SEI | WORLD BANK GROUP

Using Automated Text Mining to Align Investments with the Sustainable Development Goals:

A Case Study Analyzing World Bank Projects

Authors: Efraim Hernández-Orozco (SEI),
Mario Cárdenas-Vélez (SEI), Colleen Keenan
(World Bank), Zoe Russo (World Bank)

1818 H Street NW, Washington, DC 20433
Telephone: 202-473-1000 Internet: www.worldbank.org

This work is a product of the staff of The World Bank with external contributions. The findings, interpretations, and conclusions expressed in this work do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent.

The World Bank does not guarantee the accuracy, completeness, or currency of the data included in this work and does not assume responsibility for any errors, omissions, or discrepancies in the information, or liability with respect to the use of or failure to use the information, methods, processes, or conclusions set forth.

Nothing herein shall constitute or be construed or considered to be a limitation upon or waiver of the privileges and immunities of The World Bank, all of which are specifically reserved.

Table of Contents

Acknowledgments

This paper was prepared by a team of colleagues from the Stockholm Environment Institute (SEI) Latin America Centre and the World Bank—Efraim Hernández-Orozco and Mario Cárdenas-Vélez are from SEI, and Colleen Keenan and Zoe Russo are from the World Bank Capital Markets and Investments Department. The paper was prepared under the direction of Ivonne Lobos, Senior Expert and Leader of the Sustainable Development Goals (SDGs) Line of Inquiry at SEI Latin America, and Heike Reichelt, Head of Investor Relations and Sustainable Finance from the Capital Markets and Investments Department at the World Bank.

The authors would like to acknowledge the World Bank Group SDG Partnership Fund for supporting analytical research to raise awareness for the SDGs and the potential role they can play in building sustainable capital markets. This support is what made the partnership with SEI to develop the SDG mapping methodology and this paper possible. Our thanks go especially to Juliana Knapp and Lisa Castro at the SDG Partnership Fund.

We would like to thank those who participated in the literature review of the paper and the validation exercise, Daniela Maestre-Másmela and Juan Betancur-Jaramillo from SEI. We also would like to thank those who provided comments on the paper during the peer review process. They include, from the World Bank, Frank Wissing Madsen in Operations Policy and Country Services, Brian William Stacy in the Development Data Group, and James Seward in the Capital Markets and Investments Department, and from the International Finance Corporation, Atiyah Curmally and Blaise Wendwaoga Sandwidi. We are also grateful for feedback from others including Judith Moore, Alya Kayal and Stuart Kinnersley from ImpactArc, and Otilia Ciotau in the office of the World Bank Group Chief Finance Officer, who provided helpful input to the content of the paper. Additionally, we wish to thank David Purkey, Director of the SEI Latin America Centre; Aaron Maltais, Senior Research Fellow and Director of the Stockholm Sustainable Finance Centre; Adis Dzebo, Research Fellow and Co-leader of SEI's Initiative on Global Finance; and Georgia Savvidou, Research Associate at SEI Headquarters, for providing their expertise and advising the development of the research.

Finally, we would also like to recognize George Richardson, Director of Capital Markets and Investments at the World Bank, for his support to the project.

The paper has been funded by the World Bank Group SDG Partnership Fund. The views expressed herein are solely the authors' and should not be attributed to SEI or the World Bank or its stakeholders.

About the World Bank

The World Bank (International Bank for Reconstruction and Development, IBRD) is an international organization. Created in 1944, it is the original member of the World Bank Group and operates as a global development cooperative owned by 189 nations. The World Bank provides loans, guarantees, risk management products, and advisory services to middle-income and other creditworthy countries to support the Sustainable Development Goals and to end extreme poverty and promote shared prosperity. The World Bank has been issuing sustainable development bonds in the international capital markets for over 70 years to fund programs and activities that achieve a positive impact.

About the Stockholm Environment Institute

The Stockholm Environment Institute (SEI) is an international nonprofit research and policy organization that tackles environment and development challenges. SEI connects science and decision-making to develop solutions for a sustainable future for all. SEI's approach is highly collaborative: stakeholder involvement is at the heart of its efforts to build capacity, strengthen institutions, and equip partners for the long term. SEI's work spans climate, water, air, and land-use issues, and integrates evidence and perspectives on governance, the economy, gender, and human health. Across eight centers in Europe, Asia, Africa, and the Americas, SEI engages with policy processes, development action, and business practices throughout the world.

List of Abbreviations

AUN	Aurora Universities Network
DIE	Deutsches Institut für Entwicklungspolitik
FY	fiscal year
IAEG-SDGs	Inter-agency and Expert Group on SDG Indicators
IBRD	International Bank for Reconstruction and Development
KWIC	key word in context
NLP	natural language processing
OECD	Organization for Economic Co-operation and Development
OECD DAC	Organization for Economic Co-operation and Development’s Development Assistance Committee
PPV	positive prediction value
SDB	Sustainable Development Bond
SDG	Sustainable Development Goal
SEI	Stockholm Environment Institute
UN	United Nations
UN DESA	United Nations Department for Economic and Social Affairs

Executive Summary

Investors, issuers, and other capital market participants are using the Sustainable Development Goals (SDGs) as a reference framework to show the impacts of investments on global sustainable development. But assessing the impacts requires the development of methodologies that can help investors examine how their activities are aligning with and contributing to the SDGs.

This paper offers a case study on linking investments, which in this case are a group of 100 World Bank – financed projects, to the SDGs to understand the projects' connections to the SDGs. The paper focuses on a collaboration between the World Bank (specifically, the International Bank for Reconstruction and Development, IBRD), a pioneer in sustainable finance and a leader in promoting models for transparency and disclosure to build sustainable capital markets, and the Stockholm Environment Institute (SEI), an international nonprofit research and policy organization that connects science and decision-making to develop solutions and policies for a sustainable future.

This work aims to develop a methodology that provides a lens through which investors can see indicative linkages between their investments and the SDGs, in support of investment reporting and potentially aiding other capital markets participants in connecting project results or investments to the SDGs. The paper also adds to ongoing global efforts to map the contributions of development projects, policies, and research to the SDGs and aims to fill gaps in the existing mapping methodologies. Greater transparency around impact can help investors develop insights and inform decision-making on sustainable investment to help channel finance to sustainable purposes.

The methodology described in this paper involved the creation of a text mining protocol using Boolean queries[1] based on a controlled vocabulary of the language used by the World Bank. This included formulating keywords and phrases that can identify SDGs and their targets in the text of World Bank project documentation. For its test case, SEI applied the methodology to the 100 new IBRD-funded projects that were added to the World Bank Sustainable Development Bond (SDB) project portfolio in fiscal 2020 (July 1, 2019 – June 30, 2020) — a subset of the roughly 600 active IBRD projects in total.

In general terms, the exercise successfully highlighted which SDGs appeared the most, according to this methodology, in the subset of the fiscal 2020 projects. For instance, SDGs 17 (Partnerships for the Goals), 9 (Industry, Innovation, and Infrastructure), and 3 (Good Health and Well-Being) are those that appear most often in the mapping results. At the target level, target 17.9 (Enhancing capacity for the SDGs) appears the most frequently (in 50 projects); aside from SDG 17, targets 3.8 (Achieve universal health coverage) and 3.c (Increase health financing) appear the most, with 35 and 26 mapped projects, respectively. The results were then validated by manually mapping 20 randomly chosen projects and comparing the manual results to the text mining ones.

We note, however, that the interpretation of results needs to consider the data set used, the limitations of the methodology, the intended use and purpose of the mapping, and the interconnected nature of the SDGs when making conclusions. For instance, the SDGs include topics that cut across the entire 2030 Agenda for Sustainable Development (hereafter, 2030 Agenda), such as climate change and gender equality. Therefore, we recommend using a holistic view to interpret the results to avoid focusing on individual goals, which may lead to inaccurate interpretations. For example, the results show how World Bank projects are connected to multiple SDGs, including looking specifically at those projects connected to gender equality and climate change and environmental issues. In looking at this grouping, the results show that 66 percent of the mapped projects relate to environmental and climate change SDGs, 79 percent are related to gender- sensitive SDGs, and 49 percent are related to both. In addition, because of the prevalence of SDG 17 targets in the mapping results, which is due to the World Bank's focus on capacity building and economic development in member countries, the analysis controls for the mapping to these targets.

This paper also discusses the limitations of the proposed approach and similar methodologies. In particular, we note that this mapping methodology represents one of the many solutions available for indicative mapping of investments to the SDGs and that different approaches can produce distinct results. Moreover, this methodology and other approaches examined do not account for the differential and weighted impact that individual projects can have on the SDGs. Not all projects related to the same goal or target will have the same degree of impact. This limitation remains an important consideration and provides a direction for further research.

Finally, the proposed methodology may be adapted to the broader sustainable finance space. If scaled up and further developed, through artificial intelligence, for example, it can support capital markets' participants in meeting the growing disclosure requirements from new national and regional regulations and taxonomies setting definitions and standards around sustainable activities by assessing the alignment of investments to these emerging frameworks.

Introduction

The SDGs, which consist of a set of 17 goals and 169 targets, were adopted by the United Nations (UN) General Assembly and the international community in 2015 to address social, economic, and environmental global challenges in an integrated manner and to reach sustainable development globally by 2030 (UN 2016). However, the UN has estimated a financing gap of US$2.5 trillion annually for developing countries to achieve the SDGs by 2030 (UNCTAD 2014). Sustainable finance plays a pivotal role in filling this gap by unlocking private capital as a complement to public money, but investors need to be able to examine and understand how they are contributing to more sustainable and inclusive economic growth. Many capital markets participants, such as issuers and investors, use the SDGs as a reference framework to show links between investments and sustainable development. Disclosures on the links between the SDGs and investments can improve transparency on the impacts of sustainable development investments and potentially help channel funds to sustainable purposes.

This paper develops a methodology to efficiently and accurately map investments to the SDGs, applies the methodology to World Bank–financed projects, and shares the results. It was produced through a collaboration between the World Bank (namely the IBRD), a pioneer in sustainable finance and a leader in promoting models for transparency and disclosure to build sustainable capital markets, and SEI, an international nonprofit research and policy organization that connects science and decision-making to develop solutions for a sustainable future. The expectation is that the methodology will support investors and issuers in connecting project results or investments to the SDGs. This paper also contributes to ongoing efforts to map development projects' contributions by reviewing the most common approaches for SDG mapping and providing recommendations to fill gaps in their methodologies.

The UN system has endorsed the use of the Inter-agency and Expert Group on SDG Indicators (IAEG- SDGs) as a framework to monitor progress for the SDGs; however, most of those indicators are generated only at the country level, which complicates their application outside governmental and national-level reporting (Hernández-Orozco et al. 2021). Efforts to map activities and projects from different types of institutions to the SDGs are growing, including among governments (Hawai'i Green Growth Local2030 HUB 2020), enterprises (for example, Horne et al. 2020), scientific and academic institutions (Asatani et al. 2020; Bordignon 2021; Mistry et al. 2020; Sullivan, Thomas, and Rosano 2018), and development and investment banks (Dangelmaier 2019; Natixis 2018; World Bank 2021). This paper identifies the two main trends in the methodologies used to produce these mappings: conceptual mapping of themes and categories to the SDGs (Dangelmaier 2019; also, for example, Government of Mexico City 2019; Hawai'i Green Growth Local2030 HUB 2020; NYC Government 2019) and methodologies that analyze large volumes of text to automatically find SDG- related themes and concepts, such as text mining (for example, Bordignon 2021; Sullivan, Thomas, and Rosano 2018; World Bank 2021), content analysis (Uehara and Sakurai 2021), natural language processing (NLP) (Asatani et al. 2020), and machine learning (Pincet, Okabe, and Pawelcyzk 2019).

The methodology developed in this exercise corresponds to the second trend. It involved the creation of a Boolean queries protocol based on a controlled vocabulary of the language used by the World Bank in its project documentation for projects featured in the Impact Report 2020 on IBRD SDBs and Green Bonds and the “World Bank Theme Taxonomy and Definitions” (World Bank 2016). Then, using KH Coder software (Higuchi 2021), SEI recorded the occurrence of SDG keywords in World Bank project details—if one or more SDG keywords appeared in the text, that project was mapped to the related SDG target. SEI developed this methodology based on its NDC-SDG Connections Tool, which is developed with the Deutsches Institut für Entwicklungspolitik (DIE), that visualizes potential connections between the SDGs and countries’ climate actions toward their nationally determined contributions (NDC) (DIE and SEI 2017).

As a test case, SEI applied the resulting methodology to the 100 IBRD projects that were added to the World Bank SDB project portfolio in fiscal 2020 (July 1, 2019–June 30, 2020),[2] which is a subset of the roughly 600 World Bank projects active in fiscal 2020 (World Bank 2021). The results were then validated by manually mapping 20 projects to verify the automated results. A positive predictive value (PPV), which assesses the likelihood of the Boolean queries producing a correct mapping result, of 75 percent was obtained.

This paper reports on this exercise and discusses the limitations of and considerations for the proposed approach and similar methodologies. It is useful to note that this mapping methodology represents one of several solutions available for indicative mapping of investments to the SDGs, and different approaches can produce distinct results. Moreover, the research reveals that this methodology and most other approaches examined do not account for the differential and weighted impacts that individual projects can have on the SDGs. This limitation remains an important consideration and provides direction for further research.

Finally, when interpreting the mapping results from this exercise, and other similar analyses, it is important to consider that the SDGs are inherently interconnected and include transversal topics that cut across the entire 2030 Agenda, such as climate change and gender equality. Therefore, a focus on mapping results to individual goals may lead to inaccurate interpretations. A more holistic view is recommended, and this paper models that by highlighting some World Bank projects that are connected to gender equality and climate change and environmental issues rather than either SDG 5 or SDG 13.

The paper is structured as follows: Section 1 presents a literature review of the existing methods to map different development activities to the SDGs. Section 2 explains the creation of the keyword protocol and its application in KH Coder to map World Bank–financed projects to the SDGs. Section 3 presents the results of the exercise. Section 4 discusses the results and suggests how to improve SDG mapping methodologies based on the lessons learned from this mapping exercise. The last section concludes.

SECTION 1:

An Overview of Approaches to Map Policies & Projects to the SDGs

Section 1. An Overview of Approaches to Map Policies & Projects to the SDGs

The SDGs are being used as a framework to link the impact of investment activities and their contribution toward sustainable development, and making that link requires a deployment of approaches that can map investments, projects, and activities to the SDGs. Conventionally, to produce such an assessment, governing bodies have aimed to measure progress using the list of indicators from IAEG-SDGs (2016). However, most of these indicators are generated only at the country level, which complicates their application outside national-level reporting (Hernández- Orozco et al. 2021).

Furthermore, when assessing the contribution of various types of institutions to the SDGs, the use of the IAEG-SDGs indicators can be challenging because the concepts and terminology used by other institutions often differ from the concepts used in the SDGs and their indicators. For instance, the World Bank uses its own taxonomy of themes[3] and sectors along with Organization for Economic Co-operation and Development's Development Assistance Committee (OECD DAC) purpose codes to classify the activities of World Bank–supported projects (World Bank 2016). The themes are aligned with the SDGs but do not match them exactly, and the themes can be subjectively mapped to more than one target; for example, the World Bank theme "Adolescent health" can be thematically mapped to SDG targets 3.7 (Sexual and reproductive health), 3.8 (Achieve universal health coverage), and 5.3 (Eliminate all harmful practices against children and women).

The implication is that to know the link to or contribution of various development activities to the SDGs, researchers and practitioners need to deploy methods to map different concepts and categories to those used in the SDGs. In this regard, this paper identifies different SDG mapping methodologies that are grouped into two main trends: 1) conceptual mapping of indicators and categories[4] to the SDGs and 2) automatic detection of SDG topics in texts and documents. The first approach consists of a manual review and classification of indicators and categories conducted by researchers, policy makers, institutions, and other stakeholders, who then examine how the concepts present in their organizational frameworks overlap with the SDGs. This approach allows the rapid mapping of any project tabulated under a certain category to its corresponding SDGs. The second approach uses computer assistance to automatically detect words and phrases in texts related to the SDGs, mapping their thematic meaning (or topics) to corresponding SDGs and eliminating the need to directly review the contents of each project. This type of approach is mainly used for large data sets of text.

It is important to point out that both approaches require the researcher to subjectively map the identified topics or keywords to SDG themes, either at the category level or by assigning meaning to text mining queries. It can be argued that the subjectivity behind this process occurs because of the ambiguous nature of the SDGs: given that the SDGs relate to politically sensitive topics, the vision of what sustainable development entails and how to achieve it can depend on the vision of the implementing actors (see Mair et al. 2018). This ambiguity plus the possibility for words or phrases to have many meanings (or polysemy, see Bordignon 2021) can produce various mapping results that correspond to the different understandings of the SDGs.

Category mapping is mostly used by governmental institutions. For instance, some regional governments have used the Sustainable Development Report methodology (Sachs et al. 2021), which includes criteria to map indicators from distinct sources when the IAEG-SDGs indicators are missing. Examples of this approach include Índice de Desenvolvimiento Sustentável das Cidades – Brasil (Instituto Cidades Sustentáveis 2021) and Los Objetivos de Desarrollo en 100 ciudades españolas (Sánchez de Madariaga et al. 2020), which report SDG progress in Brazilian municipalities and Spanish cities, respectively. Other governments have mapped the categories of their government programs to the SDGs; any project, plan, or indicator assigned to a specific category is then directly mapped to its corresponding SDG. For instance, the State of Hawaii in the United States has its own sustainability framework distinct from the SDGs, called the "Aloha+ Challenge," outlining six priority areas for the state's sustainable development; in this framework, the "Local Food" priority area and any project under it are mapped to goals 1 (No Poverty), 2 (Zero Hunger), 3 (Good Health and Well-Being), 12 (Responsible Consumption and Production), 13 (Climate Action), and 17 (Partnerships for the Goals).

Comparatively fewer nongovernmental organizations than governmental institutions have used category mapping to report on SDG alignment. One organization that uses it is KfW, the German development bank, which describes in its SDGs mapping methodology document how it has assigned "markers" and "objectives" to the SDGs, and it then assigns any activity organized under specific categories to the related SDGs (Dangelmaier 2019). For instance, projects categorized with the "gender marker" are mapped to SDGs 5 (Gender Equality) and 10 (Reduced Inequalities). KfW also uses this mapping to report contributions to each SDG based on the financial resources spent on development activities. KfW does this to measure the impact of its development portfolio on the SDGs, using financial resources as a weighting factor so that all the activities mapped under a specific SDG and their corresponding budget are noted as contributing to that SDG. For instance, KfW reports that its 2018 portfolio contributed €33.4 billion to SDG 8 (Decent Work and Economic Growth), which is the SDG with the highest investment by KfW's development activities that year (Dangelmaier 2019).

The mapping of SDG categories can be convenient in terms of technical capacity and the use of time and financial resources since it does not require the examination of each project individually. Instead, the mapping is performed at the category level, and any project tabulated under a certain category is rapidly mapped to the corresponding SDG. However, the categories' definitions and concepts can have thematic differences from the SDGs, so they do not always map exactly to the goals or targets. For instance, when reviewing the "World Bank Theme Taxonomy and Definitions" (World Bank 2016), theme 436, "State-Owned Enterprise Reform and Privatization," is challenging to map to the SDGs because privatization is not explicitly covered in the SDGs. Therefore, it is necessary to define the details of the privatization process to assess SDG alignment.

The implication is that it is more effective to examine each project individually to improve the accuracy of the mapping. Nonetheless, manual examination of single documents can become time- and resource-consuming, especially when analyzing large documents or text data sets. For this reason, institutions are increasingly opting to mechanize this process, using text mining and other automated text analysis techniques, which are the second family of SDG mapping methodologies identified in this paper.

Overall, these automated document evaluation approaches predefine search queries related to the SDGs, formulating keywords and phrases that can identify SDG goals and targets in text documents. For instance, Boolean queries can be designed to detect singular words, phrases, and sentences using operators such as "OR," "AND," and "NOT." For example, applying the query "climate AND change" will look for the phrase "climate change" in a text—if there are true values, the results can then be assigned to SDG 13 (Climate Action), mapping the text to the goal.

Text mining methods have been used to map texts to the SDGs in various contexts. Still, its use is more prominent in academia, where researchers have sought to assess academic research and publications' contributions to the SDGs (for example, Armitage et al. 2020; AUN 2021; Bordignon 2021; Elsevier 2020; Hassan, Haddawy, and Zhu 2014; Jayabalasingham et al. 2019; Mistry et al. 2020; Olawumi and Chan 2018; Vanderfeesten, Otten, and Spielberg 2020). According to Bordignon (2021), the academic use of text mining for SDG mapping has had mixed results, as some queries can be too constrained because they use the original UN texts for the SDGs, while others may be too broad because of the polysemy of words, producing false positives. Nonetheless, some approaches have had promising results, such as AUN's SDG mapping of research publications that reports a "precision" of 70 percent, which was validated by the authors of the publications mapped in the analysis (AUN 2021).

In addition, text mining, namely semi-automated content analysis, was used in Germany to map SDG-related activity in 193 entrepreneurial venture competitions (Horne et al. 2020). It was also used to map media content, scientific publications, and patents to SDG 13 (Climate Action) in South Korea (Hwang et al. 2021); find the most discussed SDGs in government-issued Voluntary National Reviews (Sebestyén, Domokos, and Abonyi 2020); and align policy making in the European Union with the SDGs (JRC and INTPA n.d.).

Other text analysis approaches use artificial intelligence techniques that automatically create queries (without manual programming), learning from the data sets used as inputs to build controlled vocabularies or thematic categories. The artificial intelligence methods that have been used for SDG mapping include machine learning, NLP, and Bidirectional Encoder Representations from Transformers. The OECD (Pincet, Okabe, and Pawelcyzk 2019) used machine learning to find which SDGs are most affected by worldwide development and aid projects, and the UN Department of Economic and Social Affairs (UN DESA) used it to create topic models aligned with the SDGs for UN publications. Asatani et al. (2020) used NLP, which creates clusters of the SDG topics emerging in a data set, to find the contributions to the SDGs of over 300,000 publications. And Guisiano and Chiky (2021) used Bidirectional Encoder Representations from Transformers to label texts related to the SDGs from the OnePlanet network parentship.

Text analysis approaches can be used in a wide variety of documents to assess SDG mapping, and each approach developed can potentially be used on any text data set in English. However, the applicability of these approaches is not universal, as Bordignon (2021) notes, because of polysemy and the risk that the restrictive vocabularies used in queries will produce less accurate results. To make mapping more accurate, text mining queries have to be adapted and tailored using the specific language (also called a "controlled vocabulary”) found in the analyzed texts (for example, Krallinger et al. 2012; Spasić et al. 2008). For example, for this paper, a controlled vocabulary was created using KH Coder software to find the most used words and phrases that appear in the subset of 100 World Bank projects. This analysis was then used to create Boolean queries tailored for World Bank– specific vocabulary to map SDGs at the target level. This process is described in the next section.

SECTION 2:

Methodology

Section 2. Methodology

To map the subset of 100 projects financed by the World Bank during fiscal 2020 (World Bank 2021) to SDGs at the target level, the publicly available text mining methodologies described above could potentially have been used. But because of polysemy (Bordignon 2021), there was a risk that the queries they designed would be too general, which generates false positives, or be too narrow, which reduces the number of positive results.

To avoid this, SEI proposed the use of a controlled vocabulary developed from the type of language present in World Bank project documents and Boolean queries created from the text of the fiscal 2020 projects and the "World Bank Theme Taxonomy and Definitions" (World Bank 2016). KH Coder 3.Beta.02f version was then used to analyze the texts' vocabulary to identify the most common words and phrases in each document that can refer to SDG topics. Next, each topic identified was assigned to the SDGs at the target level. By creating queries from the vocabulary of the SDG targets, it was possible to account for other words and phrases that the KH Coder analysis did not reveal. The version of KH Coder used for this paper and new versions are free and downloadable (see Higuchi 2022), and SEI used instructions from the publicly available KH Coder manuals to build this approach (see Higuchi 2016b, 2017).

The KH Coder mapping process involves three phases. The first is a quantitative content analysis approach that automatically extracts words from the texts to analyze them statistically (Higuchi 2016a, 2017). The second is the creation of coding rules for the Boolean queries. The final phase involves intersecting the coding rules with the project documents to obtain the number of SDGs mapped per project. It is important to note that the developed approach does not account for the differential and weighted impact that individual development actions may have on the SDGs.

Phase I. Creating the Controlled Vocabulary

Data Preparation

The World Bank collected and supplied text extracts of the 100 projects' documentation, in most cases the Project Appraisal Document (to review the World Bank's project database, see the website in World Bank n.d.). These excerpts included the project development objective, project components, and target results indicators for projects currently under implementation. The information was then inserted into KH Coder's Frequency List tool to obtain the list of words used most frequently in the texts.

The next step in the data preparation was to identify all the stop words in the analysis, which are the words that frequently appear in the texts but that do not have a specific meaning, such as conjugations of the verb "to be" or prepositions. KH Coder comes preloaded to eliminate common stop words in English like "do" and "is." However, it was necessary to include other stop words that appear in the World Bank documents and do not indicate SDGs—most referred to places, nationalities, and abbreviations. Such stop words can be appropriately removed from an analysis to improve it (Rani and Lobiyal 2020). The stop words were inserted in a TXT document, which was then linked to the KH Coder in the software's settings.

Frequency List of Words

SEI created another words frequency list, again excluding the stop words, to obtain the number of times a word appeared in the complete text data set. The output, to an Excel document, is shown in table 1. It is important to note that KH Coder also isolated and counted every word by considering its root form. For instance, the word/concept "say" would include the words "say," "says," "saying," and "said" (Higuchi 2016a).

Table 1: Frequency List of Words for All 100 Projects

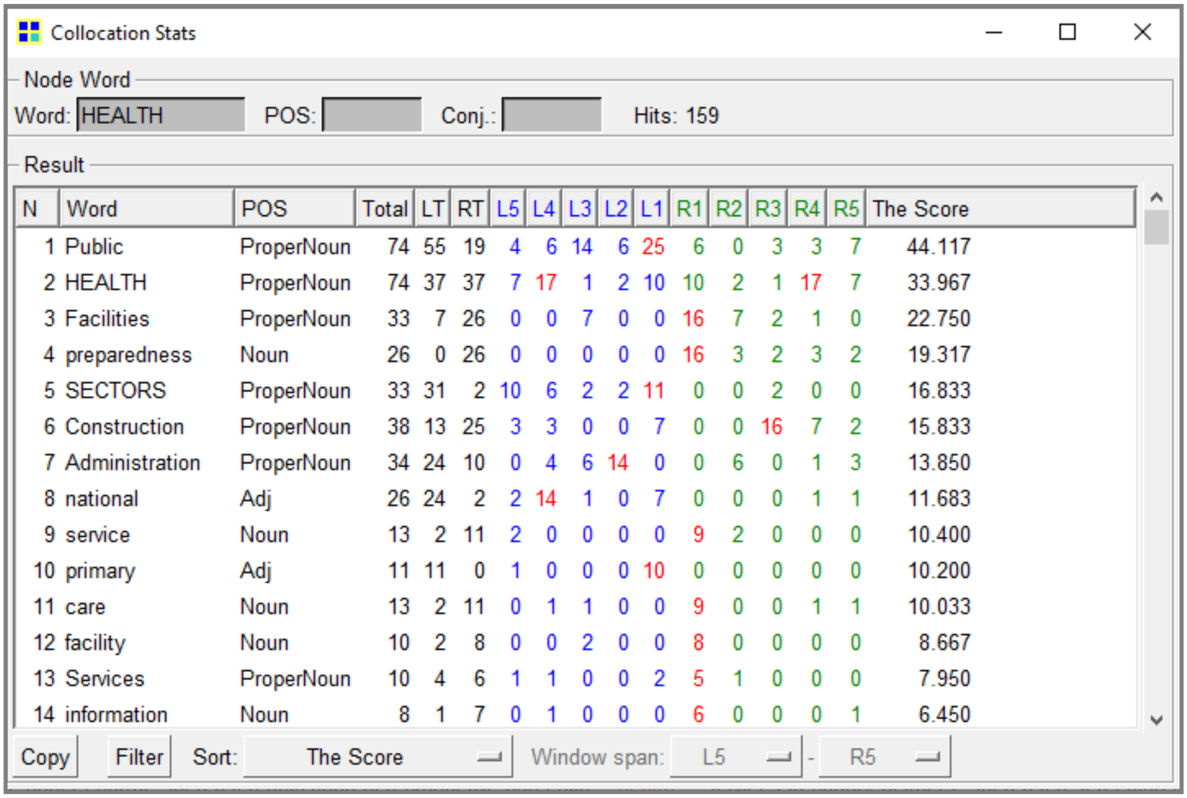
Words	Frequency
Service	198
Development	193
Public	179
Management	169
Improve	169
Percentage	169
Support	165
Health	159
Increase	153
Water	149
Strengthen	137
Access	132
Capacity	110
COVID-19	108
Program	100
Beneficiary	99
Social	96
Education	92
Provide	92
Sector	85
Female	83
Quality	80

Source: An elaboration by SEI using KH coder software.
Note: The list shows the words that appear 80 times or more in the 100 analyzed texts without considering the stop words.

Key Word in Context Concordance Analysis

Subsequently, the Key Word In Context (KWIC) Concordance tool in KH Coder was applied to the data set to analyze and identify how a specific word appeared in a sentence and show how other words related to it grammatically. The KWIC method performs a match analysis for each concept (word) within a frequency list based on the word's positional criteria (Luhn, 1960). The model assesses and identifies the pattern of words (word associations) that accompanies those primary words, or "node words" (Higuchi 2016a; Luhn 1960). The node words have more connections and centrality within the data set, which the software considers to be more critical because of their position and semantic meaning.

Figure 1: KWIC Concordance Analysis on "Health"



Node Word

Word: HEALTH POS: Conj.: Hits: 159

Result

N	Word	POS	Total	LT	RT	L5	L4	L3	L2	L1	R1	R2	R3	R4	R5	The Score
1	Public	ProperNoun	74	55	19	4	6	14	6	25	6	0	3	3	7	44.117
2	HEALTH	ProperNoun	74	37	37	7	17	1	2	10	10	2	1	17	7	33.967
3	Facilities	ProperNoun	33	7	26	0	0	7	0	0	16	7	2	1	0	22.750
4	preparedness	Noun	26	0	26	0	0	0	0	0	16	3	2	3	2	19.317
5	SECTORS	ProperNoun	33	31	2	10	6	2	2	11	0	0	2	0	0	16.833
6	Construction	ProperNoun	38	13	25	3	3	0	0	7	0	0	16	7	2	15.833
7	Administration	ProperNoun	34	24	10	0	4	6	14	0	0	6	0	1	3	13.850
8	national	Adj	26	24	2	2	14	1	0	7	0	0	0	1	1	11.683
9	service	Noun	13	2	11	2	0	0	0	0	9	2	0	0	0	10.400
10	primary	Adj	11	11	0	1	0	0	0	10	0	0	0	0	0	10.200
11	care	Noun	13	2	11	0	1	1	0	0	9	0	0	1	1	10.033
12	facility	Noun	10	2	8	0	0	2	0	0	8	0	0	0	0	8.667
13	Services	ProperNoun	10	4	6	1	1	0	0	2	5	1	0	0	0	7.950
14	information	Noun	8	1	7	0	1	0	0	0	6	0	0	0	1	6.450

Sort: The Score Window span: L5 - R5

Source: An elaboration by SEI using KH coder software.
Note: The figure lists the 14 most common words appearing together with "health" in the 100 projects in order of frequency.

Next, KH Coder's "Broader Context" option was used to observe the context of the node words. Thus, words were sorted in ascending order, utilizing the frequency value of the words that appear before or after a node word. The "frequency" value calculation for the KWIC method is shown in Equation (1),

$$f(w) = \sum_{i=1}^{5} \frac{(l_i + r_i)}{i},$$

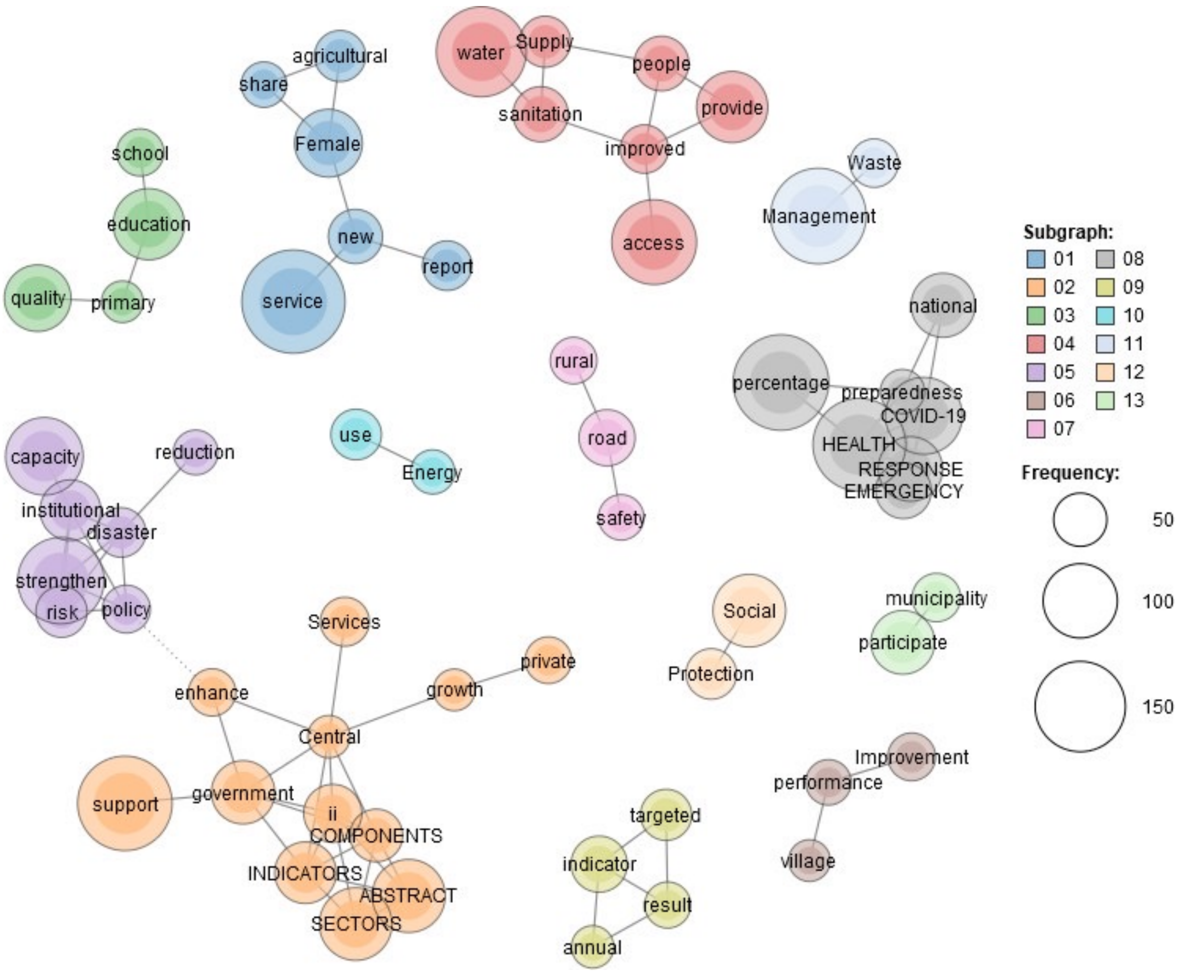
Where the frequency $'l_1'$ is the frequency of a particular word that appears just before the node word; $'l_2'$ is its frequency, two terms before the node word; $'r_1'$ is its frequency just after the node word; and $'r_2'$ is its frequency, two terms after the node word (Higuchi, 2016a, 2017).

The higher the frequency value of the particular word, w, that appears before or after the node word, the higher the value of f(w). In calculating f(w), the frequencies are divided and weighted according to their distance from the node word (Scott 2001). Thus, words that appear closer to the node word (for example, one or two words before or after the node word) have a higher weight than those located five words before or after it. SEI was therefore able to indicate the most common associations of the node words (figure 1).

Co-occurrence network

Additionally, the KH Coder Co-occurrence Network tool was applied to graphically display the associations between the most frequent words. This technique shows the node words in a chart that reveals the relative position of common words relative to others (Chen et al., 2016; Ozgur et al., 2008). Groupings of highly connected communities of words are joined by lines (or “edges”) and linked in a cluster with the words that frequently appear close to each other (figure 2). This use of lines to connect words to show a strong co-occurrence in the analyzed texts was developed by Fruchterman and Reingold (1991) and Kamada and Kawai (1989).

Figure 2: Co-occurrence Network of the 100 Projects



Source: An elaboration by SEI using KH coder software.

Note: Each circle represents words and terms that frequently appear in the 100 analyzed projects; the words are shown exactly as they appear in the database. The size of the circle shows the number of times the word appears in the texts. The circles are grouped in clusters or communities of words that frequently appear together, suggesting association. Each community is depicted in a different color.

Once the most frequent topics had been generated, SEI compared their meanings to the topics in each of the 169 SDG targets and assigned the queries to their corresponding targets. Because the themes in the World Bank documents did not exactly match the topics covered by the SDGs, a single World Bank theme could be mapped to various SDGs topics. The SEI team mapped the words and phrases from the controlled vocabulary to every appropriate SDG target. It is important to restate that this process was performed subjectively, so others doing the mapping with a different understanding of the SDGs and World Bank themes could yield different results.

In addition, SEI created Boolean queries from the official UN text that appears in the SDG targets (UN 2016) to account for the words and phrases that could have escaped the concordance and co- occurrence analyses, which show only the most frequent words and associations from the 100 projects. Each coding rule is composed of a series of Boolean queries of words or phrases that can thematically refer to the SDGs—every time the Boolean query appears in a text, the matching SDG is mapped to the project. This allows for multiple SDG mappings from a single project. The design of the code rules is described in the next section.

Phase II. Creating the Coding Rules

Writing Syntaxis

A TXT document was developed for each SDG coding rule. Coding rules are created using Boolean queries programming: before each code name, an "*" (asterisk) is added so that KH Coder recognizes it as the name of a code. In the following line, the condition is specified for that code and the kind of situation it should be applied to. Using this format, KH Coder allows as many codes as needed for text analysis.

A robust coding protocol was designed to avoid searching for and targeting single words, such as "agriculture," when appropriate. Instead, a combination of coding rules was applied, using rules such as AND, OR, NOT, "-," and "+," among others (Higuchi 2016a).

The coding used is explained below.

Single Phrase: This rule allows the specification of a phrase rather than a single word, such as "Word1+Word2+Word3+ ....“ For instance, a code for a phrase including "agricultural productivity" can be written as follows:

```
*Agricultural_productivity
 Agricultural+productivity
```

Nearby Words: This rule facilitates defining a code with specific words in an adjacent sentence structure or expression.

```
*climate_change_resilience
 near (climate-change-resilience)
```

Arithmetic Operators: As shown in table 2, there is also an option to use arithmetic operators to give a code based on how often a word occurs in a sentence structure (Higuchi, 2017).

Table 2: Standard Arithmetic Operators for the Coding Rules Document

<table>
<tr><th>Operators</th><th>Alternative</th></tr>
<tr><td>|</td><td>Or</td></tr>
<tr><td>&</td><td>And</td></tr>
<tr><td>!</td><td>Not</td></tr>
<tr><td>&!</td><td>And Not</td></tr>
<tr><td>|!</td><td>Or Not</td></tr>
</table>

Source: Obtained from Higuchi 2016b.

Note: These operators, or their alternatives, specify which words and phrases KH Coder is to use or exclude while mapping texts; in other words, they define the conditions for codes to be assigned to words or groups of words occurring in text. The operator "|" and its alternative "Or" are used to specify that a code corresponds to various individual words and that the KH coder will find a match if any of the words occurs in the text; the operator "&" and its alternative "And" are used to specify that multiple words must appear at the same time in the text; the "!" operator and its alternative "Not" are used to specify that, for a code to be mapped, one or more words most not appear in the text. Some of these operators can be combined to form complex operators, such as "&!" (And not), which is used to create the condition for mapping a code if a specific word appears in text and a second one does not; and "|!" (Or not) which is used to create the condition for mapping a code if either one word appears in text or a second one does not. Finally, operators can also be used in conjunction with single phrases and nearby words to specify mapping conditions for phrases.

Lastly, multiple conditions were used to define whether a word appears in the data set and the number of times it appears and identify synonyms. Thus, symbols such as "AND" (to determine when two words appear in the same sentence or paragraph), "OR" (to define synonyms), and "&!" ("and not" operator) were used. For instance, "TB or tuberculosis" means that "TB" and "Tuberculosis" are synonyms and should be counted as the same word when appearing in the text. Figure 3 shows the coding protocol designed for SDG 3.3.

Figure 3: Coding Protocol for SDG 3.3

***SDG 3.3**

HIV | AIDS | TB or tuberculosis | near(health-deficiency) | immune | virus | near(sexually-transmitted) | disease | epidemics or epidemic | COVID | Covid-19 | Sars-cov-2 | COVID-19 | Coronavirus | TB or tuberculosis | malaria | near(tropical-diseases) | hepatitis | near(communicable-diseases) | near(water- borne) | near(aids-epidemic) | near(end-epidemics) | miasma | contagion | plague | pest | near(neglected- tropical-diseases) | near(combat-diseases) | near(combat-hepatitis) | near(combat-communicable- diseases) | near(human-infection) | near(infection-control) | near(infectious-disease)

Source: An elaboration by SEI.
Note: These are the Boolean queries designed for SDG 3.3 based on the World Bank controlled vocabulary.

Phase III. Mapping

Crosstab of Codes

The final phase of the mapping process consisted in applying the coding rule file to the texts of the 100 projects. This was achieved by linking the TXT file with the Boolean queries to the KH Coder file using the Crosstab tool, with the SDG target specified as the “coding rule file” and the World Bank projects as the “coding units.” Then, a cross-tabulation was run to show the frequency of documents to which each code was applied, and KH Coder generated a table showing the positive or negative results for each project in terms of the SDG targets. The results from this mapping are described in section 3.

Validation of Results

To validate the results, 20 projects were randomly selected from the subset of 100 projects. To calculate the confidence interval of the validation sample, SEI applied equation (2), from Rodríguez del Águila and González-Ramírez (2014):

$$Sample\ size = \frac{Z_{\frac{\alpha}{2}}^{\ 2} * p * q * N}{(N-1) * e^{2} + \llbracket\left(Z_{\frac{\alpha}{2}}\rrbracket^{2} * p * q\right)},$$

Where:

$Z_{\frac{\alpha}{2}}^{\ 2}$: Z-score for a 90 percent confidence interval = 1.65,

p: probability of success classifying the project = 0.5,

q: probability of failure classifying the project = 0.5,

N: population size = 100,

e: margin of error = 0.165.

Using this equation, a sample size of 20 projects allows for validation with a confidence interval of 90 percent and a margin of error of 16.5 percent.

Next, the 20 projects were manually mapped to the SDG targets to obtain a conceptual mapping of the selection. Presumably, the results from the manual mapping would be similar to the ones from the methodology described above. Because the team of technicians that developed the Boolean queries conducted the manual mapping, the same understanding of the SDGs and World Bank themes used in creating the methodology was applied.

Then, the PPV was obtained to quantitively assess similarities between the manual and automatic mappings (see Chu 1999; Saito and Rehmsmeier 2015). This value is obtained by dividing the number of true positives (TP) in the data set (the SDG targets that were mapped in both the manual and automatic mappings) by the number of TP plus false positives (FP) (the mapping results that were obtained in the automatic mapping but not in the manual mapping). The PPV directly assesses the likelihood of the Boolean queries yielding a correct mapping result, as verified by the manual exercise. To see the full validation calculation, please see Appendix 3. Validation Exercise.

The PPV from the data set is calculated as:

$$PPV = \frac{\mathrm{TP}}{\mathrm{TP{+}FP}}.$$

Running this equation with the manual and automatic mapping samples produced a PPV of 75 percent.

SECTION 3:

Results

Section 3. Results

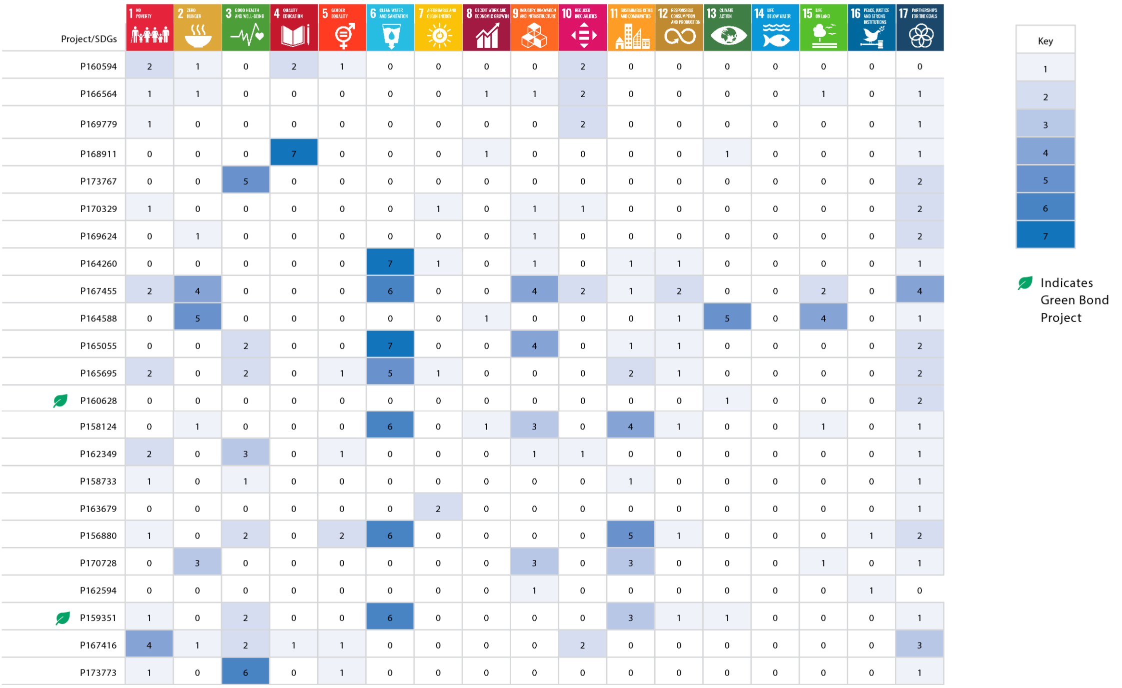
In interpreting the results on how investments in World Bank–financed projects contribute to the SDGs, it is important to remember that different methodologies are likely to yield different outcomes. This paper focuses on only a subset of World Bank–financed projects, the 100 projects added to the World Bank's project portfolio in fiscal 2020. It does not offer a complete representation of the full range of development activities performed by the World Bank across different regions and sectors or account for the differential and weighted impact that individual projects can have on the SDGs.

Additionally, given the interconnected nature of the SDGs, focusing on the results for individual SDGs may lead to inaccurate interpretations; a more holistic view of the results is recommended. For example, the results show the number of projects related to cross-cutting SDG topics such as gender equality and climate change and environmental issues. Each of the 100 projects was mapped to the SDGs at the target level, and the use of Boolean queries allowed us to identify more than one SDG target per project when applicable.

SDG Mapping Results in the *Impact Report 2020*

Impact reporting is part of the World Bank's efforts to build models for transparency and disclosure that help to develop sustainable capital markets. The Impact Report 2020, which reports on the use of proceeds from World Bank SDBs and Green Bonds, shows the results of the mapping exercise described in this paper. Table 3 shows one of the SDG mapping visualizations included in the impact report.

Table 3: Project Mapping per SDG



Project/SDGs	1 NO POVERTY	2 ZERO HUNGER	3 GOOD HEALTH AND WELL-BEING	4 QUALITY EDUCATION	5 GENDER EQUALITY	6 CLEAN WATER AND SANITATION	7 AFFORDABLE AND CLEAN ENERGY	8 DECENT WORK AND ECONOMIC GROWTH	9 INDUSTRY, INNOVATION AND INFRASTRUCTURE	10 REDUCED INEQUALITIES	11 SUSTAINABLE CITIES AND COMMUNITIES	12 RESPONSIBLE CONSUMPTION AND PRODUCTION	13 CLIMATE ACTION	14 LIFE BELOW WATER	15 LIFE ON LAND	16 PEACE, JUSTICE AND STRONG INSTITUTIONS	17 PARTNERSHIPS FOR THE GOALS
P160594	2	1	0	2	1	0	0	0	0	2	0	0	0	0	0	0	0
P166564	1	1	0	0	0	0	0	1	1	2	0	0	0	0	1	0	1
P169779	1	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1
P168911	0	0	0	7	0	0	0	1	0	0	0	0	1	0	0	0	1
P173767	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	2
P170329	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	2
P169624	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2
P164260	0	0	0	0	0	7	1	0	1	0	1	1	0	0	0	0	1
P167455	2	4	0	0	0	6	0	0	4	2	1	2	0	0	2	0	4
P164588	0	5	0	0	0	0	0	1	0	0	0	1	5	0	4	0	1
P165055	0	0	2	0	0	7	0	0	4	0	1	1	0	0	0	0	2
P165695	2	0	2	0	1	5	1	0	0	0	2	1	0	0	0	0	2
🍃 P160628	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2
P158124	0	1	0	0	0	6	0	1	3	0	4	1	0	0	1	0	1
P162349	2	0	3	0	1	0	0	0	1	1	0	0	0	0	0	0	1
P158733	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1
P163679	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1
P156880	1	0	2	0	2	6	0	0	0	0	5	1	0	0	0	1	2
P170728	0	3	0	0	0	0	0	0	3	0	3	0	0	0	1	0	1
P162594	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
🍃 P159351	1	0	2	0	0	6	0	0	0	0	3	1	1	0	0	0	1
P167416	4	1	2	1	1	0	0	0	0	2	0	0	0	0	0	0	3
P173773	1	0	6	0	1	0	0	0	0	0	0	0	0	0	0	0	1

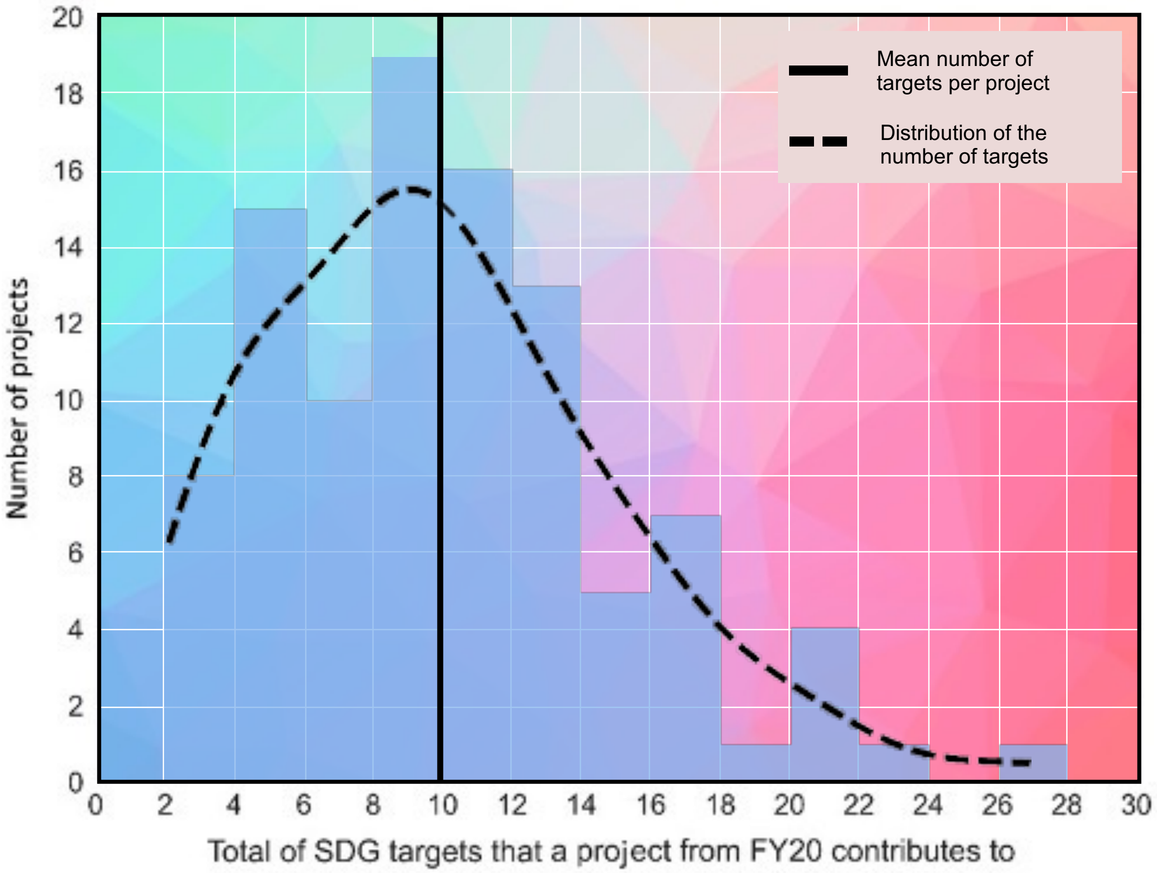
Key
1
2
3
4
5
6
7

🍃 Indicates Green Bond Project

Source: World Bank 2021.
Note: The table shows a portion of the results from the 100 analyzed texts. Each row corresponds to the World Bank project number in the far-left column. The Sustainable Development Goal (SDG) columns represent the number of times each project referred to the SDG. The darkness of the color corresponds to the number of targets mapped under that SDG, so the darker the tone, the greater number of SDG targets mapped to the project. To review the full results of this table for the 100 projects, please see "Appendix 2. Full Results of the Mapping Exercise, Project to Goal."

The results also show that every project is mapped to at least two SDG targets. On average, each project in the data set is mapped to 10 targets, and the mode is eight mapped targets per project (figure 4). Project P167455 has the most mapped targets (27), followed by P157575 with 23 targets, and then P158502, P165129, and P157245, with 20 mapped targets each. In total, the projects in the data set mapped to 109 SDG targets. This means that from the 169 Boolean queries protocols developed (corresponding to the 169 SDG targets), a sum of 109 protocols generated positive results at least in one of the 100 analyzed projects.

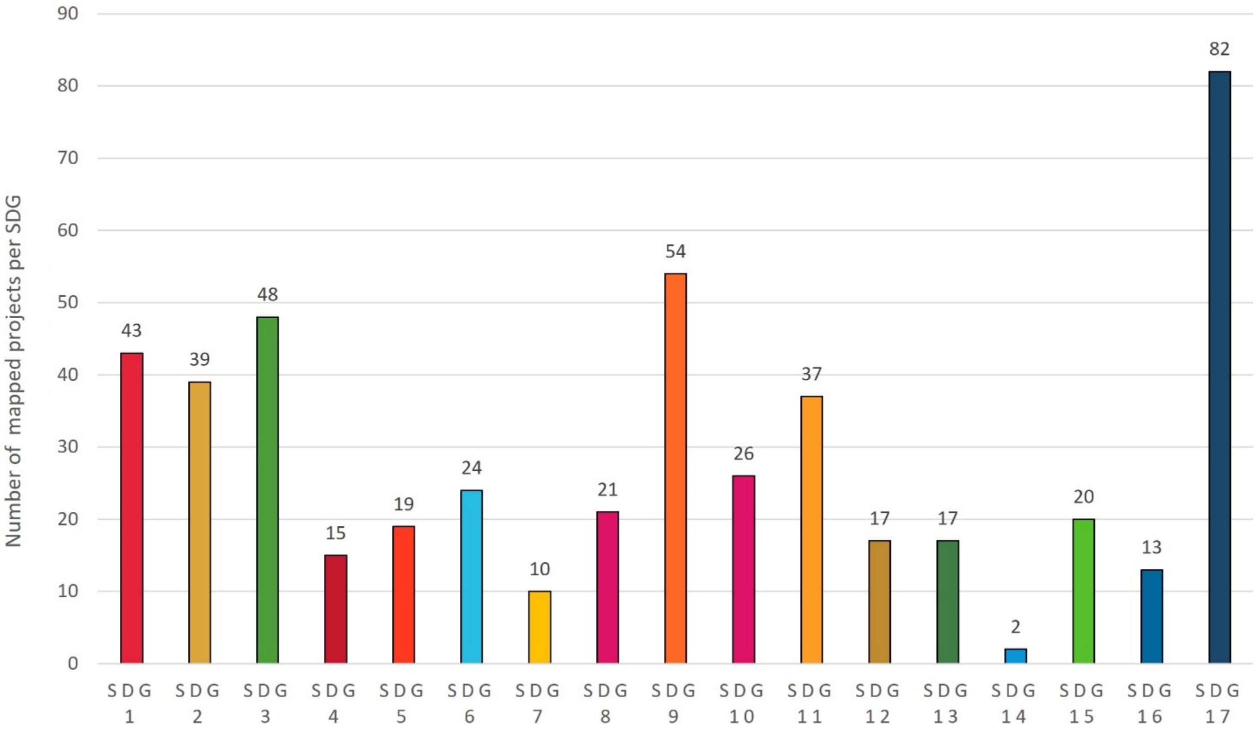
Figure 4: Distribution of SDG Targets Mapped per Project



Source: An elaboration by SEI.
Note: FY20 = fiscal year 2020; SDG = Sustainable Development Goal. On average for the 100 projects, a single project was mapped to 10 targets, while the mode is 8 targets per project. The highest number of targets per project is 27, and the lowest is 2.

For the subset of 100 projects analyzed, all 17 SDGs are matched to at least one project (figure 5). SDG 17 (Partnerships for the Goals) is the most represented, appearing in 82 projects, followed by SDG 9 (Industry, Innovation, and Infrastructure) and SDG 3 (Good Health and Well-Being), mapped to 54 and 48 projects, respectively. SDG 14 (Life below Water) appears the least, mapping to two projects, followed by SDG 7 (Affordable and Clean Energy) mapped to 10 projects.

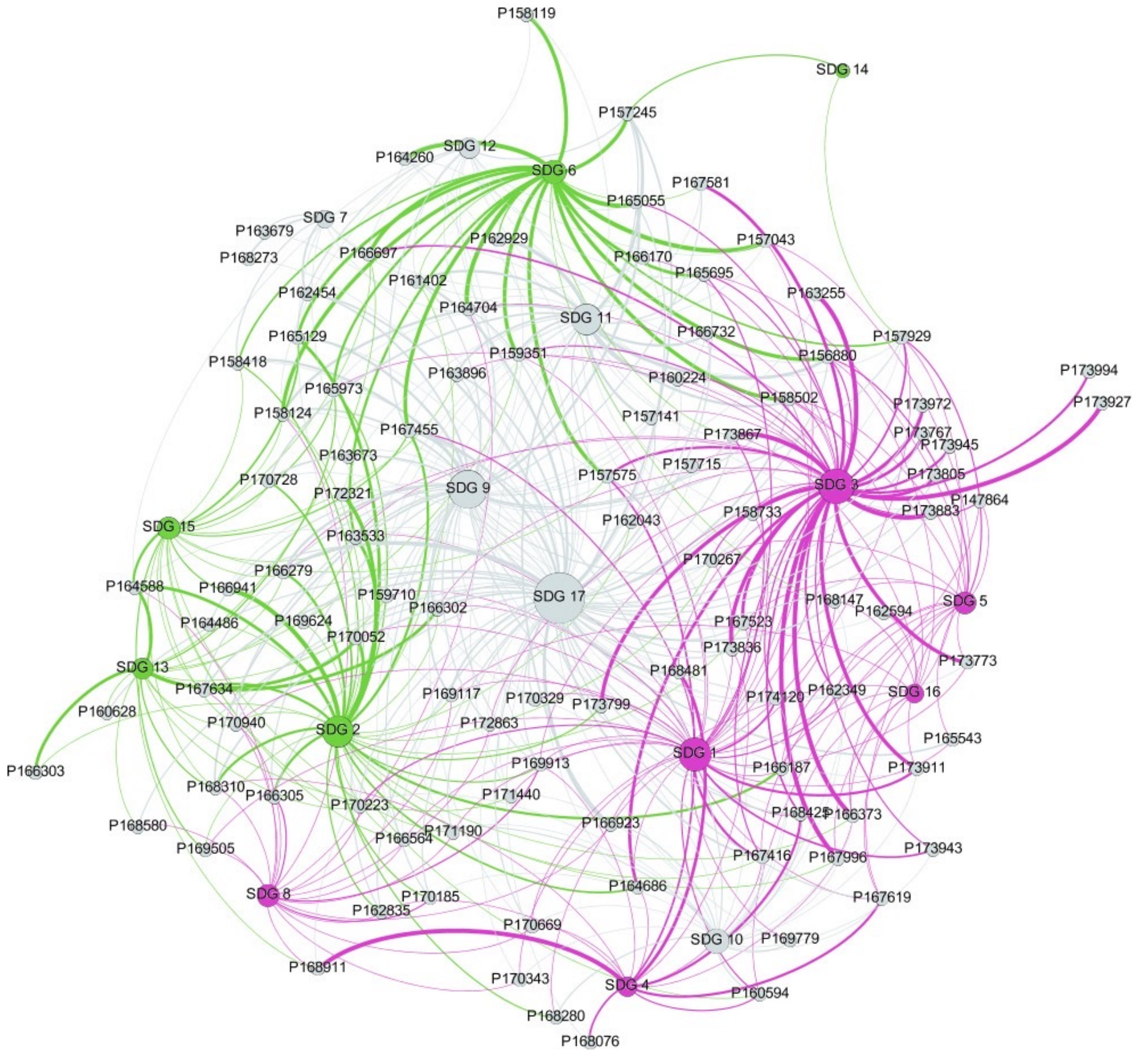
Figure 5: Number of Projects Mapped per SDG



Source: An elaboration by SEI.
Note: SDG = Sustainable Development Goal. This graph shows the total number of projects mapped under each SDG at the goal level for all 100 projects.

Of the projects mapped, some crosscut many SDGs, so a holistic view is recommended. The SDGs have an interconnected nature since they were designed to be an "integrated and indivisible" set, where every goal must be met to truly achieve sustainable development (UN 2016), so crosscutting by projects is to be expected. For instance, based on SEI's expertise and perspective on the 2030 Agenda, it can be argued that climate change and the management of ecosystems and biodiversity are predominant topics in SDGs 2 (Zero Hunger), 6 (Clean Water and Sanitation), 13 (Climate Action), 14 (Life below Water) and 15 (Life on Land). Similarly, according to UN Women (2018), the following SDGs are gender-sensitive: 1 (No Poverty), 3 (Good Health and Well-Being), 4 (Quality Education), 5 (Gender Equality), 8 (Economic Growth), and 16 (Peace, Justice and Strong Institutions). If these groupings are applied to the mapping results, 66 percent of the projects are connected with environmental and climate change SDGs, 79 percent are related to gender-sensitive SDGs, and 49 percent are connected to both groups of SDGs. Only 4 percent of projects are not connected to either environmental or gender-related SDGs (figure 6).

Figure 6: Network Graph of the Projects Mapped to Environmental and Gender-Related SDGs



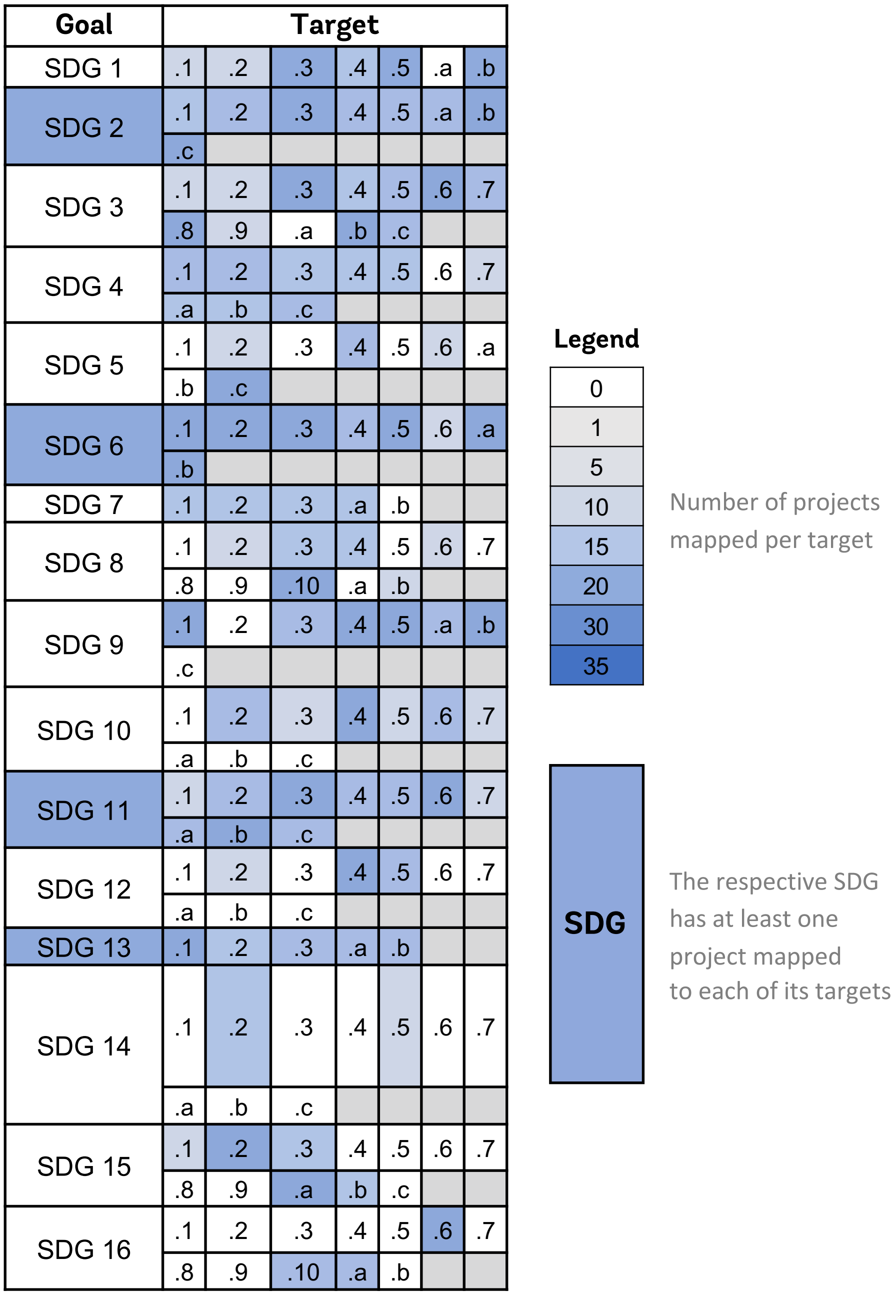
Source: An elaboration by SEI.
Note: SDG = Sustainable Development Goal. Environmental SDGs are in green, gender-sensitive SDGs are in purple, and other SDGs are shown in grey. The lines show how individual projects are connected to the SDGs at the goal level, with the line color corresponding to the type of SDG. The thickness of the line represents the number of targets that are mapped from a project to specific goals, as shown in table 3, so the thicker the line, the more connections there are with individual targets from one goal. Moreover, the size of the circles show the number of projects mapped under each SDG—larger circles depict a greater number of projects.

When examining the mapping to targets, target 17.9 (Enhancing capacity for the SDGs) appears the most (50 projects). Within that same SDG, target 17.5 (Investment promotion regimes for least developed countries) is mapped to 25 projects, and target 17.3 (Additional financial resources for developing countries) is mapped to 23. The frequent mapping of projects to SDG 17 targets may be explained by the World Bank's focus on capacity building and economic development in member countries. SDG 17 was designed within the 2030 Agenda to provide means of implementation to reach the other SDGs, such as financing, assistance, and capacity building (UN DESA n.d.).

It is interesting to note that targets 17.3, 17.5, and 17.9, while being representative of the World Bank objectives of ending extreme poverty and promoting shared prosperity, do not contain topics that may be more indicative of the type of action in the analyzed projects (such as health, food production, and climate change). In this regard, the prominence of SDG 17 does not reflect the thematic diversity of the analyzed data set, which contains actions referring to various sustainable development sectors. Accordingly, it makes sense to provide additional analysis of the mapping by controlling for SDG 17 targets in an attempt to better reveal the thematic inclination of the examined subset. This was achieved by removing all the connections to SDG 17 targets and reaggregating the data.

When controlling for the prevalence of SDG 17 in the mapping results, all targets of four SDGs are mapped to at least one project: SDG 2 (Zero Hunger), SDG 6 (Clean Water and Sanitation), SDG 11 (Sustainable Cities and Communities), and SDG 13 (Climate Action). Moreover, five SDGs have all but one of their targets mapped to at least one project: SDG 1 (No Poverty), SDG 3 (Good Health and Well-Being), SDG 4 (Quality Education), SDG 7 (Affordable and Clean Energy), and 9 (Industry, Innovation, and Infrastructure).

Figure 7: Project Frequency by SDG Target



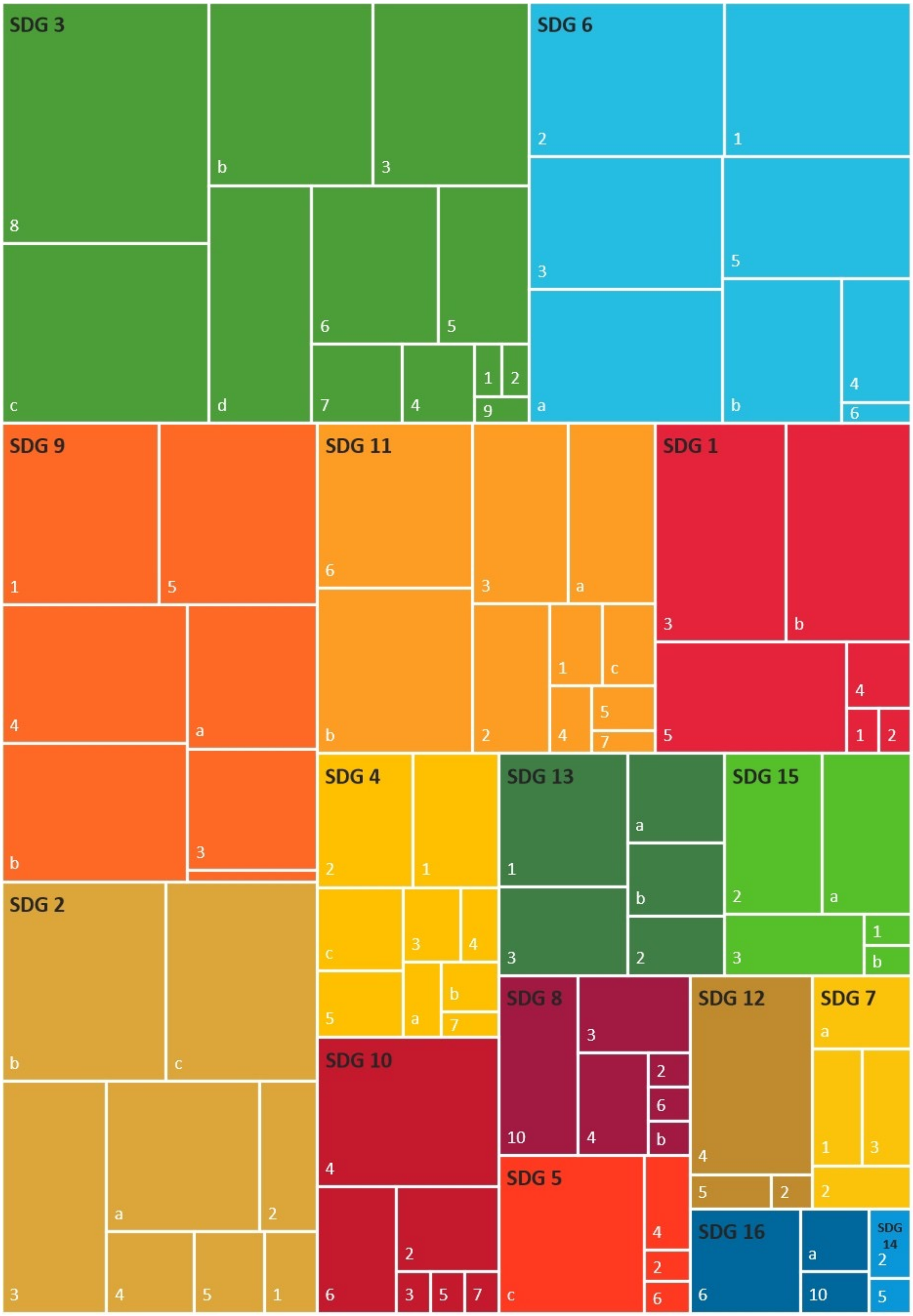
Goal	Target						
SDG 1	.1	.2	.3	.4	.5	.a	.b
SDG 2	.1	.2	.3	.4	.5	.a	.b
	.c						
SDG 3	.1	.2	.3	.4	.5	.6	.7
	.8	.9	.a	.b	.c		
SDG 4	.1	.2	.3	.4	.5	.6	.7
	.a	.b	.c				
SDG 5	.1	.2	.3	.4	.5	.6	.a
	.b	.c					
SDG 6	.1	.2	.3	.4	.5	.6	.a
	.b						
SDG 7	.1	.2	.3	.a	.b		
SDG 8	.1	.2	.3	.4	.5	.6	.7
	.8	.9	.10	.a	.b		
SDG 9	.1	.2	.3	.4	.5	.a	.b
	.c						
SDG 10	.1	.2	.3	.4	.5	.6	.7
	.a	.b	.c				
SDG 11	.1	.2	.3	.4	.5	.6	.7
	.a	.b	.c				
SDG 12	.1	.2	.3	.4	.5	.6	.7
	.a	.b	.c				
SDG 13	.1	.2	.3	.a	.b		
SDG 14	.1	.2	.3	.4	.5	.6	.7
	.a	.b	.c				
SDG 15	.1	.2	.3	.4	.5	.6	.7
	.8	.9	.a	.b	.c		
SDG 16	.1	.2	.3	.4	.5	.6	.7
	.8	.9	.10	.a	.b		

Source: An elaboration by SEI.

Note: SDG = Sustainable Development Goal. This visualization shows the number of projects mapped per target. In the goal column, the SDGs with all their targets mapped to at least one project are highlighted in blue. In the target column, the cell color shows the number of projects mapped to each target. The intensity of the blue colors shows the number of projects mapped per target (with darker colors representing higher values), while cells in white mean that there are no mapped projects to the target. The numbers and letters under the Target column correspond to the related target to each SDG. The results from SDG 17 were subtracted to produce this data.

On the other hand, by controlling for SDG 17, at the target level (figure 8), targets 3.8 (Achieve universal health coverage), 3.c (Increase health financing), and 2.b (Prevent agricultural market restrictions) are the ones that appear the most, with 35, 26, and 23 mapped projects, respectively. Then, targets 2.c (Ensure proper functioning of food commodity markets) and 6.2 (Access to sanitation) have 21 mapped projects each, followed by targets 3.3 (End communicable diseases), 6.1 (Access to safe drinking water), 9.1 (Develop resilient infrastructure), and 9.5 (Enhance technical and industrial innovation), with 20 projects each.

Figure 8: Tree map of the 100 Projects Mapped



Source: An elaboration by SEI.

Note: SDG = Sustainable Development Goal. The squares represent the SDGs, with the largest squares per color showing the goals and smaller squares and font size showing the targets. The size of each square is proportional to the number of projects mapped under the corresponding goal or target. The results mapped to SDG 17 were removed (controlled for) to produce this data.

SECTION 4:

Discussion

Section 4. Discussion

As we approach 2030, it is essential to know how different development activities are aligned and contributing to the SDGs. This is essential to foster private sector financing toward sustainable activities, and the SDGs are a widely recognized framework that can be used to increase transparency around the impact of investments.

As section 1 shows, a growing family of methodologies is providing an alternative means of linking activities or investments to the SDGs. For example, some approaches conceptually map indicators or categories to the SDGs, assigning any activity that uses such an indicator or category to the corresponding SDG or target. This process is efficient in terms of time and technical capacity. However, since the original indicators and categories used to group the non-SDG activities are different from the ones in the SDGs, these categories do not always map one to one with the SDGs and may lead to imprecise mapping; in fact, some categories are difficult to map to the SDGs because of their thematic differences.

This paper favors approaches that allow for closer inspection of the analyzed activities, mapping them directly to the SDGs without intermediate categories. However, this can be time and resource-intensive if done manually, so methods that automatically perform this process are needed. Therefore, the methodology proposed in this paper employed an automatic detection method based on text mining to map projects to the SDGs. Boolean queries based on a controlled vocabulary from World Bank projects and taxonomy were used to mitigate false results from polysemy (see Bordignon 2021).

The results show indicative linkages of a subset of World Bank–financed projects to the SDGs, and a strong connection to SDG 17 in particular, which may be explained by the World Bank's focus on capacity building and economic development in member countries. However, it is important to note that there are limitations found in the boundaries of the mapping exercise that should be considered when interpreting the results.

The first limitation is that the study analyzes only 100 projects, a subset of the roughly 600 active projects in the fiscal 2020 World Bank project portfolio. Therefore, the results cannot provide a complete picture of the range, results, and contribution of World Bank development projects and activities to the SDGs. It is worth mentioning that a broader mapping to include the entire World Bank project portfolio was published in the World Bank SDBs and Green Bonds Impact Report covering fiscal 2021 (World Bank 2022), with the Boolean queries tested against the whole set of World Bank projects active in fiscal 2021 (approximately 700 projects). The results were manually validated using the process described in section 2. Scaling up the methodology included adjusting the controlled vocabulary and text mining protocol and provides more accurate results and a more comprehensive picture of the IBRD portfolio of development activities.

Moreover, the adjustment of the Boolean queries allowed controlling for SDG 17 (Partnerships for the goals). Because many of the SDG 17 targets are about financial assistance and capacity building in developing countries, and therefore relate to most World Bank projects, the following targets were removed from the Boolean queries in the fiscal 2021 analysis: 17.2 (Assistance for least developing countries), 17.3 (Additional financial resources for developing countries), and 17.9 (Enhancing capacity for the SDGs). The remaining SDG 17 targets were revised against the vocabulary in the new subset of fiscal 2021 projects to design Boolean queries that capture specific themes such as investment promotion regimes (17.5), scientific and technological development (targets 17.6, 17.7, and 17.8), market regulations (targets 17.10, 17.11, and 17.12), policy coherence (targets 17.13, 17.14 and 17.15), public and private partnerships (targets 17.16 and 17.17), and monitoring (targets 17.18 and 17.19).

The second limitation of the mapping exercise is that the developed methodology represents one of the many solutions available for indicative mapping of investments to the SDGs. The constructed text mining protocol is subjective, and the assigning of meaning to the keywords and phrases used in KH Coder was based on the particular understanding of SDGs and World Bank themes by the team that applied the methodology. Because the SDGs relate to politically sensitive topics, their meaning and the ways to achieve them can vary (see Mair et al. 2018), and polysemy (see Bordignon 2021) affects understanding, various text mining protocols corresponding to different understandings of the SDGs are possible. Different mapping methodologies are likely to result in distinct, and perhaps divergent, mapping results.

The third limitation is the most significant of the proposed SDG mapping approach: the approach does not account for the differential and weighted impact that individual projects can have on the SDGs. For instance, the results show that there are 54 projects mapped to SDG 9 (Industry, Innovation, and Infrastructure), while only two projects fall under SDG 14 (Life below Water), but no further indicators or information are provided to understand the relative impact or contribution of the 54 projects under SDG 9 versus the two projects mapped under SDG 14. The above limitation is also true for every mapping methodology examined in the literature review, except for KfW's approach (Dangelmaier 2019), which shows SDG contributions of KfW's activities using financial resources as a weighting factor. However, like all the methodologies, this one also has limitations, as it does not account for the efficiency of the use of the financial resources deployed. Further, with this approach, the use of financial resources for weighting impact does not account for spillover effects on different sectors.

This last limitation points toward an interesting research area: the blending of SDG mapping and SDG interaction methodologies. It is worth mentioning that the latter type of methodologies conforms to a growing area of research attempting to calculate how SDGs impact one another through synergies and trade-offs in a specific geography (Bennich, Weitz, and Carlsen 2020). In this regard, the French investment bank Natixis (2018) has already called for the use of causation-based and benchmarking methodologies to model how development and investment activities impact the SDGs. Natixis also argues that SDG mapping methodologies must consider location-dependent SDG interlinkages to assess which activities can be the most synergistic in a region.

This paper argues for another possibility: connecting SDG mapping methodologies with the emerging field of SDG interactions. This would facilitate the understanding of how individual development policies interconnect with the SDGs, showing different levels of impact for different actions. For instance, the degree of how impactful one project is to the SDGs could be measured based on how many synergistic SDGs the project is mapped to. On the contrary, if the project is mapped to SDGs producing trade-offs, it could have a negative impact.

Nonetheless, previously generated interaction values must exist to calculate such SDG interactions index for any project. Therefore, the calculation of SDG interaction values in different countries and regions also requires research. However, the use of mapping approaches in SDG interaction studies would allow for a more general application of SDG interaction methodologies, as it would permit the quick conversion of actions and policies to the SDGs. Furthermore, linking SDG mapping and SDG interaction methodologies would also allow the modeling of interlinkages using financial investments, among other indicators. Some authors already suggest this (for instance, Forouli et al. 2020; Guerrero and Castañeda 2020). The relationship between SDG mapping and interaction methodologies is a topic for further research and discussion.

Conclusion

With the SDGs providing a common reference framework to measure the impact of development activities toward sustainability, the mapping of investments to the SDGs is critical with less than a decade remaining to achieve the SDGs. Future investments must be made strategically. In this regard, this paper aims to contribute to the ongoing SDG mapping efforts through the development of a methodology that uses World Bank projects to observe their indicative linkages to the SDGs and contributes to the World Bank's efforts to promote sustainable capital markets with a focus on transparency and disclosure. This work involved the creation of a text mining methodology that was applied to a subset of IBRD-funded projects that were added to the World Bank SDB project portfolio in fiscal 2020.

To obtain more accurate results, the approach uses Boolean queries based on vocabulary specific to the World Bank. As the mapping shows, the use of a controlled vocabulary for the World Bank activities allowed us to obtain results with a PPV of 75 percent, validated through a random selection of 20 projects that were manually mapped. In addition, in analyzing the results, the reader should recognize the deep interconnections between the SDGs and would benefit from using a holistic approach to interpreting results to potentially avoid any one-sided analysis.

The authors recognize three main limitations of the experiment and their methodology. Firstly, the mapping exercise analyzes only 100 projects, a small subset of the roughly 600 projects in the World Bank's fiscal 2020 project portfolio. Therefore, the results cannot provide a complete picture of the linkages of World Bank development projects and activities to the SDGs. Yet, to scale up the approach described in this paper, the methodology has been applied to map the entire World Bank project portfolio active in fiscal 2021, offering a more comprehensive picture of World Bank development activities. The new analysis also allows for adjustments in the controlled vocabulary and the resulting text mining protocol as well as control for the prevalence of SDG 17, due to the Bank's focus on capacity building and economic development, producing more accurate results. Secondly, the developed methodology is subjective, and different mapping methodologies could produce divergent results. And thirdly, and most importantly, the methodology does not calculate the different weighted impacts that individual development actions can have on the SDGs. This is an important caveat to this paper's approach that is also present in the other mapping methodologies identified in the literature review.

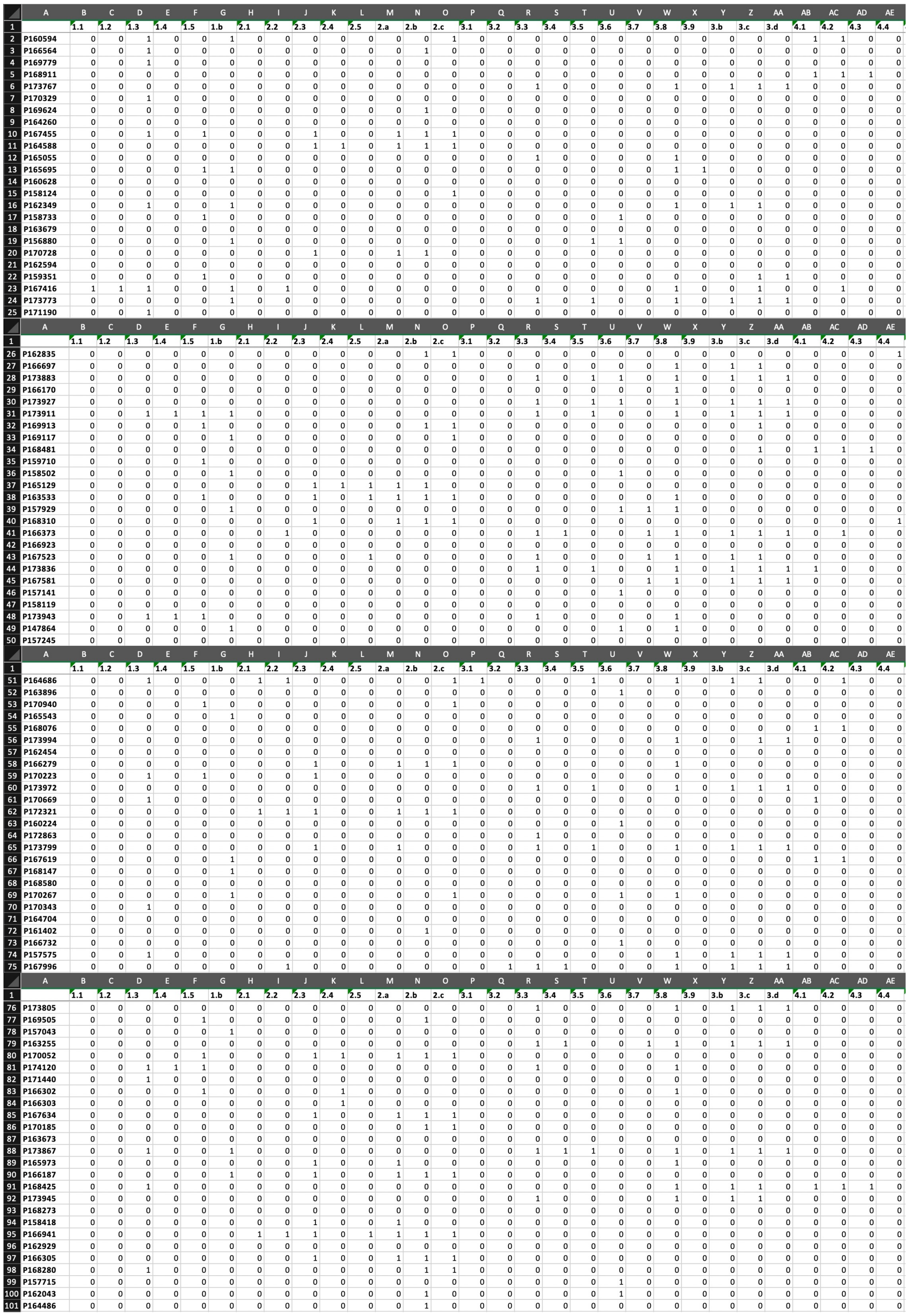
In this regard, the weighting of the impact of individual actions is perhaps the most important research opportunity in the continuing development of SDG mapping methodologies. This paper proposes connecting SDG mapping approaches with SDG interaction methodologies to address this gap. The complementation between SDG mapping and SDG interaction methodologies could allow for the calculation of SDG interaction indexes for individual projects while facilitating the conversion of non-SDG activities into SDGs, which would simplify the application of SDG interaction methodologies. This endeavor could facilitate the implementation of the SDGs globally and accelerate their achievement through the strategic allocation of investments in the remaining years of the 2030 Agenda.

Finally, the proposed methodology may be adapted in the broader sustainable finance space if scaled up and further developed, through artificial intelligence (such as, machine learning, NLP, and sentiment analysis), for example, to support the growing requirements for disclosures of investments and risks related to sustainable finance and climate. As new regulations and taxonomies emerge at the national and regional levels for setting definitions and standards around sustainable activities, capital markets participants will be increasingly subject to reporting on the alignment of their investments to these emerging frameworks. To comply with these requirements, investors will need to undertake extensive analysis of information and large data sets that will benefit from tools that can automate the process. Text mining approaches such as the one presented in this paper, which allows for the connection of activities and investments to specific topics (keywords), may be scaled up and adapted to assess the alignment of investments to sustainable finance targets to support net-zero trajectories and activities that contribute to social good.

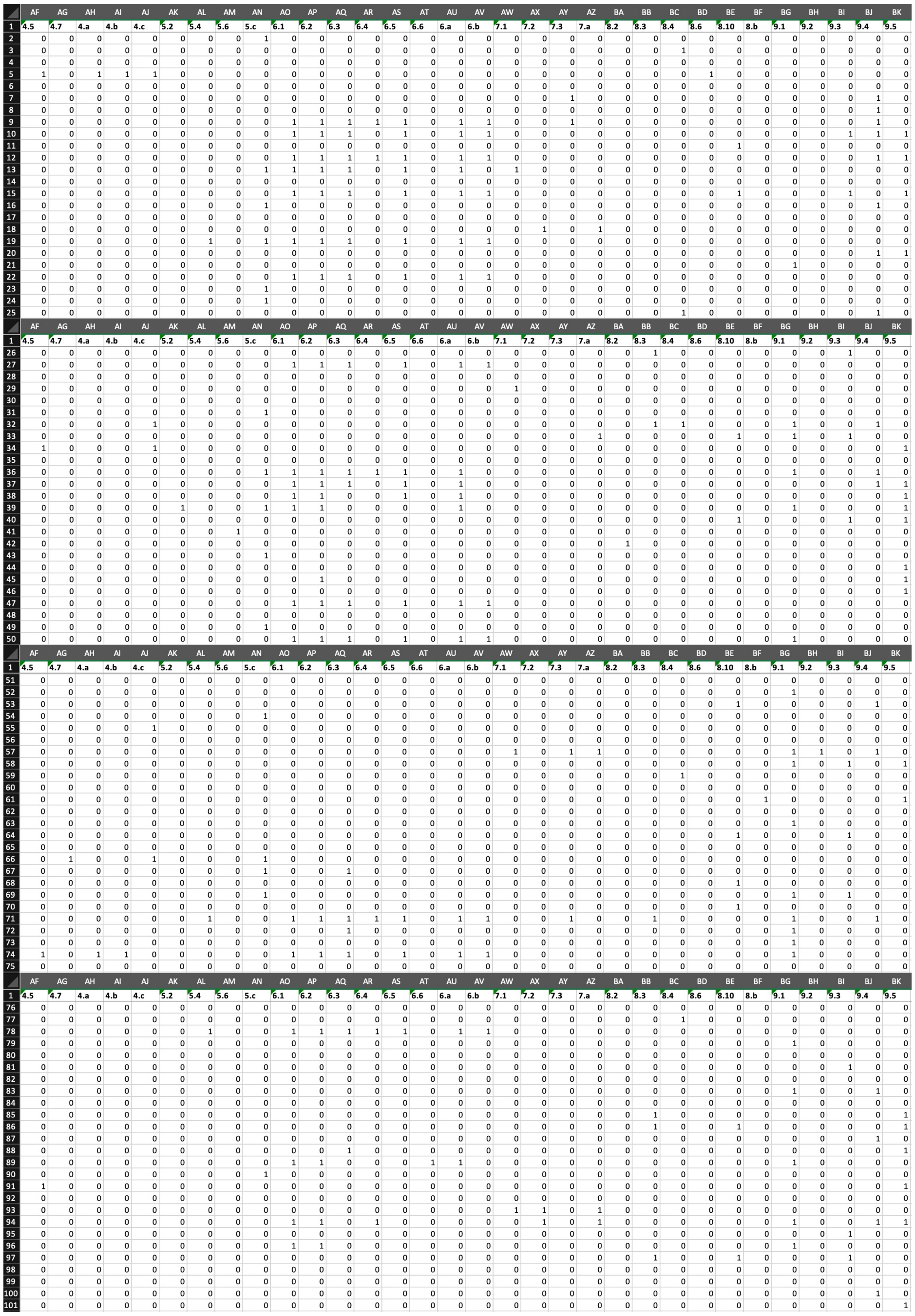
Appendixes

Appendix 1. Full Results of the Mapping Exercise, Project to Target

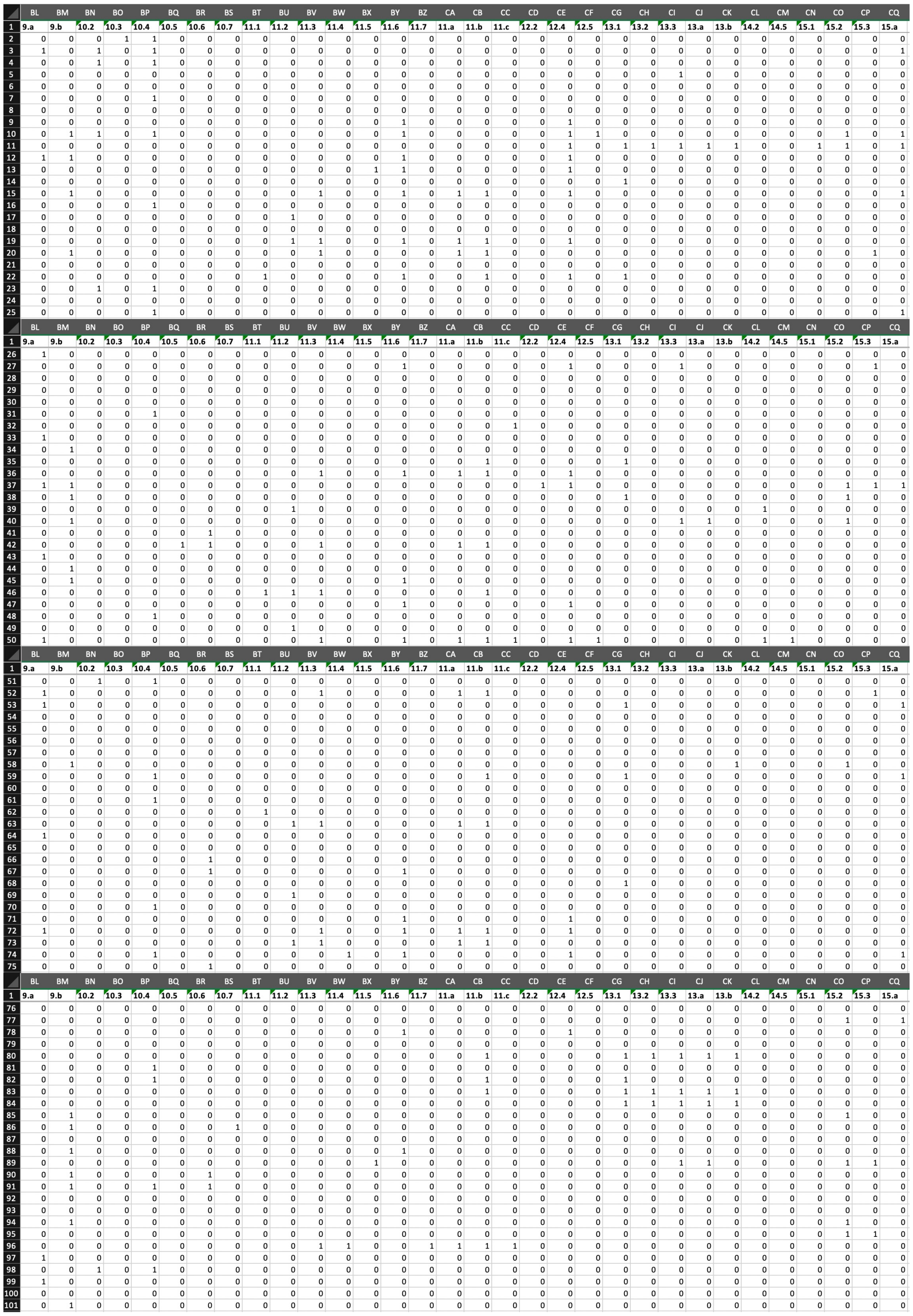
The following table shows the full mapping results of the 100 projects from fiscal 2020 at the SDG target level. Each row corresponds to the World Bank project number in the far-left column. The SDG columns show the connections between projects and SDG targets: a value of “0” indicates no mapping, whereas a value of “1” indicates that the project is mapped to the corresponding target. The final column shows the total number of mapped targets for each project.



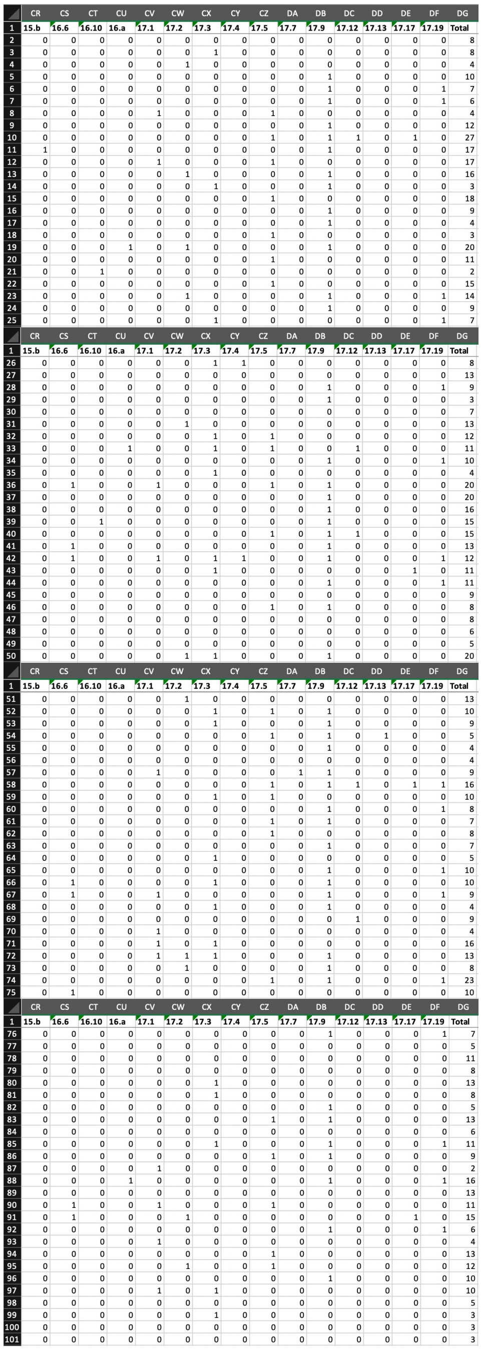
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
1		1.1	1.2	1.3	1.4	1.5	1.b	2.1	2.2	2.3	2.4	2.5	2.a	2.b	2.c	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	3.b	3.c	3.d	4.1	4.2	4.3	4.4
2	P160594	0	0	1	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
3	P166564	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	P169779	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	P168911	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
6	P173767	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	1	1	0	0	0	0
7	P170329	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	P169624	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	P164260	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	P167455	0	0	1	0	1	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	P164588	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
12	P165055	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
13	P165695	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
14	P160628	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	P158124	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	P162349	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0
17	P158733	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
18	P163679	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	P156880	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
20	P170728	0	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	P162594	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	P159351	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
23	P167416	1	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0
24	P173773	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	1	1	1	0	0	0	0
25	P171190	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	P162835	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
27	P166697	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0
28	P173883	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	0	1	1	1	0	0	0	0
29	P166170	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
30	P173927	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	0	1	1	1	0	0	0	0
31	P173911	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	1	1	1	0	0	0	0
32	P169913	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
33	P169117	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	P168481	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0
35	P159710	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	P158502	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
37	P165129	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	P163533	0	0	0	0	1	0	0	0	1	0	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
39	P157929	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
40	P168310	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
41	P166373	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	1	1	1	0	1	0	0
42	P166923	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	P167523	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	1	0	1	1	0	0	0	0	0
44	P173836	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	1	1	1	1	0	0	0
45	P167581	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0	0	0
46	P157141	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
47	P158119	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	P173943	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
49	P147864	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
50	P157245	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
51	P164686	0	0	1	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	1	1	0	0	1	0	0
52	P163896	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
53	P170940	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
54	P165543	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
55	P168076	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
56	P173994	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	1	0	0	0	0
57	P162454	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
58	P166279	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
59	P170223	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
60	P173972	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0	1	1	1	0	0	0	0
61	P170669	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
62	P172321	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
63	P160224	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
64	P172863	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
65	P173799	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1	0	1	0	0	1	0	1	1	1	0	0	0	0
66	P167619	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
67	P168147	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
68	P168580	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
69	P170267	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
70	P170343	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
71	P164704	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
72	P161402	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
73	P166732	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
74	P157575	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	0
75	P167996	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	1	1	1	0	0	0	0
76	P173805	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	1	1	0	0	0	0
77	P169505	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
78	P157043	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
79	P163255	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	1	1	1	0	0	0	0
80	P170052	0	0	0	0	1	0	0	0	1	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
81	P174120	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
82	P171440	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
83	P166302	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
84	P166303	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
85	P167634	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
86	P170185	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
87	P163673	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
88	P173867	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	1	1	1	0	0	0	0
89	P165973	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
90	P166187	0	0	0	0	0	1	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
91	P168425	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	1	1	1	0
92	P173945	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	1	0	0	0	0	0
93	P168273	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
94	P158418	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
95	P166941	0	0	0	0	0	0	1	1	1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
96	P162929	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
97	P166305	0	0	0	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
98	P168280	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
99	P157715	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
100	P162043	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
101	P164486	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH	BI	BJ	BK
1	4.5	4.7	4.a	4.b	4.c	5.2	5.4	5.6	5.c	6.1	6.2	6.3	6.4	6.5	6.6	6.a	6.b	7.1	7.2	7.3	7.a	8.2	8.3	8.4	8.6	8.10	8.b	9.1	9.2	9.3	9.4	9.5
2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
9	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0
10	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
13	0	0	0	0	0	0	0	0	1	1	1	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1
16	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	1	0	1	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
22	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0
27	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0
33	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	1	0	0
34	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
37	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
38	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
39	0	0	0	0	0	1	0	0	1	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1
41	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
43	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
45	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
46	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
47	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0
54	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
55	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	1	1	0	1	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1
59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
66	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
67	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
69	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
70	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
71	0	0	0	0	0	0	1	0	0	1	1	1	1	1	0	1	1	0	0	1	0	0	1	0	0	0	0	1	0	0	1	0
72	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
73	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
74	1	0	1	1	0	0	0	0	0	1	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
75	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
76	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
77	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
78	0	0	0	0	0	0	1	0	0	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
79	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
81	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
83	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
85	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
86	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1
87	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
88	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
89	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
90	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
91	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
92	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
93	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0
94	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	1	1
95	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
96	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
97	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0
98	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
99	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1



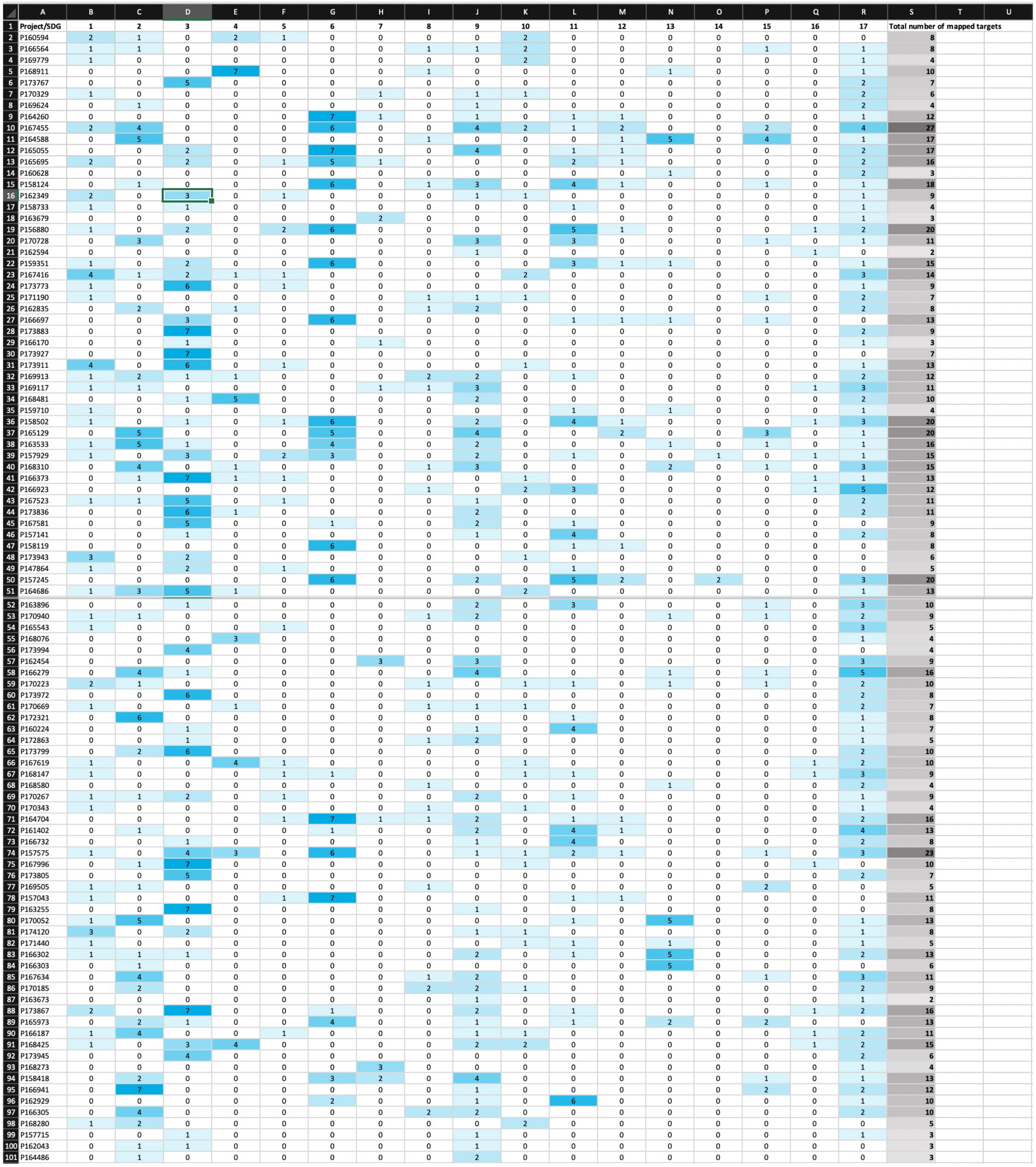
	BL	BM	BN	BO	BP	BQ	BR	BS	BT	BU	BV	BW	BX	BY	BZ	CA	CB	CC	CD	CE	CF	CG	CH	CI	CJ	CK	CL	CM	CN	CO	CP	CQ
1	9.a	9.b	10.2	10.3	10.4	10.5	10.6	10.7	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.a	11.b	11.c	12.2	12.4	12.5	13.1	13.2	13.3	13.a	13.b	14.2	14.5	15.1	15.2	15.3	15.a
2	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
4	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
10	0	1	1	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	1
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	1	0	0	1	1	0	1
12	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
15	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
16	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
20	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
23	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
26	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
32	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
34	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
36	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
37	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	1
38	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0
39	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
40	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0
41	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
42	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
43	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
44	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
45	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
46	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
47	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
48	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
49	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
50	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	1	0	1	1	0	0	0	0	0	1	1	0	0	0	0
51	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
52	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
53	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
54	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
55	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
58	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
59	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
61	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
62	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
63	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
64	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
65	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
66	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
67	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
68	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
69	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
70	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
71	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
72	1	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
73	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
74	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
75	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
76	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
77	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
78	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
79	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0
81	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
82	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
83	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0
84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0
85	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
86	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
87	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
88	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
89	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	1	0
90	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
91	0	1	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
92	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
93	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
94	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
95	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
96	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
97	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
98	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
99	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
101	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



	CR	CS	CT	CU	CV	CW	CX	CY	CZ	DA	DB	DC	DD	DE	DF	DG
1	15.b	16.6	16.10	16.a	17.1	17.2	17.3	17.4	17.5	17.7	17.9	17.12	17.13	17.17	17.19	Total
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8
3	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	8
4	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	4
5	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	10
6	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	7
7	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	6
8	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	4
9	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	12
10	0	0	0	0	0	0	0	0	1	0	1	1	0	1	0	27
11	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	17
12	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	17
13	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	16
14	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	3
15	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	18
16	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	9
17	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	4
18	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3
19	0	0	0	1	0	1	0	0	0	0	1	0	0	0	0	20
20	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	11
21	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	2
22	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	15
23	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	14
24	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	9
25	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	7
26	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	8
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13
28	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	9
29	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	3
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7
31	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	13
32	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	12
33	0	0	0	1	0	0	1	0	1	0	0	1	0	0	0	11
34	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	10
35	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	4
36	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	20
37	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	20
38	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	16
39	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	15
40	0	0	0	0	0	0	0	0	1	0	1	1	0	0	0	15
41	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	13
42	0	1	0	0	1	0	1	1	0	0	1	0	0	0	1	12
43	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	11
44	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	11
45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	9
46	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	8
47	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8
48	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
50	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0	20
51	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	13
52	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	10
53	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	9
54	0	0	0	0	0	0	0	0	1	0	1	0	1	0	0	5
55	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	4
56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
57	0	0	0	0	1	0	0	0	0	1	1	0	0	0	0	9
58	0	0	0	0	0	0	0	0	1	0	1	1	0	1	1	16
59	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	10
60	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	8
61	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	7
62	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	8
63	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	7
64	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	5
65	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	10
66	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	10
67	0	1	0	0	1	0	0	0	0	0	1	0	0	0	1	9
68	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	4
69	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	9
70	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	4
71	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	16
72	0	0	0	0	1	1	1	0	0	0	1	0	0	0	0	13
73	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	8
74	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	23
75	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	10
76	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	7
77	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
78	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	11
79	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8
80	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	13
81	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	8
82	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	5
83	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	13
84	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6
85	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	11
86	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	9
87	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2
88	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	16
89	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13
90	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	11
91	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	15
92	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	6
93	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	4
94	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	13
95	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	12
96	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	10
97	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	10
98	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
99	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3
100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
101	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3

Appendix 2. Full Results of the Mapping Exercise, Project to Goal

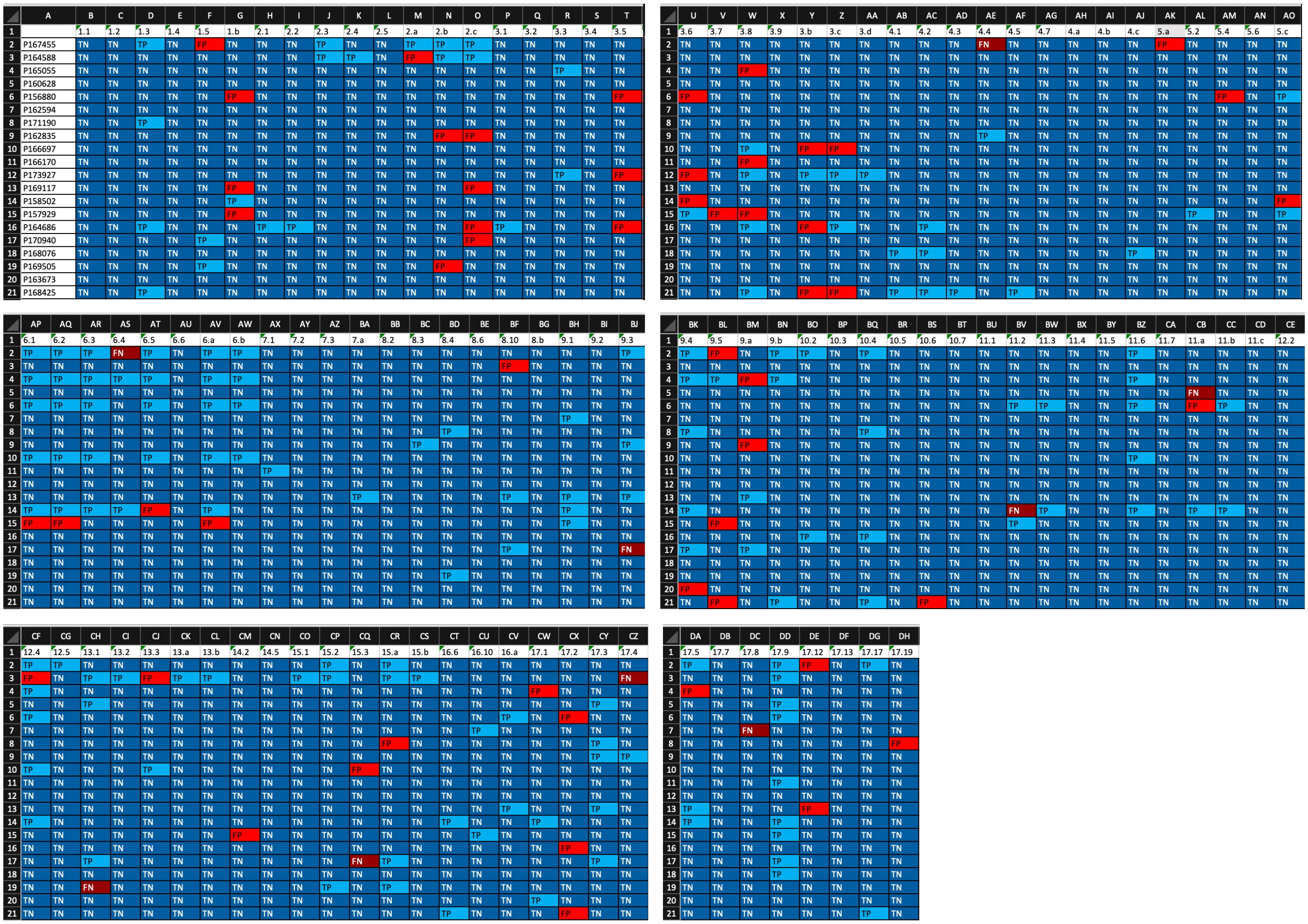
The following table shows the full mapping results of the 100 projects from fiscal 2020 at the goal level. Each row corresponds to the World Bank project number in the far-left column. The SDG columns represent the number of times each project referred to the goal. The darkness of the color corresponds to the number of targets mapped under that goal, so the darker the tone, the greater number of SDG targets mapped to the project.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Project/SDG	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Total number of mapped targets		
2	P160594	2	1	0	2	1	0	0	0	0	2	0	0	0	0	0	0	0	8		
3	P166564	1	1	0	0	0	0	0	1	1	2	0	0	0	0	1	0	1	8		
4	P169779	1	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	1	4		
5	P168911	0	0	0	7	0	0	0	1	0	0	0	0	1	0	0	0	1	10		
6	P173767	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	2	7		
7	P170329	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	2	6		
8	P169624	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	2	4		
9	P164260	0	0	0	0	0	7	1	0	1	0	1	1	0	0	0	0	1	12		
10	P167455	2	4	0	0	0	6	0	0	4	2	1	2	0	0	2	0	4	27		
11	P164588	0	5	0	0	0	0	0	1	0	0	0	1	5	0	4	0	1	17		
12	P165055	0	0	2	0	0	7	0	0	4	0	1	1	0	0	0	0	2	17		
13	P165695	2	0	2	0	1	5	1	0	0	0	2	1	0	0	0	0	2	16		
14	P160628	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2	3		
15	P158124	0	1	0	0	0	6	0	1	3	0	4	1	0	0	1	0	1	18		
16	P162349	2	0	3	0	1	0	0	0	1	1	0	0	0	0	0	0	1	9		
17	P158733	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	4		
18	P163679	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	1	3		
19	P156880	1	0	2	0	2	6	0	0	0	0	5	1	0	0	0	1	2	20		
20	P170728	0	3	0	0	0	0	0	0	3	0	3	0	0	0	1	0	1	11		
21	P162594	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	2		
22	P159351	1	0	2	0	0	6	0	0	0	0	3	1	1	0	0	0	1	15		
23	P167416	4	1	2	1	1	0	0	0	0	2	0	0	0	0	0	0	3	14		
24	P173773	1	0	6	0	1	0	0	0	0	0	0	0	0	0	0	0	1	9		
25	P171190	1	0	0	0	0	0	0	1	1	1	0	0	0	0	1	0	2	7		
26	P162835	0	2	0	1	0	0	0	1	2	0	0	0	0	0	0	0	2	8		
27	P166697	0	0	3	0	0	6	0	0	0	0	1	1	1	0	1	0	0	13		
28	P173883	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	2	9		
29	P166170	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	3		
30	P173927	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7		
31	P173911	4	0	6	0	1	0	0	0	0	1	0	0	0	0	0	0	1	13		
32	P169913	1	2	1	1	0	0	0	2	2	0	1	0	0	0	0	0	2	12		
33	P169117	1	1	0	0	0	0	1	1	3	0	0	0	0	0	0	1	3	11		
34	P168481	0	0	1	5	0	0	0	0	2	0	0	0	0	0	0	0	2	10		
35	P159710	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	4		
36	P158502	1	0	1	0	1	6	0	0	2	0	4	1	0	0	0	1	3	20		
37	P165129	0	5	0	0	0	5	0	0	4	0	0	2	0	0	3	0	1	20		
38	P163533	1	5	1	0	0	4	0	0	2	0	0	0	1	0	1	0	1	16		
39	P157929	1	0	3	0	2	3	0	0	2	0	1	0	0	1	0	1	1	15		
40	P168310	0	4	0	1	0	0	0	1	3	0	0	0	2	0	1	0	3	15		
41	P166373	0	1	7	1	1	0	0	0	0	1	0	0	0	0	0	1	1	13		
42	P166923	0	0	0	0	0	0	0	1	0	2	3	0	0	0	0	1	5	12		
43	P167523	1	1	5	0	1	0	0	0	1	0	0	0	0	0	0	0	2	11		
44	P173836	0	0	6	1	0	0	0	0	2	0	0	0	0	0	0	0	2	11		
45	P167581	0	0	5	0	0	1	0	0	2	0	1	0	0	0	0	0	0	9		
46	P157141	0	0	1	0	0	0	0	0	1	0	4	0	0	0	0	0	2	8		
47	P158119	0	0	0	0	0	6	0	0	0	0	1	1	0	0	0	0	0	8		
48	P173943	3	0	2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	6		
49	P147864	1	0	2	0	1	0	0	0	0	0	1	0	0	0	0	0	0	5		
50	P157245	0	0	0	0	0	6	0	0	2	0	5	2	0	2	0	0	3	20		
51	P164686	1	3	5	1	0	0	0	0	0	2	0	0	0	0	0	0	1	13		
52	P163896	0	0	1	0	0	0	0	0	2	0	3	0	0	0	1	0	3	10		
53	P170940	1	1	0	0	0	0	0	1	2	0	0	0	1	0	1	0	2	9		
54	P165543	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	3	5		
55	P168076	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	1	4		
56	P173994	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4		
57	P162454	0	0	0	0	0	0	3	0	3	0	0	0	0	0	0	0	3	9		
58	P166279	0	4	1	0	0	0	0	0	4	0	0	0	1	0	1	0	5	16		
59	P170223	2	1	0	0	0	0	0	1	0	1	1	0	1	0	1	0	2	10		
60	P173972	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	2	8		
61	P170669	1	0	0	1	0	0	0	1	1	1	0	0	0	0	0	0	2	7		
62	P172321	0	6	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	8		
63	P160224	0	0	1	0	0	0	0	0	1	0	4	0	0	0	0	0	1	7		
64	P172863	0	0	1	0	0	0	0	1	2	0	0	0	0	0	0	0	1	5		
65	P173799	0	2	6	0	0	0	0	0	0	0	0	0	0	0	0	0	2	10		
66	P167619	1	0	0	4	1	0	0	0	0	1	0	0	0	0	0	1	2	10		
67	P168147	1	0	0	0	1	1	0	0	0	1	1	0	0	0	0	1	3	9		
68	P168580	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	2	4		
69	P170267	1	1	2	0	1	0	0	0	2	0	1	0	0	0	0	0	1	9		
70	P170343	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	1	4		
71	P164704	0	0	0	0	1	7	1	1	2	0	1	1	0	0	0	0	2	16		
72	P161402	0	1	0	0	0	1	0	0	2	0	4	1	0	0	0	0	4	13		
73	P166732	0	0	1	0	0	0	0	0	1	0	4	0	0	0	0	0	2	8		
74	P157575	1	0	4	3	0	6	0	0	1	1	2	1	0	0	1	0	3	23		
75	P167996	0	1	7	0	0	0	0	0	0	1	0	0	0	0	0	1	0	10		
76	P173805	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	2	7		
77	P169505	1	1	0	0	0	0	0	1	0	0	0	0	0	0	2	0	0	5		
78	P157043	1	0	0	0	1	7	0	0	0	0	1	1	0	0	0	0	0	11		
79	P163255	0	0	7	0	0	0	0	0	1	0	0	0	0	0	0	0	0	8		
80	P170052	1	5	0	0	0	0	0	0	0	0	1	0	5	0	0	0	1	13		
81	P174120	3	0	2	0	0	0	0	0	1	1	0	0	0	0	0	0	1	8		
82	P171440	1	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	1	5		
83	P166302	1	1	1	0	0	0	0	0	2	0	1	0	5	0	0	0	2	13		
84	P166303	0	1	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	6		
85	P167634	0	4	0	0	0	0	0	1	2	0	0	0	0	0	1	0	3	11		
86	P170185	0	2	0	0	0	0	0	2	2	1	0	0	0	0	0	0	2	9		
87	P163673	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	2		
88	P173867	2	0	7	0	0	1	0	0	2	0	1	0	0	0	0	1	2	16		
89	P165973	0	2	1	0	0	4	0	0	1	0	1	0	2	0	2	0	0	13		
90	P166187	1	4	0	0	1	0	0	0	1	1	0	0	0	0	0	1	2	11		
91	P168425	1	0	3	4	0	0	0	0	2	2	0	0	0	0	0	1	2	15		
92	P173945	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	2	6		
93	P168273	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	4		
94	P158418	0	2	0	0	0	3	2	0	4	0	0	0	0	0	1	0	1	13		
95	P166941	0	7	0	0	0	0	0	0	1	0	0	0	0	0	2	0	2	12		
96	P162929	0	0	0	0	0	2	0	0	1	0	6	0	0	0	0	0	1	10		
97	P166305	0	4	0	0	0	0	0	2	2	0	0	0	0	0	0	0	2	10		
98	P168280	1	2	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	5		
99	P157715	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	3		
100	P162043	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3		
101	P164486	0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	3		

Appendix 3.1. Validation Exercise

The following table shows the results from the 20 fiscal 2020 projects chosen for the validation exercise. Each row corresponds to the World Bank project number in the far-left column, while the SDG columns display the correspondence between the automatic and manual mappings according to these values: true positives (TP) (the SDG targets that were mapped in both the manual and automatic mappings); true negatives (TN) (when no mapping was obtained both in the automatic and manual exercises); false positives (FP) (the mapping results that were obtained in the automatic mapping but not in the manual mapping); and false negatives (FN) (the mapping results that were obtained in the manual mapping but not in the automatic mapping). The colors help differentiate these values: TP are shown in light blue; TN, in dark blue; FP, in light red; FN, in dark red.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1		1.1	1.2	1.3	1.4	1.5	1.b	2.1	2.2	2.3	2.4	2.5	2.a	2.b	2.c	3.1	3.2	3.3	3.4	3.5
2	P167455	TN	TN	TP	TN	FP	TN	TN	TN	TP	TN	TN	TP	TP	TP	TN	TN	TN	TN	TN
3	P164588	TN	TN	TN	TN	TN	TN	TN	TN	TP	TP	TN	FP	TP	TP	TN	TN	TN	TN	TN
4	P165055	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN
5	P160628	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
6	P156880	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP
7	P162594	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
8	P171190	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
9	P162835	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP	FP	TN	TN	TN	TN	TN
10	P166697	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
11	P166170	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
12	P173927	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	FP
13	P169117	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN
14	P158502	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
15	P157929	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
16	P164686	TN	TN	TP	TN	TN	TN	TP	TP	TN	TN	TN	TN	TN	FP	TP	TN	TN	TN	FP
17	P170940	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN
18	P168076	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
19	P169505	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN
20	P163673	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
21	P168425	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN

	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO
1	3.6	3.7	3.8	3.9	3.b	3.c	3.d	4.1	4.2	4.3	4.4	4.5	4.7	4.a	4.b	4.c	5.a	5.2	5.4	5.6	5.c
2	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FN	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN
3	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
4	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
5	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
6	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TP
7	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
8	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
9	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
10	TN	TN	TP	TN	FP	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
11	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
12	FP	TN	TP	TN	TP	TP	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
13	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
14	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP
15	TP	FP	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TP
16	TN	TN	TP	TN	FP	TP	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
17	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
18	TN	TN	TN	TN	TN	TN	TN	TP	TP	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN
19	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
20	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
21	TN	TN	TP	TN	FP	FP	TN	TP	TP	TP	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN

	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH	BI	BJ
1	6.1	6.2	6.3	6.4	6.5	6.6	6.a	6.b	7.1	7.2	7.3	7.a	8.2	8.3	8.4	8.6	8.10	8.b	9.1	9.2	9.3
2	TP	TP	TP	FN	TP	TN	TP	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP
3	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN
4	TP	TP	TP	TP	TP	TN	TP	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
5	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
6	TP	TP	TP	TN	TP	TN	TP	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
7	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN
8	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN
9	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	TP
10	TP	TP	TP	TN	TP	TN	TP	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
11	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
12	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
13	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TP	TN	TP	TN	TP
14	TP	TP	TP	TP	FP	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN
15	FP	FP	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN
16	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
17	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	FN
18	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
19	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN
20	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
21	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN

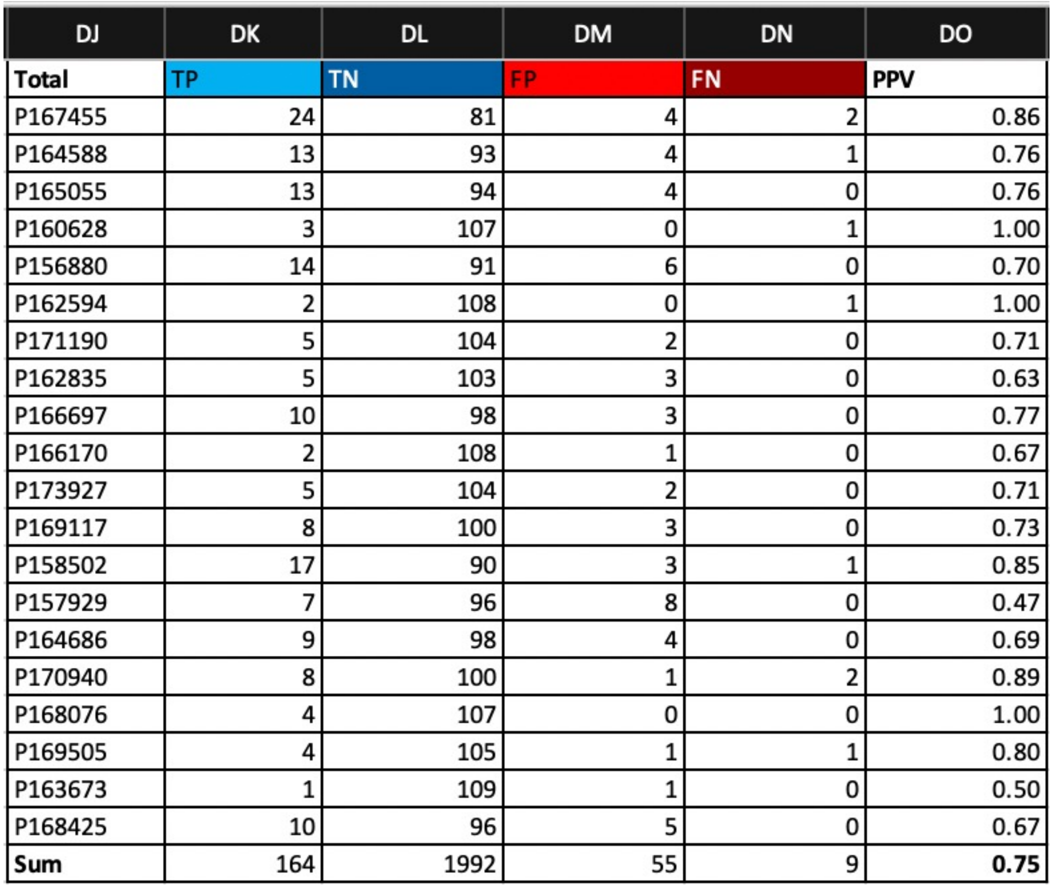
	BK	BL	BM	BN	BO	BP	BQ	BR	BS	BT	BU	BV	BW	BX	BY	BZ	CA	CB	CC	CD	CE
1	9.4	9.5	9.a	9.b	10.2	10.3	10.4	10.5	10.6	10.7	11.1	11.2	11.3	11.4	11.5	11.6	11.7	11.a	11.b	11.c	12.2
2	TP	FP	TN	TP	TP	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN
3	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
4	TP	TP	FP	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN
5	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FN	TN	TN	TN
6	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TP	TN	TN	TP	TN	FP	TP	TN	TN
7	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
8	TP	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
9	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
10	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN
11	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
12	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
13	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
14	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FN	TP	TN	TN	TP	TN	TP	TP	TN	TN
15	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN
16	TN	TN	TN	TN	TP	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
17	TP	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
18	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
19	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
20	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
21	TN	FP	TN	TP	TN	TN	TP	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN

	CF	CG	CH	CI	CJ	CK	CL	CM	CN	CO	CP	CQ	CR	CS	CT	CU	CV	CW	CX	CY	CZ
1	12.4	12.5	13.1	13.2	13.3	13.a	13.b	14.2	14.5	15.1	15.2	15.3	15.a	15.b	16.6	16.10	16.a	17.1	17.2	17.3	17.4
2	TP	TP	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN
3	FP	TN	TP	TP	FP	TP	TP	TN	TN	TP	TP	TN	TP	TP	TN	TN	TN	TN	TN	TN	FN
4	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN
5	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN
6	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	FP	TN	TN
7	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN
8	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TP	TN
9	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TP
10	TP	TN	TN	TN	TP	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TN	TN
11	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
12	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
13	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TP	TN
14	TP	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TP	TN	TN	TN
15	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	TN	TN
16	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	FP	TN	TN
17	TN	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN	FN	TP	TN	TN	TN	TN	TN	TN	TP	TN
18	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN
19	TN	TN	FN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TP	TN	TN	TN	TN	TN	TN	TN	TN
20	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN
21	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TN	TP	TN	TN	TN	FP	TN	TN

	DA	DB	DC	DD	DE	DF	DG	DH
1	17.5	17.7	17.8	17.9	17.12	17.13	17.17	17.19
2	TP	TN	TN	TP	FP	TN	TP	TN
3	TN	TN	TN	TP	TN	TN	TN	TN
4	FP	TN	TN	TN	TN	TN	TN	TN
5	TN	TN	TN	TP	TN	TN	TN	TN
6	TN	TN	TN	TP	TN	TN	TN	TN
7	TN	TN	FN	TN	TN	TN	TN	TN
8	TN	TN	TN	TN	TN	TN	TN	FP
9	TN	TN	TN	TN	TN	TN	TN	TN
10	TN	TN	TN	TN	TN	TN	TN	TN
11	TN	TN	TN	TP	TN	TN	TN	TN
12	TN	TN	TN	TN	TN	TN	TN	TN
13	TP	TN	TN	TN	FP	TN	TN	TN
14	TP	TN	TN	TP	TN	TN	TN	TN
15	TN	TN	TN	TP	TN	TN	TN	TN
16	TN	TN	TN	TN	TN	TN	TN	TN
17	TN	TN	TN	TP	TN	TN	TN	TN
18	TN	TN	TN	TP	TN	TN	TN	TN
19	TN	TN	TN	TN	TN	TN	TN	TN
20	TN	TN	TN	TN	TN	TN	TN	TN
21	TN	TN	TN	TN	TN	TN	TP	TN

Appendix 3.2. Validation Exercise

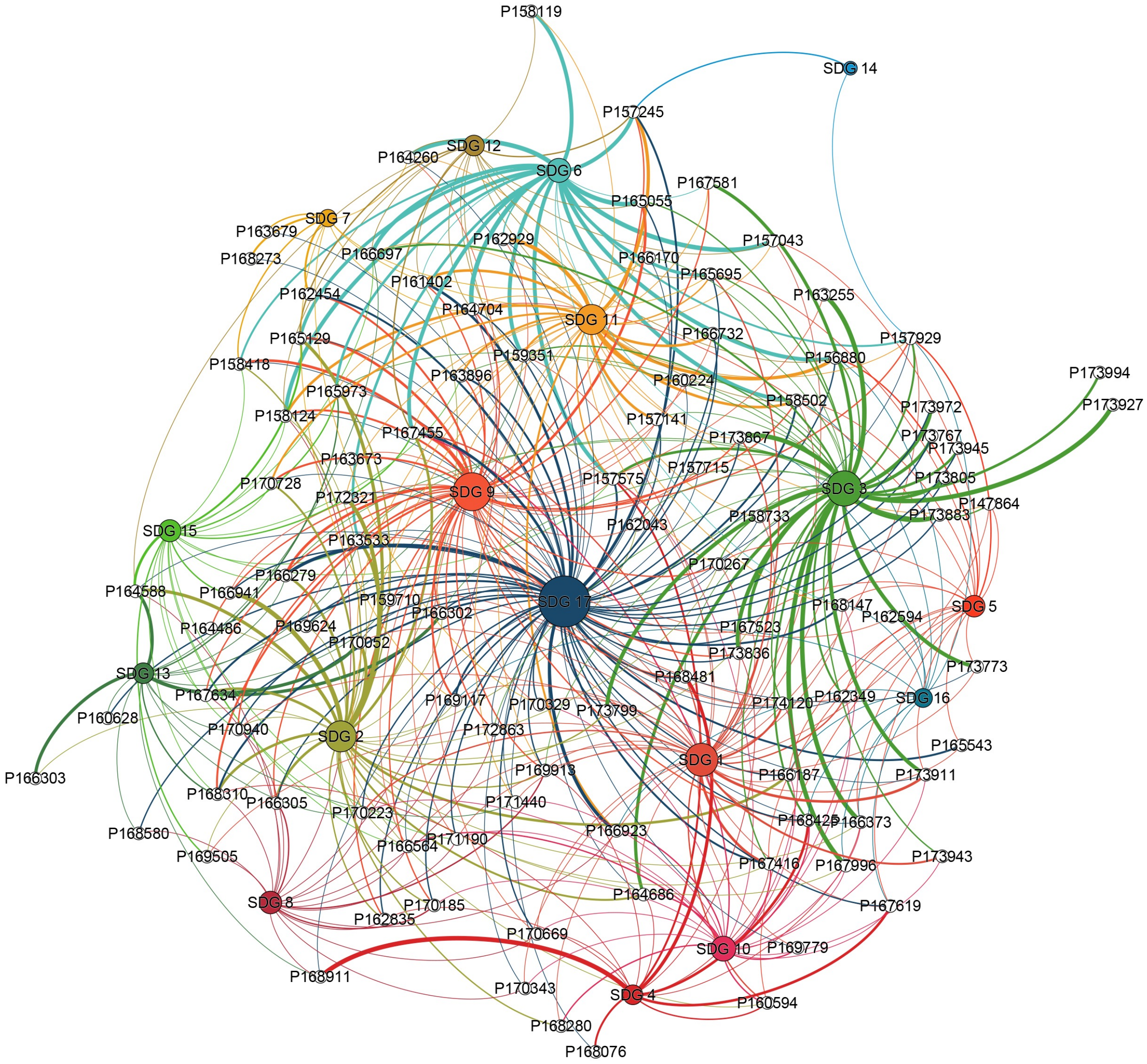
The following table shows the World Bank project number in the far-left column, and the columns show the total sum of true positives (TP) (the SDG targets that were mapped in both the manual and automatic mappings); true negatives (TN) (when no mapping was obtained both in the automatic and manual exercises); false positives (FP) (the mapping results that were obtained in the automatic mapping but not in the manual mapping); and false negatives (FN) (the mapping results that were obtained in the manual mapping but not in the automatic mapping), and PPV values for each project. The PPV is obtained by dividing the number of TP by the TP plus FP (see Chu 1999; Saito and Rehmsmeier 2015). The bottom row depicts the corresponding values for the 20 projects used in the validation exercise.



DJ	DK	DL	DM	DN	DO
Total	TP	TN	FP	FN	**PPV**
P167455	24	81	4	2	0.86
P164588	13	93	4	1	0.76
P165055	13	94	4	0	0.76
P160628	3	107	0	1	1.00
P156880	14	91	6	0	0.70
P162594	2	108	0	1	1.00
P171190	5	104	2	0	0.71
P162835	5	103	3	0	0.63
P166697	10	98	3	0	0.77
P166170	2	108	1	0	0.67
P173927	5	104	2	0	0.71
P169117	8	100	3	0	0.73
P158502	17	90	3	1	0.85
P157929	7	96	8	0	0.47
P164686	9	98	4	0	0.69
P170940	8	100	1	2	0.89
P168076	4	107	0	0	1.00
P169505	4	105	1	1	0.80
P163673	1	109	1	0	0.50
P168425	10	96	5	0	0.67
Sum	164	1992	55	9	**0.75**

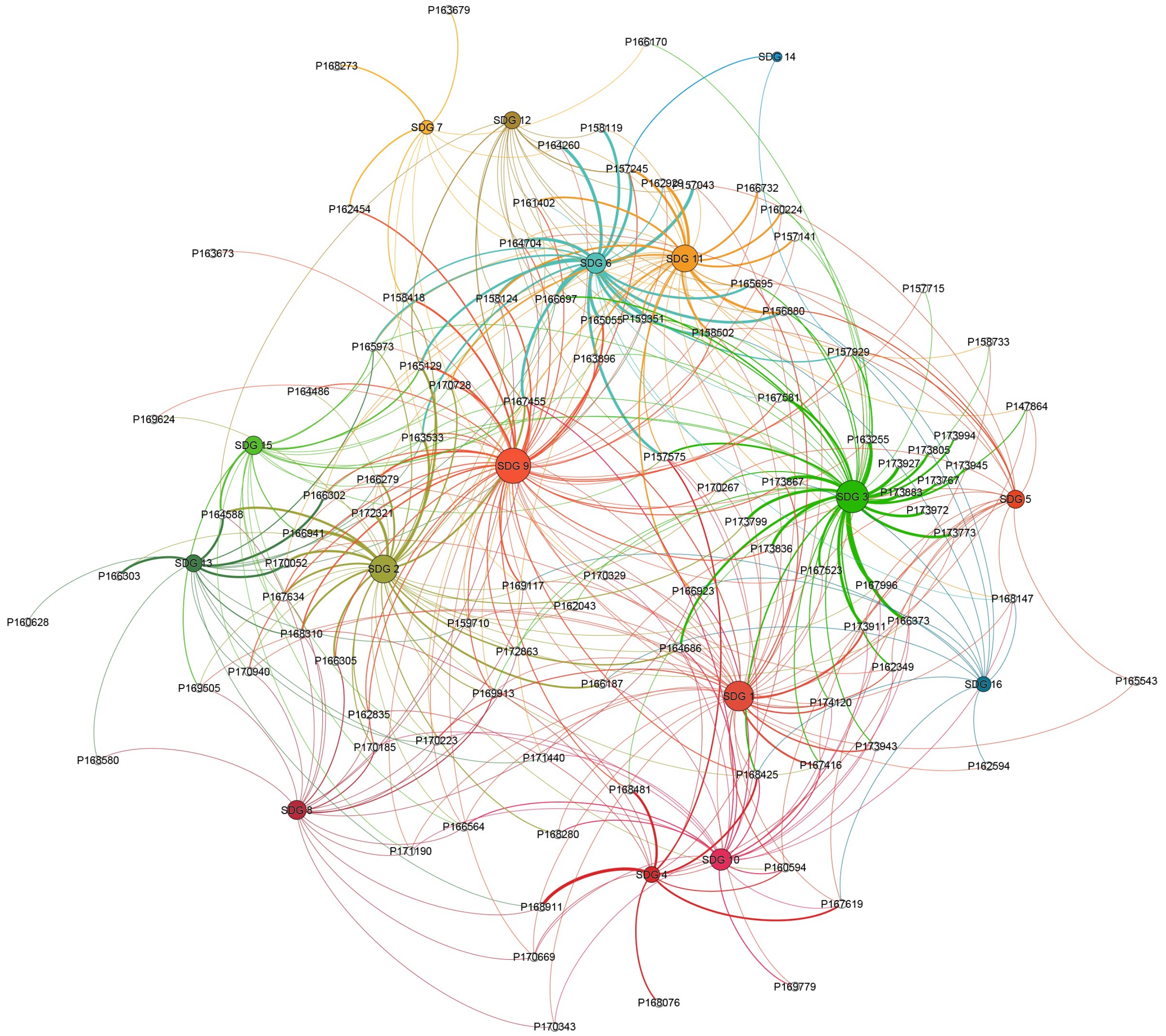
Appendix 4. Results Graph Network: Project to Goal, Full Results

This network graph shows the connection between the projects and the SDGs visually. The thickness of the connections shows how many targets from a specific SDG are connected to one project. The thicker the line, the more targets connected to a project. The size of the circle represents how many projects are mapped to each SDG: the larger the circle, the greater the number of projects mapped to the SDG. The graph shows the full results, including SDG 17.



Appendix 5. Results Graph Network: Project to Goal, Results without SDG 17

This network graph shows the connection between the projects and the SDGs visually. The thickness of the connections shows how many targets from a specific SDG are connected to one project. The thicker the line, the more targets connected to a project. The size of the circle represents how many projects are mapped to each SDG: the larger the circle, the greater the number of projects mapped to the SDG. The graph shows the results excluding SDG 17.



Endnotes

[1] This means asking only whether a keyword occurs or does not occur in the document. The number of times a word occurs is not part of the model, and neither is any measure of what words are adjacent to the keyword.

[2] To be added to the project portfolio, a project must have begun disbursement of funds in fiscal 2020.

[3] Henceforth, “themes” refers exclusively to the World Bank taxonomy.

[4] This paper uses “categories” to refer to the many taxonomies and subdivisions employed by institutions to classify their activities by sustainable development sectors and topics, such as the World Bank taxonomy (World Bank 2016).

References

Armitage, C. S., M. Lorenz, and S. Mikki. 2020. "Mapping Scholarly Publications Related to the Sustainable Development Goals: Do Independent Bibliometric Approaches Get the Same Results?" Quantitative Science Studies 1 (3): 1092–1108. https://doi.org/10.1162/qss_a_00071.

Asatani, K., H. Takeda, H. Yamano, and I. Sakata. 2020. "Scientific Attention to Sustainability and SDGs: Meta-analysis of Academic Papers." Energies 13 (4): 1–21. https://doi.org/10.3390/en13040975.

AUN (Aurora Universities Network). 2021. Search Queries for “Mapping Research Output to the Sustainable Development Goals (SDGs)” sdg-queries. Aurora Universities Network. https://doi.org/doi:10.5281/zenodo.3817445.

Bennich, T., N. Weitz, and H. Carlsen. 2020. "Deciphering the Scientific Literature on SDG Interactions: A Review and Reading Guide." Science of the Total Environment 728:1–13. https://doi.org/10.1016/j.scitotenv.2020.138405.

Bordignon, F. 2021. "Dataset of Search Queries to Map Scientific Publications to the UN Sustainable Development Goals." Data in Brief 34:106731. https://doi.org/10.1016/j.dib.2021.106731.

Chen, X., J. Chen, D. Wu, Y. Xie, and J. Li. 2016. "Mapping the Research Trends by Co-word Analysis Based on Keywords from Funded Project." Procedia - Procedia Computer Science 91:547–55. https://doi.org/10.1016/j.procs.2016.07.140.

Chu, K. 1999. "An Introduction to Sensitivity, Specificity, Predictive Values and Likelihood Ratios." Emergency Medicine Australasia 11 (3): 175–81. https://doi.org/10.1046/j.1442-2026.1999.00041.x.

Dangelmaier, U. 2019. The SDG Mapping of KfW Group. September. https://www.kfw.de/nachhaltigkeit/Dokumente/Sonstiges/SDG-Methodenpapier- DE-EN-2.pdf.

DIE (Deutsches Institut für Entwicklungspolitik) and SEI (Stockholm Environment Institute). 2017. NDC-SDG Connections: Bridging Climate and the 2030 Agenda. SEI Tool. https://klimalog.die-gdi.de/ndc-sdg/.

Elsevier. 2020. The Power of Data to Advance the SDGs Mapping Research for the Sustainable Development Goals. Elsevier. https://www.elsevier.com/__data/assets/pdf_file/0004/1058179/Elsevier-SDG-Report-2020.pdf.

Forouli, A., A. Nikas, D.-J. Van de Ven, J. Sampedro, and H. Doukas. 2020. "A Multiple- Uncertainty Analysis Framework for Integrated Assessment Modelling of Several Sustainable Development Goals." Environmental Modelling & Software 131:104795. https://doi.org/10.1016/j.envsoft.2020.104795.

Fruchterman, T. M. J., and E. M. Reingold. (1991). "Graph Drawing by Force-Directed Placement." Software: Practice and Experience 21 (11): 1129–64. https://doi.org/10.1002/spe.4380211102.

Government of Mexico City. 2019. Innovation and Rights; A Program to Advance Sustainable Development in Mexico City. Voluntary local review. https://sdgs.un.org/sites/default/files/2021-04/Mexico City VLR.pdf

Guerrero, O. A., and G. Castañeda. 2020. "Quantifying the Coherence of Development Policy Priorities." Development Policy Review 1–40. https://doi.org/10.1111/dpr.12498.

Guisiano, J., and R. Chiky. 2021. "Automatic Classification of Multilabel Texts Related to Sustainable Development Goals (SDGs)." HAL 8. https://hal.archives-ouvertes.fr/hal-03154261

Hassan, S.-U., P. Haddawy, and J. Zhu. 2014. "A Bibliometric Study of the World's Research Activity in Sustainable Development and Its Sub-Areas Using Scientific Literature." Scientometrics 99 (2): 549–79. https://doi.org/10.1007/s11192-013-1193-3.

Hawai'i Green Growth Local2030 HUB. 2020. Aloha+ Challenge 2020 Benchmark Report: Hawai'i's Voluntary Local Review of Progress on the Sustainable Development Goals. https://www.local2030.org/pdf/vlr/aloha2020.pdf.

Hernández-Orozco, E., I. Lobos-Alva, M. Cardenas-Vélez, D. Purkey, M. Nilsson, and P. Martin. 2021. "The Application of Soft Systems Thinking in SDG Interaction Studies: A Comparison between SDG Interactions at National and Subnational Levels in Colombia." Environment, Development and Sustainability. https://doi.org/10.1007/s10668-021-01808-z.

Higuchi, K. 2016a. "A Two-Step Approach to Quantitative Content Analysis : KH Coder Tutorial Using Anne of Green Gables (Part I)." Ritsumeikan Social Sciences Review 52 (3): 77–91.

Higuchi, K. 2016b. KH Coder 3 Reference Manual. Ritsumeikan University. http://khcoder.net/en/manual_en_v3.pdf.

Higuchi, K. 2017. "A Two-Step Approach to Quantitative Content Analysis : KH Coder Tutorial Using Anne of Green Gables (Part II)." Ritsumeikan Social Sciences Review 53 (1): 137147.

Higuchi, K. 2021. KH Coder. KH Coder Software. http://khcoder.net/en/.

Higuchi, K. 2022. KH Coder Releases. GitHub. https://github.com/ko-ichi-h/khcoder/releases.

Horne, J., M. Recker, I. Michelfelder, J. Jay, and J. Kratzer. 2020. "Exploring Entrepreneurship Related to the Sustainable Development Goals - Mapping New Venture Activities with Semi-automated Content Analysis." Journal of Cleaner Production 242:118052. https://doi.org/10.1016/j.jclepro.2019.118052.

Hwang, H., S. An, E. Lee, S. Han, and C. H. Lee. 2021. "Cross-societal Analysis of Climate Change Awareness and Its Relation to SDG 13: A Knowledge Synthesis from Text Mining." Sustainability (Switzerland) 13 (10): 1–21. https://doi.org/10.3390/su13105596.

IAEG-SDGs (Inter-agency and Expert Group on Sustainable Development Goal Indicators). 2016. Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (E/CN.3/2016/2/Rev.1), Annex IV. In Report of the Inter-Agency and Expert Group on Sustainable Development Goal Indicators. United Nations. https://sustainabledevelopment.un.org/content/documents/11803Official-List-of-Proposed-SDG-Indicators.pdf.

Instituto Cidades Sustentáveis. 2021. Índice de Desenvolvimento Sustentável das Cidades – Brasil. Programa Ciudades Sustentáveis. https://idsc-br.sdgindex.org/introduction.

Jayabalasingham, B., R. Boverhof, K. Agnew, and L. Klein. 2019. Identifying Research Supporting the United Nations Sustainable Development Goals. Elsevier BV; Elsevier BV. https://doi.org/10.17632/87TXKW7KHS.1.

JRC (Joint Research Centre) and DG INTPA (Department for International Partnerships). n.d. SDG Mapper. Retrieved October 15, 2021, from https://knowsdgs.jrc.ec.europa.eu/sdgmapper.

Kamada, T., and S. Kawai. 1989. "An Algorithm for Drawing General Undirected Graphs." Information Processing Letters 31 (1): 7–15. https://doi.org/10.1016/0020-0190(89)90102-6.

Krallinger, M., F. Leitner, M. Vazquez, D. Salgado, C. Marcelle, M. Tyers, A. Valencia, and A. Chatraryamontri. 2012. "How to Link Ontologies and Protein-Protein Interactions to Literature: Text-Mining Approaches and the BioCreative Experience." Database 2012, bas017–bas017. https://doi.org/10.1093/database/bas017.

Luhn, H. P. 1960. "Key Word-in-Context Index for Technical Literature (KWIC Index)." American Documentation 11 (4): 288–95. https://doi.org/10.1002/asi.5090110403.

Mair, S., A. Jones, J. Ward, I. Christie, A. Druckman, and F. Lyon. 2018. A Critical Review of the Role of Indicators in Implementing the Sustainable Development Goals, 41–56. https://doi.org/10.1007/978-3-319-63007-6_3.

Mistry, A., H. Sellers, J. Levesley, and S. Lee. 2020. "Mapping a University's Research Outputs to the UN Sustainable Development Goals. Emerald Open Research 2 (May): 61. https://doi.org/10.35241/emeraldopenres.13881.1.

Natixis. 2018. Solving Sustainable Development Goals Rubik's Cube an Impact-Based Toolkit for Issuers and Investors Executive Summary.

NYC Government (New York City Government). 2019. Voluntary Local Review: New York City's Implementation of the 2030 Agenda for Sustainable Development. Voluntary Local Review. https://www1.nyc.gov/assets/international/downloads/pdf/International-Affairs-VLR-2019.pdf

Olawumi, T. O., and D. W. M. Chan. 2018. "A Scientometric Review of Global Research on Sustainability and Sustainable Development." Journal of Cleaner Production 183: 231–50. https://doi.org/10.1016/j.jclepro.2018.02.162.

Ozgur, A., B. Cetin, and H. O. Bingol. 2008. "Co-occurrence Network of Reuters News." International Journal of Modern Physics 1–10 (February 2013). https://doi.org/10.1142/S0129183108012431.

Pincet, A., S. Okabe, and M. Pawelczyk. 2019. "Linking Aid to the Sustainable Development Goals – A Machine Learning Approach," OECD Development Co-Operation Wor. https://doi.org/10.1787/4bdaeb8c-en

Rani, R., and D. K. Lobiyal. 2020. "Performance Evaluation of Text-Mining Models with Hindi Stopwords Lists." Journal of King Saud University - Computer and Information Sciences. https://doi.org/10.1016/j.jksuci.2020.03.003

Rodríguez del Águila, M., and A. González-Ramírez. 2014. "Sample Size Calculation." Allergologia et Immunopathologia 42 (5): 485–92. https://doi.org/10.1016/j.aller.2013.03.008.

Sachs, J., C. Kroll, G. Lafortune, G. Fuller, and F. Woelm. 2021. "Sustainable Development Report 2021." In Sustainable Development Report 2021. https://doi.org/10.1017/9781009106559.

Saito, T., and M. Rehmsmeier. 2015. "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets." PLOS ONE 10 (3): e0118432. https://doi.org/10.1371/journal.pone.0118432.

Sánchez de Madariaga, I., J. Benayas del Álamo, J. García López, R. Sisto, and J. Urquijo Reguera. 2020. In Los Objetivos de Desarrollo Sostenible en 100 Ciudades Españolas, 2a edición, edited by Red Española Para El Desarrollo Sostenible. REDS. https://reds-sdsn.es/informe-ods-ciudades-2020.

Scott, M. 2001. "Comparing Corpora and Identifying Key Words, Collocations, and Frequency Distributions through the WordSmith Tools Suite of Computer Programs." Small Corpus Studies and ELT: Theory and Practice 47 (67).

Sebestyén, V., E. Domokos, and J. Abonyi. 2020. "Focal Points for Sustainable Development Strategies—Text Mining-Based Comparative Analysis of Voluntary National Reviews." Journal of Environmental Management 263: 110414. https://doi.org/10.1016/j.jenvman.2020.110414.

Spasić, I., D. Schober, S.-A. Sansone, D. Rebholz-Schuhmann, D. B. Kell, and N. W. Paton. 2008. "Facilitating the Development of Controlled Vocabularies for Metabolomics Technologies with Text Mining." BMC Bioinformatics 9 (S5): S5. https://doi.org/10.1186/1471-2105-9-S5-S5.

Sullivan, K., S. Thomas, and M. Rosano. 2018. "Using Industrial Ecology and Strategic Management Concepts to Pursue the Sustainable Development Goals." Journal of Cleaner Production 174: 237–46. https://doi.org/10.1016/j.jclepro.2017.10.201.

Uehara, T., and R. Sakurai. 2021. "Have Sustainable Development Goal Depictions Functioned As a Nudge for the Younger Generation before and during the Covid-19 Outbreak?" Sustainability (Switzerland) 13 (4): 1–18. https://doi.org/10.3390/su13041672.

UN (United Nations). 2016. Transforming Our World: The 2030 Agenda for Sustainable Development. In The 2030 Agenda for Sustainability, vol. 70, issue 1. https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_RES_70_1_E.pdf.

UNCTAD (United Nations Conference on Trade and Development). 2014. World Investment Report 2014: Investing in the SDGs: An Action Plan. https://unctad.org/system/files/official-document/wir2014_en.pdf.

UN DESA (United Nations Department of Economic and Social Affairs). n.d. 17: Strengthen the Means of Implementation and Revitalize the Global Partnership for Sustainable Development. Sustainable Development Goals. https://sdgs.un.org/goals/goal17.

UN Women (United Nations Entity for Gender Equality and the Empowerment of Women). 2018. Turning Promises into Action: Gender Equality in the 2030 Agenda for Sustainable Development. https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/Library/Publications/2018/SDG-report-Gender-equality-in-the-2030-Agenda-for-Sustainable-Development-2018-en.pdf

Vanderfeesten, M., R. Otten, and E. Spielberg. 2020. Search Queries for “Mapping Research Output to the Sustainable Development Goals (SDGs). Zenodo. https://doi.org/10.5281/ZENODO.3817445.

World Bank. 2016. Theme Taxonomy and Definitions. https://thedocs.worldbank.org/en/doc/275841490966525495-0290022017/original/NewThemeTaxonomyanddefinitionsrevisedJuly012016.pdf.

World Bank. 2021. The World Bank Impact Report. Sustainable Development Bonds & Green Bonds. https://issuu.com/jlim5/docs/world-bank-ibrd-impact-report-2020?mode=window.

World Bank. 2022. The World Bank Impact Report. Sustainable Development Bonds & Green Bonds. https://issuu.com/jlim5/docs/world_bank_ibrd_impact_report_2021_web_ready_r01?fr=sYTBhOTM4NTM3MTk

World Bank. n.d. Projects & Operations. https://projects.worldbank.org/en/projects-operations/projects-home.

Using Automated Text Mining to Align Investments with the Sustainable Development Goals:

A Case Study Analyzing World Bank Projects

Authors: Efraim Hernández-Orozco (SEI),
Mario Cárdenas-Vélez (SEI), Colleen Keenan
(World Bank), Zoe Russo (World Bank)