United Nations | Department of Economic and Social Affairs

DESA Working Paper No. 171
ST/ESA/2020/DWP/171
OCTOBER 2020

What does the United Nations "say" about global agenda? An exploration of trends using natural language processing for machine learning

Authors: Namsuk Kim[*] and Marcelo LaFleur[**]

ABSTRACT

How has the focus of the UN General Assembly changed over time and how well is the global agenda expressed in these documents? This paper presents a proof-of-concept classifier to examine the evolution of the global agenda expressed and observed in words of the UN General Assembly resolutions. Using natural language processing to identify four categories of resolutions — Sustainable Development, Justice and Law, Human Rights, and Peace and Security — the analysis of 3,765 UN GA resolutions from 2007 to 2019 reveals the changing areas of focus of the Member States and, as a result, of the UN Secretariat. Sustainable Development is slowly gaining importance in the language in UN resolutions.

JEL Classification: F53; O19

Keywords: United Nations; General Assembly Resolution; Natural Language Processing; Global Agenda; Sustainable Development

Sustainable Development Goals: 4.7; 16.1; 17.14

*kimnamsuk@un.org; **lafleurm@un.org. The authors, Economic Affairs Officers in the United Nations Department of Economic and Social Affairs, thank Elliott Harris, Roland Mollerus, Nazrul Islam, Hoi Wai Jackie Chen, Matthias Bruckner, Justin Flaminiano, Hiroshi Kawamura, Richard Roehrl, and Lin Yang for their helpful comments. Special thanks go to Bojan Grozdanic for his assistance in making the data from the UN Library available for this study. The opinions expressed in this paper do not necessarily represent the views of the United Nations. For further information, please contact the authors.

CONTENTS

Typesetter: *Nancy Settecasi*

UNITED NATIONS
Department of Economic and Social Affairs
UN Secretariat, 405 East 42nd Street
New York, N.Y. 10017, USA
e-mail: undesa@un.org
https://www.un.org/development/desa/

1 Introduction

"WE the peoples of the United Nations determined... to save succeeding generations from the scourge of war..., and to reaffirm faith in fundamental human rights..., and to establish conditions under which justice and respect for the obligations arising from treaties and other sources of international law can be maintained, and to promote social progress and better standards of life in larger freedom" (United Nations, 1945, Preamble).

The very reason for creating the United Nations was to establish a global agenda for peace, human rights, international law, and development. It represented a paradigm shift from a culture of war to a culture of peace: replacing arms with cooperation and compromise. Member States of the UN recognized that prosperity and peace are indivisible and, to be sustained, need to be shared by all (Al Khalifa, 2007).

Over the past 75 years, remarkable achievements have been made. One of the earliest achievements was to agree on the Universal Declaration of Human Rights in 1948, outlining global standards for human rights. The most recent milestone is the 2030 Agenda for Sustainable Development, in which Member States pledged to work together to eradicate poverty and hunger, protect the planet and foster peace.

Since its creation, the Member States that comprise the UN have provided global leadership in setting an agenda for the global community. It has defined a global agenda and goals to improve peace and security, strengthen human rights, fight poverty, accelerate economic growth, protect the environment, among others. The global agenda set by the UN has evolved over time to reflect the changing needs of the world, the changing understanding of these needs, and to reflect the priorities of the global community.

As with any large organization, the UN has left a record of its work in its official documents. Modern tools allow us to analyze this vast trove of UN documents and describe historical patterns. The objective of this paper is to demonstrate how this analysis can be conducted and how it may be useful to understand the history and the impact of the UN. The present paper accomplishes this objective by identifying categories of UN documents and describing timelines in the evolution of the global agenda expressed and observed in words of the UN General Assembly (GA) resolutions using the most advanced technique on language processing. Section 2 describes the methodology of analyzing the natural language by machine learning techniques. Section 3 presents the results, and section 4 concludes.

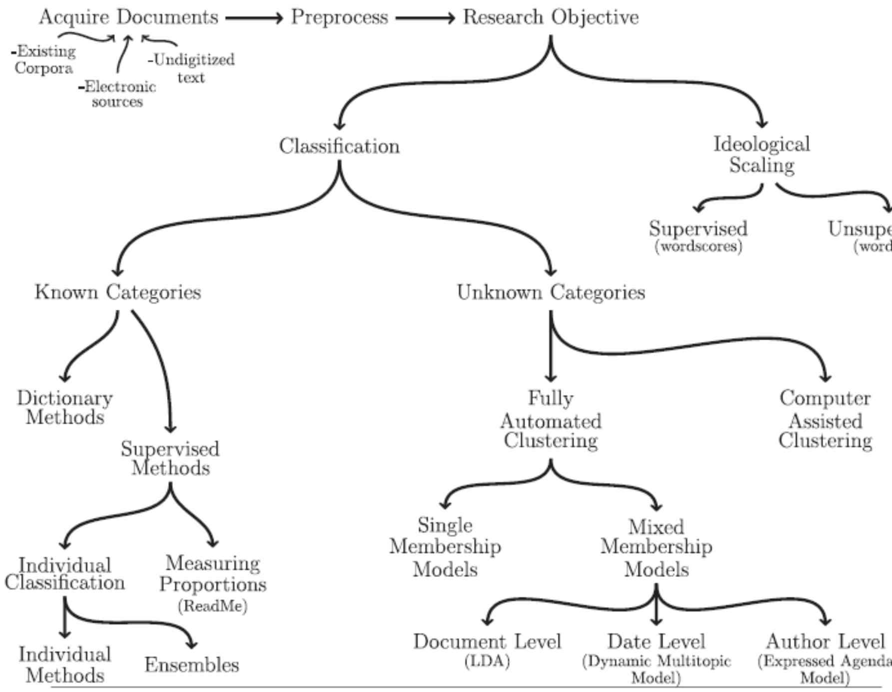
2 Methodologies for analyzing natural language

We use text as data to achieve the research objectives in this paper. The data covers a collection of GA resolutions from session 62 (2007-2008) through session 73 (2018-2019), obtained from the UN Dag Hammarskjöld Library.[1]

Literature on the field of natural language processing has been accumulating fast in recent years. Broadly, there are three groups of approaches to analyze the natural language. For documents with known categories, there are dictionary methods and supervised learning methods. For text without known categories, there are automated clustering methods. Figure 1 shows the overview of the various research approaches.

1 Only sessions 62 through 73 were available for this analysis. Future updates will include earlier GA sessions.

Figure 1
Overview of approaches to analyze text as data



Source: Grimmer and Stewart (2013).

Dictionary methods use the relative frequency of key words to measure the presence of each category in texts (Teddy 2020; Jurafsky and Martin 2009). For instance, Eshbaugh-Soha (2010) and Eshbaugh-Soha and Peake (2010) examined whether the tone of presidential rhetoric affects the tone of news coverage and public opinion, counting positive, neutral and negative words in presidential speeches, media coverage, and public opinion. In this method, a bag of words is a particular representation model used to simplify the contents of a selection of text. The bag of words model omits grammar and word order, but is interested in the number of occurrences of words within the text. Since it is a relatively simple concept, it has been frequently used in research (see Kellstedt 2000; Laver and Garry 2000; Burden and Sanberg 2003; Young and Soroka 2011; Rodman 2020). This method was recently used to extract SDG related metadata from documents, using the official SDG Taxonomy as the dictionary.[2]

Supervised learning methods are used to categorize documents into predetermined sets. The researcher first must have a large number of examples of each category. An algorithm then uses these examples to learn how to sort new documents into the predetermined categories. The basic steps are: (1) the researcher constructs a training set of categorized examples; (2) the algorithm creates a model to associate a given category with the contents of the text; and (3) the researcher validates the model output, correcting any mistakes and repeating step 2 to improve the model (Jurka et al 2012; Blei and McAuliffe 2010). The end result is a classifier that is able to “tag”, or classify new documents. This technique has many applications. Baturo and others (2017), for instance uses this form of supervised learning to create an ideological scaling to identify countries’ political positions in United Nations General Debate speeches.

2 The project is available as the LinkedSDG tool, https://sustainabledevelopment.un.org/LinkedSDGs. The full SDG Taxonomy is available at http://metadata.un.org/sdg/.

One downside of supervised learning methods is their need for a large number of classified examples to inform, or train the algorithm. In some cases, such training data is not available or is too costly to prepare. Pre-classified training data also reflects a subjective judgment that can introduce bias into the model. In such situations, automated clustering, or unsupervised learning methods can be used, letting the algorithm learn underlying features of text without the researchers explicitly imposing categories of interest. Unlike the supervised methods in which users have known categories beforehand, unsupervised learning methods use modeling assumptions and properties of the texts to estimate a set of categories and simultaneously assign documents to those categories.

Unsupervised methods open new possibilities for analyzing vast quantities of unstructured and unlabeled texts. Blei (2012), for instance, uses this method to create clusters of documents, called a "topic model", while Seiermann (2018) and Alschner (2017) use Preferential Trade Agreements text clustering to find text similarity among them. LaFleur (2019) uses an unsupervised topic model following the techniques of Blei (2012) to develop a classifier for each of the 17 Sustainable Development Goals (SDGs), allowing for quantitative analysis of UN publications, speeches, as well as Voluntary National Reviews.[3]

Methodology to analyze UN GA resolutions

In the present paper we are demonstrating a proof-of-concept classifier that can be useful for analyzing the work of the UN as reflected in the GA resolutions. UN GA documents are relatively well defined in terms of category, and therefore are well suited for analysis using the dictionary method and the supervised learning method.

As in any classifier, we must first decide on the classes or categories we wish to identify using this tool. This is determined by two considerations: 1) are the categories a good basis for understanding the UN corpus (research relevance); and 2) are they a viable structure with which to build the training data (data viability).

Research relevance

The GA is organized in six Committees, each with a well-defined area of responsibility:

1. First Committee: Disarmament and International Security;
2. Second Committee: Economic and Financial;
3. Third Committee: Social, Humanitarian and Cultural;
4. Fourth Committee: Special Political and Decolonization;
5. Fifth Committee: Administrative and Budgetary;
6. Sixth Committee: Legal.

Each Committee already reports on its own resolutions and there would be little to gain from replicating this information. Also, the Committees represent an administrative structure (how it works) and are not necessarily a good reflection of the various issues that describe the UN's mission (what it aims to do). The mission of the UN is better represented by the thematic areas of its work and therefore we analyze what the UN aims to do and how well the resolutions reflect this mission.

3 See CDP Subgroup on voluntary national reviews (2019), Voluntary National Reviews Reports – What do they (not) tell us?, *CDP Background Paper, vol. 49, July.*

There are two places where the United Nations describes and reports on what it aims to do across multiple thematic areas: 1) an annual Secretary-General report on the work of the organization, and 2) the UN's public website. The annual Secretary-General report on the work of the organization to the GA includes a summary of the substantive accomplishments and ongoing efforts of the UN.[4] Since the 2012-2013 session of the GA, this report has been organized around eight thematic areas. The UN website also includes a description of the work of the organization ("what we do").[5] This description is meant for the public at large and uses a narrower grouping compared to the annual Secretary-General report. Both of these groupings are listed in table 1.

These thematic categories have a clear research relevance and meet the first of the considerations for building a meaningful classifier mentioned above.

Table 1
Areas of work specified in SG's annual report and on UN.org website

SG Report on the work of the organization	UN.org "What we do" page
Promotion of sustained economic growth and sustainable development	Promote Sustainable Development
Maintenance of international peace and security	Maintain International Peace and Security
Development of Africa	
Promotion and protection of human rights	Protect Human Rights
Effective coordination of humanitarian assistance efforts	Deliver Humanitarian Aid
Promotion of justice and international law	Uphold International Law
Disarmament	
Drug control, crime prevention and combating terrorism	

Source: https://www.un.org/sg/en/content/reports-secretary-general-work-organization; https://www.un.org/en/sections/what-we-do/index.html

Data viability

The second consideration for deciding on the areas of research is data viability. In order to build a well performing model, it is important to have a sufficient number of examples of each category we wish to analyze. That is, we need to build a collection of samples of each category and train a model to recognize the differences between them.

The annual Secretary-General report gives us well defined examples of language that describes each of the categories we wish to analyze. We also include the contents of the "what we do" section of the UN website, which is separated by categories (see Table 1). In addition to these two sources, we include recent UN reports published for each category. For example, on the issue of terrorism, we include the recent Secretary-General report on the UN's global counter-terrorism strategy.[6]

Finally, we include the content of the websites from each UN Department responsible for each of the categories. In the area of sustainable development, for example, we use the relevant text as well as relevant documents available on the website of the Department of Economic and Social Affairs (UN DESA).

Only documents and website contents that clearly fit into one or another category were used to build the model. The dataset used for training the model comprises 211 individual text files, each tagged with one of

4 https://www.un.org/sg/en/content/reports-secretary-general-work-organization

5 https://www.un.org/en/sections/what-we-do/index.html

6 https://www.un.org/en/ga/search/view_doc.asp?symbol=A/72/840

the eight categories. The eight categories were further reduced to just four to improve the ability of the model to correctly predict the categories (see Table 2). These four categories were then used for the steps that follow.

Table 2
Categories used in the final classifier

SG Report on the work of the organization	Final categories
Promotion of sustained economic growth and sustainable development	Sustainable Development
Development of Africa	
Maintenance of international peace and security	Peace, Security and Disarmament
Disarmament	
Promotion and protection of human rights	Human Rights, Humanitarian Affairs
Effective coordination of humanitarian assistance efforts	
Promotion of justice and international law	Justice, law, drugs, crime and terrorism
Drug control, crime prevention and combating terrorism	

Source: Authors' tabulation from https://www.un.org/sg/en/content/reports-secretary-general-work-organization

Building the classifier and testing algorithms

We relied on a traditional labelled classifier following the categories described above. This supervised approach involves labeling examples of each category and applying an algorithm that can "learn" how to differentiate between the various categories using the contents of the examples. This training process is then tested using a random selection of the data where the predictions of the trained model are compared to the known classification. We also tested different classifier algorithms to compare their performance.

In addition, since one of the objectives of this work is to create a proof-of-concept solution, we tested the viability of different tools to process the data and to train the classifier. We used two separate software solutions. First we used Mallet, a software package that is widely used in this type of analysis.[7] We also replicated the work using the Python programming language. While Mallet has a number of advantages for conducting quick and high-quality analysis in only a few lines of code, Python is used extensively in this field and is the language used to access online machine learning platforms such as Amazon's Sagemaker. The development of this process using Python was done in the interest of building a flexible system that can easily be used on these platforms.[8]

We report here the performance of the selected classifier when compared to a testing dataset of 21 documents drawn randomly from the set of 211 documents used to train the model (Table 3). The results of the validation tests confirm that the classifier is efficient in correctly identifying the labels of publications both in-sample and out-of-sample. With these results we can confidently re-train the classifier using the full dataset and use it to label the target data of UN documents.

We also attempted to use an unlabeled classification method following the methodology described in LaFleur (2019) to classify UN documents according to their SDG content. This type of classifier is attractive because it indicates how much of each category is used in a given text. Unfortunately, initial tests of the classifier did

7 http://mallet.cs.umass.edu/

8 The code and dataset used in both tracks will be available upon request.

not show sufficient ability to discriminate among the categories. Achieving precise differentiation among pre-determined categories using an unsupervised approach (so-called topic models) is difficult and requires a training dataset that is balanced across the various categories. Improving the performance of this classifier is left for future work.

Table 3
Classifier confusion matrix: predicted and true labels

True label	Predicted label				Total
	0	1	2	3	
0 - sustainable development	3				3
1 - Human-rights-humanitarian		5			5
2 - justice-law-drugs-terrorism			1		1
3 - peace-security-disarmament				12	12
Training documents = 190 Testing documents = 21	Training data accuracy = 0.99 Testing data accuracy = 0.94				

Source: Authors' calculation from UN GA resolutions.

3 Evidence from the UN General Assembly resolutions

Descriptive statistics

Using the classifier described above, we apply it to categorize a collection of GA resolutions from GA session 62 (2007-2008) through GA session 73 (2018-2019). The documents and associated metadata were downloaded using the Digital Library services and APIs made available by the Dag Hammarskjold Library of the United Nations.[9] A total of 3,765 resolutions were collected (see Table 4) over the period 2007 – 2019 covering 12 GA sessions. Figure 2 shows the distribution of the word count of all resolutions during the 12 sessions, excluding the largest outliers. It shows that the bulk of the resolutions are relatively short, with less than 2,000

Table 4
Descriptive statistics of resolutions by GA session

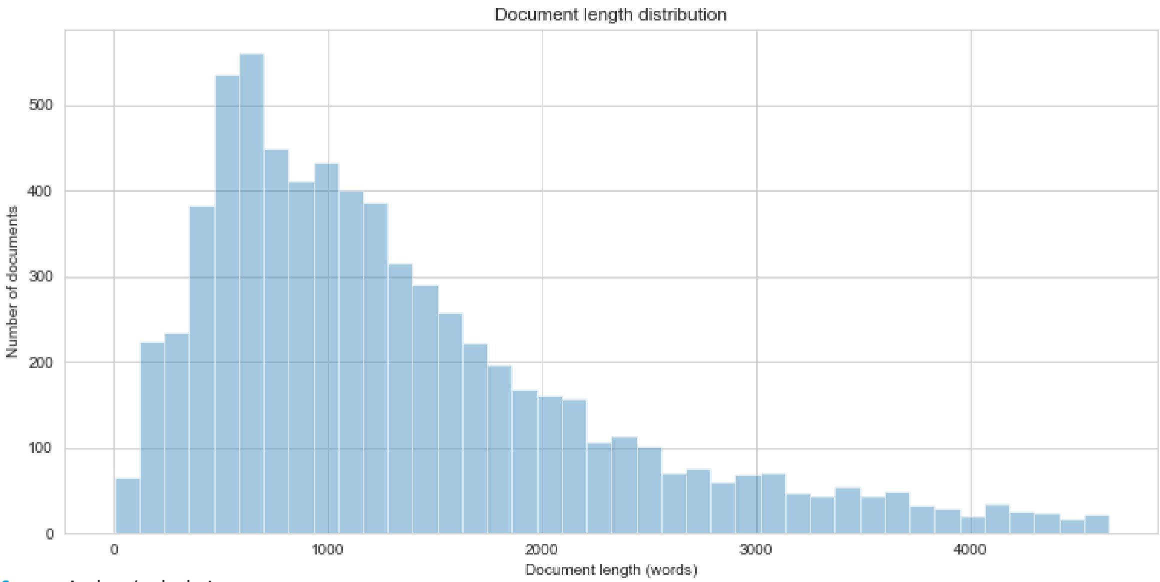
GA Session	Number of resolutions	Average word count	Largest word count
62 (2007-2008)	280	1,497	11,772
63 (2008-2009)	316	1,556	17,767
64 (2009-2010)	301	1,720	16,132
65 (2010-2011)	321	1,744	18,353
66 (2011-2012)	303	1,789	25,923
67 (2012-2013)	309	1,914	22,373
68 (2013-2014)	315	1,994	23,405
69 (2014-2015)	328	2,142	25,211
70 (2015-2016)	307	2,511	27,095
71 (2016-2017)	332	2,393	30,028
72 (2017-2018)	319	2,531	30,800
73 (2018-2019)	334	2,566	32,597

Source: Authors' calculation.

9 https://digitallibrary.un.org/

words. The average length of each resolution is relatively short, but the largest resolution has tripled in size as compared to the first session, perhaps reflecting the growing complexity of issues addressed by the General Assembly. It suggests the content and coverage of a resolution became more complicated and wider over time.

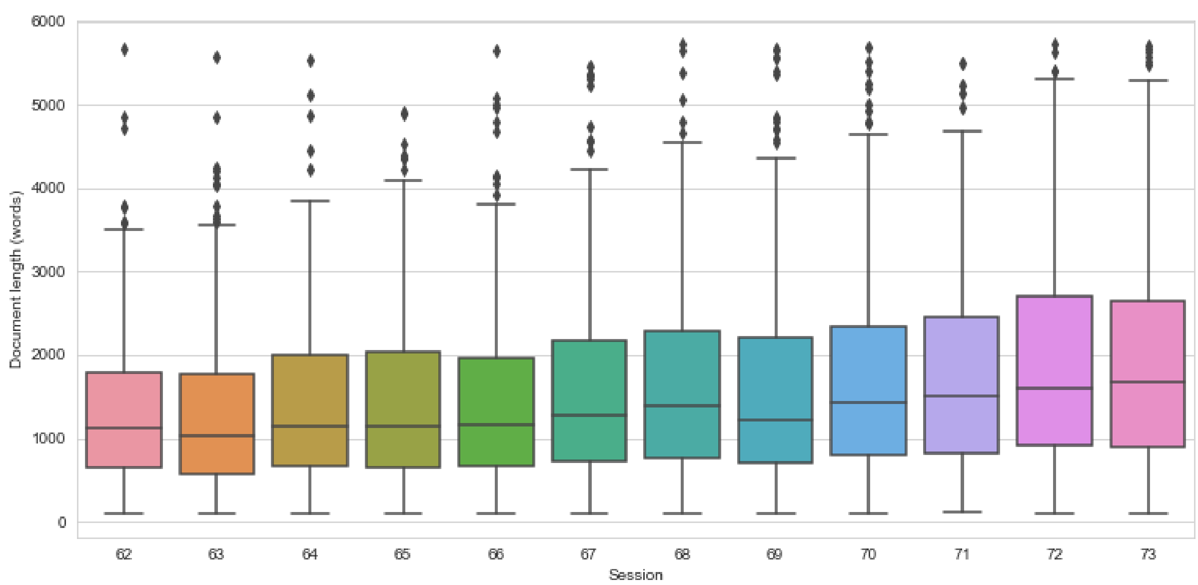
Figure 2
Distribution of resolution length (excluding outliers)



Source: Authors' calculation.

It is interesting to examine how this distribution has changed over time. Figure 3 shows the distribution of the resolution length in each session, again excluding the outliers. It shows a clear increase in the mean size and the interquartile range since the 62nd GA session.

Figure 3
Length of resolutions by GA session



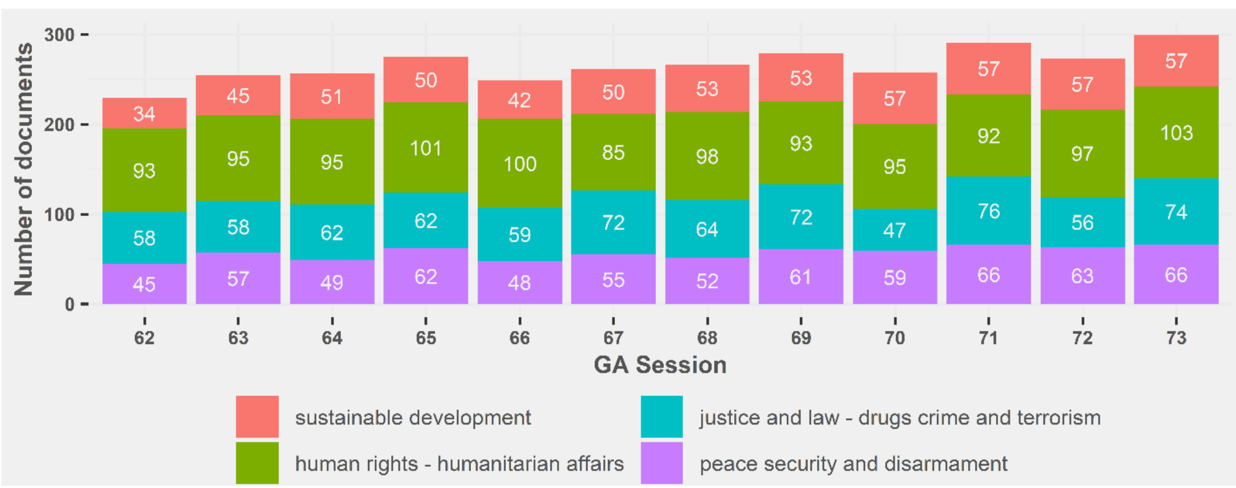
Source: Authors' calculation.
Note: The box shows the two innermost quartiles. The horizontal line represents the median.
The lines outside the box extend to 150% of the length of each box. Points above the line indicate outliers.

The scope of this analysis, described above, is limited to the thematic areas of work of the organization. It is interesting, therefore, to exclude resolutions with deal with internal administrative matters. To do this we exclude all GA resolutions listed under the Fifth Committee (Administrative and Budgetary). We also considered excluding the resolutions from the Sixth Committee (Legal), but many of these have relevance for the work on improving international law. We are left with 3,237 resolutions, which are then preprocessed in the same way as the training data set, described above. We apply the trained classifier on these processed documents and only report results if the classification score is least 80 per cent. This threshold also excludes any resolutions that may cover more than one category.

Results and trends

Figure 4 shows a summary of the results of the classification of the resolutions in each of the eleven GA sessions. In absolute numbers, each of the four categories saw increases in the number of resolutions between the 62nd session and the 73rd session. Interestingly, the number of resolutions on Sustainable Development has remained constant at 57 in the last four sessions of the GA. It is possible that this reflects structural rigidities in how this topic is handled by the GA. Examples of factors that contribute to these structural rigidities include restrictive GA mandates that are unchanged year after year, the lack of agreement among Member States on how to change resolutions, a stronger focus on implementation rather than normative changes.

Figure 4
Number of resolutions in each category, by GA session

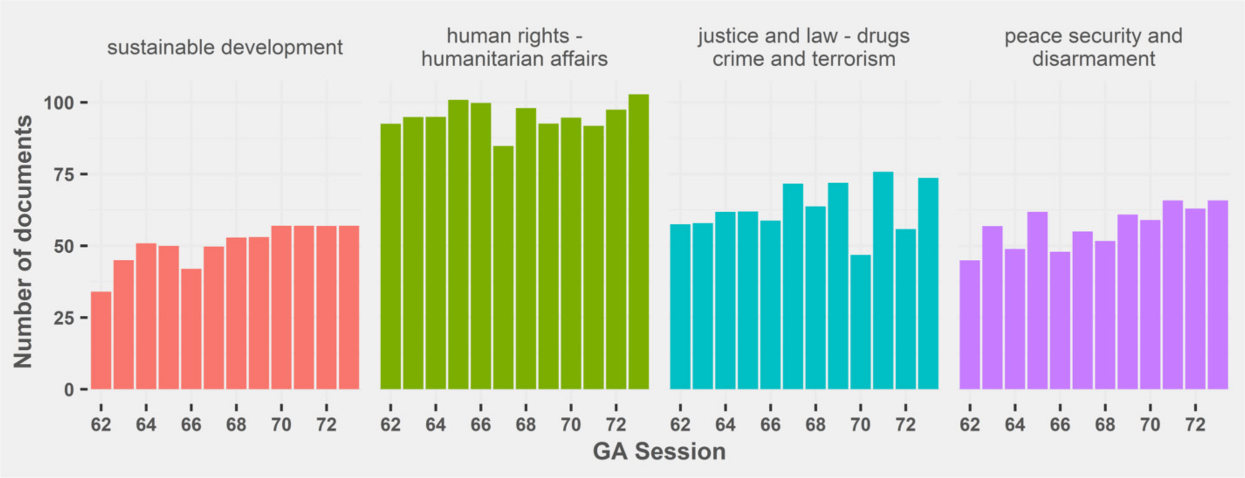


Source: Authors' calculation.

Figure 5 shows the same data grouped by category and makes a few trends immediately apparent. Except for the second category ("Human rights and humanitarian affairs"), the resolutions are balanced among the other areas of the UN' work. There is also a slight by discernable increase in focus on Sustainable Development and on Peace, security and disarmament.

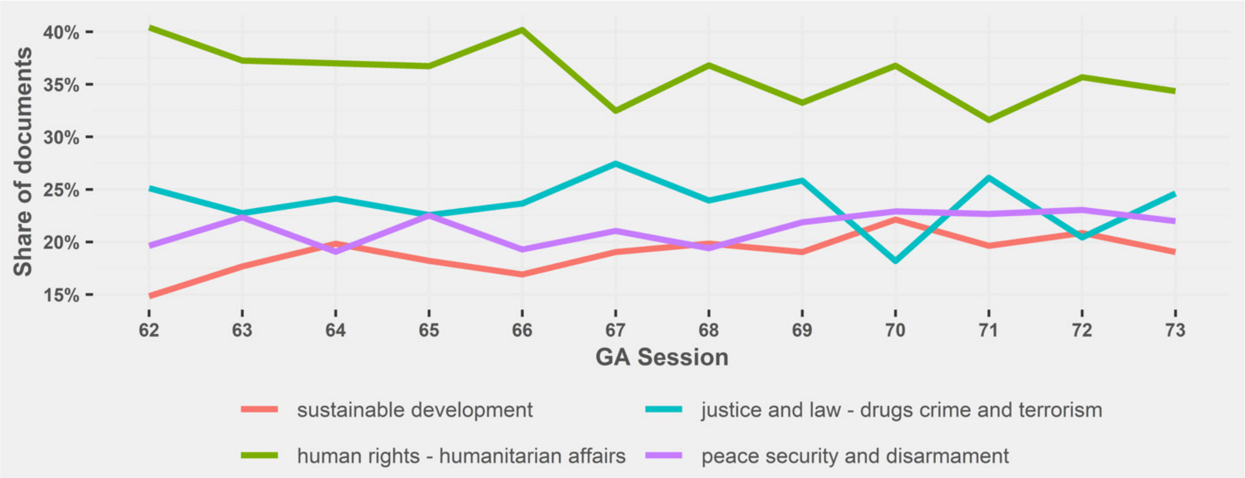
We can see this shift more clearly in Figure 6. The share of Sustainable development and Peace, security and disarmament in the total of each session's resolutions has slightly but consistently increased, mostly at the expense of the share of Human rights and humanitarian affairs. The share of "Justice, international law, drugs, crime and terrorism" has been more volatile, but has remained generally consistent over the last eleven sessions. This may reflect resolutions that mandate GA's work biennially.

Figure 5
Number of resolutions in each GA session, by category



Source: Authors' calculation.

Figure 6
Share of each category in the total of each GA session



Source: Authors' calculation.

In sum, natural language processing reveals that the UN is active in all four categories, namely, Sustainable Development, Justice and Law, Human Rights, and Peace and Security, with a growing number of resolutions over the last eleven sessions. Sustainable Development is an area that is gaining importance, not just claimed by the UN, but by the evidence from the language in the resolutions.

4 Concluding remarks

How has the focus of the UN GA changed over time and how well is the global agenda expressed in these documents? This paper shows that it is possible to start to answer these questions with the machine learning tools and techniques now available. By constructing a proof-of-concept classifier, it shows how such analysis can identify trends and answer questions about the nature of the work of the UN as reflected in UN GA documents. Using this method is clear that the issue of Sustainable Development is of growing importance not just in rhetoric, but in the text of GA resolutions, growing from approximately 15% to approximately 20% during the period of analysis.

This work is a starting point to show what kind of research is possible using natural language processing techniques and applying them to the vast collection of UN documents. Here we focused on GA resolutions, but other similar efforts have looked at substantive publications in the area of sustainable development to see how the sustainable development goals are represented in the work of the organization. Yet another used the same technique to analyze patterns in the speeches of senior officials.

There are many improvements made possible by the digitization of the organization's documents and their classification. Future work will continue in creating new and better classifier that leverage the semantic classification of UN documents made available in the UN Library catalogue.

References

Al Khalifa, Sheikha Haya Rashed (2007). The UN General Assembly Advances the Global Agenda. UN Chronicle Vol. XLIV No.1

Alschner, Wolfgang, Julia Seiermann, and Dmitriy Skougarevskiy (2017). Text-as-data analysis of preferential trade agreements: Mapping the PTA landscape. UNCTAD Research Paper No. 5.

Baturo, Alexander, Niheer Dasandi, and Slava J. Mikhaylov (2017). Understanding State Preferences With Text As Data: Introducing the UN General Debate Corpus. Research & Politics 4:2.

Blei, David M. and Jon D. McAuliffe (2010). Supervised Topic Models.

Blei, David. 2012. Probabilistic topic models. Communications of the ACM 55(4):77–84.

Bonica, Adam. 2011. Estimating ideological positions of candidates and contributions from campaign finance records. Stanford University

Burden, Barry, and Joseph Sanberg (2003). Budget rhetoric in presidential campaigns from 1952 to 2000. Political Behavior 25(2):97–118.

Eshbaugh-Soha, Matthew (2010). The tone of local presidential news coverage. Political communication 27(2):121-40.

Eshbaugh-Soha, Matthew, and Jeffery S. Peake (2010). Presidential Leadership of the Tone of News Coverage and Public Opinion on Iraq. APSA 2010 Annual Meeting Paper

Gerber, Elisabeth, and Jeff Lewis. 2004. Beyond the median: Voter preferences, district heterogeneity, and political representation. Journal of Political Economy 112(6):1364–83.

Grimmer, Justin and Brandon M. Stewart (2013). Text as Data: The Promise and Pitfalls of Automatic content Analysis Methods for Political Texts. Political Analysis, pp 1-31.

Jensen, Jacob, Ethan Kaplan, Suresh Naidu, Laurence Wilse-Samson (2012). Political Polarization and the Dynamics of Political language: Evidence from 130 Years of Partisan Speech. Brookings Papers on Economic Activity.

Jurafsky, Dan, and James Martin. 2009. Speech and natural language processing: An introduction to natural language processing, computational linguistics, and speech recognition. Upper Saddle River, NJ: Prentice Hall.

Jurka, Timothy P., Loren Collingwood, Amber Boydstun, Emiliano Grossman, and Wouter van Atteveldt. 2012. RTextTools: Automatic text classification via supervised learning. http://cran.r-project.org/web/packages/RTextTools/index.html

Kellstedt, Paul (2000). Media framing and the dynamics of racial policy preferences. American Journal of Political Science 44(2):245–60.

LaFleur, Marcelo T. (2019), Art is long, life is short: An SDG Classification System for DESA Publications, DESA Working Paper, vol. 159, May.

Laver, Michael, and John Garry (2000). Estimating policy positions from political texts. American Journal of Political Science 44(3):619–34.

Monroe, Burt, and Ko Maeda. 2004. Talk's cheap: Text-based estimation of rhetorical ideal points. Paper presented at the 21st annual summer meeting of the Society of Political Methodology. Slapin and proksch 2008

Munzert, Simon, Christian Rubba, Peter Meissner, and Dominic Nyhuis (2014). Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining. John Wiley and Sons.

Nisbet, Robert et al (2009). Handbook of Statistical Analysis and Data Mining Applications, Elsevier Science.

Rodman, E. (2020). A Timely Intervention: Tracking the Changing Meanings of Political Concepts with Word Vectors. Political Analysis, 28(1), 87-111. doi:10.1017/pan.2019.23

Seiermann, Julia (2018). Only words? How Power in Trade Agreement Texts Affects International Trade Flows. UNCTAD Research Paper No. 26

Taddy, Matt (2010). Multinomial Inverse Regression for Text Analysis

United Nations (1945). Charter of the United Nations.

United Nations (2005). In larger freedom: towards development, security and human rights for all. Report of the Secretary-General. A/59/2005. https://www.un.org/en/ga/search/view_doc.asp?symbol=A/59/2005

Wang, William Yang, Elijah Mayfield, Suresh Maidu, and Jeremiah Dittmar (xx). Historical Analysis of Legal Opinions with a Sparse Mixed-Effects Latent Variable Model.

Young, Lori, and Stuart Soroka (2011). Affective news: The automated coding of sentiment in political texts. Political Communication 29(2):205–31.