

---

**The European System of Central Banks' Centralised Securities Database**

**Prepared by the European Central Bank**

**This updates a paper presented to the  
Fifteenth Meeting of the  
IMF Committee on Balance of Payments Statistics  
Canberra, Australia, October 21–25, 2002**



## EUROPEAN CENTRAL BANK

BALANCE OF PAYMENTS STATISTICS  
AND EXTERNAL RESERVES DIVISION

Judith Daniels ☎: 7389

13 September 2001

ST/IMF/BP/BOPCOM14.UPDATE.DOC

# THE CENTRALISED SECURITIES DATABASE

**For information of the IMF BOP Committee**

## Abstract

1. Securities are the second most important class of financial instruments in the euro area (some EUR 8 trillions of debt securities issued by euro area residents are outstanding). Yet statistics on them have many deficiencies. The CSDB will be a comprehensive reference database with a range of statistical applications and also much relevance for economic analysis, financial stability issues and operations (eligible assets, risk management). The availability of information on equity and debt securities would greatly contribute to overcome shortcomings in the quality of financial statistics.

2. This paper provides an examination of the increasing importance of information on a security-by-security basis for accurate and timely (monetary, securities issues, balance of payments and flow-of-funds) statistics. The paper briefly describes the present difficulties with statistical treatment of securities, and potential sources of inconsistency and inaccuracy.

3. The development and set-up of the CSDB within the institutional framework of the European System of Central Banks (ESCB)<sup>1</sup> proved to be a complex task; hence the duration of the preparation phase of the project. This phase contained the identification of the requirements of users within the ESCB, the identification of potential data sources (their data limitations, and the legal and commercial aspects) as well as the design of the organisational and technical infrastructure of the CSDB.

4. From June 2002, the CSDB project enters its implementation phase, which will see the implementation of its technical infrastructure as well as the set-up of a “CSDB-network”. The cooperation may take the form of “clearing-house”, “network” or centralised approach; the paper discusses each and explains why the second was chosen; it stresses the importance of co-operation between the ECB, the EU – and Accession Countries’ – National Central Banks (NCBs) and the BIS in defining and implementing the CSDB. In the “CSDB network”, which will be implemented in a stepwise approach, the NCBs will

---

<sup>1</sup> The European System of Central Banks (ESCB) comprises the European Central Bank (ECB) and the 15 national central banks of EU Member States.

have the responsibility of monitoring the coverage and quality of information on the issues by residents of their jurisdictions, while the ECB and BIS, using authoritative official and commercial sources, will ensure overall consistency and contribute data on issues by residents of countries outside the EU which are internationally traded, held by euro area residents or denominated in euro.

5. The experience to be derived from the “CSDB network” may benefit in due course other areas – prudential supervision, risk and operational management – and may lead to an expansion of the network to other statistics compilers (central banks and national statistical offices) and further cooperate with international organisations (in particular the IMF and BIS) in the future, perhaps in the form of a Global Securities Database (GSDB). The results should foster transparency and efficiency of the markets.

## **The business case for the CSDB**

6. Securities represent the second largest class of financial instruments in the euro area financial system. Excluding equities and shares/units issued by collective investment institutions, some EUR 8 trillions worth of debt securities issued by euro area residents were outstanding at end-2001, an amount growing by some 7% a year. The gross issuance amount can exceed EUR 600 billion a month. Most of the outstanding amount issued by euro area residents (EUR 7.0 trillion) is denominated in euro. In addition to that there are EUR 0.9 trillion of euro-denominated securities representing claims on non-residents of the euro area. Moreover, as securities represent about 20% of the aggregated Monetary Financial Institutions (MFI) balance sheet in the euro area, securities of various kinds are an important and fast-growing component of broad money.

7. Changes in the financial structure of the main economic sectors have prompted an increasing interest in securities statistics by central bankers, market regulators, macro and micro-supervisors, and private banks. Developments in securities are of clear interest for financial stability, risk management, and operations (e.g. in connection with collateral). The ECB carries out formal analyses of developments closely related with securities issues, e.g. Monthly Bulletin articles on the financial structure of the euro area.<sup>2</sup> The scope of securities issues and trading in the euro area differs from that in the USA. It seems likely that they will grow as a source of financing<sup>3</sup>.

8. For borrowers, securities are an alternative to bank finance. Holders of financial assets may view bank deposits, negotiable instruments issued by banks and other securities as partial substitutes. Data on the outstanding amount of securities indicate the depth of capital markets. Information on securities issues in euro is relevant to assessment of the role of the euro in international financial markets.

9. In addition, due to cross-border trading/holdings in securities the importance of securities in the balance of payments (b.o.p.) and international investment position (i.i.p.) is immense. Related income has

---

<sup>2</sup> The ECB publishes aggregated information on Securities Issues according to a “short term approach”.

<sup>3</sup> Whereas, in the euro area, three-quarters of non-financial corporations’ indebtedness are vis-à-vis banks in the euro area, this share is around a third in the USA, the remainder being securitised.

become the third largest component of current account transactions and ranks just behind services<sup>4</sup> in the euro area b.o.p. Their significance in money and banking statistics and the b.o.p., together with their importance as a source of finance for non-financial corporations, non-monetary financial institutions, and governments gives securities a prominent place in the Monetary Union financial accounts (flow of funds statistics for the euro area as a whole economy).

10. It is the ECB's responsibility, under Article 5 of its Statute, to ensure that the statistics mentioned above, needed for the execution of its tasks, are harmonised where necessary and remain relevant and accurate as securities business evolves over time.

11. Securities present statistical difficulties. In the financial accounts tables currently published by the ECB, there is no sector breakdown of the holdings of securities. Thus, it is not possible to evaluate, for example, how far increasing debt of households and business is matched by holdings of financial assets.

12. The best way to ensure a correct statistical treatment of securities is to compile a reference database of those securities which euro area residents are likely to hold or transact in. Such a reference database will be a valuable tool even if reporting agents report their securities transactions and positions in the aggregate. It is indispensable if they report security by security, since the compiler needs a reference list of securities with reliable structural information about each<sup>5</sup>. (Security-by-security reporting – recording individual securities issued or held – may seem cumbersome and a heavy burden on reporting agents, but for issuers, holders or their agents who are highly automated it is deemed easier, as well as providing more flexibility to compilers, than aggregated reporting; nine euro area countries and two pre-ins already use security-by-security reporting, or plan to introduce it.)

13. This reference database should be unique (or one master list and exact copies) in order to allow the consistent construction of statistics across euro area countries. Existing sources have uneven quality and timeliness, as well as methodological inconsistencies such as the treatment of residency, sectors and double counting of international securities. To maximise the reliability of such a database one has to assemble data about individual securities issues, clean it up, and incorporate all available information about holders. Assembling the data means developing feeds from various sources<sup>6</sup>. Cleaning it up means eliminating overlaps, choosing the best sources for the large amount of supplementary data, and in general ensuring that the database is as comprehensive, reliable and up-to-date as possible. Incorporating data about holders means storing data from surveys or commercial sources.

---

<sup>4</sup> In 2001 for direct and portfolio investment income the credits were EUR 152.6 bn and the debits were EUR 184.3 bn.

<sup>5</sup> Such as a unique identifier, outstanding amount, nature of the instrument, currency, coupon and payment date(s), sector and country of residence of issuer, maturity, current market price etc..

<sup>6</sup> Apart from securities databases in some NCBs, there are several sources of data on individual securities (commercial databases; the European Commission; the BIS, which has a large database of international issues and would support and accompany work on the CSDB).

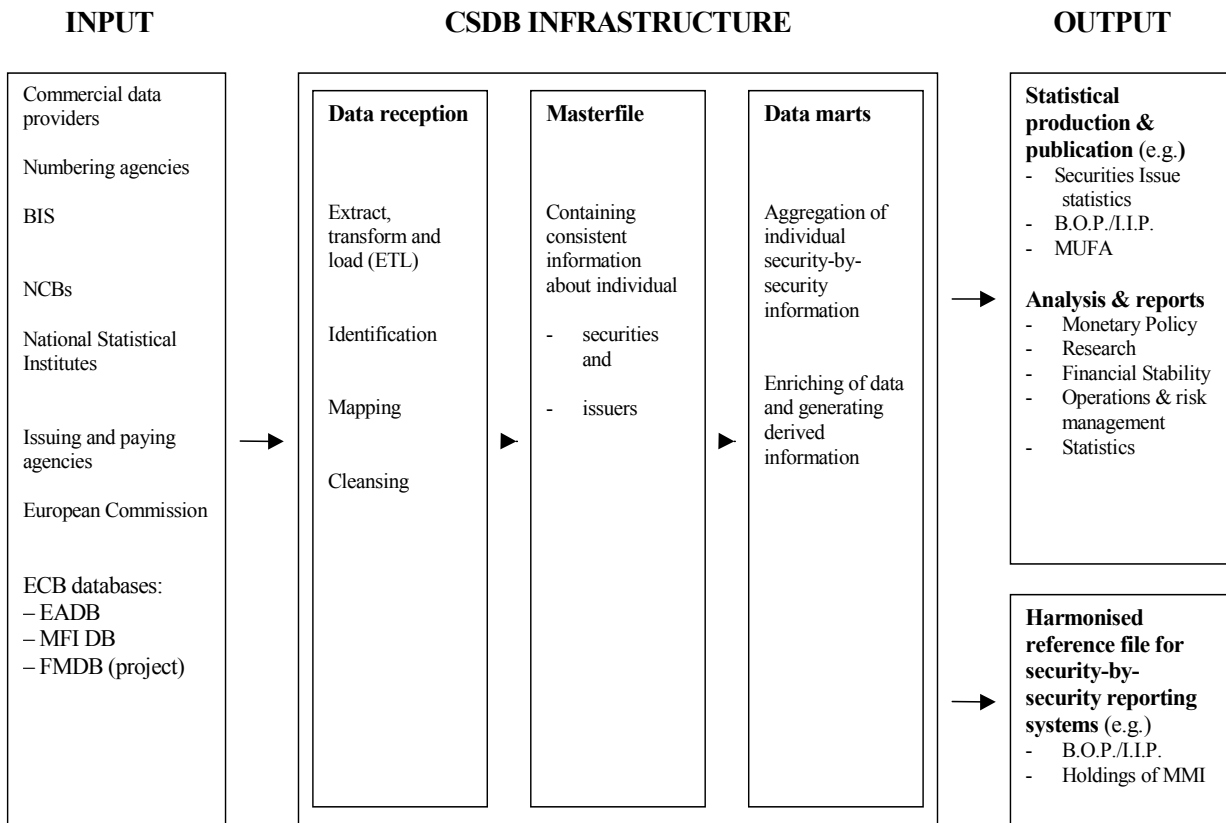
## **The CSDB Project**

14. The set-up of the CSDB within the institutional framework of the ESCB showed some complexity which is reflected in the duration of the preparation phase of the project. The main cause of such a long phase was the tender procedure for an adequate system supplier. It could only be launched once the user requirements (the main features are described in Annex 1) had been defined and main features of the feasibility of the CSDB system as an ESCB database had been outlined. This already took a year. Then the tender and the evaluation of the much differing offers received proved difficult. The final decision was taken 15 months after the publication in the OJ of the European Community of the call for tender.

15. However, tasks were undertaken in parallel to this procedure. The preparation phase contained the identification of potential data sources (their data limitations, and the legal and commercial aspects), the design of the technical infrastructure of the CSDB, and the approach toward data quality management.

16. From July 2002, the CSDB project entered its implementation phase 1 which will see the development and undertaking of its technical infrastructure as well as the set-up of a “CSDB-network” (for the data quality management). While the ECB, supported by an external developer, can set up the framework, it is impossible for the ECB alone to do all the work of setting up and maintaining the database. This is why the CSDB was designed as an ESCB project. The cooperation within the ESCB may take different forms which are detailed in the next chapter. NCBs will be major users of the information contained in the CSDB. Figure 1 gives an overview of the flow of information in the future CSDB framework.

**Figure 1: Flow of information within the CSDB - network**



**Legal and commercial aspects**

17. The database may contain confidential individual information. This is not the case, in principle, of data on issues and issuers. These will be derived from authoritative, or other public or commercial sources. This may not always be the case for some private placements or instruments which, though tradable by nature, may not be traded in effect. Data on holdings, whenever available, are not foreseen to be individual and, thus, not subject to specific protection. Protecting confidentiality, where relevant, will be ensured by the system and managed by the database administrator.

18. Data received from commercial sources are generally under license. For commercial data providers, information notably on prices is a core part of their business; often dealing with front or medium offices or market analysts, the more frequent and timely the data, the higher the price. However, in the case of the CSDB, both the system itself and the data will be covered by provisions so as to grant access to the whole ESCB for statistical purposes (monthly data, not in real time) and, under specific arrangements, to some other institutions (e.g. national statistical institutes in some EU member states, the BIS and NCBs in accession countries). As contracts are not yet ready, it is too early to define accurately which institutions will have access to what data. In addition, the data quality management is deemed to confer the property right on (at least some descriptive) data, whatever their sources, to the ECB so as to afford for the extraction of a subset and their transmission to reporting agents to allow them to make accurate breakdowns when reporting statistical data to national compilers.

## Basic organisational models to run a CSDB

19. The cooperation within the ESCB could in principle take three different forms:

20. **Clearinghouse approach:** A central clearing institution receives the contributions of participating countries and puts them into a single database without any further quality checking. The advantage of this approach is that partial results would be available fairly soon and that the resource implications for the clearing institution are limited as the participating countries remain responsible for the comprehensiveness and accuracy of their own data. Its main disadvantage, however, is that it is likely to take a long time until a sufficient degree of comprehensiveness and consistency of data was achieved.

21. **Centralised approach:** In this approach an international organisation purchases all relevant data from commercial databases, performs the quality control, harmonises the data and makes them accessible to official users. This approach is very attractive to users of the data because they do not have to contribute to the cost and have no responsibility for the quality of the data<sup>7</sup>.

22. **Network approach:** A co-ordinating institution will not only collect the relevant information from national statistical agencies and central banks, but also supplement any missing data from commercial data sources with the aim of achieving a maximum degree of comprehensiveness. In addition, the co-ordinating institution may perform quality controls on the data, fill remaining gaps using reasonable estimates and set standards for data compilation in order to ensure accuracy and consistency of the information collected.

23. The disadvantages of the clearinghouse approach does not only lie with the timing to obtain a comprehensive database, but also with the lack of harmonisation in concepts and definitions entailing a risk of “Babel tower” (e.g. homonyms may hide different meanings; similar concepts are to be found under different words; securities and prices may be described with somewhat differing sets of attributes). The centralised approach is the most demanding for the agency which develops and manages the global/centralised securities database. There also exist doubts that the quality of data from commercial databases is a priori higher than that from official sources<sup>8</sup>.

## The “network” approach of the CSDB

24. The ECSB adopted the network approach as the general framework in which to set up the administrative and technical infrastructure of the CSDB. It involves an important degree of co-operation between the ECB, the EU National Central Banks and the BIS while defining and implementing the CSDB. In a “CSDB-network” all EU countries and the ECB will share the responsibility of monitoring

---

<sup>7</sup> The centralised approach was favoured by the US. The US felt that this approach would ensure sufficient comprehensiveness and accuracy of the data at the lowest cost.

<sup>8</sup> The BIS has the impression that commercial data are often earlier available than official data, but that official data are normally more accurate than commercial data.

the coverage and quality of information on securities issued by their respective residents on their domestic markets. The BIS will provide data on securities issued on international markets on a best effort basis.

- (A) Most NCBs are already running, or intend to do so, a securities databases on issues by residents and are ready to provide data feeds to the CSDB (“data providing countries”).

A data feed by a NCB should ideally cover issues of residents of their jurisdiction on domestic markets and abroad. Where some of the information is missing, potential gaps should be identified and the CSDB operator will provide the relevant information via external sources, e.g. commercial sources or the BIS database.

- (B) All ESCB countries<sup>9</sup> are sharing the responsibility of monitoring the coverage and quality of information on the issues by their residents on the domestic markets as well as abroad. Where a country is not in a position to provide or share data on issues by residents, the operator of the CSDB will be responsible to retrieve all information available but the country would remain responsible for the monitoring of these data. In the data quality management process, all NCBs would need to check data related to issuers (e.g. sector in ESA 95) and cross-check data on securities with aggregated information from other national sources (e.g. MFI balance sheet statistics, national accounts, Other Financial Intermediaries and Central Balance Sheet Office statistics, etc.).
- (C) Finally the same rules and responsibilities apply to the operator of the CSDB for issues by the “rest of the world” which were identified (or considered) to be relevant for financial statistics within or outside the euro area.

## Data Quality Management

25. Based upon the “network” approach outlined in the preceding section, parts of the CSDB will be fed or updated “bottom-up” via participating NCBs and parts “top-down” via external sources made available by the ECB. Where the cost and effort to cover certain information in a decentralised way would be higher than via a centralised approach, the operator of the CSDB may take over the responsibility of retrieving this information<sup>10</sup>. The figure 3 in annex 2 shows the organisational structure of the CSDB data quality management. The main data quality tasks, unification (mapping) and cleaning are displayed as core tasks to be accomplished within the CSDB system.

---

<sup>9</sup> And possibly accession countries. The principle of granting them access to the CSDB and requesting them to contribute for their own residents will be considered by the STC. Modalities will be discussed at a later stage.

<sup>10</sup> A typical example for this is the information on prices, i.e. quotations that would not be provided by NCBs. A decentralised provision from NCBs would be unduly costly and they therefore may be bought by the ECB from a number of commercial data providers yet to be determined.



## **Outlook – Toward a “Global Securities Database”?**

26. The experience which will be derived from the “CSDB-network” may be extended in due course to other areas – eligible assets, prudential supervision, risk and operational management - and may lead to an expansion of the network to include other statistical compilers (international organisations, central banks and national statistical offices) in the near future, perhaps in the form of a GSDB.

27. In its last two meetings in October 2000 and 2001, the IMF Balance of Payments Committee has called for the development of a GSDB. In response to these developments, the BIS offered to explore the feasibility of creating a global securities database by expanding their existing databases on international and domestic securities issues. This approach was then considered by the IMF and BIS as less promising than the one that consists of developing the CSDB “network” within the ESCB and, once operational, considering the possible cooperation with main partners (US and Japan in particular) on data quality management for residents of their jurisdiction and sharing of information and costs. The network could be implemented with the BIS as the co-ordinating institution. The advantage of this approach would be that the BIS could learn from the ESCB project and proceed with a comprehensive and consistent database with more analytical and operational value than with any other approaches.

28. The resource implications for the BIS and other institutions involved would need careful consideration. For their part, the ECB and ESCB would only be willing to open the contents of the CSDB to other countries – given that contractual arrangements could be made – if it/they could expect a similar data quality on issues by residents from these countries.

- the number and diversity of countries is larger than in the euro area/EU, with an effect on the need and difficulty for undertaking data quality management;
- the data collection from various (e.g. commercial) sources on some markets may be of uneven quality.
- standard setting on a more global level relies on voluntary compliance.

29. In a paper on this issue, the BIS notes that “While all three approaches above entail major shortcomings, we believe that under the current circumstances the network approach of the ECB, which combines data collection from official and commercial sources and involves standard setting, represents the only reasonable way forward.”

30. Therefore, the BIS brings forward the proposal for the development of a global securities database along the following lines:

- As issuance in securities markets is highly concentrated, the countries with the largest markets, which account for almost 90% of global issuing activity, could form the core of the global securities database. The BIS could explore with these major economies (notably the US, JP and

EU) which types of data are available and under which conditions and when they could be included in the global database<sup>11</sup>.

- Once an agreement has been reached with the core countries to create the global database in principle, the IMF and the BIS could develop common standards for the data compilation jointly with the ECB, US and JP.
- Countries other than the US, JP and the EU countries, which maintain securities databases and are willing to share their data with other countries, may be invited to participate in the global database as soon as the core countries have joined.
- In analogy to the ESCB-CSDB project and in order to limit the cost for the institution maintaining the global database (e.g. most likely the BIS), each country would in principle remain responsible for the accuracy of data on securities issues of its own residents, independent on whether the data are collected from official or commercial sources.

31. Finally, the main benefits of a global securities database would be twofold:

- To improve the quality of b.o.p. and i.i.p. statistics and the from now on annual IMF Coordinated Portfolio Investment Survey;
- To enhance the analysis of financial market developments, in particular on securities markets.

32. However, other contributors may also expect using the database for fostering consistencies across a large set of financial statistics (of which flow of funds) and for other analytical purposes.

---

<sup>11</sup> For this purpose, a joint task force or working group could be formed. To the extent that sufficient coverage cannot be achieved from official sources, like in the US, the question of how to share the costs for commercial data would have to be addressed.

## Main features of the CSDB

33. Scope

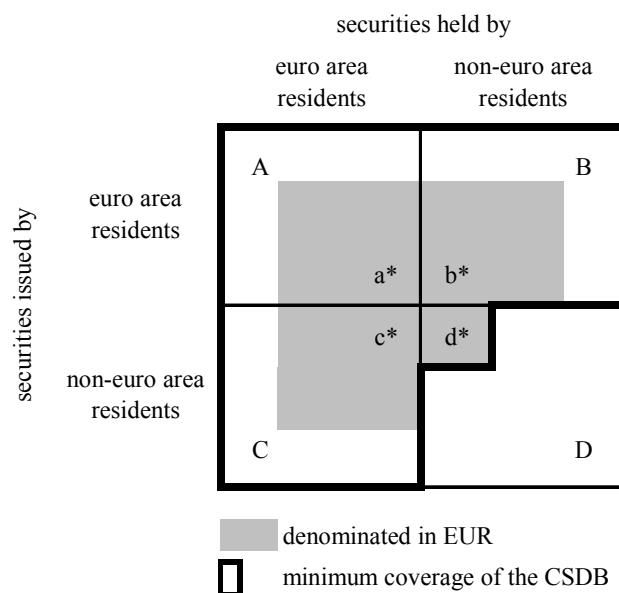
34. The objective of the CSDB is to hold complete, accurate, consistent and up-to-date information on:

1. the different types of securities relevant for the ESCB's statistical purposes, and
2. final holders of these securities (broken down by institutional sector and geographical area).

35. Thus, the CSDB should provide the ESCB with up-to-date information about the issuers and holders of securities (both debt securities and shares), stocks and their financial transactions like issue and redemption of securities and change of holders. The statistical information should be complete, coherent, and consistent. It should be available on-line to the ESCB, and possibly to other statistical agencies and users.

36. The minimum requirements - as presently identified - are marked by fat borderline. In addition the dimension of denomination (EUR/non-EUR) is highlighted. The figure illustrates that the CSDB should on the one hand cover the total set within the fat border line and on the other hand should allow for identification of sub-sets (A to C and a\* to d\*).

37. Figure 2: Scope of securities covered in the CSDB



38. Thus the envisaged coverage is all securities issued by residents of the euro area (in euro or in foreign currencies) as well as securities issued in euro by non-residents. All securities are going to be broken down by instrument type, issuer's sector, maturity or other characteristics of the securities.

39. Ideally, the CSDB should also cover (i) issues by corporations whose parent company is resident in the euro area, i.e. the location of the ultimate risk, and (ii) securities (denominated in currencies other than the euro and issued by non-euro area residents) which are held/traded by euro area residents. The CSDB is also likely to comprise issues and holdings of residents of EU countries that have not adopted the euro yet, the so-called pre-in countries. Data related to issues by residents of accession countries may also be considered upon request.

40. Furthermore, the coverage of the CSDB will be based on the definition of securities in the European System of Accounts<sup>12</sup> (ESA) 95, According to ESA 95, securities comprise the categories F.3 (Securities other than shares) and F.5 (Shares and other equity). Category F.3 includes short- and long-term securities other than shares and financial derivatives, category F.5 quoted and unquoted shares and other equity as well as mutual funds shares. However, presently the CSDB is not considered to cover the full set of instruments, in particular data on financial derivatives will not be collected. Above table 1 shows the expected number of securities in the CSDB by issuer category and instrument type.

### **Sector breakdown and valuation**

41. The CSDB will be designed in such a way that the system of accounts for securities can be derived as outlined in ESA 95, which is in compliance with BPM5<sup>13</sup>. For money and banking as well as for b.o.p./i.i.p. statistics, outstanding amounts of securities at the end of each period and, separately, new issues and redemptions during that period are required. The value of stock of securities at the end of the reporting period can be derived by adding all financial transactions, other volume changes and re-valuations to the value of stock of securities at the beginning of the reporting period.

42. According to ESA 95, the recording of financial transactions is to be on a net basis. The valuation principles for securities transactions are the valuation according to market or transaction value<sup>14</sup> as well as the nominal value. The nominal value refers to the face value of the security. In case one (or both) price information is (are) not available estimations or approximations will be obtained using state of the art statistical inference methods.

### **43. Information on holders**

44. In addition to the information on the issue of securities, the CSDB should also enable the compilation of information on the holders of securities. Data on the residency of the holders of securities issued by euro area resident Monetary Financial Institutions (MFIs) is essential. In addition to the amount

---

<sup>12</sup> This classification is considered consistent with Balance of Payments Manual 5<sup>th</sup> edition (BPM5), cf. BPM5 table 4 p.127.

<sup>13</sup> Accordance with ESA 95 implies accordance with the System of National Accounts (SNA) 1993<sup>13</sup> as well as BPM5. For euro area resident issuers/holders of securities, financial accounts and money and banking statistics need a detailed sector classification which is consistent with but goes beyond the one for b.o.p./i.i.p. Especially, the sectors 'financial corporations' and 'general government' have to be fully extended (sub-sectors).

<sup>14</sup> The relevant accounting rules are defined in ESA 95 as well as BPM5. The market value is based on the price quoted in the market, whereas the transaction value refers to the value at which the security is created, exchanged or liquidated excluding charges and taxes.

purchased/sold (flows) or held (stocks) by each holder or each category of holder two types of information should also be included:

- (i) The sector of resident holders: this is of substantial value for the compilation of quarterly MU financial accounts, individualised data set on securities need to be specified by institutional debtor and creditor sector. This breakdown will be based on the residence concept, in order to receive reliable statistical information on the liabilities and assets side of institutional units.
- (ii) The residency of non-resident holders of securities issued by euro area residents: the minimum regional breakdown has to include the USA, Japan, American countries other than USA, the pre-in countries and the accession countries. All countries within this scope have to be identified by the individual ISO code.

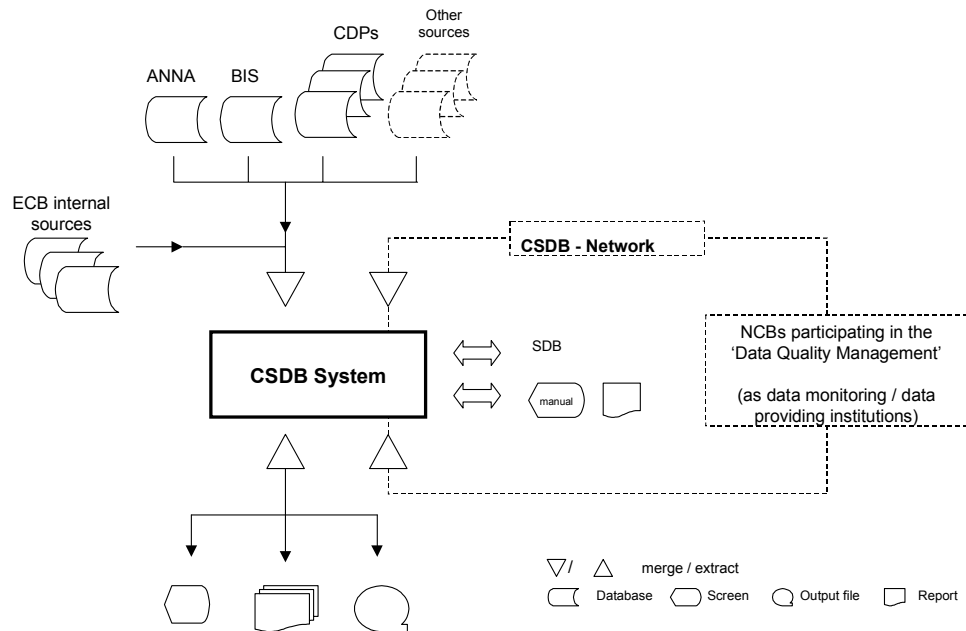
45. Incorporating information about holders will be a difficult task as current data sources are much less comprehensive than on the issuer side, and is therefore deferred to a later stage in the project, when also the technological and organisational framework will be fully available.

## Data quality management and technical infrastructure of the CSDB

### Unification and mapping

47. In combining separate databases, the mapping and cleaning process shall ensure the extraction of the best possible quality information: cleaning it up means eliminating overlaps, choosing the best sources for the large amount of supplementary data, and in general ensuring that the database is as comprehensive, reliable and up-to-date as possible. This will enable to overcome the present fragmented approach which probably contains inconsistencies and duplication as existing sources are deemed to have uneven quality and timeliness, as well as methodological inconsistencies such as the treatment of residency, sectors and double counting of international securities.

48. Figure 3: Organisational structure of the CSDB data quality management



49.

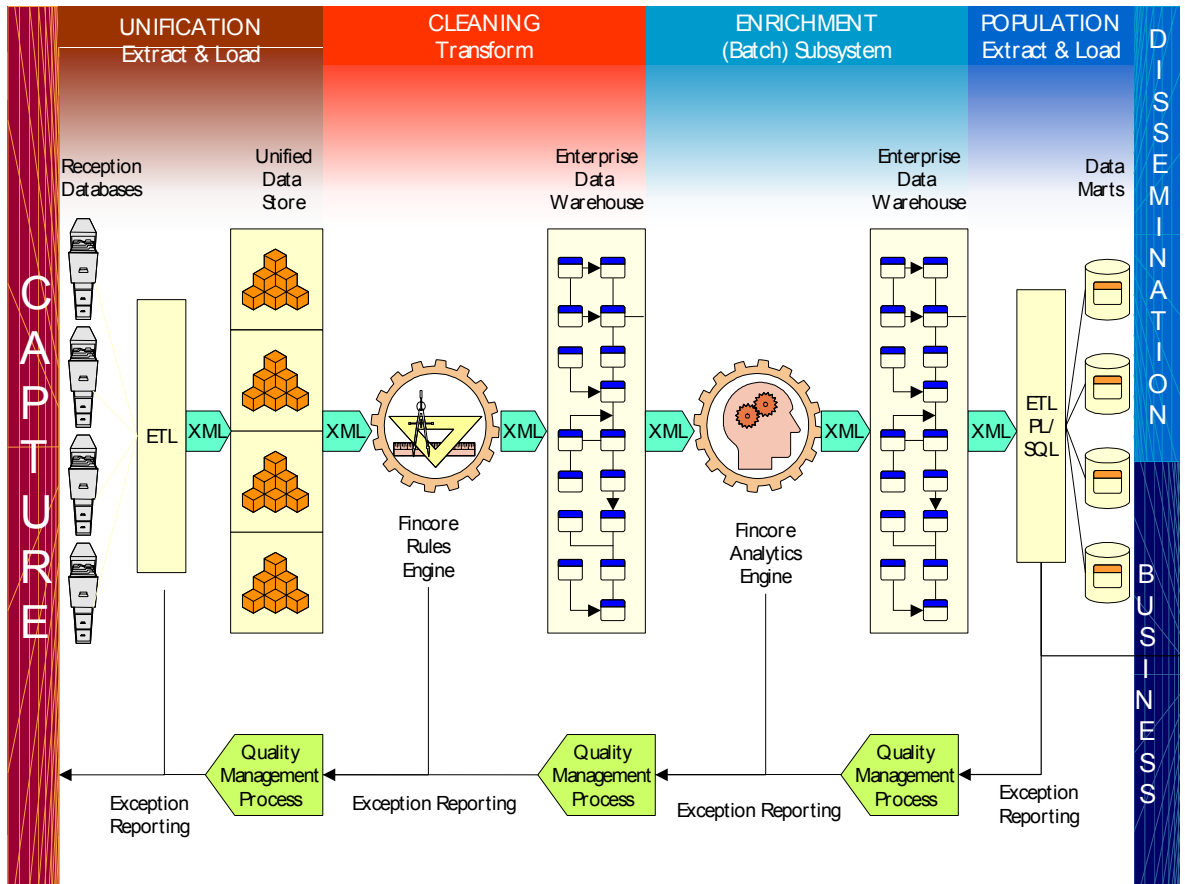
50. The objective of a mapping process is to obtain a single, master securities data file containing reference securities data from disparate sources. This master file has to meet a number of criteria, of which are data consistency and data comprehensiveness. This mapping process is done through means of a migration tool. This tool should allow the CSDB maintenance team to select data sources, identify master file data attributes, set up mapping rules and allow for compound data attribute matching. A suitable migration tool has to fulfil the following functions in order to enable the transition to an error free securities master file:

- Identification and selection of data sources

51. The identification and selection of data sources is a crucial input dependency for the mapping process. Basically two options can be considered.

52. One option is to define ex-ante the data attributes from each source available one intends to use in the master file. This implies to know beforehand which is the best source to deliver the required data attribute, hence omitting possibly better alternatives available. An additional shortcoming in this approach is that the database administrator is unaware of the degree of comprehensiveness of the master file as the end result is already determined before the actual mapping is initiated.

53. Figure 4 : Logical architecture of the CSDB system



54.

55. An alternative approach is to include all data attributes from all sources to be taken through the mapping application. This will allow the securities data reconciliation to be exhaustive and comprehensive. Additionally it permits data validation and data integrity checking by means of highlighting data differences and allowing for a specific cleansing process.

56. In the CSDB a mix of the two approaches will be applied: a strict selection of data attributes from selected sources (e.g. NCBS), complemented by a broader integration of information from the remaining data feeds (e.g. commercial sources). Experience gained in the second approach would assist in defining dynamically (i.e. by changing the rules for the data mapping without programming) the data attributes mentioned in the former.

57. Other tasks to be performed in the unification and mapping process are:

- identification along adequate keys;
- application of mapping rules;
- transformation to a common format.

### **Data cleaning and matching**

58. Once the several data sources have been unified via the outlined mapping process, the task consists in construction a unique, “cleaned” database with only one entry for each security / attribute. It consists of compounding all source data attributes for a given instrument mapped to a unique data attribute in the master file. This data cleaning is an iterative process monitoring validation, integrity checking and cleansing of the compound data attributes.

59. The most desirable and ideal situation is the one where all source data attributes yield the same value for the master file data attribute. In case the source data attributes are not equal in value a correction needs to be triggered. It has to be taken into account that a decision has to be reached on which source data attributes supply a “correct” or “most reliable” value. If it is not possible to detect a “most” reliable value a decision has to be taken not to use this specific source data attribute in the compound view. This process can be referred to as definition of “matching rules”.

60. As the “logical data differences” (i.e. data not meeting data integrity requirements) have been filtered out in the previous data mapping step, the remaining discrepancies may be described as “real data differences”. Logical errors and integrity violating data would have been set to a default value allowing to be recognised by the matching engine as void.

### **Estimating missing information**

61. Given the market valuation principle, the calculation of market prices or relevant proxies for securities for which the market value is missing is of importance<sup>15</sup>. In principle, the appropriate method to derive such approximations depends on the type of the security and the character of the missing value.

62. For **debt securities** (bonds, bills, etc.) there is a well-established set of calculation methods for securities for which a (minimum) number of basic attributes (principal, maturity, coupon, discount rate) is available. The implementation of these methods is straightforward in the sense that the necessary data and software packages are readily available in the market<sup>16</sup>.

63. However, things are less evident for **equities**. Equity market prices often reflect a whole bundle of influences - such as book values, profit expectations or mergers & acquisitions -related influence

---

<sup>15</sup> According to the CSDB – URD (User Requirement Document), securities are to be recorded with their market or transaction value as well as with their nominal value.

<sup>16</sup> For the CSDB a state-of-the-art software package will be implemented which allows the performance of a comprehensive set of financial statistical functions.



purposes - which differ significantly in their character. These factors are difficult to measure and most of them will not be sufficiently represented in the CSDB. This difficulty leads to a set of purely statistical methods to value equities and debt securities, for whom the basic set of attributes is not known (e.g. private placements).

64. Another important distinction can be made concerning the **nature of the missing value** (see figure 5). For the case that the missing value concerns only one or a limited number of observations (left-hand side) it may be possible to fill it by interpolation or by more elaborate statistical methods.

65. Figure 5: Possible Nature of missing values (white boxes) on a database

Time	Security 1	Security 2	Security 3	Security 4
t				
t-1				
t-2				
t-3				
t-4				

Time	Security 1	Security 2	Security 3	Security 4
t				
t-1				
t-2				
t-3				
t-4				

66. *Note:* the actual value of a security (Security 2) is only missing for one period, whereas it is available for the previous period (left hand-side). The value of a security is missing at all (right hand side).

More difficult cases are securities for which there is no information - except their existence – available (see right-hand side of the figure). This applies for instance to private placements or to barely traded securities. In this case one needs to identify a link between the security and other securities in the market and derive a value from here. This link will usually come through similarities in the structure of the security (in case that it is a debt security), through similarities via the issuer (comparison to other securities issued by the same issuer) or through the underlying asset / industry sector.

### Definition of quality benchmarks

68. The implementation of the data quality management requires the definition of transparent agreements (“benchmarks”) which the participants of the network have to respect. These agreements will ensure a level playing field for all participants of the network. The following types of benchmarks may be distinguished:

- Benchmarks for data transmission of each data feed (commercial, non-commercial, NCB,...) comprising a description of the elements of information (attributes, events), frequency, timeliness of data transmission and the degree of compliance with international standards.
- Benchmarks for monitoring the quality of the data finally stored (“cleaned”) in the CSDB, in particular the information under shared responsibility of NCBs and the operator of the CSDB. Such operational quality checks can either be performed (i) on attributes of individual securities or (ii) via comparison of (semi) aggregated statistics derived from independent sources;

- (i) the comparison of information for an individual (attribute of a) security derived from several sources can produce inconsistent and contradictory results;
- (ii) alternatively, quality checks can be based on (aggregated) statistics. In case results from independent sources (e.g. statistics on securities issues, MFI balance sheet statistics) or logically linked systems (e.g. recordings of stocks and flows) would be available, potential discrepancies might indicate the existence of wrong recordings in individual items of the database<sup>17</sup>.

69. The benchmark would apply to data on securities and related data issued by resident issuers of the jurisdiction of each NCB. However, as the CSDB is seen as a single set of data, the matching/cleansing process may be iterative as (authoritative or commercial) sources may show securities on international or “foreign” markets; where such securities are deemed having been issued by residents of a particular jurisdiction, they need to be checked by the concerned NCB in liaison with the operator of the CSDB. The organisational aspects of the data quality management are not yet fixed (this is an important task during phase 1 of implementation, in particular through the work of the Business Coordination Group) and may differ depending on what sources and/or tools are available in the NCBs. NCBs will either be able to process the data and send a clean set to the CSDB; or download data from the CSDB in a in-house securities database, check them and send them back; or proceed directly in the CSDB.

---

<sup>17</sup> The identification of individual problems requires of course the technical possibility to narrow down the scope of investigation step by step.