EIOPA'S REPORT ON DATA QUALITY IN SOLVENCY II REPORTING

EIOPA-BoS-22/336
06 September 2022

European Insurance and
Occupational Pensions Authority

CONTENTS

EXECUTIVE SUMMARY

Reliable data is the basis for successful data-driven supervision, evidence-based decision making and micro- and macro-prudential analysis. EIOPA is therefore constantly working to improve and assess the quality of the available insurance reporting data. After six years of Solvency II prudential reporting, this report describes the importance of data quality and the work done to improve it. A selection of individual data quality indicators together with an overall data quality score show a significant improvement in the quality of the Solvency II reporting data over the years. The data quality report also describes the intensive and wide usage of Solvency II data by EIOPA for internal analysis, supporting national supervisory authorities, publications and many other usages.
Despite the challenges of marking data quality quantitatively, it is observable a positive trend in the development of the Solvency II data quality, while the report also highlights the importance to further work improving data quality. Work on data quality is in fact an on-going and never ending task.

EIOPA received more than 130.000 individual and granular submissions between 2016 and 2020 for Solvency II reporting, including many corrective resubmissions (see chapter 3). This number underlines the importance of automated data quality processing and advanced analytic tool implemented by EIOPA to evaluate the received data.
An important tool to improve data quality and to ensure technical correctness is the XBRL taxonomy and its built-in validations. The report shows that the number of those very effective embedded validations keeps increasing and also provides examples on how taxonomy validations improve data quality significantly, even the non-blocking validations (see chapter 4).

The completeness and timeliness of the reporting to EIOPA has increased since the beginning of Solvency II in 2016. The key performance indicators (KPI) in the EIOPA annual work program are set at 97% by number of expected submissions and 99% by market share and it is recurrently fulfilled by the deadline at EEA level. Timely and complete reporting is crucial to fulfill the manifold information EIOPA's needs in time. At country level this KPI is not always fulfilled, mostly due to late submissions or outdated register information (see chapter 5).

The use of Legal Entity Identifier (LEI) ensures the correct and unique identification of insurance entities and their counterparties. The consistent use of the LEI increases the quality of the reported information, hence also its value for analysis. The use of the LEI has continued increasing significantly since 2016, but can still be further improved. The newest amendments to the Solvency II Reporting Implementing Technical Standard should contribute to this improvement in the future (see chapter 6).

A key area for data quality assessment is the unambiguous reporting of investments in the list of assets. This report describes the outcomes of comparing investments reported with the Central Securities Data Base (CSDB) and two additional data quality checks implemented by EIOPA. Important for the assessment of the CSDB comparison is as well the share of assets reported with ISIN code. The overall trend for all those data quality indicators is positive signaling an increased quality in the investments reporting. (see chapter 7).

The most relevant quality indicators described in this report are compacted into a single one based on EIOPA's judgement. Following a weighted combination of indicators[1] it provides an EIOPA's reference to the overall development of the Solvency II reporting quality. As a matter of fact, quality checks at EIOPA level are more available on the assets side and hence those are outweighed. This synthetic indicator gives for the first time a concentrated, measurable and comparable indication on the overall data quality in Solvency II reporting from EIOPA's perspective. In particular, this quality benchmark for annual solo reporting increased significantly from 82% in 2016 to 94% in 2020 (see chapter 8).

While the focus remains on EIOPA's view on insurance reporting quality, the report still gives a concise overview on the current situation on data quality in pensions reporting (see box on page 18).

[1] 40% Completeness ratio (by entity and by market share) + 25% CSDB score (by value and by count) + 15% LEI score + 10% ISIN check (by count and by value) + 5% CIC check + 5% rating check.

1. INTRODUCTION

With its introduction in 2016, Solvency II harmonised the supervisory reporting requirements for European insurance and reinsurance companies. The "third" pillar of Solvency II is since then fundamental for a risk-based and proportionate insurance supervision.

Since the beginning of Solvency II, data quality is a major topic for the supervisory community. It is a joint effort by undertakings, national competent authorities (NCAs) and EIOPA to improve the quality. EIOPAs single programming document 2022-2024 sets the Annual Operational Objective (3.13) to "Ensure harmonised reporting for the insurance and pensions sectors allowing for timely availability of adequate granularity, fit-for-purpose and high quality data to support NCAs and EIOPA's work on crisis prevention, financial stability, oversight, policy and consumer protection."

Based on the decision of the Board of supervisors, EIOPA receives the full set of quantitative reporting templates (QRTs) from all entities subject to Solvency II through the national supervisors. This puts EIOPA in the position to use a unique centralised database. This database is used for numerous purposes, including quality assurance, creating material added value for EIOPA, the national supervisors and external users.

EIOPA uses the Solvency II data for very different types of analyses, such as individual insurer or group assessments, statistical publications, financial stability and consumer protection studies and to substantiate policy technical advice and impact assessments. A tangible and regular added value of the centralised database for national competent authorities are the standardised reports on colleges, peer group analysis, cross-border business and data quality produced regularly by EIOPA.

Further to the use within the supervisory community, EIOPA supports external data users, like the European Commission, the ECB, the ESRB and the IAIS, with aggregated reports and the general public, industry, academia, etc. with the publication of sectorial reports and statistics based on Solvency II reporting data.

The main responsibility on the quality of the data reported for supervisory purposes lies on the supervised entities and it is intimately linked to their systems of governance in place. Persistent poor reporting quality can only be interpreted by supervisors as a sign of weak governance.

Solvency II reporting foremost purpose is insurance supervision, hence the core users of this data are the national competent authorities and oversight and supervisory experts in EIOPA. Therefore, national supervisors have a major interest in and a core responsibility for ensuring

high data quality. This report focuses on data quality of Solvency II reporting specifically from EIOPA's perspective.

The report focuses mainly on annual prudential solo reporting. Nevertheless, the same indicators can and are mostly implemented for group and quarterly reporting as well and are available internally at more granular level, whether at country or entity level.

This report shows that even if it is difficult to accurately measure data quality, the implemented indicators prove a significant increase in the data quality since 2016.

2. DATA QUALITY WORK BY EIOPA

Since the beginning of Solvency II reporting, data quality was identified as a major task for EIOPA. Such task has been undertaken from the understanding that a proactive use of the data constitutes a powerful impulse to improve its completeness and quality, provided that the appropriate processes are in place for issues detection and follow up. It is an ongoing effort by undertakings, national supervisors and EIOPA to improve completeness, reliability and usability of the Solvency II reporting data. In 2016, EIOPA introduced a first internal data quality project that described and identified processes and indicators to improve and measure data quality.

Important data quality checks and reports are EIOPA's completeness reports and additional data quality checks implemented in EIOPA's database. Data quality checks that can clearly identify erroneous reporting are added to the taxonomy whenever possible (see chapter 4). EIOPA has implemented additional data quality checks and outlier detection methods that do not necessarily flag erroneous reporting, but can also detect peculiar results only plausible for very specific business models or valuation situations which often deserve supervisory judgement and attention. For example, negative values for total liabilities or total assets on the balance sheet are reported and possible although rare. Adding non-blocking validations in the taxonomy and improving overall data quality, help detecting unusual data structures which leads to identifying possible erroneous reporting but also detecting potentially relevant observations for supervisory assessment.

In addition to the specific data quality work by data experts, all data users need to validate it against their individual expert knowledge, e.g. national supervisors. To follow up on those findings, EIOPA implemented a ticketing tool for structured sharing of potential data issues with national authorities which hold the direct contact with the reporting entities. The communication with undertakings and therefore the assessment on whether a value is indeed erroneously reported is done by national supervisors. EIOPA does not undertake unilateral amendments of the data reported by NCAs, instead will report back to the NCAs and expect a resubmission or an explanation from their side

An important process to verify completeness and correctness of part of the Solvency II data runs regularly involving all national competent authorities before each publication of EIOPA's insurance statistics on its website[1]. This regular comparison after EIOPA processing and aggregation allows spotting differences between the national and EIOPA's database, which are then investigated, explained or resolved before the publication. Hence, the publication of

[1] https://www.eiopa.europa.eu/tools-and-data/statistics-and-risk-dashboards/insurance-statistics_en

insurance statistics is not only an important source of information for the public, but its quality assurance process also serves as a basis for internal data quality verification. The publication of statistics progressively sooner after the reporting deadlines and with the clear motivation for improving granularity also leads to increasing data quality.
Given the integrated reporting approach followed for supervisory and statistical reporting to EIOPA and the ECB, a common understanding is required if and when a correction of data is necessary in case a report or parts of it are identified as erroneous[2]. EIOPA and ECB therefore conciliated and published Common Minimum Standards for Data Revisions.

An important and growing area for EIOPA's data work are methods to detect outliers in reporting datasets and unexpected changes in a time series (novelties) which benefit of statistical properties of larger reported datasets. Given the larger number of submissions to EIOPA (see chapter 3) compared to national samples, EIOPA is in an optimal position to developed tools to detect those types of outliers and novelties with machine learning and other advanced analytics algorithms. For this kind of applications, the larger centralised EIOPA database may bring additional value to the implementation of techniques eventually developed at national level. EIOPA is therefore in close contact with national supervisors and established an internal code sharing platform to further improve the collaboration and use of SupTech methods. The programs implemented by EIOPA, including machine learning algorithms in Python for outlier and novelty detection, are as well shared on this internal platform.

[2] https://www.eiopa.europa.eu/document-library/standards/common-minimum-standards-revisions-of-reported-data_en

3. KEY FIGURES ON REPORTING

The level of necessary automation on data quality assurance process and the constant nature of this work, is determined by the volume of the Solvency II reporting submissions EIOPA receives.

Between 01.01.2016 and 31.12.2021 EIOPA received 133,699 Solvency II individual submissions[3], i.e., on average 61 submissions per day. While there are obviously more incoming submissions before the reporting deadlines, NCAs can and do regularly send submissions and corrective resubmissions for all types of reporting and reference dates to EIOPA.

Table 1: Number of submissions received by EIOPA in 2016 to 2021, split by type of reporting.

Submissions received by EIOPA 2016-2021		**Solo**	**Group**	**Third country branches**	**SPVs**
Prudential reporting	**Day 1**	4,456	549	41	
	Quarterly	82,312	10,063	601	
	Annually	28,129	3,293	155	26
FS reporting	**Quarterly**	563	2,774	21	
	Annually	128	582	6	

Most of the submissions (82,312 concretely) refer to quarterly solo prudential reporting. For "Day 1" reporting, i.e., the on-time reporting obligation at the beginning of Solvency II, EIOPA received 4,456 submissions for solo and 549 for group reporting.

Of those submissions, 36,391 (28%) are resubmissions. Resubmissions correct and replace an already received reporting at EIOPA. Therefore EIOPA receives more submissions for one reporting type than the number of reporting entities obliged to report (Figure 6 in chapter 5 shows the number of resubmission received also after the reporting deadline).

[3] One submissions means the reporting file from one entity for one reference date and reporting type.

Resubmissions are, on one side, an important sign for improving data quality, as they show that a data quality issue was detected and corrected. On the other side, they show incorrect reporting in the initial submission. The distinction between submission and resubmission for this analysis is made from EIOPA's perspective[4]. Before sending the Solvency II reporting submissions to EIOPA, national authorities validate and check the submission received by the undertakings. Many data quality issues are therefore identified by national supervisors and corrected through new submissions before EIPOA is receiving the information. This report does not inform about the data issues solved by the national competent authorities before the initial submission to EIOPA[5], which are presumably abundant and based on their own national data quality assurance processes.

[4] The first submission for one reporting type, one reference date and entity ID is automatically flagged as initial submission, all following submissions for the same type reference date and entity ID are resubmissions. It is not based on template S.01.02 R0210 "Initial submission or re-submission", as this flag refers to the reporting between undertaking and NCA. The NCA might not send all submissions to EIOPA, in particular not if identified as erroneous.

[5] This is due to the lack of evidence at EIOPA's side.

4. TAXONOMY VALIDATIONS

EIOPA implemented the Data Point Model (DPM)[6] methodology and XBRL[7] taxonomy for reporting data to EIOPA. This methodology ensures technical correctness of the data, e.g. when a monetary value is expected, only a number is allowed and does not allow that a string can be reported instead. By blocking and impeding technically wrong inputs, it also ensures the uniqueness and validity of the incoming reporting data.

A key part in improving data quality are the validations embedded in the taxonomies. Those validations are published[8] and are used by undertakings to validate the reporting before sending them to the national supervisors. To improve data quality, the business validations are of upmost importance. EIOPA uses detected data quality issues as a trigger to introduce new or further improve the existing validations. With the implementation of those checks as blocking business validations, it is guaranteed that no submissions failing such a check can be submitted.

Figure 1: Number of business validations by taxonomy release

[6] The data point model (DPM)

A structured representation of the data, identifying all the business concepts and their relations, as well as validation rules. The data point model (DPM) implements the uniform and consistent definitions included in the implementing technical standards (ITS), guidelines and Board of Supervisors decisions on reporting and disclosure. It is composed of the annotated templates for the Solvency II and for the pension funds with the common DPM dictionary.

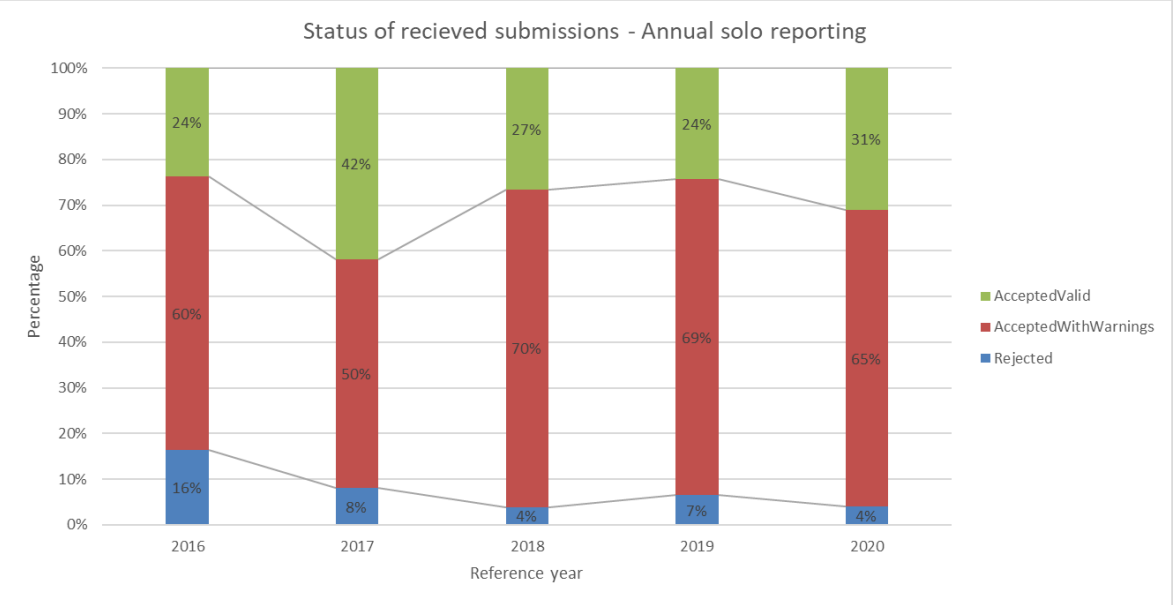
[7] The XBRL taxonomies and related artefacts

Implementing the technical data requirements and validation rules described by the data point model in the technical format of XBRL. The XBRL standard and software solutions are implemented by external parties (XBRL International and software vendors) to provide the software solutions to meet the defined requirements.

[8] https://dev.eiopa.europa.eu/Taxonomy/Full/2.6.0/S2/EIOPA_SolvencyII_Validations.xlsx

Most of those business validations are blocking. A submission violating one of those validations is automatically rejected. However, around one quarter of the validations in the 2.6.0 version of the EIOPA's solvency II taxonomy are only warning validations. A submissions violating one of the warning validations will be accepted, but the sender and receiver gets a notification that the reported data conflicts with the validation. Table 2 shows that for annual solo reporting most submissions are accepted with warnings, so at least one warning is triggered.

Figure 2: Share of submissions rejected, accepted and accepted with warning at EIOPA – based on annual solo prudential reporting.



The decrease of submissions accepted without any warning between 2017 and 2018 (resp. from 42% to 27%) is not necessarily a sign of decreasing data quality. It is impacted by the significant increase of validations from taxonomy version 2.2.0 to 2.3.0 (see Figure 1).

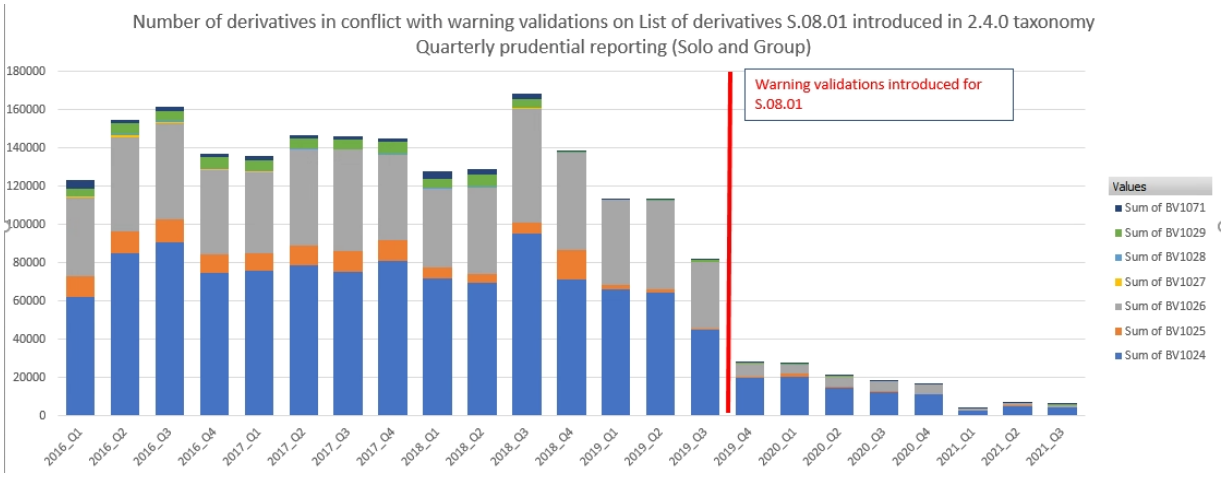
The relatively low number of rejected submissions is a consequence of the fact that EIOPA is receiving the submissions from the national supervisors. Before sending submission to the national supervisors, undertakings can validate the submission with the published taxonomy. In addition, the national supervisors assess and check the data quality before sending it to EIOPA. By nature, blocking validations work very well in stopping the submission of data not fulfilling the rule. However, the example in Figure 3 shows that also warning business validations have a significant positive impact on the data quality.

Taxonomy version 2.4.0 introduced new warning (non-blocking) validations on the list of derivatives.[9] Figure 3 shows the number of derivatives violating each of the validations, based

[9] BV1024: The item "Premium received to date" in S.08.01 - Open Derivatives should be positive.
BV1025: The item "Number of contracts" in S.08.01 - Open Derivatives should be positive.

on quarterly prudential reporting. The impact of the validations is clearly visible with their introduction for Q4 2019 reporting. Interestingly, the validations showed already positive impact before they were active, following the publication of the new taxonomy (drop in Q3 2019). This highlights the importance of validations (even if "only" warning) to improve data quality.

Figure 3: Example - Impact of warning validations



BV1026: The item "Contract size" in S.08.01 - Open Derivatives should be positive.

BV1027: The item "Maximum loss under unwinding event" in S.08.01 - Open Derivatives should be positive or equal to zero.

BV1028: Item "Swap outflow amount" should be positive or equal to zero.

BV1029: Item "Swap inflow amount" should be positive or equal to zero.

BV1071: The item "Contract size" in S.08.01 - Open Derivatives should be reported for futures and options (CIC '##A#' or '##B#' or '##C#' or '##F2').

5. COMPLETENESS & TIMELINESS

A key dimension on data quality is ensuring completeness and timeliness of the reporting data. Completeness has multiple aspects. It means that all undertakings with a reporting obligation must report and that EIOPA receives these submissions. Completeness also means that all necessary items within the submission are reported (e.g. all derivatives that the undertaking holds are reported or that cross-border business is reported when relevant). The focus of this chapter is on the first aspect, i.e. measuring if an undertaking with reporting obligation reported.

Timeliness measures if EIOPA received the submission before the respective reporting deadline. For data users it is important to know when the data is ready for the analysis. The completeness ratio is a key factor to decide when a data extraction for analysis can be launched. Multiple data reports are scheduled according to the deadline. It is a major improvement that EIOPA can plan with sufficient completeness, at least on EEA level according to the KPI, at the deadline.

Since the beginning EIOPA measures completeness and timeliness. National supervisors regularly receive completeness reports, informing them about the completeness ratio and the missing entities. The basis to identify the entities with reporting obligation is the insurance undertakings[10] and insurance group register at EIOPA. The measured completeness therefore also depends on the quality of the register. The content of the EIOPA register is maintained by the information communicated by national supervisors. Frequently there is a delay between the business status changes of a company, e.g. stop of operation, and the change in the EIOPA register. Such a delay in updating the register can lead to measured incompleteness. It is therefore important for the data quality that national authorities update the EIOPA register on time.

The correctness of the register for entities subject to Solvency II is checked for each reporting period by the completeness report described above. For Freedom of Services (FoS) and Freedom of Establishment (FoE Branches) entries, EIOPA implemented a data quality check comparing the entries in the register with the reporting in Solvency II[11]. The results are shared with national supervisor for revision and correction. Frequent mismatches are "reported branch in SII without branch in register", "reported FoS business without entries in the register" or branches that exist in the register but did not write any business according to the Solvency II reporting. The last is

[10] https://www.eiopa.europa.eu/tools-and-data/registers_en

[11] Based on cross-border business reported in template S.04.01

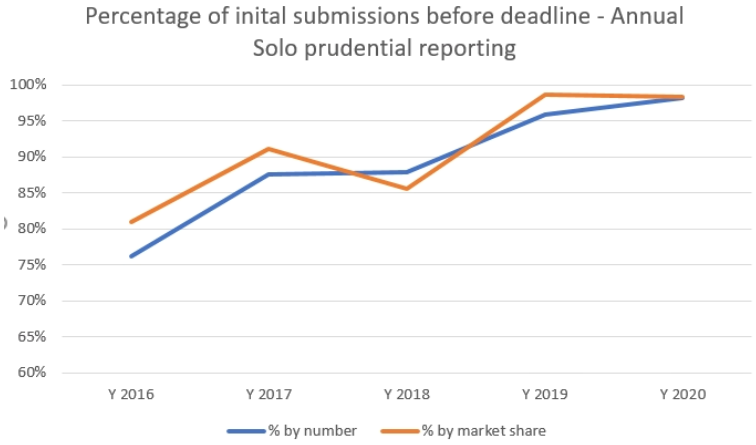
not necessarily a proof of erroneous reporting, but should be validated by the national supervisors.

For Solvency II reporting EIOPA has a key performance indicator (KPI) to measure completeness and timeliness in the EIOPA work program[12]:

Indicator	Data Source	Target
Data available in the Central Repository calculated as a percentage of number of insurance undertakings reported on with valid data by NCAs and market share of each country, as evidenced by the completeness ratio of technically valid reports of the Quantitative Reporting Templates (QRTs).	EIOPA Registers and information received and stored in the EIOPA Central Repository	2022-2024: 97% of the number of reporting entities or 99% of market share

The completeness at the reporting deadline shown in Figure 4 present the evolution of the KPI defined in the EIOPA programming document. The percentage by number shows for how many of the entities with reporting obligation EIOPA received the initial submission before the reporting deadline. The KPI of 97% by number of expected valid submissions or 99% by market share is reached at EEA level since 2019 for solo annual prudential reporting.

Figure 4: Completeness at the deadline by number of expected reporting entities for annual solo prudential reporting

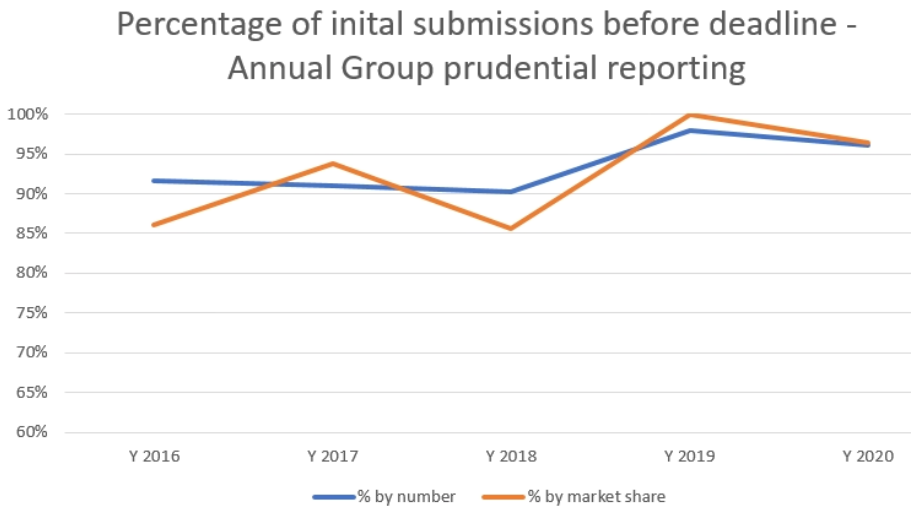


[12] https://www.eiopa.europa.eu/sites/default/files/publications/administrative/eiopa-bos-21-419-single-programming-document-spd-2022-2024.pdf

The completeness at the reporting deadline has significantly increased over the years, allowing EIOPA to use a complete set of annual reporting immediately after the deadline. EIOPA monitors the completeness degree at country level by number of submissions and market share. The market share is calculated using the total assets. While the KPI is reached on EEA level, it is not always fulfilled by all countries at the reporting deadline. The status of the completeness, including incompleteness, is regularly shared and presented at the EIOPA IT and Data Committee and the EIOPA Board of Supervisors.

Also for groups, an improvement of the timeliness of the annual prudential reporting can be observed. Given the significantly lower number of groups, the volatility of the completeness ratio is higher than for solos.

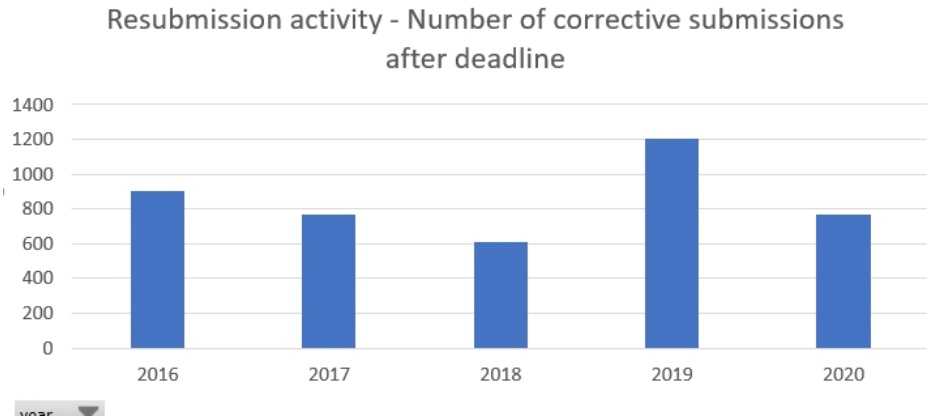
Figure 5: Completeness at the deadline by number of entities for annual group prudential reporting



EIOPA monitors the completeness and timeliness of reporting for all types of reporting, including Financial Stability reporting and all types of quarterly reporting. While on EEA level the KPI is reached in quarterly reporting, incompleteness on country level are regularly detected. The figures 4 and 5 show the completeness ratio at the deadline (timeliness). In case of incompleteness at the deadline, EIOPA contacts the national authorities and monitors changes of the completeness ratios at country level regularly after the deadline.

After the reporting deadline, EIOPA constantly receives corrective resubmissions (see chapter 2). For users it is therefore important to know that even if the reporting for a given type and date is flagged as complete, the data are not stable but can change due to corrective resubmissions.

Figure 6: Number of corrective resubmissions for annual solo prudential reporting after the reporting deadline.



Data quality in pension reporting

In addition to the insurance reporting, EIOPA receives since 2020 reporting in a structured format on Institutions for occupational retirement provision (IORPs). Although the present report focuses on the data quality of Solvency II insurance reporting, this box describes some key aspects on data quality in pensions reporting. Different to insurance reporting, the data received in IORP reporting can take the forms of individual or aggregated reports[1].

EIOPA's first data quality objective was to ensure that it actually receives the data: timely and complete.

- *Timeliness*: on the day one reporting, 60 percent of the NCAs failed to meet the set deadlines. Two years later, this ratio has dropped to 17 percent despite the deadlines being two weeks earlier.

- *Completeness*: the number of NCAs that were not able to deliver any information has halved. Unfortunately, this means that three NCAs[1] have not been able to report any data since implementation.

Equal to the insurance sector, **validations** embedded in the taxonomies are key to improve data quality. Based on assessment of the individual data points, how they relate in a single report and their plausibility as discussed with NCAs, EIOPA continues to add new validations to taxonomy. Furthermore, in the original taxonomy many validations were non-blocking. This should trigger a reaction from NCAs but avoids that data could not be submitted. Now a few years after implementation, EIOPA intends to change the severity of certain validations into blocking[1].

In addition to verifying the data quality of the fields in a single report, EIOPA also analyses the development of the data points over time and across countries. If sudden changes or **outliers** are detected, EIOPA follows up with those NCAs concerned. This procedure takes place every quarter before EIOPA confirms the data it intends to publish with all NCAs. If approved, EIOPA publishes these occupational pension statistics on its website[1].

Going forward, EIOPA aims to focus its additional data quality efforts on **IORPs investment** reporting and to develop a **data quality indicator for IORPs**. The main difference with the insurance sector data quality indicator shall be the interpretation of the quarterly completeness, considering the different requirements.

6. LEGAL ENTITY IDENTIFIER (LEI)

To link information coming from different data sources, to combine information and to build connections and networks, it is crucial to uniquely identify entities. For the use of Solvency II reporting data it is therefore important that the Legal Entity identifier (LEI) is used whenever existing.

To further increase the use of LEI, EIOPA published in 2021 revised guidelines on Legal Entity identifier. The Guidelines[13] continue to establish consistent, efficient and effective supervisory practices by harmonizing the identification of legal entities in order to ensure high quality, reliable and comparable data. Having such data contributes to:

a) better supervision and oversight of financial institutions as well as improved regulatory policies and decision making process;

b) identifying, assessing, monitoring and reporting risks to the financial stability of the European insurance and pensions sectors;

c) supporting overall EIOPA's work on crisis prevention, financial stability, oversight, policy and consumer protection.

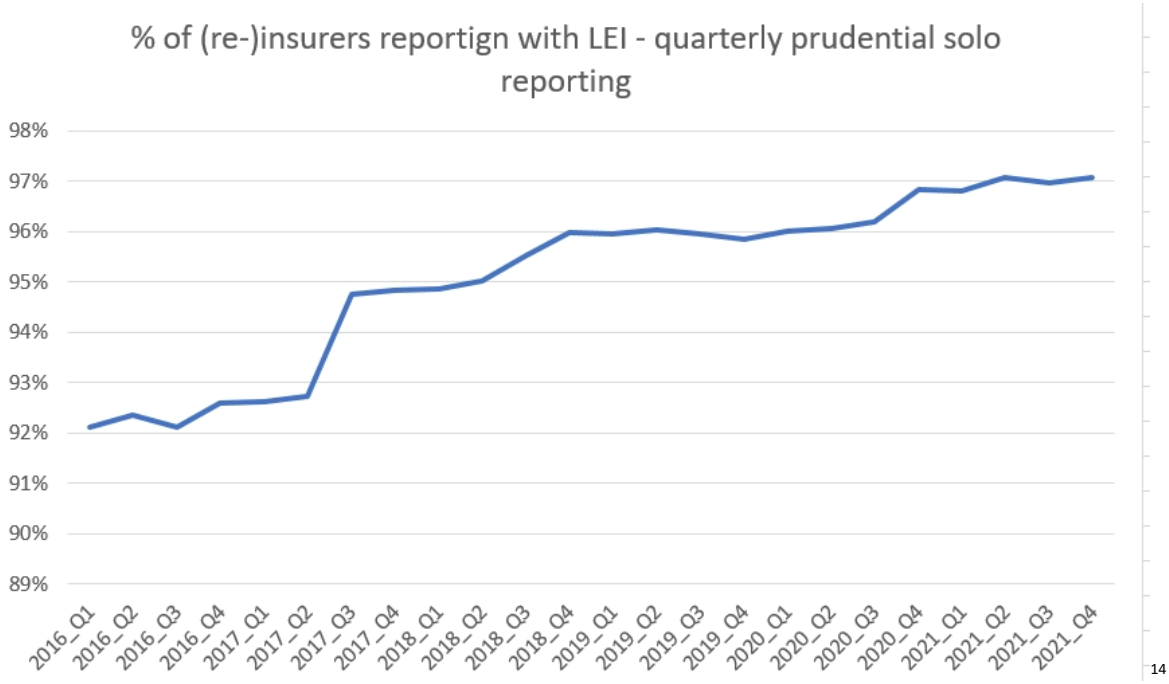
From end 2023 on, the Implementing Technical Standards on reporting will make the use of LEI mandatory when addressing entities covered by EIOPA LEI Guideline.

The use of LEI codes to identify asset issuer, reinsurer and other counterparts significantly increases the usability and value of the information. Concentrations, interlinkages and networks can be more reliable, in a higher quality and more automatically analysed. The use of the name of an entity for identification is very resource intensive (as mostly manual interaction is required) and not reliable.

One aspect to assess the value, and thereby quality of the provided information for supervision, is to compare the use of the LEI over time, between undertakings, countries and templates. Upmost, this is important for the reporting insurance undertakings and groups itself. The share of Solvency II entities that report with their LEI code increased significantly since 2016.

[13] https://www.eiopa.europa.eu/document-library/guidelines/revised-guidelines-legal-entity-identifier_en

Figure 7: Share of insurance and reinsurance undertakings reported with LEI in quarterly solo prudential reporting, Q1 2016 to Q4 2021:



The use of the LEI is increasing constantly. However, still not all insurance undertakings subject to Solvency II report their LEIs. The remaining ~3% without LEI are nearly all from one country. EIOPA developed an "LEI score indicator" to monitor the use of the LEI in the Solvency II QRTs. It is based on the following items:

- S.01.02 R0020 Undertaking identification code and type of code
- S.06.02 (if CIC category not in 8, 9 or CIC subcategory in 71 and 75)
  C0210 Issuer Code
  C0250 Issuer Group Code
- S.08.01 C0270 Counterparty Code
  C0340 Counterparty_Group_code
- S.10.01 C0080 Counterparty code and type of code
- S.30.02 C0280 Code and type of code of the reinsurer
- S.30.04 C0180 Code and type of code of the reinsurer
- S.31.01 C0160 Code and type of code of the reinsurer

[14] Without UK for full timeline for comparability.

This list does not include all cells where the LEI needs to be reported if available. For some cells the use of LEI is too rare and the template used by only a limited number of undertakings making the comparison imbalanced[15].

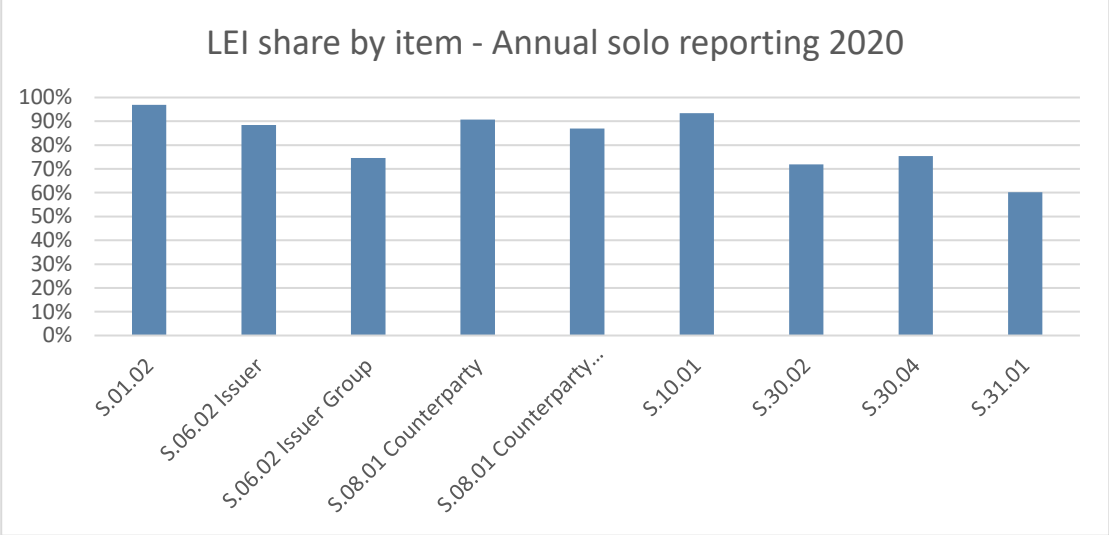
For all items listed above the LEI share is calculated as follows:

LEI share (template) = $\frac{Number\ of\ reported\ items\ with\ LEIs}{Number\ of\ reported\ items}$, where LEI is identified by the Code starting with the characters "LEI". The correctness of the LEI is not validated.

For the list of assets, the nominator and denominator are filtered to exclude CIC 71, 75, 8 and 9, as issuer code is not applicable for those asset categories.

The total LEI score indicator is then calculated as the unweighted mean of the individual LEI shares.

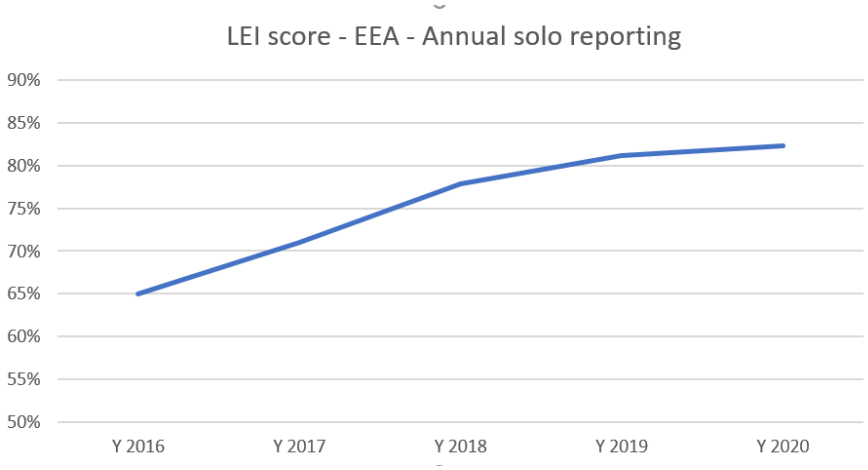
Figure 8: LEI share by cell annual reporting solo 2020



For some cells, the LEI coverage is quite high, for others, namely the identification of the reinsurer, it needs to be further improved. For example, in S.31.01 - C0160 Code and type of code of the reinsurer, there is still a significant number of reinsurers with location in the EEA reported without an LEI. This limits the use for supervision, e.g. when looking at reinsurance networks and concentrations.

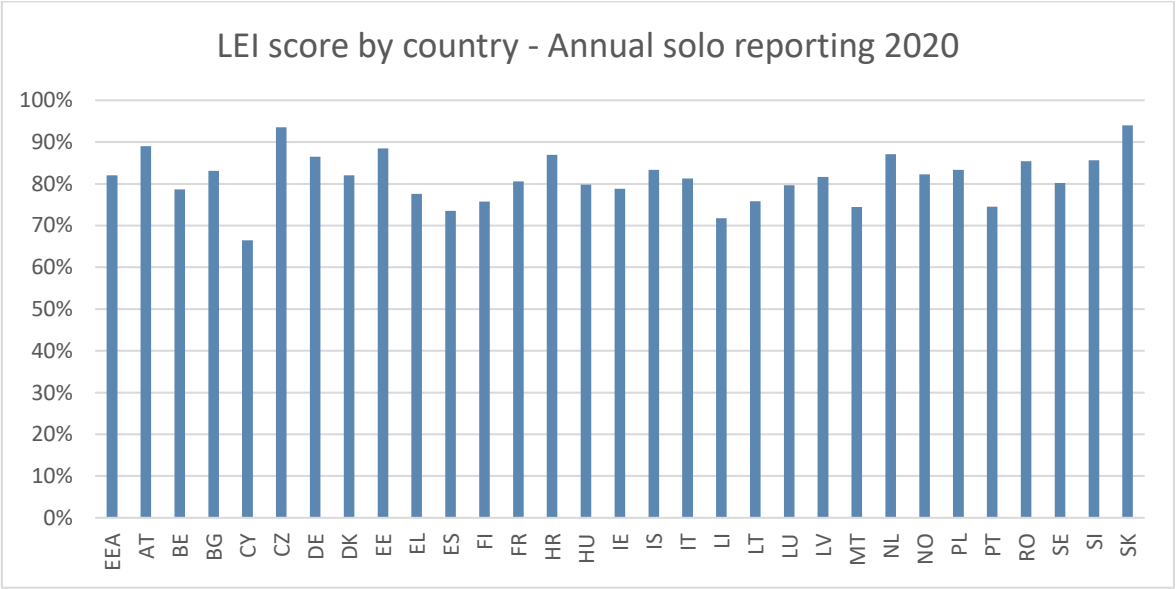
[15] One example of such a cell is S.03.02 Off-balance sheet items - List of unlimited guarantees received by the undertaking – C0030 Code and type of code of provider of guarantee.

Figure 9: Development of LEI score from 2016 to 2020



The LEI score – the unweighted average of the LEI shares listed above- has increased significantly since 2016 (from 65% to over 82%), reflecting an important improvement in the quality of Solvency II reporting.

Figure 10: LEI score by country



The more granular split of the LEI score by undertaking and by country (Figure 10) shows relevant differences. However, an improvement in the LEI scores from 2016 to 2020 has been noted for all countries.

For the interpretation of the LEI score, it is important to highlight that it is not mandatory that the LEI is reported for all items. Only if an LEI is available, it needs to be reported. It should

therefore not be expected that a score of 100% can be reached. The share of entities, not only insurances, with an LEI, diverges heavily by size, market and home country. A percentage different from 100% does not mean erroneous reporting. However, a lower ratio means that the usefulness of the reported information is weaker in this specific aspect.

7. CHECKS ON INVESTMENT REPORTING

An important area for data quality is the list of assets, template S.06.02. To assess the data quality of the reported assets, EIOPA has access to the Centralised Securities Database (CSDB) from the ECB. The CSDB contains information on over five million debt securities, equities and mutual fund shares/units.[16] Every month EIOPA receives a copy of the CSDB. EIOPA uses the CSDB to assess the correctness of the received Solvency II reporting. The CSDB score measures the scale of matches between the attributes received in SII reporting with the ones from CSDB.

Assets reported in Solvency II are linked to the CSDB by using the ISIN code of the asset. Before looking in detail at the results of the CSDB score, it is helpful for the interpretation of the results - in particular when looking at undertaking level - to measure the share of assets reported with and without the identification by ISIN code.

ISIN COVERAGE

The ISIN coverage is the share of assets from the list of assets template S.06.02 with an ISIN[17], measured both by number of assets and by value[18]. It is calculated as the count (or sum by value) of all assets with ISIN code, divided by the count (or sum by value) of all assets. Assets with CIC category 71, 75, 8 and 9 are not included in the calculation. [19]

The ISIN coverage by number of instruments slightly increased from 2016 to 2020 from 81% to 83% and a bit more attending to the value of those instruments (from 73% in 2016 to 78% in 2020).

There are multiple reasons why an undertaking, or a group of undertakings, are investing in assets without ISINs. A lower ratio does not imply data quality issue by its own and it does not necessarily mean wrong or incomplete reporting. However, it does mean lower usefulness to assess data quality. For example, it cannot be validated against external data sources like CSDB, or combined with the reporting from other companies. The ISIN coverage is instrumental providing additional information on the list of assets which exponentially increases the explanatory power of the CSDB score.

[16] https://www.ecb.europa.eu/pub/pdf/other/centralisedsecuritiesdatabase201002en.pdf

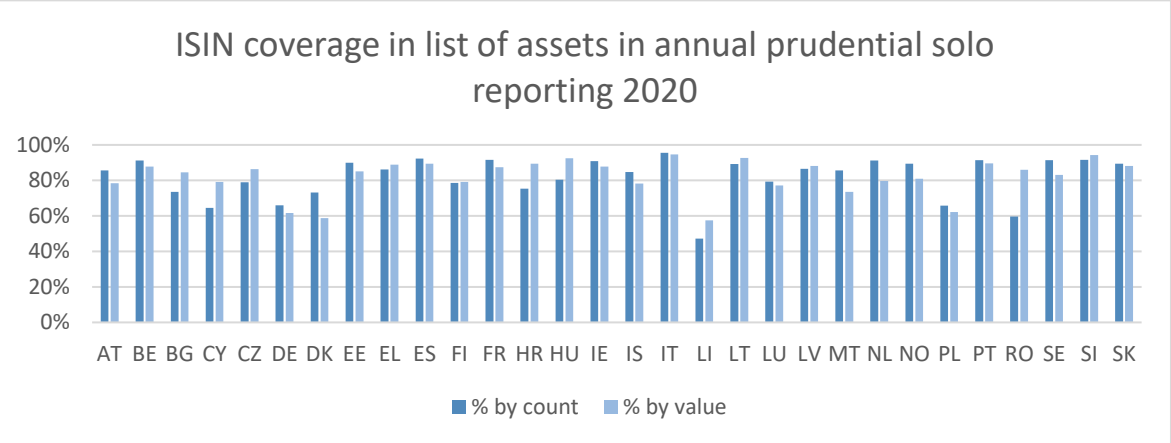
[17] C0040 "Asset ID Code and Type of code"

[18] Using the "Solvency II amount" C0170

[19] The indicator for annual reporting includes automatically the reporting of the list of assets from Q4, in case the undertaking did not submit the list of assets in annual prudential reporting due to no material changes to Q4 reporting.

One well justified example of significant investments in asset type without an ISIN are registered bonds (Namensschuldverschreibung) and bonded loans/promissory notes (Schuldscheindarlehen), mainly used in Germany. Those assets are correctly classified as Government bonds or Corporate bonds (CIC 1 and 2), but are not listed and can therefore not have an ISIN.

Figure 11: ISIN share by country



The ISIN share differs heavily by undertaking level and country. As described above, this is caused by specific asset categories in some markets and is thus not a sign of erroneous reporting. The possibilities for supervisors to monitor the correctness of the reported values, and to link the information with other data sources, is limited for undertakings with very low ISIN coverages.

The consistency of the identification code of an investment over time is also important for supervision. For assets reported with a code created by the reporting insurer this time series is more difficult to validate.

CSDB COMPARISON

The CSDB helps EIOPA to better measure and improve the quality of the reported attributes in the list of assets. As described, this is done for all assets reported with an ISIN code.

The CDSB comparison is a key performance indicator in EIOPA's work programme[20]:

Indicator	Data Source	Target
Data quality of Solvency II information available in the Central Repository, calculated based on the agreed methodology.	EIOPA Central Repository supported by Central Securities Data Base (CSDB)	2022-2024: List of assets (using assets value): ratio of good quality data higher than 93.5% (2021: 93%) Overall quality (using market share): ratio of good quality data higher than 93.5% (2021: 93%)

The CSDB score compares the match between the reporting in S.06.02 List of assets and the CSDB. The CSDB score is provided as a ratio of matches by number of assets and by value of the assets (C0290 Solvency II value). While also CSDB might not be in all cases correct, it is a very valuable database for comparison.

CSDB score calculation:

Simple average of the following 8 checks for matching the list of assets (S.06.02) with CSDB:

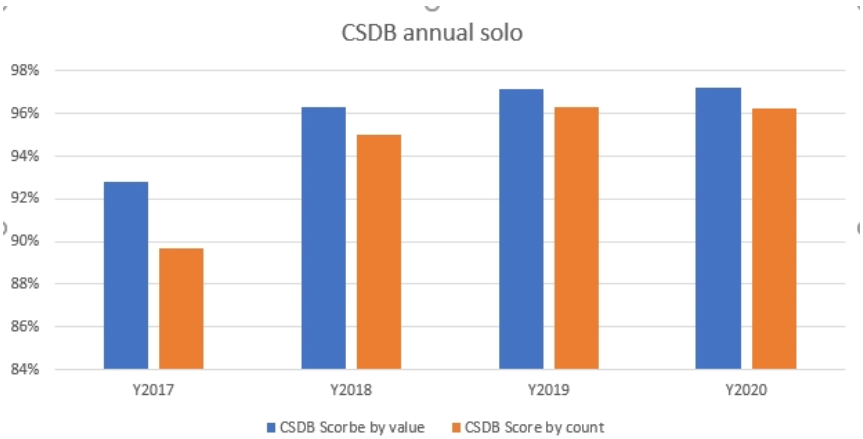
- ISIN (share of ISIN found in CSDB / all ISINs reported)
- Issuer LEI (excluded for CIC Category = 4),
- Issuer country (excluded for CIC Category = 4),
- Asset currency,
- Bond Duration (with 10% tolerance and more than a 3-month deviation gap) – half weight,
- Maturity date (with 10% tolerance) – half weight,
- Accrued interest (with 100 bps tolerance),
- Price (with 5% tolerance).

[20] https://www.eiopa.europa.eu/sites/default/files/publications/administrative/eiopa-bos-21-419-single-programming-document-spd-2022-2024.pdf

Figure 12: CSDB development in quarterly reporting from 2016 to 2021, by count and value

The increase in quarterly reporting in the CSDB score is another important indicator showing the steady improvement of Solvency II data quality. [21]

Figure 13: Development of the CSDB score in annual solo prudential reporting

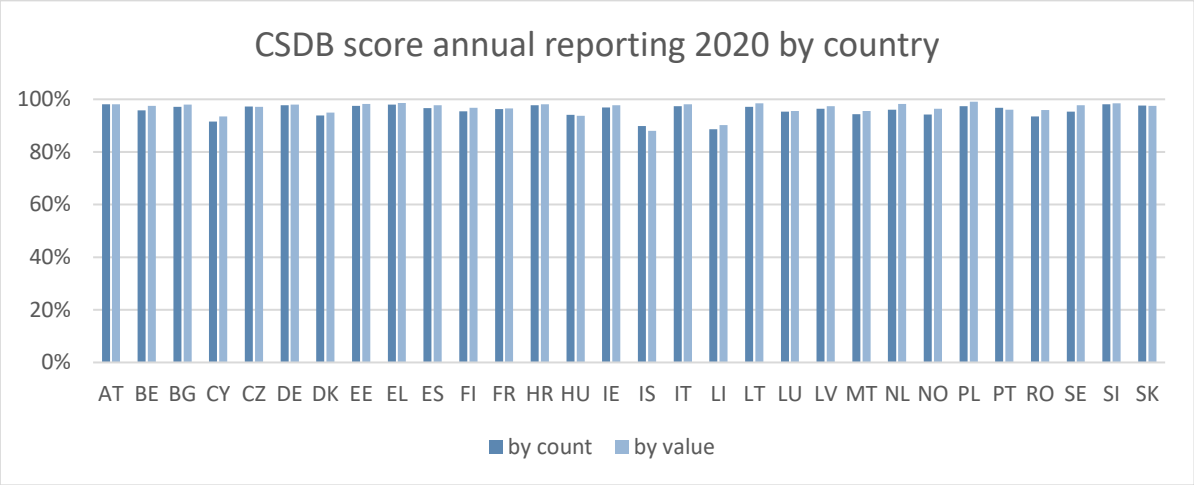


The KPI for the CSDB match is fulfilled for quarterly and annual reporting at EEA level. At undertaking and country level, the CSDB score diverges. Some are close to 100% match between

[21] The benchmark in this chart is the value when the indicator was calculated for the first time, based on end-year 2018 values, for visualization purposes. It is different to the KPI in EIOPA annual work plan.

SII and CSDB, for others the comparison with CSDB can trigger further assessment on the underlying reasons.

Figure 14: CSDB score for annual solo prudential reporting 2020 by country



EIOPA calculates the CSDB score at EEA, country and undertaking level. The outcome of such assessment at country and undertaking level are regularly shared with the national supervisors.

EXAMPLES OF ADDITIONAL DATA QUALITY CHECKS

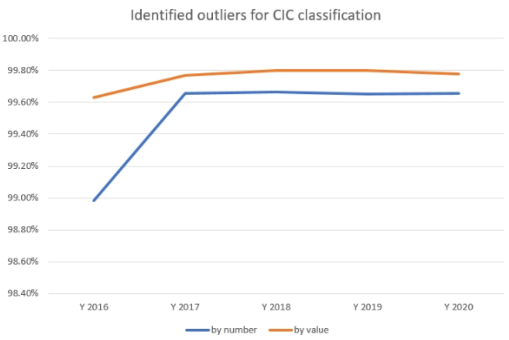
Besides the CSDB comparison, EIOPA is carrying out multiple other data quality checks on the list of assets. As described in Chapter 2, the focus is also on the implementation of data quality checks with an added value of using the central EIOPA database.

Comparison on the CIC Asset category

Undertakings need to identify for each asset the Complementary Identification Code (CIC code), as set out in Annex VI — CIC Table of the Implementing Technical Standard (ITS) on reporting. When classifying an asset using the CIC table, undertakings shall take into consideration the most representative risk to which the asset is exposed to. This CIC category (3-rd digit)[22] classification is the basis for the allocation on the balance sheet, but also for nearly all supervisory analysis on the list of assets. As a data quality check to detect possible misallocation of an asset to the CIC category, EIOPA implemented a program that compares for each ISIN reported by multiple different undertakings, if the CIC classification is aligned or if there are outliers. An outlier is identified if at least 75% of the other undertakings report the same asset with another CIC category.

[22] CIC Categories: Government bonds, Corporate bonds, Equity, Investment funds CIU, Structured notes, Collateralized securities, Cash and deposits, Mortgages and loans, Property, Other investments.

Figure 15: Result of CIC comparison check

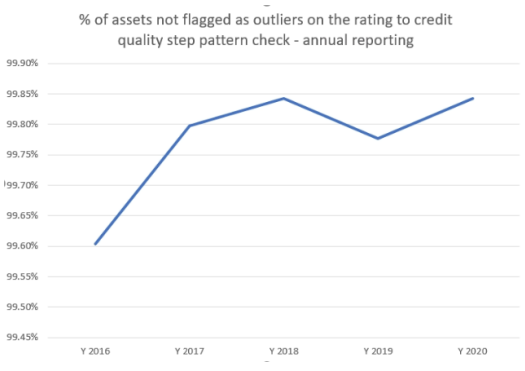


As with many data quality checks, the findings should not be considered automatically as erroneous reporting, but might identify cases of wrong classification. The results of this check at undertaking and asset level are regularly shared with NCAs.

To measure the development over time the findings are compared to the number of all reported assets, by number[23] and by value[24]. The main change at EEA level is a major improvement from 2016 to 2017, since then the level shown in Figure 15 is quite high with only 0.02% of the assets by value flagged as potential outliers by the described method.

Another example for additional data quality check, using the added value of having a larger centralised database at EIOPA, is a pattern recognition code identifying automatically outliers in the credit quality step assignment, as a result of the combination of the rating[25] and rating agency[26]. An asset is flagged as outlier by this check, if at least 90% of assets with the same rating and rating agency have another credit quality step allocation[27]. The identified rules and outliers for each country are shared regularly with the national supervisors. The number of identified exceptions (outliers) in comparison with the total number of reported assets can be used to monitor this aspect of data quality, at undertaking, country and EEA level.

Figure 16: Result of CQS pattern check



At EEA level, the number of identified exceptions in comparison with total number of assets is decreasing and showing still a number of assets with potentially erroneous data quality steps, but overall an improving data quality.

For both checks described above, the percentage of assets identified are very small. The number (> above 10,000 assets for both annual and quarterly reporting) shows how important automatisation is for EIOPA and how difficult an individual assessment of the result would be.

[23] Count of identified outliers / Count of all reported assets in list of assets

[24] Sum of Solvency II value of identified outliers/ Sum of Solvency II value of all assets in list of assets

[25] External rating (C0320)

[26] Nominated External Credit Assessment Institutions ('ECAI') (C0330)

[27] The python code used by EIOPA is based on code developed by DNB: https://github.com/DeNederlandscheBank/data-patterns

8. OVERALL DATA QUALITY INDICATOR

Chapter three to seven in this report explain data quality measures and indicators for specific aspects. The combination of the described data quality indicators and scores is an overall data quality indicator. While the absolute value of the overall data quality indicator is less meaningful by itself, the change over time gives a good indication of the development of the data quality. EIOPA implemented the overall data quality indicator with the following weighting:

Overall data quality indicator =

40% Completeness (20% completeness by entity + 20% completeness by market share) +

25% CSDB (12.5% CSDB score by value + 12.5% CSDB score by count) +

15% LEI score +

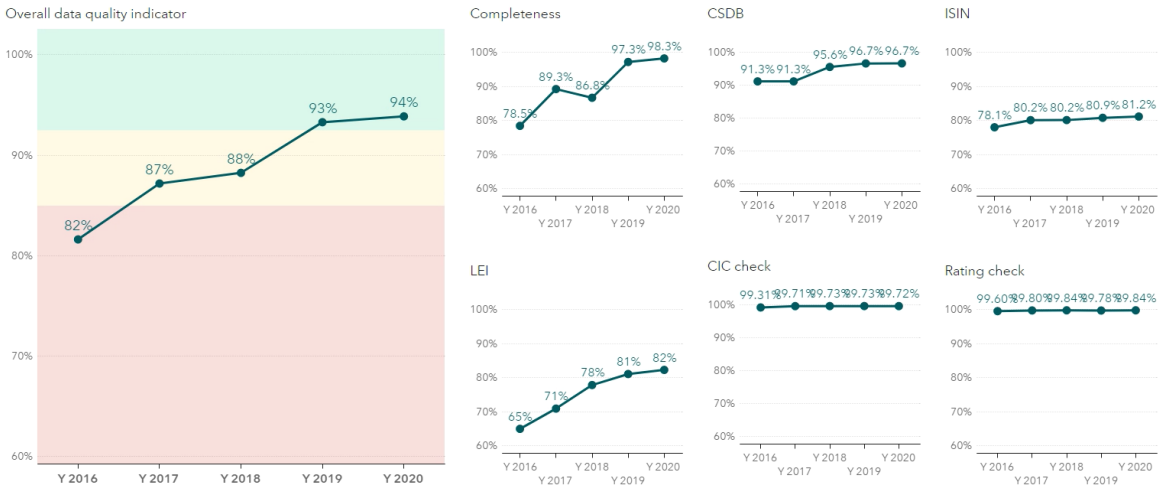
10% ISIN (5% by count + 5% by value) +

5% CIC check +

5% rating check

The weighting is based on expert judgement. The completeness has the highest weight as non-reporting or submissions rejected by taxonomy is the weakest data quality. The CSDB score - as a comparison against a reliable benchmark database – has the second highest weight. The ISIN share has a lower weight, as it cannot directly be considered as a sign of erroneous reporting (see chapter 6). The weighting and composition in detail can be adjusted in future.

Figure 18: Screenshot from EIOPA data quality dashboard – Overall data quality indicator and modules



The positive observation is that all aspects in the data quality indicator improved from 2016 to 2020. As a result also the overall data quality indicator improved significantly.

A weakness of this overall data quality indicator is the focus on the investment side. The CSDB score, ISIN share, CIC and rating check and part of the LEI score are all based on the list of assets. This is a consequence of the granularity and comparability of the reporting in list of assets. For the liability side, similar data quality indicators are more difficult to define. The data quality checks that can reliably be automatically calculated in closed templates[28] can in most cases be implemented in taxonomy. This has clearly the best impact on the data quality, but can then not be used for data quality indicators in a timeline anymore. The open tables on the liabilities side are more prone for expert judgment analysis rather than quantitative data quality reports. In any case EIOPA will continue exploring new technologies to assess more automatically also quality of information reported liabilities.

To conclude, the indicators and the day-to-day experience working with data show increasing data quality in Solvency II reporting. However, data quality is still a major issue and needs to be further enhanced. Permanent intensive effort is needed by reporting entities, national supervisors and EIOPA to further improve the quality of the Solvency II reporting data.

[28] Closed templates are templates with defined number of rows and columns, Open templates, like the list of assets, are templates with open number of rows.

EIOPA

Westhafen Tower, Westhafenplatz 1
60327 Frankfurt – Germany
Tel. + 49 69-951119-20
info@eiopa.europa.eu
https://www.eiopa.europa.eu