The Interrelation of Sustainable Development Goals in Publications and Patents: A Machine Learning Approach

Arash Hajikhani[1[0000-0003-2032-9180]] and Arho Suominen[1,2[0000-0001-9844-7799]]

[1] Quantitative Science and Technology Studies, VTT Technical Research Centre of Finland, Tekniikantie 21, 02044 Espoo, Finland
[2] Tampere University, P.O. Box 541, Tampere FI-33014, Finland

Abstract. The Sustainable Development Goals (SDGs) are the blueprint for achieving a better and more sustainable future for all by defining priorities and aspirations for 2030. In this paper, the attempt was to expand SDGs' definition by performing a comprehensive literature review. Furthermore, the descriptions of SDGs were utilized to compile a Machine Learning (ML) model so to automate the detection of SDG relevancy in other types of artefacts. The model was employed for identifying the SDG relevancy of patents as well-known proxies for innovation. The ML model was then used to classify a sample of patent families registered in the European Patent Office (EPO). The analysis revealed the extend to which SDGs were addressed in patents and the interrelations between SDG definitions. The findings guide how to align patenting strategies as well as measurement and management of their contribution to the realization of the SDGs when it comes to Intellectual Property (IP) strategies.

Keywords: United Nations; Sustainable Development Goals; SDGs; Innovation; Intellectual Property; Patents; Natural Language Processing; Machine Learning Model; Patenting Strategy

1 Introduction

Our planet faces massive economic, social and environmental challenges. To combat these, the Sustainable Development Goals (SDGs) define global priorities and aspirations for 2030. They represent an unprecedented opportunity to eliminate extreme poverty and put the world on a sustainable path.

Science, technology and innovation (STI, as referred to in the UN and OECD contexts) have been recognised as one of the main drivers behind productivity increases and a key long-term lever for economic growth and prosperity [1]. STI is a fundamental tool to implement the new agenda, as it allows improving efficiency in both economic and environmental senses, developing new and more sustainable ways to satisfy human needs, and empowering people to drive their own future [2]. In the SDGs framework, STI features strongly both in Goal 17, as well as a cross-cutting one to achieve several sectoral Goals and Targets. Fostering innovation is part of Goal 9 related to resilient infrastructure and inclusive, sustainable industrialisation, while Target 9.5 elevates the role of research and innovation policy well beyond STI as one of the Means of Implementation.

From United Nations general assembly briefing materials, the importance of Science, Technology and Innovation (STI) for the SDGs has been numerously mentioned in yearly forums 1. A direct quote from Marie Chatardová President of the Economic and Social Council, at the 2018 New York STI Forum "No one can ignore the vital role of science, technology and innovation" (STI) in "advancing the transformative impact"2. Another quote from Technology Adviser to the US Secretary of State, said that the integrated nature of the SDGs requires multi-disciplinary and holistic science, technology and innovation approaches that break silos and take into account different sources of knowledge, at the concluding session of the Forum.

Innovation in general, but also innovation in the context of sustainable development affects many parts of human life and should thus be treated with more concern. Goal nine needs to be highlighted in the context of innovation. This goal states that the United Nations' objective is to "promote inclusive and sustainable industrialization and foster innovation" (p. 17). The importance of innovation to reach a sustainable development is also recognized by Ashford and Hall [3].

STI is recognized as a democratizing tool for transferring science to society by innovation and technology and it can show capacities to mobilize science, technology and innovation for the achievement of SDGs. The first stage to set the path for direct efforts to address SDGs is to understand what the track records have been so far. One of the main ways of STI's oriented efforts manifested is through scholarly literature and intellectual property protected in the form of patents. The ability to take stock of and analyse the output of science and technology has increased tremendously in the past decade due to the increasing degree of digitalization of research article and intellectual property databases (e.g., Web of Science, Scopus, PATSTAT, Google Patents). To a significant extent, this literature still focuses on descriptive values rather than focusing on creating an in-depth feature to data that create additional vantage points to evaluate innovation systems.

This research is set to identify the SDG oriented artefacts which to a large extent materialize the efforts and outcomes of science, technology and innovation. The research is trying to identify and capture values-based and sustainability-oriented Science, Technology and Innovation. A rigorous methodology will be utilized to extract definitions and criteria of sustainable development goals. This knowledge will then be used to perform a classification task and build a machine-learning model to identify publication and patents' relevancy to SDGs. The direct research questions from this exercise are formulated as what the distribution of SDG relevancy on patents and to what extent there is a relationship between distinct SDGs. Next part of the outlines describes the research design and methodological approach. Analysis and discussion will follow.

---

[1] UN, Science, Technology and Innovation for the SDGs 2018 Forum: https://www.un.org/development/desa/indigenouspeoples/science-technology-and-innovation-for-the-sdgs.html

[2] Science, technology and innovation crucial to 'transformative impact' of Development Goals, UN 2018 forum hearing: https://www.un.org/development/desa/en/news/sustainable/sti-forum-2018-opening.html

2 Background

STI and the interaction between different actors is seen as the core indicator for the economic growth [4–6]. The increase in the production of scientific and technological knowledge acts as the key source of innovation and competitive advantage [7]. This has been central to our understanding of the competitiveness of nations, but also the competitiveness at the firm level. The centrality of the concept of productivity and its increase to sustain the long-term competitiveness of nations has been the central paradigm of economic policy for decades. This also explains that much of the literature on the STI process focus on innovation outcomes [8]. Within this literature much of the focus is centred around scientific work and research and development within the innovation system. These are seen as vehicles to enable job creation, firm performance and ultimately increases in the gross domestic product.

However, there is an on-going debate focused to extend our focus beyond productivity or gross domestic product to other impact measures [9]. These developments have been fortified by global challenges like climate crisis that has opened a dialogue to re-evaluate the role of pure productivity as the objective for governments or businesses. In the public sector. the climate crisis has strengthened the call for additional outcome measures for innovation system activities. This has been clear in the broadening of public impact assessment [10]. Policy discussion has also actively discussed the role on grand challenges and the role of governments to take an active role in facilitating transitions unlikely to happen through other means but creating significant overall benefits (e.g. Mazzucato [11]). We have also seen significant transitions in company leaders' positions to the role of companies in grand challenges. The call from large company CEOs to extend firm's objectives beyond shareholder value [12] can be regarded as major transition beyond the current paradigm towards looking towards a sustainable economy.

In this transition the work of the United Nations on the creation of SDG has been central. SDGs offer one of the first holistic taxonomies of grand challenges. The emergence of the SDG framework, and the overall shift in discussion, has required transformative changes to the overall innovation system [13]. This again is particularly discussed in the policy domain, but literature also looks at the role of industry and innovation activities relationship the SDGs. It is clear that there is a need to adjust all aspects be they economic, governance and public policy at all levels if science, technology and innovation to reorient to the SDG agenda [14]. One of the central questions is, if it is fundamentally unrealistic to fit the corporate expectation to maximize profits to ensuring equitable and sustainable development [15].

This said, we can still estimate the impacts of industrial activity on the SDG. This requires the development of practical proxy measures to establish a practical measurement of corporate activities impact on the goals. Approached towards measuring societal impact of innovation has been done for example in the context of frugal innovation [16]. However, literature has not shown practical approaches for creating a proxy measure for large scale analysis of STI impact on the SDGs. Our attempt in this study is to extent the measurement of SDG relevancy to intellectual property (IP) type of document such as patents. Patent documents are technical in nature but our approach will benefit from equivalent scientific publication document to extend the lexical query regarding SDG and once the approach compiled to a machine learning model then utilize it for identifying the SDGs in patent documents.

3 Methodology (Query crafting and data collection)

With the overall aim to identify SDG related science and technology and innovation artefacts, the study design grounds on creating a lexical query by utilising an iteratively developed database of SDG terminology (Figure 1). The keyword search was applied to map scientific publications from the last decade. This period was chosen because it encompasses the most recent activities, partly impacted by the SDG agenda. The outcome of this process results in research publications concerning SDG focus. Meanwhile, patent documents, due to their content and descriptive nature, have not given representative results regarding their relevancy to SDGs by use of SDG lexical searching queries. While we have received representative results from publication data, we can extrapolate the machine-learning model based on publication data to detect relevant patents to SDGs.

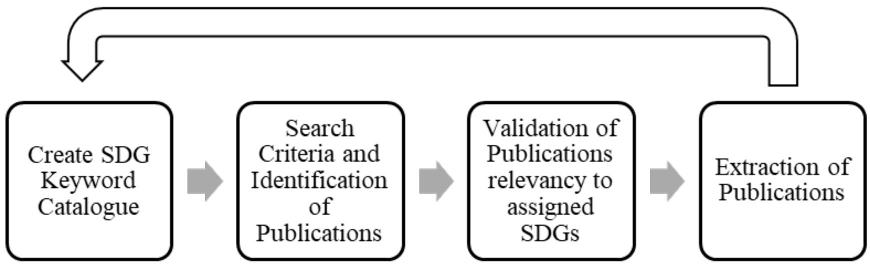


Figure 1. Workflow process of identifying and retrieving SDG related publications

A detailed taxonomy has been developed for mapping the SDG relevant publications. The process of curating the lexical keywords involved analysing the UN Sustainable Development Goals documents [17]. From the semantics perspectives, each word or concept has been expanded to lexically similar terms. Besides, the extracted list of keywords was matched with existing taxonomies [18–21]. After this process, with the keywords, searching queries are compiled for each SDG and then initiated on the publication database. For this study, the SCOPUS database was used to identify potential sources. SCOPUS is the largest abstract and citation database of peer-reviewed literature. It includes books, scientific journals, and conference proceedings. Compared to other scientific databases such as the Web of Science, SCOPUS has broader coverage, and it is a widely used database to create datasets for systematic reviews of research [22]. SCOPUS has already identified publications relevant to SDGs [23]. The resulting publications were observed carefully to validate the relevancy of the resulted records to the corresponding SDG.

The bibliometric data for the resulting publication were extracted for each SDG. We made sure to extract the textual content of the publications such as Title, Abstract and

Keywords within the bibliometric data. Harvesting the textual content, we are able to train a model for automating the detection of unseen SDG related document. The model is useful to pick up the hidden semantics of SDGs and the commonalities among the SDGs. In particular, the model will be utilized for the more technical nature documents such as patents. In our experiment, scientific publication metadata and the previous studies on identifying SDG related publication was a starting point to engineer the detecting mechanism of SDG related patent documents.

This research has benefited from text classification and machine learning algorithms to facilitate the SDG detecting model's construction. Text classification is one of the research hotspots in the field of Natural Language Processing (NLP). Originated from computer science and evolved from pattern recognition, the automated process of categorization (or classification) of an object such as text has become one of the growing interests of utilizing machine learning. Due to the increased availability of documents in digital form, ensuring the need for flexible ways to access them [24]. Therefore, the activity of labelling natural language text (text classification or topic modeling) by machine learning algorithms gives an opportunity to process a large amount of text automatically for better insights.

The python programming language was used to handle the data structuring and ML model building. The SDG related publications identified in the previous step will be utilized as a training set for the classification algorithm. The classification methods for performing a multi-class text classification model are many (e.g. Naive Bayes, Maximum Entropy, SVM). Naive Bayes is the algorithm used in our research to train the machine learning classification model. Naive Bayes is a probabilistic model which works well on text categorization [25]. In validating the accuracy and reliability of the model, we use a test set (new publications) to confirm the usability of the ML model. Figure 2 illustrates the workflow of the methodological steps.

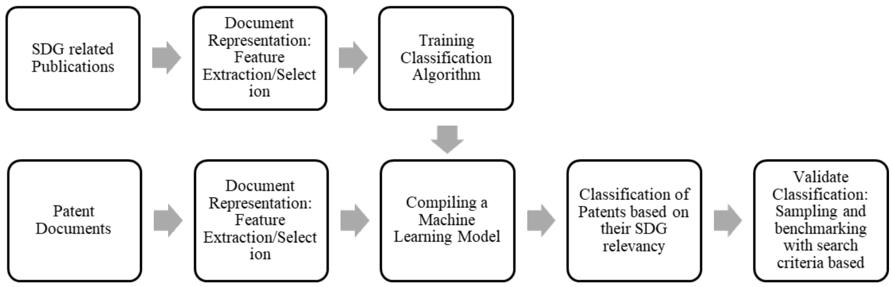


Figure 2. Workflow and study design for the Sustainable Development Goals (SDG) detection and mapping of Intellectual property documents

After compiling the machine learning model, we can rely on its judgment to classify an unseen document concerning its relevancy to SDG themes. In the case of our study, we are interested in identifying patents that are addressing the SDGs. The assignment of SDG relevancy to a patent document will be a probabilistic distribution. We can see the patent document relevancy by a percentage point to all SDG themes.

4 Analysis

To identify the relatedness of patents to sustainable development goals, we are operationalizing the strategy described in the methodology part. With the help Python programming language, the steps will be carried out to compile the SDG identification task model to classify the new text (in our case patent text). A prerequisite step before compiling the Machine Learning model is text pre-processing of the initial text. Text pre-processing is traditionally an essential step for natural language processing (NLP) tasks. It transforms text into a more digestible form so that machine learning algorithms can perform better. The pre-processing text phases include stop word removal, stemming, and lemmatization, which aims to normalize all the text on a level playing field. The text will also be tokenized, splitting strings of text into smaller pieces or "tokens". We also take care of noise removal from the initial text, such as cleaning up text from extra whitespaces, lowercase of all text and removing special characters. Finally, we will convert our text documents to a matrix of token counts, then transform a count matrix to a normalized ratio such as TF-IDF (term frequency-inverse document frequency) a numerical representation of the text. After that, we train several classifiers using Python's Scikit-Learn library and Gensim library.

Once the features are generated from the text, machine learning classifier can be trained with the labelled publication data to finally predict the SDG relevancy of a patent document. Six different classification strategy has been tested and the performance of each model is reported in Table 1. The models' overall accuracy did not reach above 60% (based on the first prediction of the model). This benchmark is not a high standard accuracy for accepting the model for useful classification of all the classes. Our model experimentations deliver acceptable accuracy (above 60%) for most of the SDG classes such as: SDG 1, 2, 3, 4, 5, 6, 7, 9, 10, 13 and 16.

Table 1. Classification Models Performance Comparison

	NAIVE BAYES CLASSIFIER FOR MULTINOMIAL MODELS			**LINEAR SUPPORT VECTOR MACHINE**			**LOGISTIC REGRESSION**			**WORD2VEC AND LOGISTIC REGRESSION**			**DOC2VEC AND LOGISTIC REGRESSION**			**MULTI-LAYER PERCEPTRON CLASSIFIER**		
	precision	***recall***	***f1-score***	***precision***	***recall***	***f1-score***	***precision***	***recall***	***f1-score***	***precision***	***recall***	***f1-score***	***precision***	***recall***	***f1-score***	***precision***	***recall***	***f1-score***
SDG1	0.53	0.75	0.62*	0.67	0.72	0.69*	0.60	0.56	0.58	0.65	0.68	0.66*	0.58	0.62	0.60*	0.68	0.60	0.64*
SDG2	0.57	0.60	0.58	0.57	0.67	0.62*	0.52	0.61	0.56	0.61	0.62	0.62*	0.57	0.60	0.58	0.55	0.57	0.56
SDG3	0.82	0.88	0.85*	0.69	0.93	0.79*	0.87	0.89	0.88*	0.86	0.87	0.86*	0.81	0.89	0.85*	0.85	0.92	0.88*
SDG4	0.75	0.74	0.74*	0.71	0.87	0.78*	0.78	0.77	0.77*	0.78	0.75	0.76*	0.76	0.77	0.77*	0.86	0.79	0.82*
SDG5	0.55	0.72	0.63*	0.59	0.79	0.68*	0.60	0.57	0.58	0.65	0.61	0.63*	0.62	0.61	0.61*	0.62	0.68	0.65*
SDG6	0.57	0.63	0.60*	0.58	0.76	0.66*	0.62	0.59	0.61*	0.62	0.66	0.64*	0.64	0.63	0.63*	0.63	0.55	0.59
SDG7	0.83	0.78	0.80*	0.69	0.86	0.76*	0.80	0.81	0.81*	0.82	0.84	0.83*	0.83	0.82	0.82*	0.82	0.85	0.83*
SDG8	0.50	0.46	0.48	0.59	0.40	0.48	0.41	0.41	0.41	0.52	0.46	0.49	0.50	0.48	0.49	0.46	0.48	0.47
SDG9	0.49	0.66	0.56	0.61	0.70	0.65*	0.59	0.56	0.57	0.59	0.66	0.62*	0.59	0.54	0.56	0.66	0.63	0.65*

SDG10	0.69	0.45	0.54	0.76	0.56	0.64*	0.62	0.57	0.59	0.63	0.56	0.60*	0.52	0.49	0.50	0.57	0.61	0.59
SDG11	0.51	0.44	0.47	0.55	0.54	0.55	0.42	0.48	0.45	0.53	0.54	0.53	0.53	0.55	0.54	0.51	0.50	0.51
SDG12	0.56	0.40	0.47	0.65	0.36	0.46	0.45	0.43	0.44	0.54	0.49	0.51	0.47	0.42	0.44	0.50	0.46	0.48
SDG13	0.50	0.64	0.56	0.55	0.60	0.58	0.55	0.50	0.53	0.56	0.59	0.58	0.60	0.66	0.63*	0.54	0.53	0.54
SDG14	0.16	0.16	0.16	0.13	0.10	0.11	0.10	0.11	0.11	0.31	0.37	0.34	0.30	0.29	0.29	0.09	0.11	0.10
SDG15	0.10	0.07	0.09	0.09	0.06	0.07	0.11	0.11	0.11	0.25	0.19	0.22	0.31	0.31	0.31	0.05	0.05	0.05
SDG16	0.73	0.49	0.58	0.74	0.51	0.61*	0.60	0.64	0.62*	0.60	0.69	0.65*	0.56	0.57	0.57	0.65	0.57	0.61*

According to the model comparison in Table 1, the highest overall accuracy (f-score) is achieved by "Word2vec and logistic regression" models. Next, we will utilize the classifier model to perform the patents' classification task to the selected SDGs. For this step a sample patent dataset is retrieved from Clarivat's Derwent innovation database. For the duration of 3 years between 2017 and 2019, all the granted patent families were retrieved along with their textual content (title and abstract), patent number, assignee and patent country code. A set of 31 thousand patent family was collected. In order to apply the "Word2vec and logistic regression" classifier for assigning the patents family relevancy to any of the SDGs, the same pre-text cleaning procedures applied to the textual content of the retrieved patents. After passing the patent's process texts into the ML model, we achieved the relevancy of each patent document to any of the SDGs with the distribution of probability percentage to each SDG. Considering the ML model's better performance into classifying a specific category of SDGs, we have trimmed the results for that specific SDGs. Figure 3 presents the accumulative highest to lowest relevant SDGs which was addressed in the patent application texts.

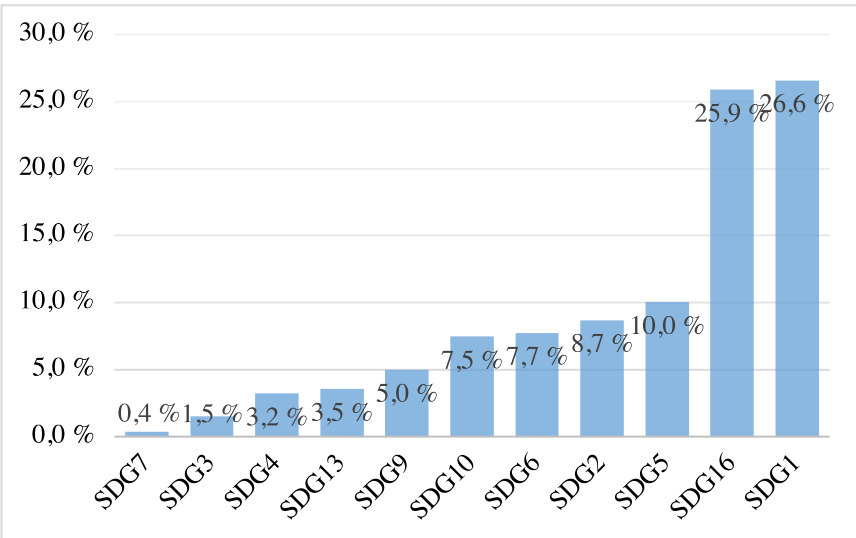
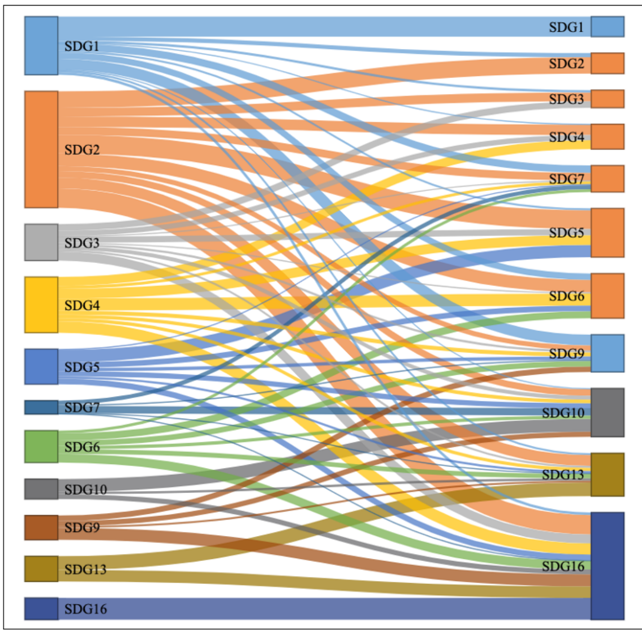


Figure 3. Relevancy of patent documents to SDGs

Figure 3 indicates that almost half of the patents address SDG 1 and 16, then the 35% of the patents address SDG 5, 2, 6 and 10. The rest of SDGs were not picked significantly in patents textual information.

One interesting observation could be looking at the language structure of patent and publications across various SDGs. To quantify the difference of language structure between patent and publications in the same SDG category, we utilize the vector representation of the tokens of text. Back in the text preparation phase we used the "word2vec" method, a type of word representation that allows words with similar meaning to be understood by classifiers. In technical terms, it maps words into vectors of real numbers using the neural network, probabilistic model, or dimension reduction on word co-occurrence matrix. Once we transformed our textual content into their numerical representation, we are able to perform a general arithmetic operation such as measuring the similarity or calculating the different structure of word when comparing two bags of text. Knowing the background for the method, we did estimate the difference in language structure of patents and publications for each SDG and reported the measure as similarity measure. The value was then normalized so to show in a percentage point. The higher the percentage point the similar the text would be between patent and publications. Figure 4 illustrates the interconnectedness of SDG oriented patent and publications with similarity measure. Obviously, the same SDG category indicates a higher similarity percentage (i.e. Publication: SDG1 – Patent: SDG1 60.6% similarity).



Publication	Patent	Similarity
SDG1	SDG1	60.6%
SDG2	SDG2	49.8%
SDG3	SDG3	19.6%
SDG4	SDG4	26.8%
SDG5	SDG5	36.9%
SDG6	SDG6	21.5%
SDG7	SDG7	12.1%
SDG9	SDG9	16.8%
SDG10	SDG10	38.7%
SDG13	SDG13	41.2%
SDG16	SDG16	65.7%

Figure 4. Flow diagram illustrating the interconnectivity in Publication (left side) and Patent (right side) document regarding SDG relevancy.

Interestingly, in this observation we noticed textual similarities between patent and publications in unexpected categories such as “Publication: SDG2 – Patent: SDG16 58.60% similarity”, “Publication: SDG2 – Patent: SDG5 56.70% similarity”, “Publication: SDG2 – Patent: SDG6 38.10% similarity”. The extensive table of text relationships with a similarity ratio of above 30% between non-similar categories of SDGs is visible in Table 2.

Table 2. The interrelatedness of different SDG categories within Publication and Patents

Publication	Patent	Similarity
SDG2	SDG16	58.60%
SDG2	SDG5	56.70%
SDG2	SDG6	38.10%
SDG4	SDG6	36.60%
SDG9	SDG16	36.20%
SDG13	SDG16	35.10%
SDG4	SDG16	34.10%
SDG1	SDG9	32.80%
SDG2	SDG13	32.80%
SDG2	SDG4	31.70%

5 Discussion & Conclusion

As an international developmental framework to achieve a better and more sustainable future, the United Nations Sustainable Development Goals (SDGs) offers a gridline to activate a development-oriented approach. Governments worldwide have already agreed to these goals. Now the pressure to take action is on a global as well as on local levels. Scientific and technological innovations are necessary but enabling them to impact requires an understanding of their utility to the sustainable positive economy. Our study is capable of makes several contributions on systematically comprehending sustainability-oriented science, technology and innovation. First, it offers a systematic path for creating a catalogue for sustainable development goals requirement and objectives. Second, based on the publications with the highest relatedness to SDGs, the study train and develop a machine learning model in order to detect the relatedness of another type of Scientific and technological innovations such as Patents. Development intersects with IP policies as creativity and innovation are either fostered or frustrated by an economy’s chosen development policy. Therefore, including consideration of the SDGs in IP policy could lead to more significant and more lasting success.

Based on the UN’s SDG definition, we queried for relevant publications which address the 16 SDGs. This way we extend our vocabulary to capture the breadth and depth of SDGs within scientific publications. The application of comprehensive lexical design for SDGs was compiled in a machine learning model so to classify other textual content regarding their relevancy to the SDGs. Then we utilized the trained machine learning model to identify the applicability of intellectual property (IP) documents concerning SDGs.

Within a sample set of patents collected from the European patent office for 3 years (2017-2019), the model performed with high accuracy in detecting 11 of the SDGs. Among the identified SDGs, there was high relatedness on some of the SDGs compares to others (e.g. SDG 1 and SDG 16). Based on the similarity measure, we learned the interconnectedness between SDG categories. For example, SDG2 between SDG 5 and 16. While patent texts are technical in nature, we could quantify if any notion of SDGs were addressed in the patent's textual content and to what extent.

The study's implicit implications can provide an overview of the STI so far contributed to SDGs on a macro level. On a micro level, it guides companies on how they can align their strategies and measure and manage their contribution to the realization of the SDGs when it comes to IP strategies.

Acknowledgments

This project has received funding from Business Finland Innovation Research 2020 under project name INNOSDG.

References

1. Daejeon Declaration: Daejeon Declaration on Science, Technology and Innovation Policies for the Global and Digital Age. (2015).
2. Giovannini, E., Niestroy, I., Nilsson, M., Roure, F., Spanos, M.: The Role of Science, Technology and Innovation Policies to Foster the Implementation of the Sustainable Development Goals (SDGs) Report of the Expert Group " Follow-up to Rio + 20 , notably the SDGs ." (2015). https://doi.org/10.2777/485757.
3. Ashford, N.A., Hall, R.P.: The importance of regulation-induced innovation for sustainable development. Sustainability. 3, 270–292 (2011). https://doi.org/10.3390/su3010270.
4. Nelson, R.R., Sidney, G.: An Evolutionary Theory of Economic Change. Harvard Business School Press, Cambridge, MA, USA (1982).
5. Dosi, G.: Innovation, organization and economic dynamics: selected essays. Edward Elgar Publishing (2000).
6. Freeman, C.: The "National System of Innovation" in historical perspective. Cambridge J. Econ. 19, 5–24 (1995). https://doi.org/Article.
7. TIdd, J., Bessant, J., Pavitt, K.: Managing Innovation: Integrating Technological, Market and Organizational Change. John Wiley & Sons (2005).
8. Kahn, K.B.: Understanding innovation. Bus. Horiz. 61, 453–460 (2018). https://doi.org/10.1016/j.bushor.2018.01.011.
9. Simonova, M.D.: Beyond GDP. Measuring What Counts for Economic and Social Performance. MGIMO Rev. Int. Relations. 12, 172–180 (2019). https://doi.org/10.24833/2071-8160-2019-5-68-172-180.
10. Nieminen, M., Hyytinen, K.: Future-oriented impact assessment: Supporting strategic

decision-making in complex socio-technical environments. Evaluation. 21, 448–461 (2015). https://doi.org/10.1177/1356389015606540.
11. Mazzucato, M.: The entrepreneurial state. Soundings. 49, 131–142 (2011). https://doi.org/10.3898/136266211798411183.
12. Gelles, D., Yaffe-Bellany, D.: Shareholder Value Is No Longer Everything, Top C.E.O.s Say. New York Times. Business (2019).
13. Schot, J., Steinmueller, W.E.: Three frames for innovation policy: R&D, systems of innovation and transformative change. Res. Policy. 47, 1554–1567 (2018). https://doi.org/10.1016/j.respol.2018.08.011.
14. Walsh, P.P., Murphy, E., Horan, D.: The role of science, technology and innovation in the UN 2030 agenda. Technol. Forecast. Soc. Change. 154, (2020). https://doi.org/10.1016/j.techfore.2020.119957.
15. Scheyvens, R., Banks, G., Hughes, E.: The Private Sector and the SDGs: The Need to Move Beyond 'Business as Usual.' Sustain. Dev. 24, 371–382 (2016). https://doi.org/10.1002/sd.1623.
16. Altgilbers, N., Walter, Lo., Moehrle, M.G.: Frugal invention candidates as antecedents of frugal patents — the role of frugal attributes analysed in the medical engineering technology. Int. J. Innov. Manag. 2050082 (2020). https://doi.org/10.1142/S1363919620500826.
17. UN: Transforming Our World: The 2030 Agenda for Sustainable Development; United Nations: New York, NY, USA, 2015. , New York, NY, USA. (2015).
18. Elsevier: Sustainability Science in a Global Landscape. (2015).
19. Jia, Q., Wei, L., Li, X.: Visualizing sustainability research in business and management (1990-2019) and emerging topics: A large-scale bibliometric analysis. Sustain. 11, (2019). https://doi.org/10.3390/su11205596.
20. Vatananan-Thesenvitz, R., Schaller, A.A., Shannon, R.: A bibliometric review of the knowledge base for innovation in sustainable development. Sustain. 11, 1–22 (2019). https://doi.org/10.3390/su11205783.
21. UNSDG: sustainable development report. (2019).
22. Mongeon, P., Paul-Hus, A.: The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics. 106, 213–228 (2016). https://doi.org/10.1007/s11192-015-1765-5.
23. Jayabalasingham, B., Boverhof, R., Agnew, K., Klein, L.: Identifying research supporting the United Nations Sustainable Development Goals. Mendeley Data. (2019).
24. Sebastiani, F.: Machine Learning in Automated Text Categorization. ACM Comput. Surv. (2001).
25. Weikum, G.: Foundations of statistical natural language processing, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.9.8780, (2002). https://doi.org/10.1145/601858.601867.