DEPARTMENT OF INTERNATIONAL AND EUROPEAN ECONOMIC STUDIES

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Uncovering the SDG content of Human Security Policies through a Machine Learning web application

Phoebe Koundouri

Panagiotis-Stavros Aslanidis

Konstantinos Dellis

Georgios Feretzakis

Angelos Plataniotis

Working Paper Series

24-06

February 2024

Uncovering the SDG content of Human Security Policies through a Machine Learning web application

Phoebe Koundouri[1], *Panagiotis-Stavros Aslanidis*[2], *Konstantinos Dellis*[3], *Georgios Feretzakis*[4], *Angelos Plataniotis*[5]

[1] School of Economics and ReSEES Research Laboratory, Athens University of Economics and Business; Department of Technology, Management and Economics, Denmark Technical University (DTU); Sustainable Development Unit, Athena Research Centre; UN SDSN (Global Climate Hub, European Hub, Greek Hub); phoundouri@aueb.gr

[2] Laboratory of Operations Research, Department of Economics, University of Thessaly; ReSEES Research Laboratory, Athens University of Economics and Business; Sustainable Development Unit, ATHENA Research Centre; paslanidis@uth.gr

[3] ReSEES Research Laboratory, Athens University of Economics and Business; Sustainable Development Unit, ATHENA Research Centre; kdellis@aueb.gr

[4] School of Science and Technology, Hellenic Open University, 26335 Patras, Greece. georgios.feretzakis@ac.eap.gr

[5] Sustainable Development Unit, ATHENA Research Centre; angplat@athenarc.gr

Abstract

This paper introduces a machine learning (ML) based approach for integrating Human Security (HS) and Sustainable Development Goals (SDGs). Originating in the 1990s, HS focuses on strategic, people-centric interventions for ensuring comprehensive welfare and resilience. It closely aligns with the SDGs, together forming the foundation for global sustainable development initiatives. Our methodology involves mapping 44 reports to the 17 SDGs using expert-annotated keywords and advanced ML techniques, resulting in a web-based SDG mapping tool. This tool is specifically tailored for the HS-SDG nexus, enabling the analysis of 13 new reports and their connections to the SDGs. Through this, we uncover detailed insights and establish strong links between the reports and global objectives, offering a nuanced understanding of the interplay between HS and sustainable development. This research provides a scalable framework to explore the relationship between HS and the Paris Agenda, offering a practical, efficient resource for scholars and policymakers.

Keywords: *Artificial Intelligence in Policy Making, Data Mining, Human-Centric Governance Strategies, Human Security, Machine Learning, Sustainable Development Goals*

INTRODUCTION

Human security (HS), integral to sustainable development (SD), faces significant challenges in today's multi-crisis era. HS has gained renewed attention due to the COVID-19 pandemic, energy and food crises, and geopolitical conflicts, especially in Eastern Europe and the Middle East. Aligning HS with the Sustainable Development Goals (SDGs) is crucial for achieving Agenda 2030's objectives. HS and the SDGs collectively aim to enhance well-being, resilience, and sustainable development globally.

Initially, the United Nations identified two core aspects of HS: freedom from fear and freedom from want, focusing on individual rather than state security (UNDP 1994). Modern HS concerns have evolved echoing globalization, to encompass 'new wars' characterized by changing actors, goals, methods, and financing (Kaldor 2013a, 2013b). Traditional national security measures are insufficient to address the complex human and environmental challenges worldwide. HS now centers on human development and well-being, extending beyond physical violence and destruction (Gasper 2005).

The International Institute for Applied Systems Analysis (IIASA 2018) emphasizes human potential as central to resilience and societal change by 2050. Factors influencing this transition include social organization, technological advancements, and environmental events (UNDP 2022). Social structures must address the economic impacts of violence, technological breakthroughs may pose new HS risks, and climate resilience is crucial for HS in the face of extreme weather events (FFP 2022).

HS adopts both human-centered and state-centered approaches. Human-centered issues, such as poverty and inequality, along with state-centered challenges like political dominance and resource control, significantly impact HS (UNESCO 2008; Human Security Unit United Nations 2009; Jacobs 2016). Effective HS strategies require both protection, a top-down approach involving government and international cooperation, and empowerment, a bottom-up approach driven by citizen engagement.

Differentiating HS from state security is vital for policymakers to develop appropriate agendas and streamline initiatives germane to sustainable development. While HS primarily focuses on citizens, it remains interlinked with state security, encompassing human rights, peace, and sustainable development (UNDP 1994, p. 22-23; CHS 2003; Liotta et al. 2007; Owen 2007; Fukuda-Parr and Messineo 2012; Koundouri and Dellis 2023). Environmental integrity and the dual impact of climate change on HS – deprivation of basic needs and erosion of livelihoods – are also essential considerations (IPCC 2014).

This study introduces a novel, practical tool – a machine learning-based web application – to reveal the connections between HS-related reports and the SDGs. This tool simplifies understanding of the interplay between specific reports and SDGs, filling a research gap in HS analysis. The paper is structured as follows: Section 2 discusses the seven HS pillars; Section 3 explores HS-SDG interlinkages; Section 4 outlines our methodological approach using advanced NLP and machine learning to align data with SDG targets; Section 5 presents our findings; Section 6 details the web application's utility for policymakers and researchers; and Section 7 concludes the research.

The Commission on Human Security emphasizes universal protection for individuals, focusing on achieving dignity and fundamental freedoms without severe threats (CHS 2003). Scholars of human security (HS) highlight the need to eliminate warfare, foster international cooperation, and

promote stability at both individual and community levels, prioritizing human safety (Fuentes and Rojas Aravena 2005; Ogata 2011; JICA 2022). HS, as defined in the Human Development Report of 1994, comprises seven key pillars: economic, food, health, environmental, personal, community, and political security (UNDP 1994).

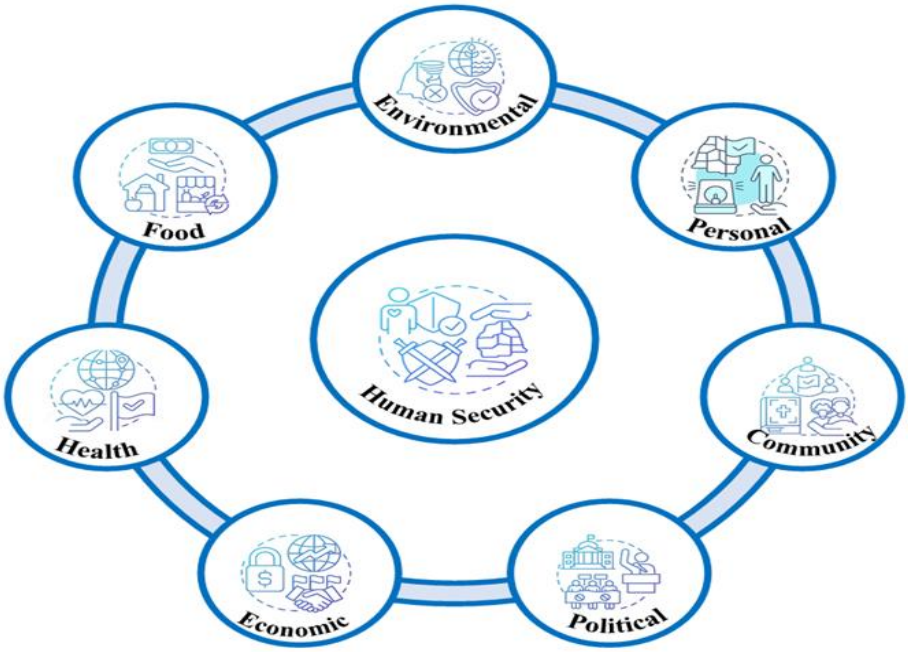

Figure 1 *The seven aspects of Human Security.*

Source: Authors' elaboration based on UNDP (1994)

Economic Security involves ensuring a basic income and having access to robust public and environmental resources, and social safety nets, particularly for vulnerable groups. Economic instability can lead to poverty, job insecurity, and long-term unemployment (OECD 2020b). Inflation, labor market inequality, and recent geopolitical events like the Ukraine war exacerbate increases in costs of living and economic disparities (CHS 2003; ILO 2022; FSIN and GNAFC 2023). Furthermore, addressing economic fragility is essential for mitigating debt crises and fostering sustainable competitiveness (SolABILITY 2022; WEF 2022).

Food Security entails accessible, sufficient, and nutritious food for everyone. Challenges include population growth's impact on food and water resources and climate change exacerbating food insecurity and potential social unrest (IIASA 2018; OECD 2020a; Boehm et al. 2022; FAO 2022; FSIN and GNAFC 2022; IEP 2022a; Wolf et al. 2022; UNDESA 2023).

Health Security encompasses affordable healthcare access and protection from diseases. Urbanization has increased pollution and health risks, particularly for vulnerable groups (WHO 2022). The COVID-19 pandemic, alongside other diseases, has highlighted the intersection of health issues with biodiversity loss and ecosystem degradation (CBD 2020; WEF 2022; George E Halkos and Aslanidis 2023)(CBD 2020; WEF 2022; Halkos and Aslanidis 2023).

Environmental Security focuses on maintaining healthy ecosystems and mitigating natural and climate-induced disasters (IAP). Conflicts over natural resources and the effects of human-induced climate change demonstrate the link between HS and environmental sustainability (CHS 2003; IPCC 2014; Almond et al. 2022; Chancel et al. 2023)

Personal Security involves protecting individuals from violence, human rights abuses, and domestic violence. It encompasses societal resilience to crises, the importance of education, and addressing gender inequality and minority suppression (Gasper and Gómez 2013; IEP 2022b; OECD 2022b; UNDP and OPHI 2022; WHO 2022; Chancel et al. 2023; George E. Halkos and Aslanidis 2023; IAP forthcoming)

Community Security addresses safety from ethnic conflict and preserving cultural identity. Urbanization and climate change-induced agricultural challenges contribute to conflicts, particularly in low-income regions (UNDESA 2018; Chancel et al. 2023).

Political Security emphasizes freedom from state oppression and civil rights protection. Political suppression, accountability issues, polarization, and corruption are also key notions for this aspect (Transparency International 2021; Gwartney et al. 2022; OECD 2022c; UNDP 2022). Political security also relates to sustainable energy and legal systems.

Following recent developments and challenges, HS has been broadened to encompass cybersecurity threats underpinning digital and technological developments. More specifically, the prevalence of fake news in digital platforms following the widespread use of generative artificial intelligence (AI) poses new HS concerns that warrant our attention and analysis. The latter indicates that HS is a living concept and that the work of scholars and policymakers needs to be versatile and adaptive to evolving global challenges.

LITERATURE REVIEW

Human Security and the Sustainable Development Goals

The 17 SDGs are aimed at addressing critical challenges and promoting sustainability and equity by 2030 (UN 2015). They cover a breadth of universal issues including poverty, hunger, health, education, gender equality, clean water, and climate action, offering a framework for collective action. The SDGs are integral to fostering inclusive growth, and environmental protection, and ensuring no one is left behind in the green and digital transformations. Agenda 2030's holistic, interdisciplinary approach marks a shift from the siloed Millennium Development Goals, focusing on international cooperation (Nilson 2018).

Human security (HS) is inherently linked with the SDGs, with Agenda 2030 underpinning policies to safeguard individuals and societies from various threats (Koundouri and Dellis 2023). The UN Trust Fund for Human Security views HS as crucial to achieving the SDGs, placing individual

needs at the forefront (UNTFHS 2017). Crabtree and Gasper (2020) argue that HS and the SDGs are mutually reinforcing, sharing a commitment to global well-being and sustainable development. Each SDG aligns with HS dimensions, with goals like "No Poverty" (SDG1), "Good Health and Well-being" (SDG3), and environmental targets (SDGs 13, 14, 15) reflecting HS aspects. Economic security also links with SDGs focused on work, innovation, and reduced inequality (SDGs 8, 9, 10). Events like the COVID-19 pandemic and energy crises highlight these connections, emphasizing the need for integrated policy approaches.

However, SDGs often center around national and regional governments, sometimes overlooking marginalized groups due to entrenched power structures and inadequate institutional functioning (Gasper and Gómez 2013). An HS perspective can address this by integrating bottom-up approaches with top-down policies, enhancing policy efficiency, and adhering to the "leave no person behind" principle (Gasper et al. 2020). This approach also empowers individuals and communities in sustainable development initiatives (Martin and Owen 2014; Gasper et al. 2020).

In conclusion, the synergy between HS and the SDGs is crucial for building a just, equitable, and sustainable world. Recognizing their interconnectedness enables a comprehensive approach to security and development, prioritizing the well-being and resilience of individuals and communities for a sustainable future.

Tools for Mapping text to SDG

Machine Learning (ML) techniques are increasingly used to link policy documents with Sustainable Development Goals (SDGs). Koundouri et al. (2023) present two ML models for this purpose. The first uses Information Retrieval and Bag-of-Words to correlate EU energy policy documents with SDGs, simplifying document representation for easier identification of

connections. The second model employs BERT (Bidirectional Encoder Representations from Transformers), a Deep Learning technique, for a more intricate understanding of the text and precise mapping to SDGs. This research demonstrates significant progress in using ML and natural language processing for policy document analysis, offering new insights into policy alignment with sustainability goals.

The Joint Research Centre (JRC) of the European Commission developed the "SDG Mapper," an online platform aligning EU policies with SDGs. The methodology, detailed in Borchardt et al. (2022), initially involved manual scanning of policies, which was time-consuming and subjective. To overcome this, an automated text mining method was introduced, employing classic text mining and natural language processing techniques. This included creating a comprehensive set of SDG-related keywords, text preprocessing, and keyword string matching, focusing on practicality and transparency over advanced AI methods. The paper discusses the potential and challenges of automated text mining in SDG policy analysis.

Open Sustainable Development Goals (OSDG) has also developed an open-source tool for SDG text classification, as described by Pukelis et al. (2022). OSDG 2.0, an upgrade from its predecessor, uses ML models and multilingual support for better classification. The methodology involves training ML models with data from the OSDG Community Platform, where volunteers label text snippets for SDG relevance. Each SDG has a dedicated model using a one-vs-rest approach, cross-verified by an ontology-based keyword map. OSDG 2.0's multilingual capabilities enable analysis of a broader range of documents, with improved user interface and functionality.

These tools represent significant advancements in text classification for sustainable development and highlight the importance of collaborative projects in global challenges. Our paper later

compares these tools with our web app, examining their effectiveness in linking documents to SDGs, to provide an overview of ML applications in policy document analysis for SDGs.

METHODOLOGY

Dataset

In this study, we conducted a structured analysis of the relationship between HS reports and the SDGs using advanced text analytics and machine learning methods. Our set of examined documents, which serves also as the training dataset of our machine learning algorithm, consists of 44 reports (Table 1 in the Supplementary Material) covering aspects of HS in varied degrees. Table 3 in the supplementary file contains the specific HS aspects covered by each report, as well as their objectives. All reports are publicly available and were retrieved from various websites.

Methodological Framework and Data Alignment

We started by mapping 44 PDF documents, mostly reports, related to HS aspects, supported by an Excel file containing expert evaluations of each document's relevance to the SDGs. These evaluations used a three-tier categorization system: '0' for no relevance, '1' for partial relevance, and '2' for high relevance. This allowed for a detailed assessment of the reports' alignment with the SDGs.

Our methodology (Figure 2) involved collecting and preparing the dataset, which consisted of HS reports in PDF format and an Excel file with SDG relevance labels. We extracted text from the PDFs using a Python function compliant with data processing standards Bird et al. (2009) and aligned this data with SDG labels using the Pandas library (McKinney 2010; Wickham and

Grolemund 2017). The data was then transformed into a numerical format using the Term Frequency-Inverse Document Frequency (TF-IDF) method, preparing it for machine learning analysis (Salton and McGill 1986; Ramos 2003).

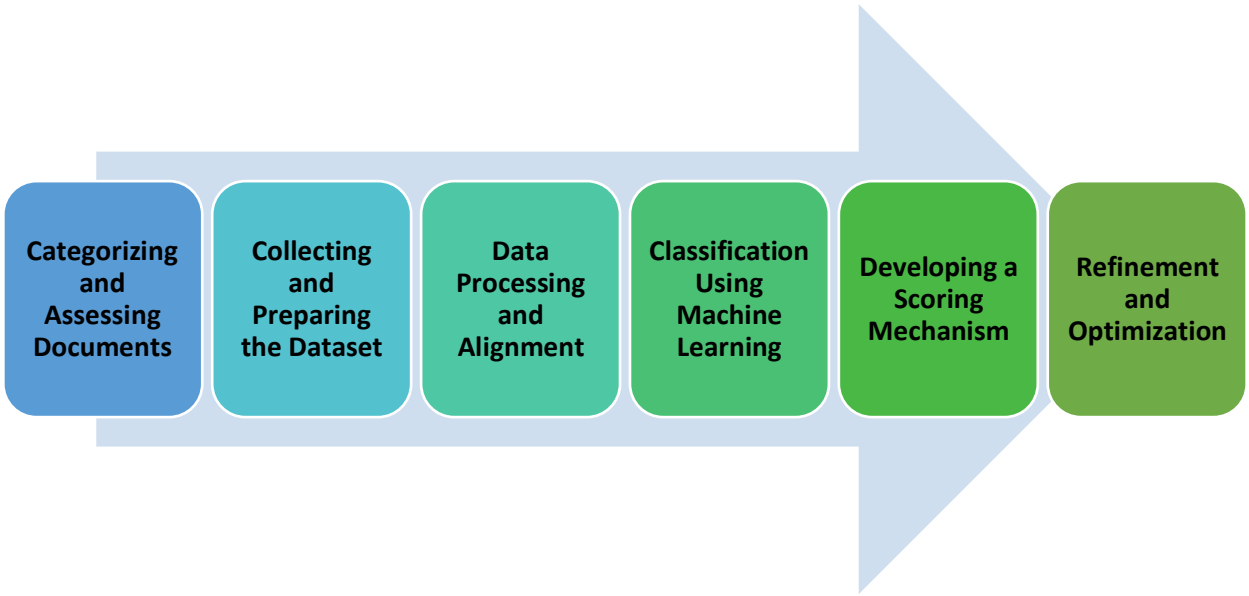

Figure 2 A six-step approach for ML algorithm development. Authors' elaboration.

For classification, we employed the Random Forest Classifier, chosen for its effectiveness with complex data and its resistance to overfitting (Breiman 2001; Liaw and Wiener 2002). We analyzed the top 20 most influential features for each SDG to understand the thematic focuses of the reports and the factors affecting classification results (Guyon and Elisseeff 2003; Strobl et al. 2007). This approach offered a quantitative, data-driven understanding of the documents' alignment with the SDGs, aiding policymakers and stakeholders in sustainable development.

We also developed a scoring mechanism to evaluate each report's relevance to the SDGs based on predefined keywords. This involved calculating keyword frequencies for each SDG, normalizing these against the total word count, and comparing them against serialized thresholds to classify the

documents. This quantitative approach provided a comprehensive view of the reports' SDG alignment, supporting further analysis and statistical examinations.

The results of this scoring process are crucial for policy development, offering a data-focused method for prioritizing initiatives within the SDG framework. The adaptability of our method supports rapid analysis of a growing number of reports, highlighting the evolving nature of policy document analysis.

Additionally, we recognized the importance of accurate threshold settings for the 17 SDGs. We fine-tuned these thresholds using an optimization algorithm, which adjusted them based on the distribution of keyword frequencies in reports with established SDG labels. This refinement ensured that the thresholds accurately reflected the content of the reports, not just the presence of keywords. The optimized thresholds were then serialized and integrated into our computational framework, ensuring a nuanced and precise scoring process.

Exploration with Advanced NLP Models

In the progression of our methodological approach, we explored the potential of state-of-the-art Natural Language Processing (NLP) models such as BERT (Devlin et al. 2018), DistilBERT (Sanh et al. 2019), and ELECTRA (Clark et al. 2020). These advanced models are at the forefront of text analysis and language understanding, offering breakthrough capabilities in processing natural language data.

Our exploration was driven by the goal of augmenting our analytic capabilities and addressing the intricacies inherent in the HS reports related to the Sustainable Development Goals (SDGs). Despite the sophisticated nature of these NLP models, the limited scope of our labeled dataset, which comprised 44 PDF documents, posed a significant challenge. The specialized context of our

research, intersecting policy analysis and sustainable development, necessitated the models to operate with high precision across a compact data sample.

The complexity was further exacerbated by the need to tailor individual models for the 17 SDGs. Each SDG encapsulates specific thematic attributes, necessitating a nuanced learning and predictive capacity from the models across three distinct relevance categories. However, the constrained dataset did not provide the volume or variety of examples needed for the models to learn effectively and discern the subtle distinctions between the categories.

Our foray into the application of these NLP models underscored a pivotal limitation—while these techniques hold immense promise, their performance is contingent on the availability of a rich and diverse training dataset. The depth and breadth of the training data are critical for these models to internalize and reflect the complex patterns and themes associated with the SDGs.

This phase of our research highlighted the constraints of applying advanced NLP models within specialized research domains, particularly when faced with limited data. The insights gained from this exploratory endeavor emphasize the need for expansive and varied datasets to train NLP models to capture the full spectrum of thematic complexities within areas such as the SDGs.

Deployment of Methodology in a Web Application

To further our study's impact and reach, we developed a web application, creating an accessible platform for users to interact with our findings. This web application employs Flask, a Python microframework known for its simplicity and effectiveness in web development (Grinberg 2018). It enables users to upload documents and receive an assessment based on the SDG-related keywords, utilizing the fine-tuned thresholds from our machine learning analysis to provide immediate, practical feedback (Ronacher 2010).

The web app's interface was designed with HTML, CSS, and Bootstrap to ensure it was user-friendly and accessible on various devices. JavaScript was used to enhance the functionality and provide a smooth user experience (Pilgrim 2010). Visualization tools were incorporated into the web app, using libraries like D3.js for dynamic and interactive data visualization. This feature allows users to see graphical representations of the scoring data, offering intuitive insights into the numerical evaluations (Bostock et al. 2011). The Python scripts that form the backbone of the web app's backend ensure consistency with the methodological approach used in our initial research phase. Docker was used to containerize the application, aiding in deployment and ensuring uniformity across different computing environments (Merkel 2014). A requirements.txt file was used to manage the web app's dependencies, ensuring that the necessary Python packages were available and up-to-date for all users (Pip. 2021). Cloud deployment provided the flexibility to scale computational resources to meet user demand, maintaining the web app's responsiveness and availability (Fox 2009).

The web app thus serves as a testament to the potential of combining data science with web technology, making research findings accessible and actionable. It represents a significant step in the digitalization of policy analysis, where insights can be interactively explored and utilized by a broad audience.

In sum, this web application is an extension of our analytical endeavors, forging a link between our study's meticulous research and the practical requirements of decision-makers. It stands as an innovative tool in the advancement of sustainable development, reflecting the commitment to making sophisticated, data-driven insights readily available for the betterment of global initiatives.

RESULTS

Our research employed a two-phase approach to evaluate the relevance of PDF documents to SDGs, providing a comprehensive assessment of our methodology's effectiveness and adaptability.

In the first phase, we tested our machine learning (ML) tool against 13 recent reports related to various HS aspects (which were not used in the training phase of the ML tool), as listed in Table 4 of the supplementary file. Figure 3 demonstrates the tool's proficiency in identifying SDG interlinkages. For instance, the Eurofound report on economic and social inequalities in Europe shows strong connections to SDGs 5 and 10, which address inequality, and to SDG 8, highlighting the report's emphasis on decent work and economic growth. Conversely, the Global Climate Action's COP28 summary, though only 9 pages long, is linked to 12 SDGs due to its comprehensive coverage of multi-disciplinary issues discussed at COP28 in Dubai. The Freedom in the World report (FIW 2022), spanning 37 pages, is primarily associated with SDG 16 for its focus on institutional quality and democratic governance, with a secondary link to SDG 4, underlining the importance of education in bridging institutional gaps.

In the second phase, we compare our model with two other advanced models: the SDG Mapper API by the European Commission (EC 2024) and the OSDG by PPMI, UNDP SDG AI Lab, and collaborators (Pukelis et al. 2022). Each model displays results distinctively: SDG Mapper in percentage terms of relevance to each SDG, OSDG by specifying related SDGs, and our model by mapping report content to the 17 SDGs with a three-tier relevance scale (0, 1, and 2 for no, moderate, and high relevance respectively).

The accuracy assessment of these models encompasses multiple factors such as the number of linkages identified, the authenticity of these connections, and the quality of the source PDFs. Our model strikes a balance, highlighting the most pertinent relationships without overstatement and basing its results on keyword analysis detailed in the supplementary file. Table 2 presents a

comparative analysis of the 13 HS-related reports across the three models, offering insights into each model's capabilities and precision in mapping document content to the SDGs.

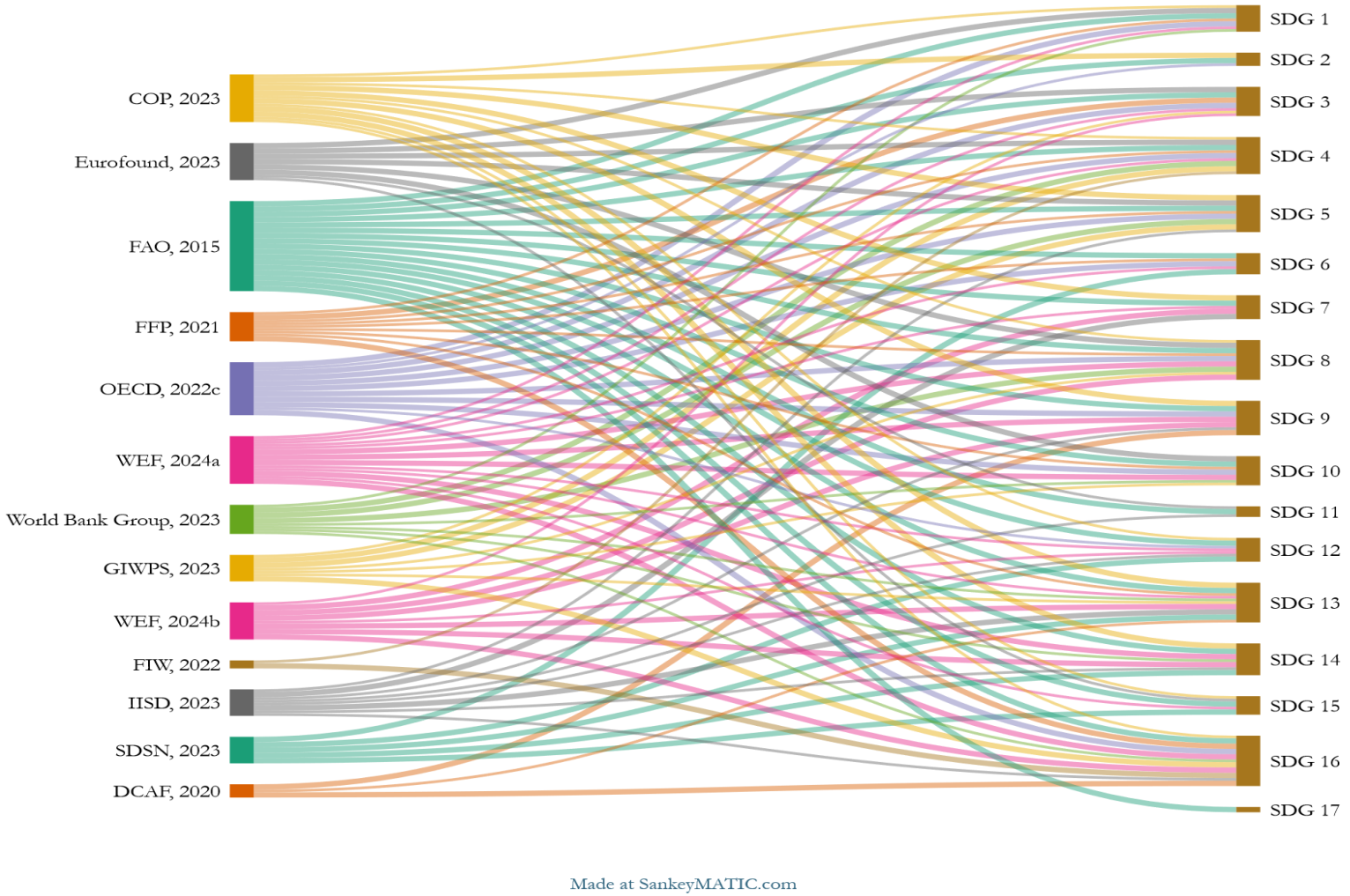

Figure 3 HS Reports and SDG Relevance. Authors' elaboration.

Table 1, illustrating the linkage of PDF files to the SDGs by three different models, reveals commonalities in how these models associate reports with specific SDGs. By analyzing these common parameters, we can discern key areas of agreement among the models.

For the "Fragile States Index Annual Report 2021" (FFP 2021), all three models concur on its association with SDGs 1 (No Poverty), 8 (Decent Work and Economic Growth), and 16 (Peace, Justice, and Strong Institutions). This consensus suggests a clear link between the report's content and these particular goals, highlighting issues of poverty, labor, and institutional integrity. Similarly, the "Women Peace and Security Index 2023/2024" (GIWPS 2023) is unanimously linked by the models to SDGs 5 (Gender Equality) and 16 (Peace, Justice, and Strong Institutions). This uniformity indicates a strong focus in the report on gender equality issues and the reinforcement of legal and institutional frameworks.

These findings demonstrate the effectiveness of the models in identifying key themes within the reports and their relevance to specific SDGs. Such insights are valuable for understanding the focus areas of these reports and how they align with global sustainable development objectives.

Table 1 Comparative analysis of model performance in identifying linkages between HS reports and the SDGs

Report	**Models**		
	OSGD	***SDG Mapper***	***AE4RIA SDG tracker***
(FFP 2021)	1, 8, 16	1, 2, 3, 5, 6, 8, 9, 10, 11, 13, 15, 16	1, 3, 4, 5, 6, 8, 10, 13, 16

(GIWPS 2023)	5, 16	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 15, 16, 17	3, 4, 5, 8, 10, 13, 16
(COP 2023)	7, 8, 13	1, 2, 3, 5, 7, 10, 11, 12, 13, 15, 17	1, 2, 4, 5, 7, 8, 9, 12, 13, 14, 15, 16
(IISD 2023)	7, 13, 17	1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 13, 15, 16, 17	5, 7 ,9, 11, 12, 13, 14, 16
(Eurofound 2023)	1, 3, 4, 5, 8, 10	1, 2, 3, 4, 5, 8, 9, 10, 11, 15, 16	1, 3, 4, 5, 8, 10, 11, 15
(SDSN 2023)	-	1, 2, 6, 8, 9, 12, 13, 14, 15, 17	6, 12, 13, 14, 15
(FIW 2022)	11, 16	1, 3, 4, 8, 9, 10, 16, 17	4, 16
(FAO 2015)	1, 2, 10, 17	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
(OECD 2022a)	1, 3, 5, 10	1, 2, 3, 5, 6, 8, 9, 10, 16	1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 16
(DCAF 2020)	16, 17	4, 8, 12, 16	9, 13, 16
(WEF 2024a)	3, 13, 16, 17	1, 3, 7, 8, 9, 10, 13, 14, 15, 16, 17	3, 7, 8, 9, 12, 13, 14, 16

(WEF 2024b)	8, 13, 16	1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 13, 14, 15, 16, 17	1, 3, 4, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16
(World Bank Group 2023)	1, 8, 10, 16, 17	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17	1, 4, 5, 8, 10, 13, 14, 16

The analysis of COP reports, "Summary of Global Climate Action at COP 28" (COP 2023) and "Summary of the 2023 Dubai COP" (IISD 2023), revealed a consensus among the three models on their connection to SDGs 7 (Affordable and Clean Energy) and 13 (Climate Action). This uniformity underscores the reports' focus on energy and climate-related issues. The Eurofound (2023) report displayed notable agreement across the models, linking it to SDGs 1 (No Poverty), 3 (Good Health and Well-being), 4 (Quality Education), 5 (Gender Equality), 8 (Decent Work and Economic Growth), and 10 (Reduced Inequalities). This harmony indicates a comprehensive coverage of various SDGs in the report.

In contrast, the (SDSN 2023) report, "The Global Commons Stewardship Index," showed the least connection, with no SDG linkages identified by the OSDG model. However, both our model and the SDG Mapper found connections to SDGs 6 (Clean Water and Sanitation), 12 (Responsible Consumption and Production), 13 (Climate Action), 14 (Life Below Water), and 15 (Life on Land), while the SDG Mapper also linked it to SDGs 1, 2, 8, 9, and 17.

When comparing the SDG Mapper and our model, it's crucial to consider the user's objectives. The SDG Mapper's percentage-based analysis offers a detailed breakdown of SDG-related content, providing a nuanced understanding. Our model, using a three-level relevance scale (0, 1, and 2), offers a quicker, heuristic assessment beneficial for swift preliminary evaluations.

The visualization of results also varies. The SDG Mapper's bar chart, with its visual differentiation, allows for easy distinction of SDG relevance levels and quick comparisons between documents. Conversely, our model's simpler green bar output, while less detailed, offers user-friendly accessibility, especially useful for interdisciplinary stakeholders with diverse backgrounds.

The choice between these tools depends on the users' needs. The SDG Mapper's detailed percentage-based approach is informative for measuring a document's alignment with each SDG and for comparative analysis. However, for entities needing a straightforward indication of SDG relevance, our binary model might be more efficient. It's important to note that percentage-based models may not reflect the varying significance different organizations place on each SDG, potentially leading to misinterpretations.

CONCLUSION

Our research journey, merging machine learning with policy analysis, focuses on the intricate relationship between HS reports and Sustainable Development Goals (SDGs). We mapped 44 HS reports against the 17 SDGs, using advanced machine learning models like RandomForestClassifier, tailored to each SDG. This method has illuminated the versatility of machine learning in policy document analysis and opened new possibilities for its application.

For policymakers, our study provides a nuanced tool to classify HS reports in relation to SDGs, enabling them to gain deeper insights into HS for informed decision-making aligned with sustainable development. Our approach aids in designing targeted interventions for specific HS aspects, ensuring efficient resource allocation and impactful solutions. Researchers benefit from this study's novel intersection of computational techniques and social science. It advances

interdisciplinary knowledge in HS and sustainable development, encouraging further exploration and cross-disciplinary collaboration.

Our methodology contributes significantly to understanding global HS and sustainable development trends. It demonstrates how machine learning can yield insights into societal issues, potentially inspiring more technology-social science hybrid research. A key outcome is our user-friendly web application, accessible at *https://sdgtracker.ae4ria.org*, which revolutionizes how users interact with HS reports. It provides instant classifications from our models, democratizing access to advanced data analysis and facilitating the practical application of theoretical findings. The application's design ensures customization, scalability, and relevance in the evolving security and development landscape.

Future research will focus on expanding training datasets and refining models with new data for enhanced performance. As empirical literature and data availability on HS and SDGs grow, our techniques will become more specific and reliable. This advancement will support decision-makers across sectors in identifying critical HS areas based on SDG criteria.

Our research bridges machine learning and policy analysis, offering new avenues for tackling global issues. The development of our web application marks a significant step from theoretical research to practical applications, playing a key role in interdisciplinary research and global challenge resolution. Our methodologies and insights are vital for developing comprehensive strategies for a sustainable future.

References

Almond, R. E. A., M. Grooten, D. Juffe Bignoli, and T. Petersen. 2022. Living Planet Report 2022: Building a Nature-positive Society. *WWF*.

Bird, S., E. Klein, and E. Loper. 2009. *Natural Language Processing with Python*. O'Reilly Media, Inc.

Eurofound. 2023. Promoting social cohesion and convergence: Economic and social inequalities in Europe in the aftermath of the COVID-19 pandemic. *Publications Office of the European Union.*

Boehm, S., L. Jeffery, K. Levin, J. Hecke, C. Schumer, C. Fyson, A. Majid, J. Jaeger, et al. 2022. State of Climate Action 2022. *World Resources Institute*. doi:10.46830/wrirpt.22.00028.

Borchardt, S., G. Barbero Vignola, D. Buscaglia, M. Maroni, and L. Marelli. 2022. Mapping EU Policies with the 2030 Agenda and SDGs. *EUR 31347 EN, Publications Office of the European Union, Luxembourg*.

Bostock, M., V. Ogievetsky, and J. Heer. 2011. D3: Data-Driven Documents. IEEE Transactions on Visualization and Computer Graphics (Proc. InfoVis).

Breiman, L. 2001. Random Forests. *Machine Learning* 45: 5–32. doi:https://doi.org/10.1023/A:1010933404324.

CBD. 2020. Global Biodiversity Outlook 5. *Secretariat of the Convention on Biological Diversity*.

Chancel, L., P. Bothe, and T. Voituriez. 2023. Climate Inequality Report 2023. *World Inequality Lab, United Nations Development Programme*.

CHS. 2003. Human Security Now. *Commission on Human Security*.

Clark, K., M.-T. Luong, Q. V. Le, and C. D. Manning. 2020. ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators.

COP. 2023. Summary of Global Climate Action at COP 28. *Conference of the Parties.*

Crabtree, A., and D. Gasper. 2020. Conclusion: The sustainable development goals and capability and human security analysis. In *Sustainability, capabilities, and human security*. Palgrave Macmillan, Cham.

DCAF. 2020. Promoting Coherence between the OECD Guidance and the Voluntary Principles on Security and Human Rights. *Geneva Center for Security Sector Governance*.

Devlin, J., M. W. Chang, K. Lee, and K. Toutanova. 2018. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.

EC. 2024. SDG Mapper. *European Commission*.

FAO. 2015. Sustainable Development Goals: 17 Goals to Transform our World. *FAO Regional Office for Asia and the Pacific*.

FAO. 2022. The State of Food Security and Nutrition in the World: Repurposing Food and Agricultural Policies to make Healthy Diets more Affordable. *Food and Agriculture Organization of the United Nations*.

FFP. 2021. Fragile States Index Annual Report 2021. *Fund For Peace*.

FFP. 2022. State Resilience Index Annual Report 2022. *Fund For Peace*.

FIW. 2022. Freedom in the World, 2022. The Global Expansion of Authoritarian Rule. *Freedom House*.

Fox, A. 2009. Cloud Computing—What's in It for Me as a Scientist? *Science* 327: 419–420.

FSIN and GNAFC. 2022. Global Report on Food Crises 2022. *Food Security Information Network and Global Network Against Food Crises*.

FSIN and GNAFC. 2023. Global Report on Food Crises 2023. *Food Security Information Network and Global Network Against Food Crises*.

Fuentes, F. C., and F. Rojas Aravena. 2005. Promoting Human Security: Ethical Normative, and Educational Frameworks in Latin America and the Caribbean. *UNESCO*.

Fukuda-Parr, S., and C. Messineo. 2012. Human Security: A critical review of the literature. *Center for Research on Peace and Development* CRPD Worki.

Gasper, D. 2005. Securing Humanity: Situating 'Human Security' as Concept and Discourse. *Journal of Human Development* 6: 221–245. doi:10.1080/14649880500120558.

Gasper, D., and O. A. Gómez. 2013. Evolution of thinking and research on human security and personal security 1994-2013. In *Safeguarding Human Progress: Reducing Vulnerabilities, Building Resilience*, ed. K. Malik. New York: UNDP.

Gasper, D., R. Jolly, G. Koehler, T. Kool, and M. Simane. 2020. Adding human security and human resilience to help advance the SDGs agenda. *ISS Working Paper Series/General Series, 665.*

GIWPS. 2023. Women Peace and Security Index 2023/2024: Tracking sustainable peace through inclusion, justice, and security for women. *Georgetown Institute for Women, Peace and Security*.

Grinberg, M. 2018. Flask Web Development: Developing Web Applications with Python. *O'Reilly Media, Inc.*

Guyon, I., and A. Elisseeff. 2003. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research* 3: 1157–1182.

Gwartney, J., R. Lawson, J. Hall, and R. Murphy. 2022. Economic Freedom of the World: 2022 Annual Report. *Fraser Institute*.

Halkos, George E., and P.-S. C. Aslanidis. 2023. Addressing Multidimensional Energy Poverty Implications on Achieving Sustainable Development. *Energies* 16: 3805.

doi:10.3390/en16093805.

Halkos, George E, and P.-S. C. Aslanidis. 2023. Causes and Measures of Poverty, Inequality, and Social Exclusion: A Review. *Economies* 11: 110. doi:10.3390/economies11040110.

Human Security Unit United Nations. 2009. *Human Security in Theory and Practice: Application of the Human Security Concept and the United Nations Trust Fund for Human Security*. *United Nations Trust Fund for Human Security*. New York, NY 10017, USA.

IAP. forthcoming. IAP Statement on Human Security. *The Inter-academy Parntership.*

IEP. 2022a. Ecological Threat Report 2022: Analysing Ecological Threats, Resilience, and Peace. *Institute for Economics & Peace*.

IEP. 2022b. Global Peace Index 2022: Measuring peace in a complex world. *Institute for Economics & Peace*.

IIASA. 2018. Transformations to Achieve the Sustainable Development Goals. *International Institute for Applied Systems Analysis*.

IISD. 2023. Summary of the 2023 Dubai Climate Change Conference: 30 November - 13 December 2023.

ILO. 2022. Global Wage Report 2022-23: The Impacts of Inflation and COVID-19 on Wages and Purchasing Power. *International Labour Organization*.

IPCC. 2014. Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. *Intergovernmental Panel on Climate Change*.

Jacobs, G. 2016. Integrated Approach to Peace & Human Security in the 21st Century. *Cadmus* 3: 48–71.

JICA. 2022. Human Security Today. *Japan International Cooperation Agency*.

Kaldor, M. 2013a. In Defence of New Wars. *Stability: International Journal of Security and Development* 2: 4. doi:10.5334/sta.at.

Kaldor, M. 2013b. *New and Old Wars: Organised Violence in a Global Era (3rd Edition)*. Polity.

Koundouri, P., and K. Dellis. 2023. Human Security: Concepts and Measurement. *Cadmus* 5. Cadmus: 28–44.

Koundouri, P., N. Theodossiou, C. Stavridis, S. Devves, and A. Plataniotis. 2023. The European Green Deal and the 17 SDGs: Uncovering their Connection with a ML-based Approach. In *Decision Making Using AI in Energy and Sustainability: Methods and Models for Policy and Practice*, 21–48. Springer International Publishing.

Liaw, A., and M. Wiener. 2002. Classification and Regression by RandomForest. *R News* 2: 18–22.

Liotta, P. H., D. A. Mouat, W. G. Kepner, and J. M. Lancaster. 2007. *Environmental Change and Human Security: Recognizing and Acting on Hazard Impacts*. Dordrecht, The Netherlands: Springer Science + Business Media, B.V.

Martin, M., and T. Owen. 2014. *Routledge Handbook of Human Security*. Abingdon Oxon.: Abingdon: Routledge.

McKinney, W. 2010. Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference.*

Merkel, D. 2014. Docker: Lightweight Linux Containers for Consistent Development and Deployment. *Linux Journal* 239.

Nilson, C. 2018. Community safety and well-being: Concept, practice, and alignment (LEPH2018). *Journal of Community Safety and Well-Being* 3: 96. doi:10.35502/jcswb.81.

OECD. 2020a. Environment at a Glance 2020. *Organisation for Economic Co-operation and Development Publishing*.

OECD. 2020b. How's Life? 2020: Measuring Well-being. *Organisation for Economic Co-operation and Development Publishing*.

OECD. 2022a. Adding the human dimension to the OECD fragility framework. *OECD Development Co-operation Directorate, OECD Publishing*.

OECD. 2022b. Disability, Work, and Inclusion: Mainstreaming in All Policies and Practices. *Organisation for Economic Co-operation and Development Publishing*.

OECD. 2022c. The Short and Winding Road to 2030: Measuring Distance to SDG Targets. *Organisation for Economic Co-operation and Development Publishing*.

Ogata, S. 2011. Hitobito o Torimaku Kyoi to Ningen no Anzen Hosho no Hatten. (In Japanese). *Kokusai Mondai. The Japan Institute of International Affairs* July-Augus.

Owen, T. 2007. Measuring Human Security: Methodological Challenges and the Importance of Geographically Referenced Determinants. In *Environmental Change and Human Security: Recognizing and Acting on Hazard Impacts2*, ed. P. H. Liotta, D. A. Mouat, W. G. Kepner, and J. M. Lancaster, 35–66. Dordrecht, The Netherlands: Springer Science + Business Media, B.V.

Pilgrim, M. 2010. HTML5: Up and Running: Dive into the Future of Web Development. *O'Reilly Media, Inc.*

Pip. 2021. PipUser Guide. *Python Packaging Authority*.

Pukelis, L., N. Bautista-Puig, G. Statulevičiūtė, V. Stančiauskas, G. Dikmener, and D. Akylbekova. 2022. OSDG 2.0: a multilingual tool for classifying text data by UN Sustainable Development Goals (SDGs).

Ramos, J. 2003. Using TF-IDF to Determine Word Relevance in Document Queries. *Proceedings of the First Instructional Conference on Machine Learning.*

Ronacher, A. 2010. Flask: A Python Microframework. *Flask Documentation*.

Salton, G., and M. J. McGill. 1986. Introduction to Modern Information Retrieval. *McGraw-Hill, Inc*.

Sanh, V., L. Debut, J. Chaumond, and T. Wolf. 2019. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter.

SDSN. 2023. Global Commons Stewardship Index 2022. *SDSN, Yale Center for Environmental Law & Policy, and Center for Global Commons at the University of Tokyo.*

SolABILITY. 2022. The Global Sustainability Competitiveness Index 2022. *SolAbility Sustainable Intelligence*.

Strobl, C., A.-L. Boulesteix, A. Zeileis, and T. Hothorn. 2007. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* 8: 25. doi:10.1186/1471-2105-8-25.

Transparency International. 2021. Corruption Perceptions Index 2021.

UN. 2015. The 17 goals. *United Nations*.

UNDESA. 2018. 2018 Revision of World Urbanization Prospects. *United Nations Department of Economic and Social Affairs*.

UNDESA. 2023. World Economic Situation and Prospects. *United Nation Department of Economic and Social Affairs*.

UNDP. 1994. Human Development Report 1994: New Dimensions of Human Security. *United Nations Development Programme*.

UNDP. 2022. New threats to human security in the Anthropocene: Demanding greater solidarity.

United Nations Development Programme.

UNDP and OPHI. 2022. Global Multidimensional Poverty Index 2022: Unpacking deprivation bundles to reduce multidimensional poverty. *United Nations Development Programme & Oxford Poverty and Human Development Initiative*.

UNESCO. 2008. *Human Security: Approaches and Challenges.* Paris, France: United Nations Educational, Scientific and Cultural Organization.

UNTFHS. 2017. Human Security and Agenda 2030.

WEF. 2022. The Global Risks Report 2022: 17th Edition. *World Economic Forum*.

WEF. 2024a. The Global Cooperation Barometer 2024. *World Economic Forum.*

WEF. 2024b. The Global Risks Report: 19th Edition. *World Economic Forum.*

WHO. 2022. World Health Statistics 2022: Monitoring health for the SDGs. *World Health Organization*.

Wickham, H., and G. Grolemund. 2017. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. *O'Reilly Media, Inc.*

Wolf, M. J., J. W. Emerson, D. C. Esty, A. de Sherbinin, Z. A. Wendling, and et al. 2022. Environmental Performance Index: Ranking country performance on sustainability issues. *New Haven, CT: Yale Center for Environmental Law & Policy*.

World Bank Group. 2023. World Development Report: Migrants, Refugees, and Societies. A World Bank Group Flagship Report. *World Development Report.*

Appendix I - Application of Results: Development of a Web App

Building on the insights and methodologies established in our research, we ventured into applying our results in a practical, user-friendly tool: a web application developed using Flask and Docker. This application represents the culmination of our research efforts, transforming the theoretical and analytical aspects of our study into a tangible, interactive platform.

Overview of the Web Application

User Interface: The application features a straightforward user interface, where users can upload PDF documents. The primary function is to assess these documents for their relevance to the Sustainable Development Goals (SDGs) based on the keyword frequency analysis established in our research.

Backend Processing: Upon uploading a PDF, the application employs the `pdfplumber` library to extract text from the document. This text is then processed through our scoring system, which evaluates the frequency of pre-defined SDG keywords, normalized by the total word count.

Scoring Mechanism

Keyword Frequencies: The core of the scoring system lies in calculating the keyword frequencies, a process adapted from our research methodology. For each SDG, the application counts the occurrences of relevant keywords within the text, offering a nuanced understanding of the document's alignment with the SDGs.

Utilization of Thresholds: The scoring is guided by the thresholds established in our research. These thresholds, stored in the `thresholds.pkl` file, determine the score for each SDG based on the keyword frequency. The scores range from 0 (low relevance) to 2 (high relevance).

Deployment and Accessibility

Flask and Docker Integration: The web application is built using Flask, a lightweight web framework that provides the necessary tools for web app development. Docker is used for

containerization, ensuring that the application runs smoothly across different environments. This setup allows for easy deployment and scalability.

Practical Application: The application serves as a practical tool for various stakeholders, including researchers, policy-makers, and educators, who seek to understand the relevance of specific documents to the SDGs. It offers a quick and efficient way to gauge this relevance, backed by the research and methodologies from our study.

Closing Remarks

The development of this web application is a significant stride in bridging the gap between academic research and practical applications. It not only validates our research methodology but also extends its utility to a wider audience, facilitating greater engagement with and understanding of the Sustainable Development Goals. The application stands as a testament to the potential of integrating advanced research methodologies with modern web technologies to create impactful and accessible tools.