IFC Workshop on "Addressing climate change data needs: the global debate and central banks' contribution"

6-7 May 2024

Central bank's approaches to fill data gaps on green debt securities: case of Indonesia[1]

O Aruda Lisjana, H Prasnawaty, N Aisyah Safitri, D Sulliati and
A Widjanarti,
Bank Indonesia

---

[1] This contribution was prepared for the workshop. The views expressed are those of the authors and do not necessarily reflect the views of the Central Bank of the Republic of Türkiye, the BIS, the IFC or the other central banks and institutions represented at the event.

Central Bank's Approaches to Fill Data Gaps on Green Debt Securities: Case of Indonesia

Dini Sulliati [1], Oktefvia Aruda Lisjana[2], Nur Aisyah Safitri[3], Anggraini Widjanarti[4], Herina Prasnawaty D.[5]

Climate change has become one of the main issues in Indonesia. Low Carbon Development and climate resilience in Indonesia has incorporated into the National Long-Term Development Plan (RPJPN) for 2025-2045, making emissions reduction a concrete target for sustainable development. In accordance with the Paris Agreement, Indonesia commits to achieve net zero emission in 2060 or earlier. The Ministry of Finance has issued Green Sukuk and Sustainable Development Goals Bond as a form of the government's commitment in financing sustainable development. Moreover, The Financial Services Authority (FSA) supports the government policy by issuing some regulations related to the issuance of debt securities in financing the green project (green bond). However, currently in Indonesia, there are data gaps on securites-by-securities data for green, sustainability, and sustainability-linked bond in terms of type of issuer-holder, currency, maturity, and market issuance.

Central bank plays a critical role in collecting green financing data to support policy recommendations. Thus, we need to answer the challenges on these data gaps to address climate change data that policy maker needs. This paper proposes a different approaches to fill data gaps by utilizing data mining to classify and identified bonds into three categories: green, sustainability, and sustainability-linked bonds. We also proposes a framework to create data partnership to enrich green debt securities data. We believe with combination of these approaches will help the acceleration to fill data gaps needed.

Keywords: Green Debt Securities, Green Finance, Central Bank, Machine Learning, Text Mining, Data Partnership, Institutional Collaboration

---

[1] Statistics Department, Bank Indonesia, email: dini_s@bi.go.id

[2] Statistics Department, Bank Indonesia, email: oktefvia_aruda@bi.go.id

[3] Statistics Department, Bank Indonesia, email: safitri@bi.go.id

[4] Statistics Department, Bank Indonesia, email: Anggraini_widjanarti@bi.go.id

[5] Statistics Department, Bank Indonesia, email: herinapd@bi.go.id

I.   Introduction

In the last few decades, climate change has become a main topic in environmental issues and a very serious global threat after the pandemic. Economic growth is often used as an excuse for development by ignoring environmental aspects. The pressure to create high economic growth drives industrialization, excessive use of fossil fuels, and massive deforestation which creates carbon emissions. Therefore, a sustainable development strategy is needed, by pursuing economic growth in line with environmental conservation and responding to climate change issues. Globally, sustainable financing is considered as a key instrument to support the transition to a more resilient economy. One of instrument of sustainable financing is debt securities.

The availability of a valid, reliable, and granular data of sustainable financing is needed to support the formulation of appropriate policy to achieve the sustainable financing goals and address climate-related risks and opportunities. G20 Data Gaps Initiatives (DGI) phase 3 has identified that so far there has been a lack of reliable and comparable indicators for tracking the growth in green debt securities and green equity listed shares instruments across the G20 economies.

In Indonesia, low carbon development and climate resilience has incorporated into the National Long-Term Development Plan (RPJPN) for 2025-2045, making emissions reduction a concrete target for sustainable development. In accordance with the Paris Agreement, Indonesia commits to achieve net zero emission in 2060 or earlier. In the latest Nationally Determined Contribution (NDC) document, Indonesia increased its emission reduction target to 31.89% in 2030 with an international support target of 43.20%. Based on Indonesia's Green Climate Fund Country Programme Document published by Ministry of Finance (2021), there will be mitigation and adaption financial gap in achieving NDC targets during 2020-2030 in amount of USD 206 billion. Therefore, more financing sources are needed to close this gap either from the government or private sector. Banga (2019) also stated that developing countries require to explore new financing mechanisms, such as green bonds in order to fulfill its commitments on NDC target.

In 2017, the Indonesian government issued regulations on green bonds and green sukuk, then launched a $3 billion green sukuk in 2018. The value of sustainable finance in Indonesia as of third quarter of 2021 reached USD 55.9 billion or Rp809.75 trillion (exchange rate Rp14,440/USD). The issuance of green bonds in the domestic market was $35.12 million (Rp500 billion), representing 0,01% of the total outstanding bonds issued in the domestic market (Supriyanto, 2023). However the issuance of green bonds in Indonesia seems fluctuated with the latest figures on 1st semester of 2023 in the amount of US$1.15 billion, after achieving US$1.83 billion on 1st semester of 2022 (Triawan and Nagari, 2023)

The Indonesian government via Ministry of Finance has issued Green Sukuk and Sustainable Development Goals Bond as a form of the government's commitment in financing sustainable development. Moreover, the Financial Services Authority (OJK) has been encouraging sustainable financing practices through sustainable financial policies contained in the Sustainable Finance Roadmap Phase 1 (2015-2019) and Phase 2 (2021-2025), OJK Regulation no. 51/2017 concerning the Implementation of Sustainable Finance and the Indonesian Green Taxonomy (THI). In line with the THI program, OJK also issued a policy regarding securities issuance with the latest regulation no. 18/2023 concerning Issuance and Requirements of Sustainable Debt Securities and Sukuk. The categories of the debt securities in the regulation includes green bond, sustainability bond, and sustainability-linked bond.

However, currently in Indonesia, there are data gaps on securites-by-securities database for climate-related financing i.e. green, sustainability, and sustainability-linked bond in terms of sector of issuer, currency, maturity, and market issuance. Moreover, there has been no flagging yet on those bonds categories issued in the domestic market, and those issued in foreign markets have not been captured perfectly. Meanwhile the 2015 Paris Agreement encourages the immediate need for related sustainable financing statistics, especially green bonds, equities, and invesment funds.

Central Bank plays a critical role in collecting granular sustainable financing data to support policy recommendations. According to the Network of Central Banks and Supervisors for Greening the Financial System (NGFS, 2022), climate-related data gaps include several dimensions: the availability (i.e. coverage, granularity, accessibility), reliability (i.e. quality, auditability, transparency), and comparability. In some cases, the relevant data are not available, while other cases, the data exist but not granular. Thus, Central Bank need to answer the challenges on these data gaps specifically in the green debt securities to address climate change data that policy maker needs. The purpose of this paper are:
1. Propose a different approach to fill data gaps by utilizing text mining to classify and identified bonds into categories: green bonds, sustainability-bonds, and sustainability-linked bonds (the case of Indonesia).
2. Proposes a framework to create data partnership to enrich green debt securities data (the case of Indonesia).

II. Literature Review

II.1 ESG and Green Financing

Broad concept of Environmental, Social, and Governance (ESG) financing is financing for "activities or projects that sustain or improve the condition of the environment or society of governance practices" (OECD). According to Höhne at al. (2012), green financing is a broad term that can refer to financial investments flowing

into sustainable development projects and initiatives, environmental products, and policies that encourage the development of a more sustainable economy. It includes climate finance but is not limited to it. The main instruments of ESG and green financing are debt securities, loans, and equity and investment fund shares. Debt Securities can be in the form of bills, bonds, notes, negotiable certificate of deposit, commercial paper, and similar instruments normally traded in the financial markets that serve as evidence of a debt (Handbook of Securities Statistics - IMF, 2015). Equity can refer to company shares that are oriented towards environmental sustainability in the long term, this can be assessed by the output produced by the company which has an impact on the environment.

In the issue note of sustainable finance definition in the 2025 SNA and BPM7 (2024), ESG (green) debt securities defined as negotiable financial instruments serving as evidence of debt in which the use of the debt security is restricted to finance or refinance activities or projects or where the issuer agrees to achieve performance objectives that improve the condition of the environment or society or governance practices (of the environment). These include social, green, sustainability, sustainability-linked, and other ESG debt securities. Social debt securities are not included in the DGI-3 Rec 4 because the focus is on climate-related financing. OECD (2020) emphasizes the importance of including sustainable finance statistics especially green bonds, equities and investment funds in the financial accounts and balance sheets of the national accounts for tracking investment in the green and climate/ transition economy and informing decisions on monetary and fiscal incentives.

II.2 Green Bond, Sustainability Bond, and Sustainability-Linked Bond

According to Banga (2019), in developed and emerging countries, the issuing of green bond are quickly growing to finance the adaption and mitigation projects due to factors such as yield to maturity similarity with the conventional bond, investors' climate awareness, policy makers' commitment to counter climate change and the current macroeconomic environment. Meanwhile in developing countries, a set of institutional and market barriers are preventing from appropriating the full benefits of green bonds.

Ehlers & Packer (2017) defined green bonds as fixed income securities which finance investments with environmental or climate-related benefits. This specific purpose of issuing green bonds differentiates them from the conventional bonds. In Indonesia, the definition of ESG debt securities for climate-related financing are stated in OJK Regulation No. 18/2023 as follows:
1. Green bonds and/or green sukuk are debt securities and/or sukuk whose proceeds from the issuance are used to finance or refinance environmentally sound business activities.
2. Sustainability bond and/or sustainability sukuk, are debt securities and/or sukuk whose proceeds from the issuance are used to finance or refinance environmentally sound business activities and socially minded business activities.

3. Sustainability-linked bond and/or sustainability-linked sukuk, are debt securities and/or sukuk whose issuance is linked to the achievement of certain key sustainability performance indicators.

II.3 Task Force on Climate-related Financial Disclosures (TCFD)

TCFD created by Financial Stability Board, provides a guideline for organizations to report information about climate-related financial information, including risks and opportunities. By adopting TCFD, organizations could enable stakeholders to make informed decisions in regards of organization's exposure to climate-related risks. The framework consists of four core elements namely Governance, Strategy, Risk Management, and Metrics and Targets each with specific recommended disclosures that should be explained in financial filings. TCFD relates to green bond as financial assets is a part of climate-related market opportunities.

Figure II.1 Core elements of TCFD recommendations
(TCFD Overview, 2022)

According to TCFD Status Report 2023, there are more than 4,850 supporters of TCFD consists of corporations, industry association, and governments with 51% of the supporters reside in Asia-Pacific. 19 authorities including European Commission, Singapore Exchange, US Securities and Exchange Commission have imposed or have proposed TCFD-aligned disclosure requirements for various types of company, such as listed issuers, large corporations, or financial institutions. Moreover, International Financial Reporting Standards (IFRS) in 2023 has completed IFRS S2 Climate-related Disclosures that is consistent and in line with TCFD recommended disclosures with some additional requirements (IFRS, 2023).

Although TCFD report is not yet mandatory for organizations in Indonesia, OJK Regulation no. 51/2017 mandates financial services, issuers, and listed corporations to published Sustainable Finance Report. PWC (2023) reports that in 2022, only 10% of Indonesia listed corporations incorporated TCFD framework for their sustainability report, including bank, processing industry, and oil and gas company.

II.4 Classification Modelling

The advancement of Information Technology has led to the creation of extensive databases and vast amounts of data across different fields. Research in databases and

information technology has introduced methods for storing and managing this valuable data to facilitate decision-making. Data mining involves extracting useful information and identifying patterns from large datasets.

According to (Ramageri, 2010), Data mining is a systematic procedure utilized to sift through extensive datasets to unearth valuable information. Its objective is to uncover previously unidentified patterns, which can subsequently inform decision-making processes crucial for business advancement. It is alternatively referred to as the knowledge discovery process, knowledge extraction from data, knowledge mining, or data/pattern analysis.
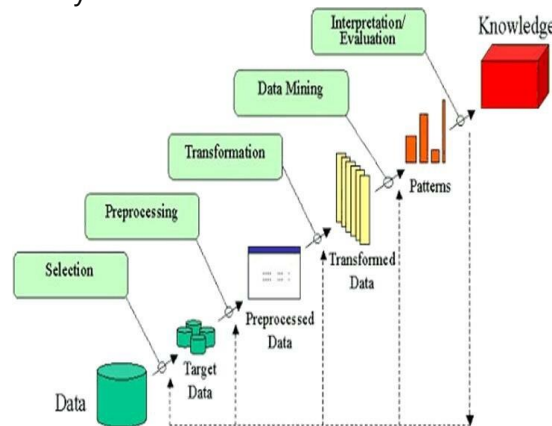

Figure 2.1 Knowledge Discovery Process
Source: (Saxena, 2020)

Data mining has various techniques like Association, Classification, Clustering, and Regression. Classification is the most commonly applied data mining technique, which employs a set of pre-classified examples to develop a model that can classify the population Ahmed, 2014). Several examples of classification algorithms include Support Vector Machine, Decision Tree, and Logistic Regression.

II.4.1 Support Vector Machine

Support Vector Machine (SVM) is a supervised learning system that employs a hypothetical space represented by linear functions in a high-dimensional feature space. This space is trained using an optimization-based learning algorithm, allowing SVM to excel in classifying data compared to prior techniques (Ovirianti, 2022). In the case of classification, SVM aims to find the hyperplane that separates two data classes with the largest margin. This hyperplane is chosen in such a way that the distance between the hyperplane and the closest point from each class, called support vectors, is maximized. In this way, SVM not only looks for a linear separator but also looks for a separator with maximum margin, making it highly resistant to overfitting. SVM is also capable of handling non-linearly separable data classes through the concept of the "kernel trick". Kernels allow SVM to map data to higher dimensions where a linear separator might be found. Commonly used kernels include linear kernel, polynomial kernel, and radial basis function (RBF) kernel.

During training, SVM minimizes the loss function that includes margin and regularization terms to find the best hyperplane. This optimization is often done through convex optimization techniques such as stochastic gradient descent. The advantages of SVM include its ability to handle classification problems with high-dimensional and relatively small sample sizes, as well as its tolerance for non-linearly separable data. However, SVM tends to be less efficient in handling very large datasets.

II.4.2 Decision Tree

The decision tree methodology is a frequently employed technique in data mining for creating classification systems utilizing multiple covariates or for crafting prediction algorithms for a target variable. It works by recursively partitioning the feature space into subsets based on the values of input features, with the goal of minimizing impurity or maximizing information gain at each split. This approach categorizes a population into branch-like segments, forming an inverted tree structure with a root node, internal nodes, and leaf nodes (song, 2015). Decision trees have the capacity to perform effectively on smaller datasets due to their simplicity, reduced susceptibility to overfitting, lower computational demands, ability to discern feature importance, and inherent interpretability.

II.4.3 Logistic Regression

Logistic regression is a data analysis technique that uses mathematics to find the relationship between two data factors. It then uses this relationship to predict the value of one factor based on the other. The logistic regression model is based on the sigmoid function that maps the outcome of the linear equation into the range [0, 1], representing probabilities. During the training phase, the logistic regression model minimizes the loss function, usually the cross-entropy loss function, to find the optimal values of the regression coefficients. This is done using optimization algorithms such as gradient descent. Logistic regression is often used in various applications, including credit risk prediction, spam detection, and medical diagnosis. The advantages of logistic regression include its high interpretability and its ability to handle binomial classification problems well, despite assuming linearity between independent variables and the log-odds of the dependent variable. Thus, logistic regression is a powerful and effective analytical tool for modeling the relationship between variables in various data analysis contexts (Boateng, 2019).

II.5 Data Partnership for Data Collection

Digitalization and globalization have accelerated social and economic dynamics, including the abundance of data produced by the digital platform. From the COVID-19 pandemic lessons, there is a need for new data sources to support policymaking and ensure the relevance of the official statistics. With these challenges, Bank Indonesia needs to explore and utilize new data sources to fill the gap in our data,

including the Green Debt Securities data. The different methodological approach using Data Mining that we propose on II.4 needs to be completed with a data source that is sustainable from its source to ensure its certainty to produce regularly. Thus, it is important to expand access to new data sources, as stated by G20 Finance Ministers and Central Bank Governors in their April 2021 communiqué: "*We also recognize that improving data availability and provision, including on environmental issues, and harnessing the wealth of data produced by digitalization, while ensuring compliance with legal frameworks on data protection and privacy, will be critical to better inform our decisions.*" However, these new data sources are usually kept by the private sector.

As the central bank, Bank Indonesia has a well-established legal framework and mandate for data collection by means of surveys and regulatory reporting for our primary data collection, but this is not limited to these methods alone. Our regulatory reporting comprises banks and non-banks. Bank Indonesia is also utilizing transactional data, which is a product of its Payment Systems and Financial Market Infrastructure, as primary data collection sources.

For secondary data collection[6], Bank Indonesia has practiced data partnerships based on mutual recognition, trust, social responsibility, and the premise that the multi-use of the data will benefit the whole of society, including reducing the regulatory reporting burden on reporters. Besides data partnerships, other methods of collecting secondary data include, e.g., data subscriptions.
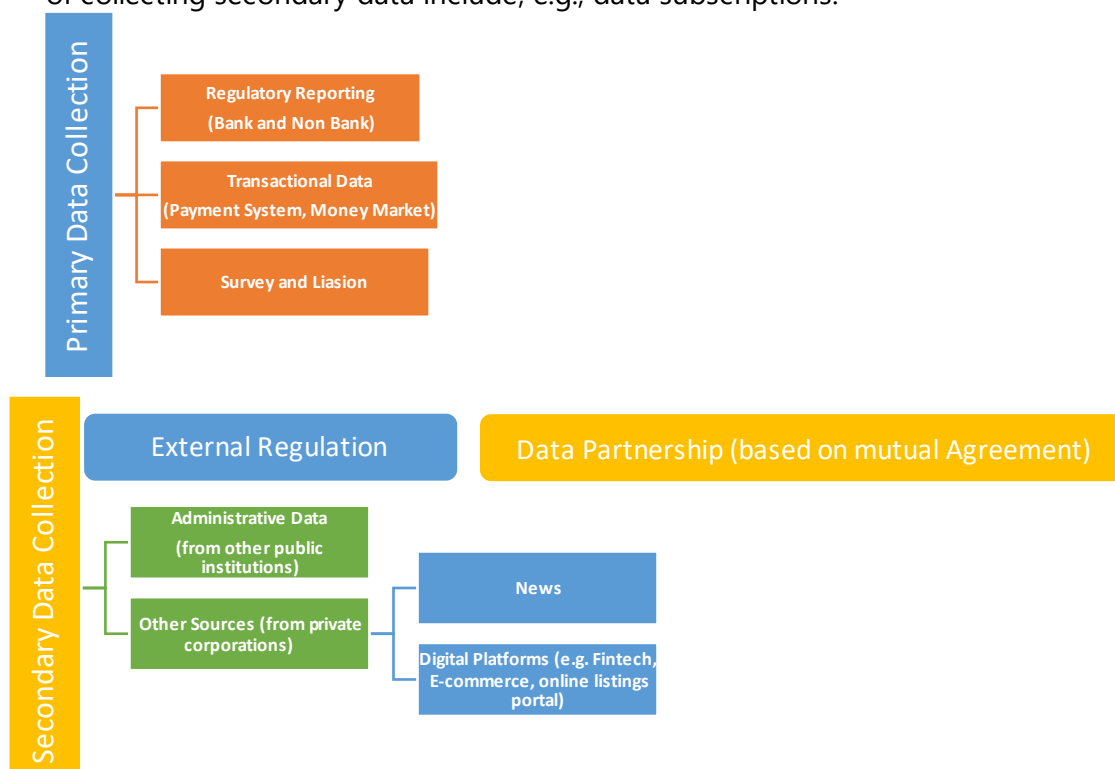

Figure 2.2 Primary and Secondary Data Collection Framework

---

[6] Secondary data is the data that Bank Indonesia collected from other sources (e.g. other authorities or private institution)

Since 2019, Bank Indonesia has transformed most of its bank's regulatory reporting system into a national metadata-based integrated reporting application whose main goal is to capture standardized and metadata-based regulatory reporting data from Bank and reduce the reporting burden. The metadata of this reporting system is mutually agreed upon and used together by Bank Indonesia, OJK (Indonesia's Financial Services Authority), and LPS (Indonesia's Deposit Insurance Corporation). Nowadays, the metadata-based integrated reporting system covers financial, risk, and payment system activity information from Bank.
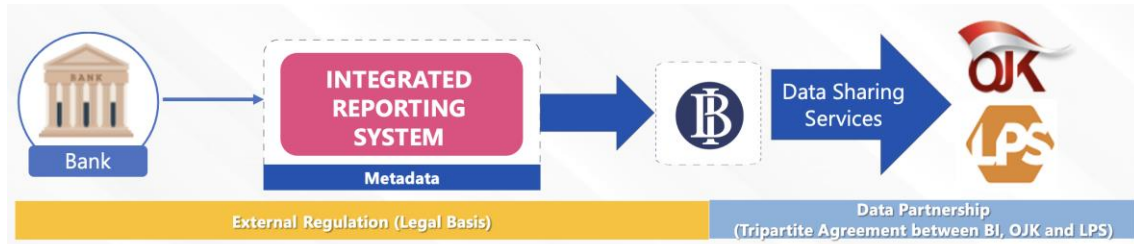
With OJK and LPS, Bank Indonesia established a data partnership in the form of a tripartite agreement between these three organizations. These data partnership goals are based on mutually agreed-upon operational modalities of data reuse. Combining the mechanisms of integrated reporting and data sharing with the basis of data partnerships is an efficient way to reduce the reporting burden of the industry to the authorities.

III. Data and Methodology

III.1 Data

The textual data used for our analysis was sourced from the prospectus of all listed corporation that issued domestic bond in Indonesia Stock Exchange[7]. Moreover, we conducted web scraping to obtain series of government bonds that have been categorized as green bonds, sustainability bonds, or sustainability linked bonds from press releases. Additionally, we also employ web scraping to gather data on corporations that have issued bonds abroad.

III.1.1 Dataset Preparation

In this stage, dataset preparation is used to collect data from various data sources and undergo a series of specific steps to produce a structured dataset with labels corresponding to the desired categories, namely green bond, sustainability bond, and sustainability linked bond. This is done because in the case of Indonesia, there has not yet database with the official flagging into the three bond mentioned categories. The dataset preparation stages can be seen in Figure III.1.

---

[7]

The Indonesia Stock Exchange (IDX) is the country's principal stock exchange. It facilitates the trading of stocks, bonds, mutual funds, and other financial instruments.
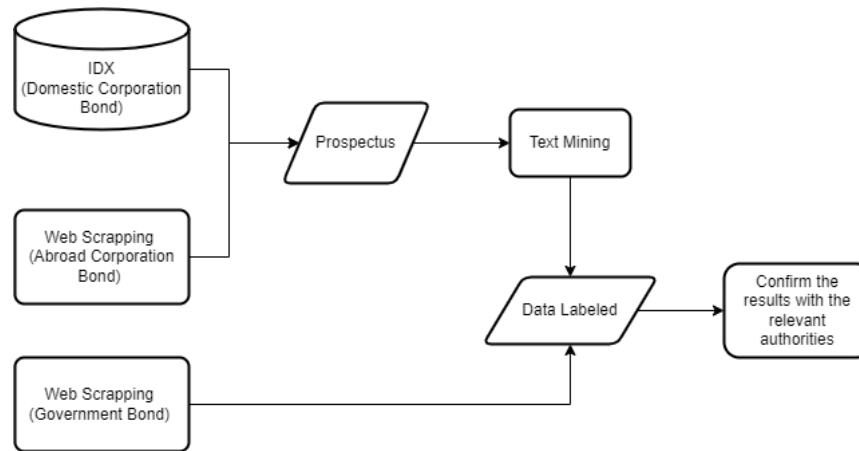
Figure III.1 Data Preparation Stages

After obtaining a list of corporations that issued bond in domestic market from IDX website, the next step is to download prospectuses from these corporations. For corporations bond that issued abroad, we conduct web scraping from news portal. Subsequently, we compile a separate list of corporations to download their prospectuses. Then, text mining is carried out by conducting text-based searches from the prospectuses based on keywords from each class. Before that, we preprocess the text by performing tasks such as text cleaning, text normalization, and stopword removal. These keywords refer to the definitions of the OJK Regulation No. 18/2023, which can be found in Table III.1. After obtaining the labeled data, the next step is to confirm the results with the relevant authorities.

Table III.1 List of Keywords for each Class

Keywords	Class
"green bond", "green sukuk", "kubl", "*kegiatan usaha berwawasan lingkungan*" (the environmentally conscious business activities), "*berwawasan lingkungan*" (environmentally conscious), "*efek bersifat utang berwawasan lingkungan*" (the environmental debt-induced effects).	Green Bond
"sustainability bond", "*efek bersifat utang keberlanjutan*" (the effects are debt-sustainable), "sustainability sukuk", "*ebus keberlanjutan*" (sustainable EBUS), "*lingkungan dan sosial*" (environment and social), "*kegitan usaha berwawasan lingkungan dan kegiatan usaha berwawasan sosial*" (environmentally conscious business activities and socially conscious business activities)	Sustainability Bond
"sustainability linked bond", "*efek bersifat utang terkait keberlanjutan*" (debt-related sustainability effects), "ebus terkait keberlanjutan" (sustainable related EBUS), "*indikator kinerja utama keberlanjutan*" (key performance indicators for sustainability), "sustainability linked sukuk"	Sustainability Linked Bond

Web scraping is also employed to retrieve bond series issued by the Indonesian government by press release. After completing the processes above, we obtain the labeled data of the name of corporation and the series of government bonds that fall into each three type categories of sustainable debt securities.

Furthermore, in order to obtain the securities-by-securities database of sustainable debt securities consist of issuer sector, series number/code, issuance and maturity date, currency, interest type, market of issuance, and the outstanding amount, further data collection is required. In case of Indonesia, this can be collected form the database of securities-by-securities that compiled by the Central Bank in order to fulfill the DGI2 recommendation 7. Securities Statistics. The database data sources are from integrated commercial bank reports, Bank Indonesia – Scripless Securities Settlement System, Indonesia External Debt Statistics, the Indonesia Central Securities Depository, the Directorate General of Financing and Risk Management – Ministry of Finance.

III.2 Classification Modelling

The purpose of our research is to automate the input of company prospectuses and generate their class outputs automatically. To accomplish this, classification modeling is performed. The steps involved in the process can be seen in Figure 4.1.
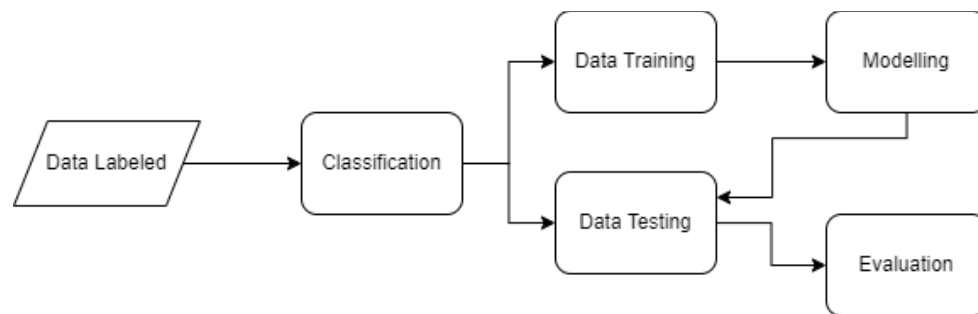

Figure 4.1 Classification Modeling Process

Once the labeled data is obtained from the dataset preparation process, the next step is classification. Following this, the dataset is divided into two subsets: training data and testing data, with a split ratio of 80:20. The training dataset is employed to train three different models: Support Vector Machine, Decision Tree, and Logistic Regression. Subsequently, these models are applied to the testing data. To assess the performance of the models, an evaluation process is conducted, involving the calculation of accuracy. Accuracy provides an overall measure of correct predictions while F1-score is a balance between precision and recall, helpful when there is an imbalance in classes or differing costs for false positives and false negatives.

IV. Results

IV.1 Classification Model Result

In this paper, analysis of different classifiers shows different results that can be shown in table IV.1.

Table IV.1 Classification Model Result

Methods	Accuracy	F1-Score
Support Vector Machine	70,9%	30,1%
Decision Tree	91,3%	68,6%
Logistic Regression	72,1%	40,3%

Among those three methods, the decision tree is superior. This is because the simple concept of the decision tree enables it to capture small data and small process effectively. The F1-score results indicate that the model is satisfying but also suggest a need for improvement.

IV.2 Classification of Corporations and Government Bonds: Categorization within the Class

From the processes conducted, several issuer names from corporation and government along with their series were obtained. In this paper, we have categorized them based on type of sustainable debt securities, issuer sector, currency, maturity, interest type, and market of issuance with position in Q1 2024 as seen in figures IV.l to IV.6.

Based on Figure IV.1, the method used to acquire sustainable debt securities in Indonesia reveals that the majority consist of green bonds (93.82%), with sustainability bonds comprising a smaller portion (6.18%). Moreover, the majority of sustainable debt securities are issued by the Central Government sector (88.6%), followed by Other Depository Corporations (8.06%), Non-Financial Corporations (3.21%), and Other Financial Corporations (0.13%). Additionally, within the ODC sector, one corporation has implemented the TCFD concept in its sustainability reports.

From the perspective of currency, the US dollar still dominates (70.95%) compared to the Indonesian rupiah (29.05%). In addition, the characteristics of sustainable debt securities in Indonesia entail long-term durations and mostly fixed interest rates (94.09%). Meanwhile, when viewed from the perspective of market of issuance, the majority of sustainable debt securities are issued abroad (70.95%) and the rest are issued in domestic market (29.05%).
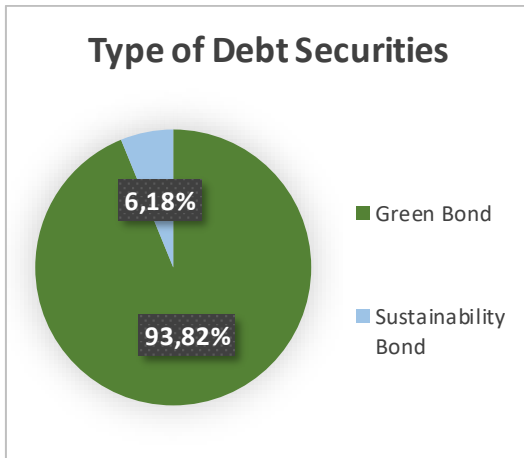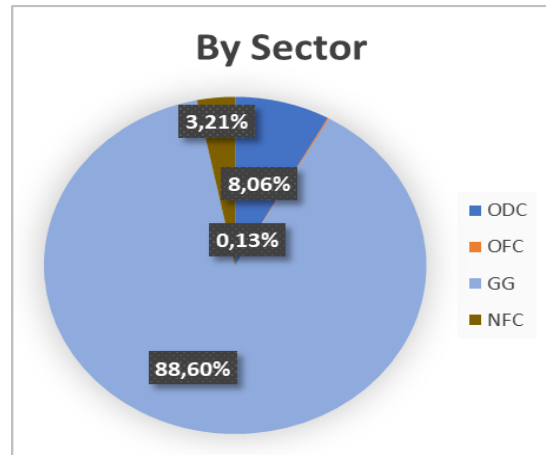
Figure IV.1 Type of Debt Securities

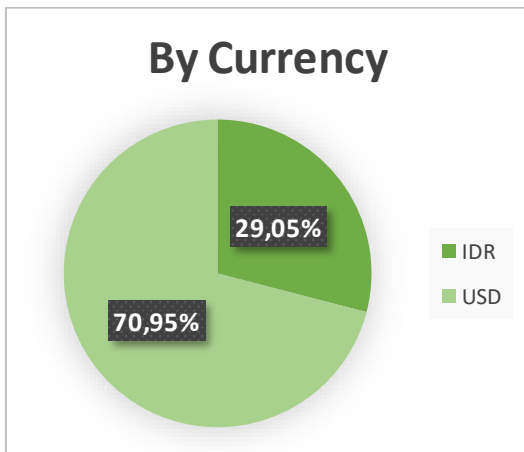Figure IV.2 By Issuer Sector[8]

Figure IV.3 By Currency

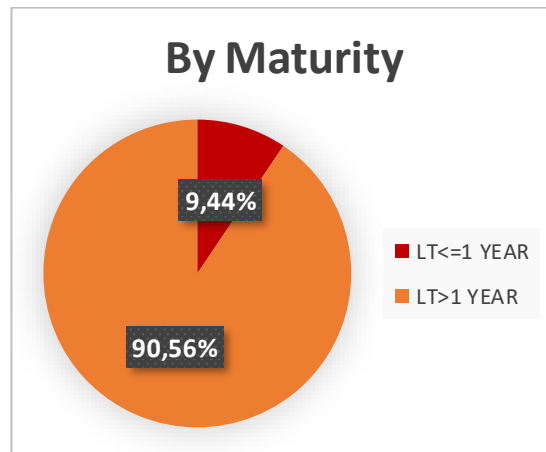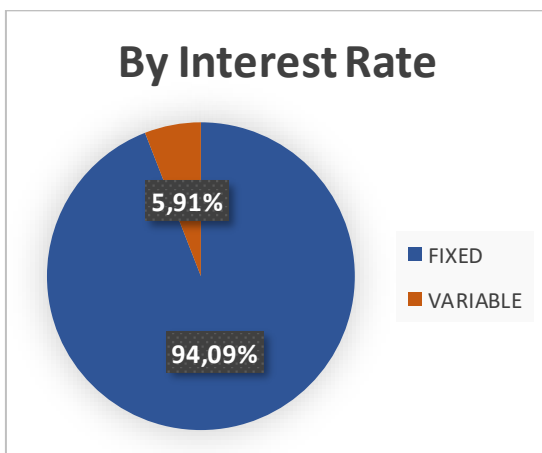
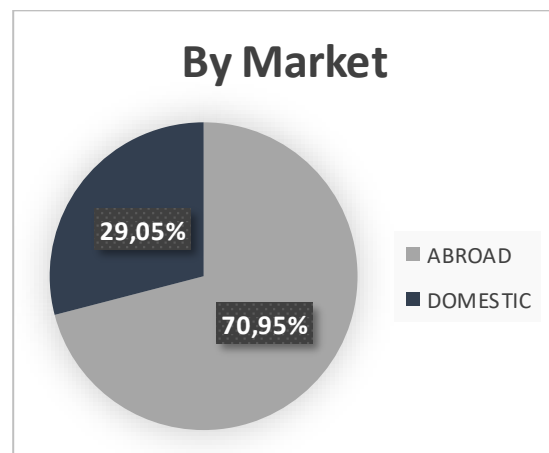Figure IV.4 By Maturity

Figure IV.5 By Interest Rate

Figure IV.6 By Market of Issuance

---

[8] NFC = Non-Financial Corporations, ODC = Other Depository Corporations, OFC = Other Financial Corporations, CG = Central Government

IV.3 Framework of Data Partnership for Data Collection

From our experienced, we learned that the metadata-based integrated reporting that build from the principles of efficiency, flexibility, consistency, and clarity, is one way to enhance data quality and significantly reduce the reporting burden. But this method must be based on and supported by the collaboration/data partnership G2G between related authorities.

Based on mandate of the new omnibus law for Financial Sector, Bank Indonesia with OJK and Ministry of Finance need to coordinate and collaborate for harnessing sustainable f inancial data to support its policy. This collaboration should not be limited to G2G data partnership but also with industry related as well (G2B). Therefore, we propose a framework of data collection that combine the collection of primary data that received through integrated regulatory reporting and the collection of secondary data that received through data partnership with the industry. With the advancement of frontier technologies in storage and analytics, primary and secondary data can be integrated into one format data that can be used for multi purposes policy making.
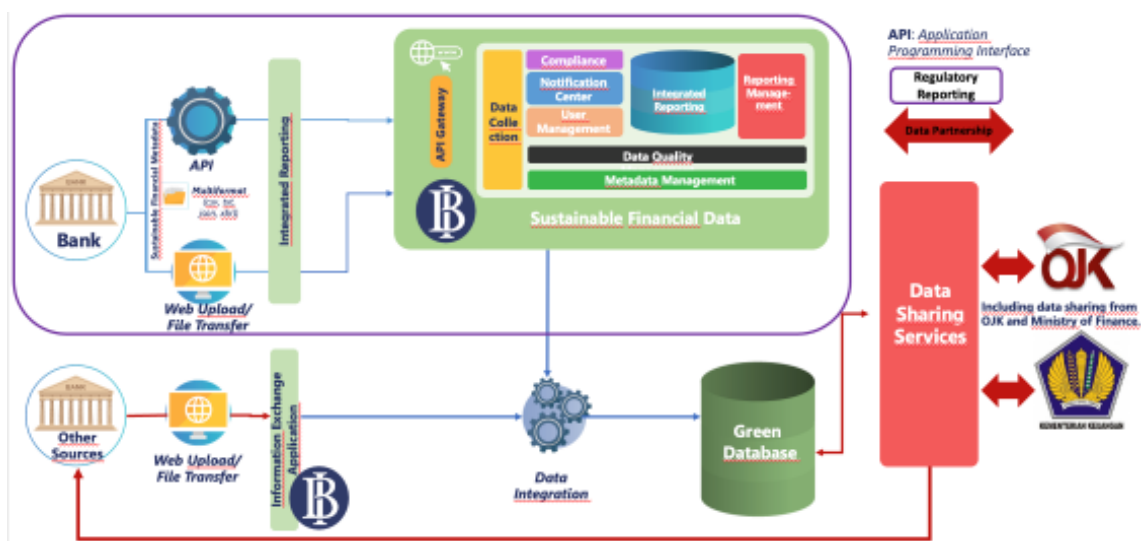
This implementation this framework is not a trivial task, because the first thing that must be agreed upon the authorities is the metadata of sustainable financial data that required to support its policy. Then, we should ensure data availability from its sources, proportionate cost and benefit, and established data sharing mechanisms that ensure data protection and confidentiality.

These technical processes need to elaborate well on the agreement of data partnership that stated the principles of data partnership which consists of but not limited to:

- Data and metadata availability: all parties need to identify, list, and make available all the data and metadata that necessary to the stated purpose of data partnership.

- Technicality mechanism: all parties need to use standard technology, security and IT solution of data sharing and data protection.
- Best practices standard: in ideal environment, all parties should refer to global benchmarks of metadata to ensure its interoperability and use standard metadata format.
- Purpose of use and utilization of data: all parties can define its use case that can be implemented and derive from data partnership.

V. Conclusion and Future Works

1. Flagging sustainable debt securities is important in gathering data to support the analysis of sustainable finance. Data mining is one of the methods that can be utilize to flag and identify the sustainable debt securities in Indonesia. Based on our paper, the decision tree method achieved the highest accuracy among the other two methods, namely Support Vector Machine and Logistic Regression, at 91.3%. However, all three models struggle to learn the data effectively due to its imbalance. Therefore, for future research, the oversampling method can be employed.

2. As per Q1 2024, the sustainable debt securities in Indonesia obtained from this method shows that most of the type is green bond (93.8%), followed by sustainability bond (6.2%) and none falls into sustainability-linked bond. From the sector issuer side, Government sector dominated the issuance of sustainable debt securities (88.6%). From the market of issuance, most of them are issued abroad (70.9%) and the rest are issued in domestic. The characteristics of most of sustainable debt securities in Indonesia are long term, fixed interest rate, and in foreign currency (USD).

3. Out of the corporations identified issued sustainable debt securities, only one company uses TCFD framework for sustainability report. This is in line with the finding on PWC report that only 10% of Indonesia listed company already used TCFD framework. Considering IFRS S2 Climate-related Disclosures is aligned to TCFD, in the future we can expect that the number of companies using TCFD framework will arise. If the authority mandates TCFD framework, the implementation will be more extensive.

4. A framework of data collection that combine the collection of primary data that received through integrated regulatory reporting and the collection of secondary data that received through data partnership with the industry is one way to fulfilled data gap on sustainable finance data. But this framework is not a trivial task. Because, to achieve its efficiency and optimalization in the industry, the most important prerequisite is to ensure an agreement between institutions regarding sustainable finance metadata. Then, we should ensure data availability from its sources, proportionate cost and benefit, and data sharing mechanisms that continue to ensure data confidentiality. Thus, with the data partnership, it needs to have a transparent methodological and quality framework that links directly to

the future production of official statistics from all the parties involved, including the same understanding of the metadata of the data sources. All parties also need technical and organizational measures that should safeguard the security and confidentiality of the data being reused.

VI. References

2023 TCFD Status Report. (2023, October 12). Www.fsb.org. https://www.fsb.org/2023/10/2023-tcfd-status-report-task-force-on-climate-related-financial-disclosures/

PWC. (2023). Sustainability Counts II: State of sustainability reporting in Asia Pacific. https://www.pwc.com/gx/en/issues/esg/asia-pac-esg/sustainability-counts-2023.pdf

Ramageri, B. M. (2010). DATA MINING TECHNIQUES AND APPLICATIONS . Indian Journal of Computer Science and Engineering , 301.

Saxena, A. (2020). A Comparative Analysis of Association Rule Mining Algorithms. IO*P Conference Series: Materials Science and Engineering*, 1-10.

Banga, J. (2019). The green bond market: a potential source of climate finance for developing countries, Journal of Sustainable Finance & Investment, 9:1, 17-32, DOI: 10.1080/20430795.2018.1498617c

Cendekiawan, M., Firmansyah, A. (2024), Pengembangan Green Bonds di Indonesia: Upaya Pemerintah untuk Mewujudkan Keuangan Berkelanjutan, Journal of Law, Administration, and Social Science Volume 4 No. 1, 2024

Finance, Competitiveness, and Innovation Global Practices, World Bank Group, (2021). Toolkits for Policymakers to Green the Financial System

Höhne / Khosla / Fekete / Gilbert, (2012). Mapping of Green Finance Delivered by IDFC Members in 2011, Ecofys

OECD (2020), Developing Sustainable Finance Definitions and Taxonomies, OECD Publishing, Paris.

Network for Greening the Financial System, (2022). Final Report on Bridging Data Gaps

Supriyanto, E. E. (2023). Green and renewable energy financing policy in Indonesia. In S. Ariyanto & S. I. Heriyanti (Eds.), Renewable energy: Policy and strategy (115–144). BRIN Publishing

Triawan, I., Nagari, L. P., (2023), Zero Emission & Pasar Modal, Bisnis Indonesia 17 Oktober 2023.

Med, A. B. (2014). Data Mining: A prediction for Student's Performance Using Classification Method. World Journal of Computer Application and Technology, 43-47.

Boateng, E. Y. (2019). A Review of the Logistic Regression Model with Emphasis on Medical Research. Journal of Data Analysis and Information Processing.

Ovirianti, N. H. (2022). Support Vector Machine Using A Classification Algorithm. Jurnal dan Penelitian Teknik Informatika.

Song, y. y. (2015). Decision tree methods: applications for classification. Shanghai Archives of Psychiatry, 130.