ORIGINAL RESEARCH

Validity of Machine Learning in Assessing Large Texts Through Sustainability Indicators

Juan A. García-Esparza[1] · Javier Pardo[2] · Pablo Altaba[1] · Mario Alberich[3]

Abstract

As machine learning becomes more widely used in policy and environmental impact settings, concerns about accuracy and fairness arise. These concerns have piqued the interest of researchers, who have advanced new approaches and theoretical insights to enhance data gathering, treatment and models' training. Nonetheless, few works have looked at the trade-offs between appropriateness and accuracy in indicator evaluation to comprehend how these constraints and approaches may better redound into policymaking and have a more significant impact across culture and sustainability matters for urban governance. This empirical study fulfils this void by researching indicators' accuracy and utilizing algorithmic models to test the benefits of large text-based analysis. Here we describe applied work in which we find affinity and occurrence in indicators trade-offs that result be significant in practice to evaluate large texts. In the study, objectivity and fairness are kept substantially without sacrificing accuracy, explicitly focusing on improving the processing of indicators to be truthfully assessed. This observation is robust when cross-referring indicators and unique words. The empirical results advance a novel form of large text analysis through machine intelligence and refute a widely held belief that artificial intelligence text processing necessitates either accepting a significant reduction in accuracy or fairness.

Keywords Indicators Optimization · Machine Learning · Software Development · Natural Language Processing

✉ Juan A. García-Esparza
  juan.garcia@uji.es

  Javier Pardo
  javier.pardo.sfg@gmail.com

  Pablo Altaba
  paltaba@uji.es

  Mario Alberich
  malberich@gmail.com

[1] School of Technology and Experimental Sciences, Universitat Jaume I, Avda. Vicent Sos Baynat s/n, 12006 Castellón de la Plana, Spain

[2] Department of Business, Universitat Autònoma de Barcelona, Barcelona, Spain

[3] Department of Statistics, Universitat de Barcelona, Barcelona, Spain

 Springer

1 Introducción

Over the past decades, the development of indicators has greatly matured with several theoretical and empirical scholarly works focused on sustainable urban development (Bienvenido-Huertas et al., 2020; Kubiszewski et al., 2022; Ruan & Yan, 2022; Verma & Raghubanshi, 2018). Research through indicators has informed an integral part of urban governance and sustainability by signalling the relevant aspects of a society or a place (Elgert, 2016; Sharifi, 2020). The number of existing indicators for analysis is constantly increasing (La Rosa et al., 2016; Jain & Tiwari, 2017; Mapar et al., 2017; Borsekova et al., 2018; Dawodu et al., 2018; Hatakeyama, 2018; Ameen & Mourshed, 2019; Frare et al., 2020; among others). Indicators, particularly in urban studies, help to systematize and categorize data (Akuraju et al., 2020). Nonetheless, any set of indicators may seem rigid and unable to reflect on specific complexities (Holden, 2013). However, the importance of measuring sustainable aspects through indicators lies in providing an accurate and summarised analysis of factors when dealing with textual data (De Sherbinin et al., 2013).

An underpinning principle for urban sustainability is to place society at the core of valuation processes. Culture valuing and enhancement is a social product that comes with development processes (Bandarin, 2019; ICCROM, 2015; García-Esparza & Altaba, 2022; Jones & Leech. 2015;). The New Urban Agenda (NUA) (United Nations, 2017; UNESCO, 2017) established the reasons for a new urban culture approach in which conceptual transitions are still underway. Cultural ecosystems are perceived as a system of values intrinsically linked to the social, environmental, and economic dimensions of sustainability. The implications of this are not yet fully analysed and will mark the nature of the cultural realm in the twenty-first century (García-Esparza, 2022).

Labadi & Logan (2016) already exposed the need for culture to reduce poverty, mitigate social inequalities, and increase security and health. ICOMOS (2017, 2019) endorsed this approach with an *Action Plan* for cultural heritage and the UN Sustainable Development Goals (SDGs) and a later *Concept Note* as policy guidance for implementing the Action Plan. The Plan and Note explicitly recommended linking culture and sustainable local socio-economic development by ensuring that all four spheres contribute to sustainable development. This framework outlines how critical it is to put culture and individuals in the best context to leverage policies and interventions and how necessary it is to analyze this minimizing potential biases.

Artificial intelligence (AI) has proved to be helpful in a holistic set of tasks related to text mining. Digital processing may serve to comprehend the importance of indicators' composition and versatility, for example, in examining their practical impact according to the polyvalent terms they contain (Sciandra et al., 2021). Through AI, indicators would reflect greater leverage policies and principles that address pressing societal challenges where social and cultural determinants (defined as the preconditions of places and people) account for disadvantages and inequality (Guitton, 2020; Chen et al., 2018; Ramos et al., 2018). Through Machine Learning (ML), the analysis of multiple parametres such as basic needs, access to essential services, accesibility, housing adequacy, environmental pollution, access to green areas, or well-being, are correlated to present directions for future work to leverage synergies in machine learning and text analysis(Mhasawade et al., 2021; Yeung & Fernandes, 2022).

To better understand the elements impacting society's functioning, empirical and traditional statistical methods such as principal component analysis (PCA), clustering methods, regression, and other linear approaches have been employed previously (Rivera, 2014). ML has successfully overcome the limitations of statistical approaches. These advanced analytics are known to yield greater or at least equal accuracy results compared to previous approaches (Lima et al., 2015; Shortridge et al., 2016). Besides, ML techniques have several advantages that include the capacity to deal with data of various types, structures, and quantities (i.e., big data) (Molnar, 2019; Ren et al., 2020; Viana et al., 2021).

ML models have been successfully applied to date in many science studies (Rivera et al., 2014; Schober et al., 2018). However, in this study, rather than doing a "bag of words" (TF-IDF) search (Park & Okudan Kremer, 2017), what the model does is to encode words in vectors to evaluate indicators' distance in terms of affinity and occurrence. In this way, the study explores the use of an ML model coupled with a test of different algorithms to increase understanding of how indicators' composition and preconditions can deal better with analytical challenges and thus provide novel insights into artificial text assessment.

Current challenges require considering indicators as a network formed by linked categories that interact and merge spatiality (Egilmez et al., 2015; Phillis et al., 2017; Zinatizadeh et al., 2017). Understanding indicators' connections and their application to texts facilitate methodological AI developments (Spadon et al., 2019). In this regard, indicators help problem selection and formulation for judging processes and outputs (Akhanova et al., 2020; Dornelles et al., 2020). Context-appropriate indicators are useful at many scales; however, they require prior work on limitations in data collection, adaptability to problem selection and formulation (Valencia et al., 2019). To do so, Machine Intelligence evaluates data and helps researchers understand how, after sorting indicators by range and affinity, with particular attention to the cultural realm of indicators, words' occurrence explains affinity and indicators' effectiveness in the analysis of large texts.

The research objectives are twofold. On the one hand, this study aims to understand the relevance of indicators and unique words through textual analysis. This objective pretends to ease measurement and incorporate complex social determinants in AI models. On the other hand, another objective is to build up reliable algorithmic assessments to trace indicators' composition and interrelation patterns to better assess and predict texts' compositions. These objectives may go beyond to what a specific field of enquiry comprehends, and invades alien areas of knowledge, looking for more plural, multidisciplinary and integrated forms of analysis that help scientists programme and apply machine learning processes.

Researchers have gathered 1082 indicators from previous scholarly works. Indicators are divided into 798 general indicators (Annex A, spreadsheet 1) that cover the environmental, social, and economic spheres of sustainability and 284 cultural indicators (Annex A, spreadsheet 2) that cover the fourth dimension. Behind the subdivision, researchers intend to cross-refer them to comprehend how the model works towards affinity and occurrence. To what extent can we trust these general and cultural indicators to rely on each other and summarise and understand content-related texts? Are cultural indicators more specific than their general counterparts? If so, to what extent are they?

Through a machine intelligence assessment of indicators, researchers will answer these questions and apply the results to the analysis of the Agenda 2030. The research process will be split into four phases, documented and explained in detail in the Methodology, Results section and Annexes A to D. The Method section outlines the entire process to convey

Table 1 Selected journals and articles for the general indicators

Rank	Journal	Articles	Selected articles
Regional and Urban Planning	Technological Forecasting and Social Change	0	0
	Landscape and Urban Planning	5	1
	Journal of the American Planning Association	0	0
	Computers Environment and Urban Systems	7	1
	Sustainable Development	2	2
	Journal of Rural Studies	0	0
	Regional Studies	0	0
	TOTAL	**14**	**4**
Urban Studies	Cities	8	4
	Habitat International	7	3
	Urban Forestry & Urban Greening	4	1
	Journal of Urban Technology	2	0
	Environment and Urbanization	3	0
	Journal of Planning Education and Research	1	0
	Urban Geography	0	0
	International Journal of Urban and Regional Research	0	0
	TOTAL	**25**	**8**
Others	Ecological Indicators	47	11
	Sustainable Cities and Society	14	7
	TOTAL	**61**	**18**
TOTAL ARTICLES		**100**	**30**

the project's structure efficiently. Afterwards, the Results section documents the different phases according to the outputs of the previous one.

2 Methodology

2.1 Data Collection Indicators' Obtention and Classification

Researchers elaborated the list of indicators using the Web of Science (WoS) search engines. As a result, journals were selected within the first quartile of the Regional and Urban Planning and Urban Studies categories. In addition, the journals Ecological Indicators and Sustainable Cities and Society, not classified within those categories, were also considered for their connection with the scope of this study.

The search for indicators in WoS database journals followed these criteria: the title must include the word "indicator", keywords as part of the topic must be "urban" or "city-cities", while the publishing period was limited to the period 2015–2020. From 100 papers, we selected four articles from the Regional and Urban Planning category, eight articles from Urban Studies, and 18 articles from the two journals belonging to Environmental science and Green and sustainable science and technology categories (Table 1, Annex A). The articles

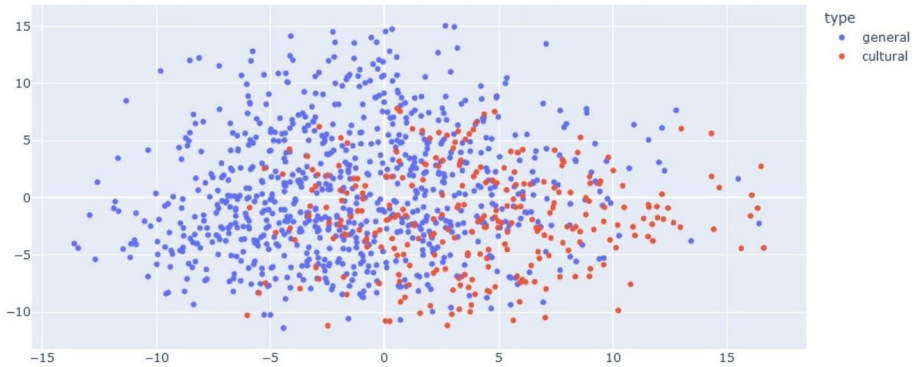2 Components PCA Visualization for indicator type

Fig. 1 Two components PCA. Visualisation for indicator type. General indicators (blue), Cultural indicators (red)
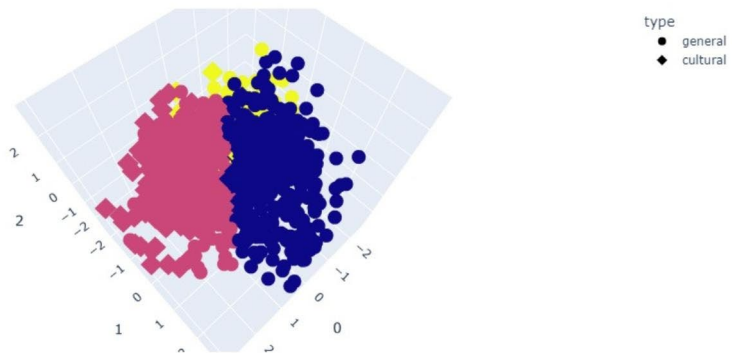
3 Components PCA and 3 Clusters visualization

Fig. 2 Image of the Indicators Match Visualizer, website version

provided a total of 855 indicators, and we classified them into dimensions and categories according to their nature (Andries et al., 2019; Huovila et al., 2019; Sowińska-Świerkosz, 2017). Researchers discarded some indicators as they were outside the scope of the study. The final amount reached 798 evaluable indicators.

The search of cultural indicators was carried out using the WoS search engines as well and based on the following criteria: the title must include the word "cultural indicator", keywords as part of the topic must be, "framework", while the publishing period was limited to the period 2015–2020. As a result, we obtained 27 articles. Of these articles, a total of 359 indicators were categorised, but 75 were discarded due to repetition or because they were outside the scope of the article, resulting in 284 indicators.

Table 2 Selection of best matches

	indicator	closest_indicator	similarity
1	Access to public transport	Access to public transportation	0.998026
3	Renewable Energy Production	Renewable energy	0.981198
5	Google hits for the string city name & climate change & flood (hits per million inhabitants)	Google hits for the string city name & climate change & sea level rise (hits per million inhabitants)	0.977662
7	Water consumption	Water use	0.96699
9	Proportion of single-parent families;	Proportion of households that are lone-parent households	0.962271

Table 3 Full and batch test of Housing, Marginalized Communities & Inequality text

	top_match_0	top_match_1	top_match_2	similarity_0	similarity_1	similarity_2
Full	Housing quality (area per capita), informal housing and slum reduction	Using smart solutions to enhance accessibility to services and amenities	Quick response system to municipal Health, safety and enivironmental problems	0.630492	0.577984	0.513161
Batch	Affordable housing	Accessibility for disadvantaged	Adaptation for social inclusion	1.64125	1.5291	1.24735

2.2 Training Set and Data Processing

2.2.1 Indicators Cleaning

After indicators' collection and classification, the heterogeneous lexical composition in some cases, the inclusion of non-alphanumeric characters in others, the specificity of others containing dates, and the eventual inclusion of non-English terms or unknown encoding symbols make indicators not purely objective. Therefore, if not appropriately cleaned the model application to analyse appropriateness and relevance would have led to errors and misunderstandings. The cleaning process consists of three steps. First, remove wrong characters such as commas, semicolons, dots, quotation marks, etc. Second, to translate non-English words and indicators. And third, to discard duplicates of both types of indicators. Researchers use RegEx, a popular tool that allows modifying (including *replacing* and *removing*) characters through standard encoding Python methods to work with Latin encoding (ISO-8859-1).

2.2.2 Model Selection, Fine Tuning and Indicators Matching

The objective of this stage of the research is to select an appropriate algorithmic model to perform text similarity. This research means to filter and cross-refer sustainability and culture-related topics, fine-tune them to have the most accurate result, and then to match indicators and large texts. When working with Natural Language Processing and text simi-

larity, algorithms count the words present in a sentence and check if they are present in other sentences under comparison, so the more words exist in both texts, the more similar they will be. The most well-known method for doing so is the TF-IDF Vectorizer. The researchers apply a pre-trained transformer (word embeddings) from HuggingFace called *all-roberta-large-v1*.

With the algorithmic approach, the researchers want to retrieve an interpretation of relationships based on an average of biases. Using the algorithm, the linkage between general and cultural indicators and vice versa is analysed employing similarity indices to their counterparts. The relevant question the researchers want to solve is whether the model is consistent in terms of affinity between indicators and their keywords, and at a final instance, with more extensive texts.

The assessment of the model includes the following steps (Annex B):

- Model Exploration: selection of an appropriate algorithmic model,

  – Analysis of Components PCA Visualisation for clusters of indicators,

- Validation: clustering method to represent indicators,

  – KMeans Optimization (algorithm) to obtain the best number of clusters,
  – Elbow Method to justify the differences between clusters,
  – Chi-Square contingency: checks the hypothesis behind clustering,

- Fine-Tuning (optimization): of components' number for the encoding matrix,

  – Analysis of variance according to the Nº of components,

- Error Analysis: of indicators matching through *cosine similarity*,

  – Lexical variations, long sentences limitation and synonym recognition,

- Indicators Matching: test of best and worst matches to evaluate the model,

  – Soft Voting Classifier for the model.

2.2.3 Model Application Word Embedding Limitations

Once the model is fine-tuned and ready to work, the authors realise that the transformer's performance (word embeddings like Roberta) decreases when lengthier texts pass through it. To avoid this bias, researchers shortcut extensive sentences by reducing the number of words by removing stop-words and not applying the model once but applying it for every sub-sentence defined by all words between dots and commas characters. In this way, researchers reduce the vagueness of the model. The process is most code-intensive but simplistic in procedural terms. The stages are as follows: To remove stop-words, to apply the

model to a vector of sentences (Hadamard Product), and finally, to create a Soft Voting Classifier for the model.

3 Results

3.1 General and Cultural Indicators Matching

One of the initial objectives of the project is to match general with cultural indicators. The model of indicators matching exports two CSV files. One contains the indicator we want to match and the top 5 similar indicators from the other type (indicator_matches.csv). The other file contains the encoded matrix of the clean indicators to be used directly in the next phase of the project, where the model and the PCA will be initialized again then (See Annex C).

3.2 Indicators Match Visualizer¶

The Software exposes graphically and interactively the 3D visualization of indicators matching (Fig. 11). With the visualizer, it is possible to navigate through the spatial distribution of points-indicators and observe the top 5 matches for the selected indicator and their positions in the three components of PCA. Note that this visualizer is not available through the HTML file as the filter cannot be embedded as it is. Therefore, the graphic source is only available in the software version.

3.3 Model Application for Indicators and Large Texts

Training the model against long sentences entails checking how it behaves with small samples of texts extracted from some commitments of the Agenda 2030. When these texts are assessed, the top indicator similarity value declines dramatically compared to the previous assessments with indicators (median of 0.74) due to the length constraints. But even though this happens, the model can still recognise the essential meaning of the sentences overall even if related to some ambiguous indicators (e.g., Health, safety and environmental initiatives and innovations at municipality level).

After the first trial application, researchers assess Full and Batch models. This test evaluates the base model (Full) vs. the soft-voting classifier (Batch). The soft-voting classifier takes every sentence, splits them in between every coma and dot, and applies the model separately.Here, it is unclear which model, the batch or the full one, performs better (Annex B). Therefore, the improvement is unclear when applying the batch model over the full model, but it will be worth trying when processing large documents. As a result of this analysis, the software retrieves a Python file (*module*) containing the necessary functions to work with the model at any time. This will be used in the last step, the PDF Reader Software.

3.4 Test of Indicators Affinity and Unique Words Occurrence

Following the organisation and classification of indicators and the subsequent cleaning process, researchers employed the results of the indicators' matching (Annex C) to analyse the

Table 4 Occurrence of unique words among indicators by affinity ranks. Legend: (ind) indicator

	Cultural	Heritage	Protection	Preservation	Conservation	Building	Landscape
Rank 1	92.31%	38.46%	7.69%	23.08%	7.69%	15.38%	7.69%
13 ind	12 ind	5 ind	1 ind	3 ind	1 ind	2 ind	1 ind
Rank 2	82.86%	62.86%	28.57%	25.71%	8.57%	8.57%	8.57%
35 ind	29 ind	22 ind	10 ind	9 ind	3 ind	3 ind	3 ind
Rank 3	96.00%	76.00%	44.00%	20.00%	12.00%	4.00%	12.00%
25 ind	24 ind	19 ind	11 ind	5 ind	3 ind	1 ind	3 ind
Rank 4	100.00%	84.21%	47.37%	26.32%	15.79%	5.26%	5.26%
19 ind	19 ind	16 ind	9 ind	5 ind	3 ind	1 ind	1 ind
Rank 5	100.00%	100.00%	62.50%	37.50%	12.50%	0.00%	0.00%
8 ind	8 ind	8 ind	5 ind	3 ind	1 ind	0 ind	0 ind

occurrence of key internal words in the most recurring indicators (cultural) and their counterparts (general). The CSV file comprises a first column containing the cultural indicators to which the resemblance is sought with general indicators. The next five columns contain the general indicators with the highest affinity to the cultural one. The next five columns show the quantitative similarity values of each general indicator with respect to the cultural one.

This test of unique words' matching redounds on understanding how to build or complement more complete indicators or indicators that serve different purposes when effectively applied to specific fields of inquiry. To do the test, researchers analyse the cultural indicators by conducting a quantitative weighting of similar unique words. From Annex C, researchers organised the similarity values for the 284 cultural indicators by ranks, from lowest affinity to highest affinity with the general indicators. Researchers establish Rank 1 for the lowest values of affinity ($\geq 0.45 < 0.55$), containing a total of 13 indicators. Rank 2 ($\geq 0.55 < 0.65$) contains 35 indicators, Rank 3 ($\geq 0.65 < 0.75$) 25 indicators., Rank 4 ($\geq 0.75 < 0.85$) 19 indicators, and Rank 5 ($\geq 0.85 < 0.95$) 8 indicators; see Table 4.

Within the test of affinity between indicators, the occurrence of unique words within each indicator was sought to understand better how the Roberta algorithm works and the possible reasons for the affinity between indicators. The unique words that were searched for, both in the cultural and general affine indicators within the established ranges, were: *Cultural, Heritage, Protection, Preservation, Conservation, Building*, and *Landscape*. Thus, Table 4 reflects whether the word is ever mentioned in the indicators of each rank, either in the cultural indicator or in the general indicators and quantified in terms of percentage of the total number of times it could be mentioned, only once for each set of indicators within ranks.

Therefore, and as an example, the word *Cultural*, within rank 1, is included in twelve indicators, of which two are cultural and ten generals. Similarly, the word *Heritage* is explicitly referred to in five rank 1 indicators, one cultural and four generals. This single word analysis returns a relevant result for the word *Cultural* in all affinity ranks; in practically all of them it achieves between a 90 and a 100% of occurrence.

In summary, the table returns an intensity in percentage of occurrence of each word (column) correlated with the affinity ranks of indicators (rows). Thus, the highest presence of these unique words among cultural and general indicators occurs in ranks four (4) and five (5), which contain the twenty-seven (27) indicators with the highest affinity ($\geq 0.75 < 0.95$). Among the unique words selected for the affinity analysis, *Cultural* and *Heritage* stand out.

Table 5 Another example of Agenda 2030 text, top matches, and similarity (Annex D)

Sentence	top_match_1	top_match_2	top_match_3	top_match_4	top_match_5
62. This Agenda, including the Sustainable Development Goals, can be met within the framework of a revitalized Global Partnership for Sustainable Development, (…), which is an integral part of the 2030 Agenda for Sustainable Development.	Events successfully implemented and evaluated	Management and action plan	Shared architecture for multi-level governance and inter-agency collaboration	Participatory master plan	Strategies and infrastructure for timely data communication, sharing, and reporting
	similarity_0	similarity_1	similarity_2	similarity_3	similarity_4
	3,4613	3,3029	3,0454	3,0413	2,3403

This is because they are the most recurrent in searching for the affinity between cultural indicators and their generic counterparts.

3.5 PDF Reader Software

This phase of the research analyses the model's learning through significant texts. Researchers developed a PDF file reader to extract the relevant information from the desired files and apply the model. By correlating the text with general and cultural indicators, the model selects those prevalent or with the highest affinity to the document. The text employed as an example is the NUA. This text is of particular interest for the researchers to ascertain the extent to which the text considers culture and sustainability and whether the software can summarise and match documents properly.

The PDF Reader software employs the PDF parser of the tika library. The software is tested with the 2030 Agenda (using *RegEx*). Once the PDF reader is analysed, researchers conduct a series of checkings with some pieces of text to ascertain its accuracy. As this is the final step, the PDF Reader software exports a CSV file with a sentence from the text for every row and the top 5 indicators, general and cultural merged together (Annex D).

One sample of the rendered text is: "One in which development and the application of technology are climate sensitive, respect biodiversity and are resilient. One in which humanity lives in harmony with nature and in which wildlife and other living species are protected". The Reader returns five top matches. Top match 1: Biodiversity and habitat
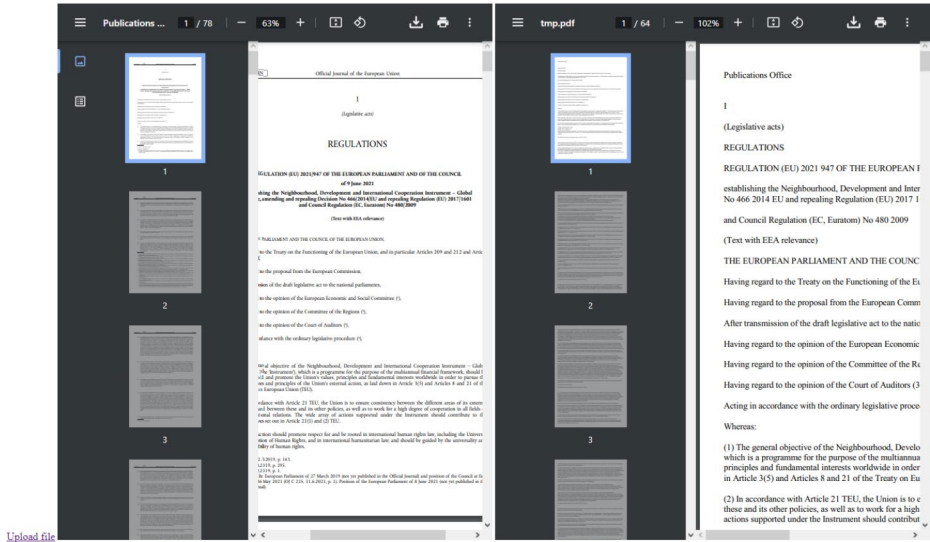
Home

Fig. 3 Indicators detector software. Developer: Javier Pardo.

conservation/restoration; scoring 1.2745 in similarity. Top match 2: Nature conservation and biodiversity, scores 1.26302. Top match 3: Areas of protected natural spaces, scores 1.24429. Top match 4: Protection of culturally important species, scores 1.19592. Top match 5: Well-preserved village, scores 1.1812.

In the same line as outlined in the affinity of words section, the software finds an affinity between culture and sustainability dimensions in a manner that is not that obvious when simply reading the text. In this case, while Top match 4 explicitly refers to the intersection of environmental and cultural dimensions of indicators when referring to the protection of culturally important species, Top match 5 links the text to culture (villages) preservation. Therefore, the software comes to demonstrate the pertinence of the analysis and the importance of retrieving appropriate indicators that help comprehend openly and sensitively the meaning behind large texts.

In brief, matches need further attention since the order of some sentences is switched. However, the exact position of sentences does not affect the results of this project, affinity and occurrence between indicators, words, and texts. Overall, the software works very well as it can extract all the text without almost any error. Minor errors in text extraction have to do with the random structure of pdf documents and the internal codings employed to edit them.

The indicators detector software is developed under license X11 MIT. The premise for its elaboration was to be available not just to the scientific community but to everyone; however, its development is taking longer than expected. The idea is that anyone can input a PDF file and see the result of the PDF Parser together with the final indicator detected for each phrase (see Fig. 3). Availability and online documentation are expected in 2023. The software required is Python and Django, and the programming language is Python.

Regarding potential final users, no level of software development is required to apply this software and its methodology. It only requires intermediate Python and basic knowledge about machine learning, understanding the concepts of supervised-unsupervised learning and how they are trained, the expected results, and where to find material to help users do it by themselves. Nonetheless, this has been the case for developing the full software. If the model can be applied directly with no manipulation, which is not recommended, following a tutorial may be enough; it depends on how far users want to get.

4 Conclusion

Researchers examine and summarize areas where ML innovation might synergize with, advance, and improve on research and practice in the field of text analysis through a discussion of sustainable indicators principles. The authors expose main areas of challenge, such as the data used, the methodologies developed, and the questions posed, all of which are critical in the realm of sustainable development. These issues involve obtaining crucial questions for the measurement and incorporation of social determinants in AI models, as well as a reliable algorithmic assessment to comprehend indicators' composition and to avoid eventual biases that affect data evaluation.

The authors show how algorithmic analysis and unique words tests must be addressed in the context of the data and systems in which they are used, demonstrating how they could otherwise perpetuate or enhance culturerelated issues. When discussing ML's social responsibility in terms of urban governance and fair AI procedures in socio-cultural terms, these have to do with the límits of ethics. These ideas for shaping data, measures, and questions of ML efforts in indicator analysis should be drawn from domains such as public and population rights, which are central to the study of sustainability.

The analysis reflects unique and ambivalent core terms within indicators, and in turn, it explains how fine-tuned indicators help match text analysis accurately through ML. The study exposes an interconnection between the different dimensions of sustainability and culture. From the analysis of unique words, both indicator's types, cultural and general, are strongly linked to the social dimension of sustainability and, less prominently, to the environmental and economic ones.

Although studies of this type have limitations of analysis and representation mainly, the authors prove that algorithmic models to assess eventual text biases by using large datasets mined from existing literature could improve accuracy in addressing the critical drivers of social change and justice. This approach is important as academics attempt to improve the public realm through policies and processes for all people in the face of changing climates, priorities, and data. In conclusion, the goal of this position is to stimulate the ML community's imaginary and stimulate discussion about the types of data and problems researchers address when thinking about AI applied to population concerns, particularly to those of trustworthy AI for democratic futures.

Further development of this research may include a greater systematization of data processing. From the selection of indicators until the model assessment, a more complete and systematized machine learning model with better prediction of errors and improvement abilities may produce a more comprehensive and complex analysis. Therefore, this is something to be addressed in future research. The methodology developed to obtain this software

can also be applied to almost anything. In this case, it is used for a detailed analysis of sustainability and urban-based natural language. Still, it can be applied to images (e.g., to find similar images, google search, etc.), sound (e.g., to match the actual song with a saved song, shazam, etc.) and everything as long as users are capable of describing it adequately.

References

Akhanova, G., Nadeem, A., Kim, J. R., & Azhar, S. (2020). A multi-criteria decision-making framework for building sustainability assessment in Kazakhstan. *Sust Cities Soc*, *52*, 101842.

Akuraju, V., Pradhan, P., Haase, D., Kropp, J. P., & Rybski, D. (2020). Relating SDG11 indicators and urban scaling–An exploratory study. *Sust Cities Soc*, *52*, 101853.

Ameen, R. F. M., & Mourshed, M. (2019). Urban sustainability assessment framework development: the ranking and weighting of sustainability indicators using analytic hierarchy process. *Sust Cities Soc*, *44*, 356–366.

Andries, A., Morse, S., Murphy, R., Lynch, J., Woolliams, E., & Fonweban, J. (2019). Translation of Earth observation data into sustainable development indicators: an analytical framework. *Sustainable Development*, *27*(3), 366–376.

Bandarin, F. (2019). "Reshaping urban Conservation". In F. Bandarin (Ed.), *Reshaping Urban Conservation. The historic Urban Landscape Approach in Action* (pp. 3–20). Singapore: Springer.

Bienvenido-Huertas, et al. (2020). Comparison of artificial intelligence algorithms to estimate sustainability indicators. *Sust Cities Soc*, *63*, 102430.

Borsekova, K., Koróny, S., Vaňová, A., & Vitálišová, K. (2018). Functionality between the size and indicators of smart cities: a research challenge with policy implications. *Cities*, *78*, 17–26.

Chen, J. D., Zhou, X. J., & Tang, X. J. (2018). An empirical feasibility study of societal risk classification toward BBS posts. *Journal Of Systems Science And Systems Engineering*, *27*(6), 709–726.

Dawodu, A., Cheshmehzangi, A., & Akinwolemiwa, B. (2018). The systematic selection of headline sustainable indicators for the development of future neighbourhood sustainability assessment tools for Africa. *Sust Cities Soc*, *41*, 760–776.

De Sherbinin, A., Reuben, A., Levy, M. A., & Johnson, L. (2013). *Indicators in practice: how environmental indicators are being used in policy and management contexts*. New York; Yale Centre for Environment Law & Policy, New Haven: Center for International Earth Science Information Network.

Dornelles, L. D. L., Gandolfi, F., Mercader-Moyano, P., & Mosquera-Adell, E. (2020). Place and memory indicator: methodology for the formulation of a qualitative indicator, named place and memory, with the intent of contributing to previous works of intervention and restoration of heritage spaces and buildings, in the aspect of sustainability. *Sust Cities Soc*, *54*, 101985.

Egilmez, G., Gumus, S., & Kucukvar, M. (2015). Environmental sustainability benchmarking of the US and Canada metropoles: an expert judgment-based multi-criteria decision-making approach. *Cities*, *42*, 31–41.

Elgert, L. (2016). The double edge of cutting edge: explaining adoption and nonadoption of the STAR rating system and insights for sustainability indicators. *Ecol Indic*, *67*, 556–564.

Frare, M. B., Clauberg, A. P., Sehnem, S., Campos, L. M., & Spuldaro, J. (2020). Toward a sustainable development indicators system for small municipalities. *Sustainable Development*, *28*, 1148–1167.

García-Esparza, J.A. (2022). Urban Scene Protection and Unconventional Practices – Contemporary Landscapes in World Heritage Cities of Spain. Land 11(3): 324. https://doi.org/10.3390/land11030324

García-Esparza, J.A., & Altaba, P. (2022). Identifying habitation patterns in World Heritage areas through social media and open datasets. Urban Geography, https://doi.org/10.1080/02723638.2022.2140971

Guitton, M. J. (2020). Cybersecurity, social engineering, artificial intelligence, technological addictions: societal challenges for the coming decade. *Computers In Human Behavior*, *107*, 106307.

Hatakeyama, T. (2018). Sustainable development indicators: conceptual frameworks of comparative indicators sets for local administrations in Japan. *Sustainable Development*, *26*(6), 683–690.

Holden, M. (2013). Sustainability indicator systems within urban governance: usability analysis of sustainability indicator systems as boundary objects. *Ecol Indic*, *32*, 89–96.

Huovila, A., Bosch, P., & Airaksinen, M. (2019). Comparative analysis of standardized indicators for Smart sustainable cities: what indicators and standards to use and when? *Cities*, *89*, 141–153.

ICCROM. (2015). *People-Centred Approaches to the conservation of Cultural Heritage: living Heritage*. Rome: ICCROM.

ICOMOS (2019). Concept Note: Producing the 'ICOMOS Policy Guidance on Cultural Heritage for the Sustainable Development Goals'. Paris: ICOMOS. Available at: https://www.icomos.org/images/PG_Concept_Note_Eng.pdf (accessed 25 November 2021).

Jain, D., & Tiwari, G. (2017). Sustainable mobility indicators for indian cities: selection methodology and application. *Ecol Indic*, *79*, 310–322.

Jones, S., & Leech, S. (2015). *Valuing the historic environment: a critical review of existing approaches to social value [AHRC Cultural Value Project Report]*. Manchester: University of Manchester.

Kubiszewski, I., Mulder, K., Jarvis, D., & Costanza, R. (2022). Toward better measurement of sustainable development and wellbeing: a small number of SDG indicators reliably predict life satisfaction. *Sustainable Development*, *30*(1), 139–148.

Labadi, S., & Logan, W. (2016). "Approaches to urban heritage, development and sustainability". In W. Logan (Ed.), *Urban heritage, development and sustainability. International framework, national and local governance* (pp. 1–20). Abington & New York: Routledge.

La Rosa, D., Spyra, M., & Inostroza, L. (2016). Indicators of Cultural Ecosystem Services for urban planning: a review. *Ecol Indic*, *61*, 74–89.

Lima, A., Cannon, A. J., & Hsieh, W. W. (2015). Nonlinear regression in environmental sciences using extreme learning machines: a comparative evaluation. *Environmental Modelling And Software*, *73*, 175–188.

Mapar, M., Jafari, M. J., Mansouri, N., Arjmandi, R., Azizinejad, R., & Ramos, T. B. (2017). Sustainability indicators for municipalities of megacities: integrating health, safety and environmental performance. *Ecol Indic*, *83*, 271–291.

Mhasawade, V., Zhao, Y., & Chunara, R. (2021). Machine learning and algorithmic fairness in public and population health. *Nat Mach Intell*, *3*, 659–666.

Molnar, C. (2019). Interpretable Machine Learning. A Guide for Making Black Box Models Explainable. Available online: https://christophm.github.io/interpretable-ml-book/ (accessed on 16 February 2022).

Park, K., & Kremer, O., G.E (2017). Text mining-based categorization and user perspective analysis ofenvironmental sustainability indicators for manufacturing and servicesystems. *Ecol Indic*, *72*, 803–820.

Phillis, Y. A., Kouikoglou, V. S., & Verdugo, C. (2017). Urban sustainability assessment and ranking of cities. *Computers, Environment And Urban Systems*, *64*, 254–265.

Ramos, S. B., de Paula Silva, J., Bolela, C. A., & de Andrade, M. (2018). Prediction of Human Development from Environmental indicators. *Social Indicators Research*, *138*, 467–477.

Ren, X., Mi, Z., & Georgopoulos, P. G. (2020). Comparison of Machine Learning and Land Use Regression for fine scale spatiotemporal estimation of ambient air pollution: modeling ozone concentrations across the contiguous United States. *Environment International*, *142*, 105827.

Rivera, S. J., Minsker, B. S., Work, D. B., & Roth, D. (2014). A text mining framework for advancing sustainability indicators. *Environmental Modelling And Software*, *62*, 128–138.

Ruan, F. L., & Yan, L. (2022). Challenges facing indicators to become a universal language for sustainable urban development. *Sustainable Development*, *30*(1), 41–57.

Schober, A., Kittel, C., Baumgartner, R. J., & Füllsack, M. (2018). Identifying dominant topics appearing in the Journal of Cleaner Production. *Journal Of Cleaner Production*, *190*, 160–168.

Sciandra, A., Surian, A., & Finos, L. (2021). Supervised machine learning methods to disclose action and information in "U.N. 2030 Agenda" Social Media Data. *Social Indicators Research*, *156*, 689–699.

Sharifi, A. (2020). A typology of smart city assessment tools and indicator sets. *Sust Cities Soc*, *53*, 101936.

Shortridge, J. E., Guikema, S. D., & Zaitchik, B. F. (2016). Machine learning methods for empirical stream-flow simulation: a comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds. *Hydrology And Earth System Sciences*, *20*, 2611–2628.

Spadon, G., de Carvalho, A. C., Rodrigues-Jr, J. F., & Alves, L. G. (2019). Reconstructing commuters network using machine learning and urban indicators. *Scientific Reports*, *9*(1), 1–13.

Sowińska-Świerkosz, B. (2017). Review of cultural heritage indicators related to landscape: types, categorisation schemes and their usefulness in quality assessment. *Ecol Indic*, *81*, 526–542.

UNESCO (2017). UNESCO moving forward the 2030 Agenda for Sustainable Development. UNESCO: Paris. Available at: https://en.unesco.org/creativity/sites/creativity/files/247785en.pdf (accessed 25 November 2021).

United Nations. (2017). *New Urban Agenda. Habitat III*. UN: Quito.

Valencia, S. C., Simon, D., Croese, S., Nordqvist, J., Oloko, M., Sharma, T., Buck, T., N., and, & Versace, I. (2019). Adapting the Sustainable Development Goals and the New Urban Agenda to the city level: initial reflections from a comparative research project. *Int J Urban Sustain Dev*, *11*(1), 4–23.

Verma, P., & Raghubanshi, A. S. (2018). Urban sustainability indicators: Challenges and opportunities. *Ecol Indic*, *93*, 282–291.

Viana, C. M., Santos, M., Freire, D., Abrantes, P., & Rocha, J. (2021). Evaluation of the factors explaining the use of agricultural land: a machine learning and model-agnostic approach. *Ecol Indic*, *131*, 108200.

Yeung, R. C., & Fernandes, M. A. (2022). Machine learning to detect invalid text responses: validation and comparison to existing detection methods. *Behavior Research Methods*. doi: https://doi.org/10.3758/s13428-022-01801-y.

Zinatizadeh, S., Azmi, A., Monavari, S. M., & Sobhanardakani, S. (2017). Evaluation and prediction of sustainability of urban areas: a case study for Kermanshah city, Iran. *Cities*, *66*, 1–9.

Zinatizadeh, S., Azmi, A., Monavari, S.M., and Sobhanardakani, S. (2017). Evaluation and prediction of sustainability of urban areas: A case study for Kermanshah city, Iran, *Cities*, *66*, 1–9.