptional strike prices.
- E.g. FX: some reporting entities may report 780 as the strike price for USD/HKD options, which looks abnormal given the range of USD/HKD FX rate (7.75 – 7.85).

Well-established follow-up actions in place to deal with data quality issues.

- A designated operational team at the HKMA reaches out to the reporting entities for clarification such that they can rectify any reporting errors if necessary.
- Simultaneously, data users at the HKMA may make necessary adjustments for timely analysis (e.g. outlier removal pending rectification by reporting entities).

"SESA" framework – Aggregation of exposures

- To **deal with potential duplication** of trades due to double-sided reporting* and to **transform the data** for entity-level exposure analysis.
- Identify duplicate trades using **UTI** and **adjust** the notional amount by applying a **linking factor (0.5)** to avoid overstating total notional amounts.

Trade Reference	Unique Transaction Identifier	Reporting Entity	Counter-party	Base Product	Notional Amount (USD bn) *(i)*	Linking Factor *(j)*	Adjusted Notional Amount (USD bn) *(=i x j)*
T1	X12345	A	B	Option	2	0.5	1
T2	X12345	B	A	Option	2	0.5	1
T3	X54321	A	C	Option	1	1	1

- By summing up "Adjusted Notional Amount", total outstanding option position is equal to USD 3 bn without double counting notional amount of Trade 1 and Trade 2.

*Double-sided reporting regime: Both sides of the trade need to report the trade to the HKTR if they fall under the reporting requirement. So an entity can be the reporting entity in one trade and the counterparty in another trade.

"SESA" framework – Aggregation of exposures

Trade Reference	Unique Transaction Identifier	Reporting Entity	Counterparty	Adjusted Notional Amount (USD bn)	Option Buyer	Option Seller
T1	X12345	A	B	1	A	B
T2	X12345	B	A	1	A	B
T3	X54321	A	C	1	C	A

Trade Reference	Unique Transaction Identifier	Reporting Entity	Counterparty	Adjusted Notional Amount (USD bn)	Role	Entity
T1	X12345	A	B	1	Option buyer	A
T1	X12345	A	B	1	Option seller	B
T2	X12345	B	A	1	Option buyer	A
T2	X12345	B	A	1	Option seller	B
T3	X54321	A	C	1	Option buyer	C
T3	X54321	A	C	1	Option seller	A

By summing up "Adjusted Notional Amount" by "Entity", we can compute the gross positions of Entity A, B and C (USD 3 bn, USD 2 bn, USD 1 bn respectively).

- **Restructure the data** to facilitate the computation of an entity's total position by **pivoting "Option Buyer"** and **"Option Seller"** into **"Role"** and **"Entity"**.
- Compute the **entity-level net exposures** by determining the **long vs short** positions based on the role of the entities.

Base product	Entity	Role	Adjusted Notional Amount (USD bn)	Long/Short HKD Position (USD bn)
Option	A	Option buyer	2	-2
Option	A	Option seller	1	+1
			Total	**-1**

Note: Negative value refers to a short HKD position.

Entity A's net position: Apply a negative sign to its short position.
The sum gives us its net short position of USD 1 bn.

Conclusion and looking ahead

Three developments have enabled the HKMA to make best use of the HKTR data:

- **SESA** framework: (i) **Standardisation** of entity and securities names by LEI and other identifiers; (ii) **Enrichment** of entity and trade information with market data; (iii) **Scanning** for anomalies; and (iv) **Aggregation** of exposures.
- **Timely follow-up with reporting entities** for clarification and rectification if needed.
- **Development of use cases** with useful insights for financial stability surveillance.

The HKMA and the SFC intend to implement the new derivative reporting regime for **Critical Data Elements** (CDE) in Sep 2025, providing several key benefits for HKTR analysis:

- Standardising the definitions, formats and allowable values of data elements, to streamline the reporting process and reduce reporting burden.
- Aligning Hong Kong's OTC derivative reporting regime with international reporting frameworks and extending it to important data elements that are not under Hong Kong's current reporting requirements.
- Harmonising Hong Kong's reporting regime with international reporting standards.

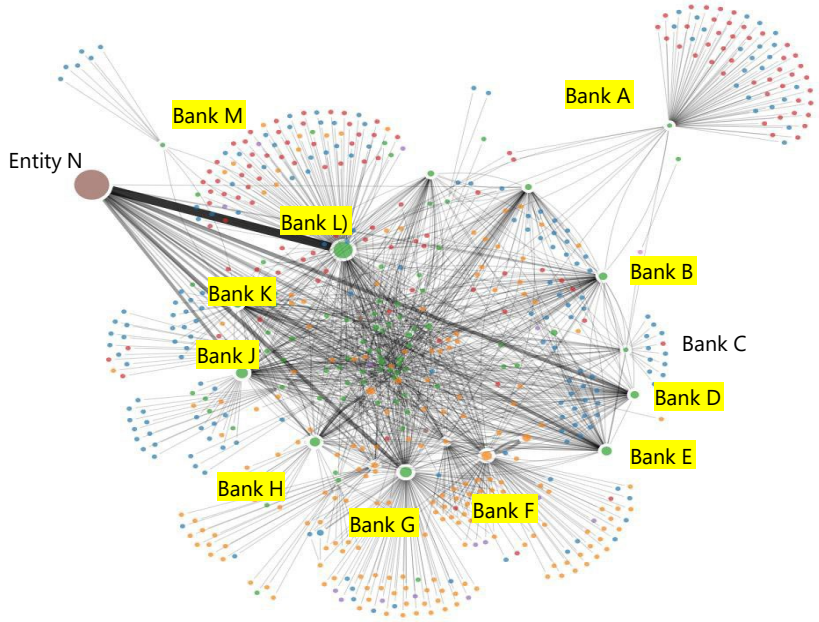
Annex 1: Development of use cases – Interconnectedness during stress events

Gauging **exposures** and **interconnectedness** of target entities during **stress events** (e.g. Credit Suisse failure)

Network analysis with HKTR data

- Analysis of concentration, interconnectedness, channels of financial contagion, and network structure (concentrated, core vs periphery)
- Risk assessment: how shocks propagate through the network

Structure of HKD interest rate swap market



Network map:

nodes = financial institutions

links between them = trading relationships

Annex 2: Development of use cases – Non-bank financial institutions (NBFI) surveillance framework

- The NBFI framework incorporates diverse data sets from HKTR and granular banking data to news, to compute a set of **risk indicators** to quantify potential **impact** and **vulnerability** of an NBFI.
- Risk indicators are standardised and aggregated to produce impact and vulnerability scores for each NBFI.
- Output is a watchlist, ranking the top NBFIs based on their scores.
- **Red-amber-green illustrate the NBFIs' riskiness** (high-medium-low).