SDG ACCELERATION DIAGNOSTIC
DIAGNOSTIC SIMULATOR METHODOLOGICAL NOTE & USER GUIDE

The purpose of this document is to provide an overview of the Sustainable Development Goals (SDG) Diagnostic Simulator. More specifically, it can provide a guide on how to use the three steps of the simulator along with methodological background used to produce them. The three steps are Current Gaps, Current Priorities, and SDG Push, which can be used sequentially as part of the SDG Acceleration Diagnostic process or as standalone investigations.

Overview

The SDG Diagnostic Simulator is a key element of the SDG Acceleration Diagnostic, a process that combines analytical capabilities with qualitative methods to assess context, options, and bottlenecks to determine effective policy choices, investments and pathways. The Simulator forms part of the Scoping Work, during which gaps, priorities, interlinkages, and opportunities are identified.

The three steps of the Simulator focus on the following:

- **Current Gaps:** Visualizing data to easily identify SDG gaps, progress and interlinkages that allows for a holistic dialogue
- **Current Priorities:** Supporting faster real time learning and identification of priorities using advanced text-analytics; and
- **SDG Push:** Mapping the projected impact of the SDG Push scenarios to provide an example of an ambitious set of actions that can accelerate SDG and national development goals.

Current Gaps

The Current (SDG) Gaps section gives users a straightforward visualization of current SDG progress for selected countries. The user can select the country of interest by clicking the 'Change Country' option, and the default country is South Africa.

The Current Gaps tab contains three sections: Current Gaps, Target Overview, and SDG Overview. All these sections use four color-coded, categorical scales to classify the progress of SDGs, Targets, or Indicators:

- Green for "On Track" - The country is on track to fulfill the SDG by 2030,
- Yellow for "For Review" - With current progress the country will miss the SDG by 2030 by a small margin,
- Red for "Identified Gap" - With current progress the country will miss the SDG by 2030 by a large margin,

- Grey for "Gap Unidentified" - Country doesn't have enough data to identify the progress of the SDG.

Data Source

All the time series data for all the indicators in all the goals are either fetched using UNStats API: https://unstats.un.org/sdgapi/swagger/#!/Goal/V1SdgGoalDataGet or the data is provided by the government to the national country team.

The aim is to have a dialogue with the national country teams and governments; and work with them to get the most up to date data and also define which timeseries they deem important to use for the simulator.

METHODOLOGY TO IDENTIFY THE CATEGORIES FOR EACH SDG

A single **SDG** has multiple **Targets;** each **Target** has multiple **Indicators**; and each **Indicator** can have multiple time series**.** The data that is fetched from UNStats is for these time series. Some of the time series have very granular disaggregation which might not be required in each case; therefore, key time series for each indicator are identified and filtered.

Once the key time series are identified and the data is available, following the methodology given by UNStats (technical note here), the indicators are classified into 5 categories:

- Target achieved (if the country has already reached the target value)
- On track,
- Fair progress but acceleration needed,
- Limited or no progress and
- Deterioration

Please note that we are not able to classify all the indicators in these categories as for some indicators we might not have enough data for a country, or we might not have the target values for some indicators. This limits us to only use the time series which are mentioned in the technical note by UNStats. In some cases, different UNDP country offices after consulting with the governments may provide target values for indicators for which we might not have a global target value.

All the time series within an indicator (as a single indicator can have multiple time series) are assigned scores of 1 to 4 (1 being "*Target achieved*" or "*On track*", 2 "*Fair progress but acceleration needed*", 3 "*Limited or no progress*", and 4 "*Deterioration*"). All the time series for which the categories cannot be identified because of lack of sufficient data are ignored.

An average of the score of all the time series within an **indicator** (ignoring all the time series for which the categories cannot be identified because of lack of sufficient data) is calculated. This average value ranges from 1 to 4. Based on this average score, the **indicators** are classified into four categories (defined earlier) as follows:

- On Track (for scores between 1 - 1.5),
- For Review (for scores between 1.5 - 2.5), and
- Identified Gap (for scores between 2.5 - 4)
- Gap unidentified (for indicator with no time series with a category)

Using the above categorization **indicators** within a **target** are assigned scores of 1 to 3 (1 for "*On Track*", 2 for "*For Review*", and 3 for "*Identified Gaps*"). An average of the score of all the indicators within a target (ignoring all the indicators belonging to the category Gap Unidentified) is calculated. This average value ranges from 1 to 3. Based on this average score, the **targets** are classified into four categories (defined earlier) as follows:

- On Track (for scores between 1 - 1.5),
- For Review (for scores between 1.5 - 2.5), and
- Identified Gap (for scores between 2.5 - 3)
- Gap unidentified (for target with no indicators with a category)

Using the above categorization all the **targets** within an **SDG** are assigned scores of 1 to 3 (1 for "*On Track*", 2 for "*For Review*", and 3 for "*Identified Gaps*"). An average of the score of all the targets within an **SDG** (ignoring all the targets belonging to the category "Gap Unidentified") is calculated. This average value ranges from 1 to 3. Based on this average score, the **SDGs** are classified into four categories (defined earlier) as follows:

- On Track (for scores between 1 - 1.5),
- For Review (for scores between 1.5 - 2.5), and
- Identified Gap (for scores between 2.5 - 3)
- Gap unidentified (for target with no indicators with a category)

We explain each of the three sections and tabs below.

CURRENT GAPS

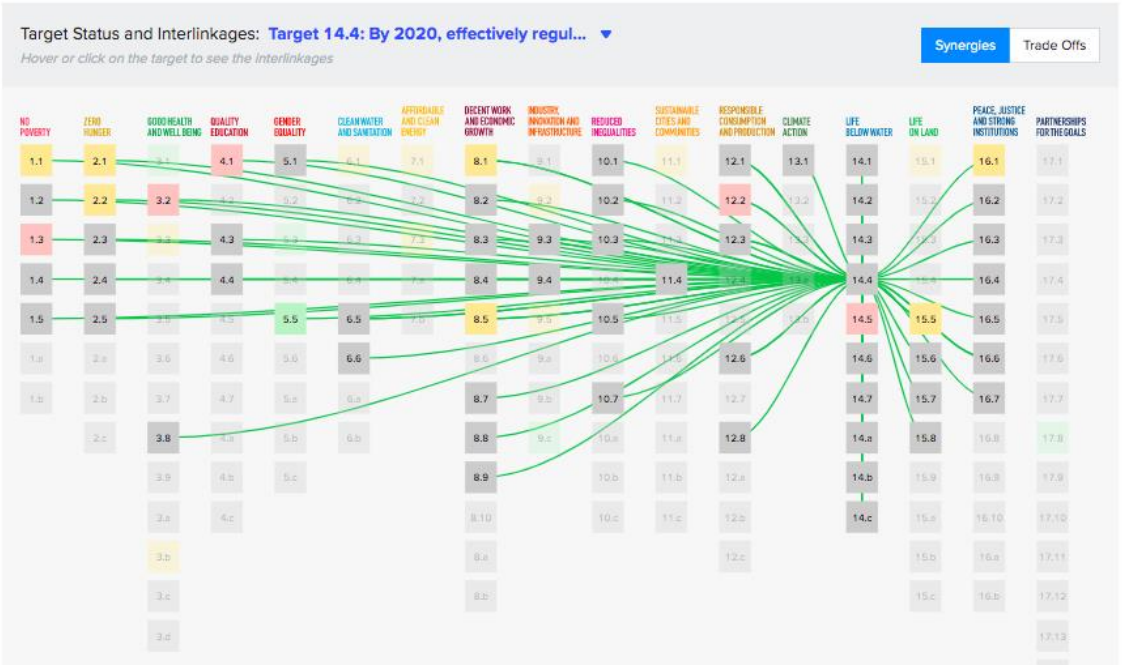
The section has a sub header that provides a quick summary of out of 17 SDGs, I.e., how many fall under either On Track, or Identified gaps, or For Review categories. For example, for South Africa, it states, "2 are On Track, 8 are Identified Gaps, and 5 are For Review," as demonstrated below:

Followed by the sub header, is a color-coded doughnut chart formed by the four categories that each SDG goal belongs to as stated earlier. On the right of the chart is a list of the same SDG categories that form the chart.

The doughnut chart and the accompanying color-coded, visual classification of a country's progress towards the 2030 SDG goal is displayed per the chosen country/region by the user. Additionally, with the righthand side overview list, users can identify where the SDGs fall into each of the four classifications, including whether they are unidentified due to lack of data. If users wish to further explore the Targets and Indicators specifications for each SDG, they can click on the SDG to the right of the chart, and they will be directed to SDG Overview section of the page (more on this later).

TARGET OVERVIEW

The Current Gaps section is followed by Target Overview section, presenting each of the 17 SDGs in column form with the respective targets under each SDG. The sub header of this section gives a summary of the SDG target with most synergies/interlinkages with other SDG targets as well as top three target interlinkages with the highest acceleration potential. For example, for the default country South Africa, target 14.4 under SDG 14 is listed as the one with the most synergies, however, targets 14.7 under SDG 17, 6.1 under SDG 6, and 1.3 under SDG 1 are listed as top three target interlinkages with highest acceleration potential. This is followed by the main color-coded figure where users can use the drop-down button to select any target, for which a visual display of the target and its synergies with other targets (represented by green lines) is presented in default mode. Additionally, users can click on the tradeoff tab to obtain a visual presentation of the target and its tradeoff (instead of synergies), represented by red lines. Alternatively, users can click on any of the Target directly in the figure and it will display their synergies or tradeoffs depending on what tab users choose to explore.



The data for the synergies and tradeoff is from the European Commission's KnowSDGs webpage.

SDG OVERVIEW

The Target Overview section is followed by the SDG overview section which provides a drop-down menu of all 17 SDG options that users can select from. For the given country/region earlier, the subsection below the drop down is followed by a box that provides a quick summary of status of targets for that SDG. Then, it breaks down the selected SDG into targets that constitute the SDG. Additionally, it describes the target and color-codes it to indicate to the user where the country stands in terms of achieving its 2030 SDG agenda. For example, for South Africa, when the user chooses SDG1, it lists seven targets, 1.1, 1.2, 1.3, 1.4, 1.5, 1.a, and 1.b in three colors, where yellow indicating "For Review" for target 1.3, red indicating "Identified Gap" for target 1.1, and grey indicating missing data for the rest. Then it lists each of these targets with the corresponding color code and a clickable link with an option to expand further to explore all the indicators that comprise each of these targets. When the user clicks on the indicator, depending on data availability, it provides the user with a graphical representation of time series data for that indicator.

Current Priorities

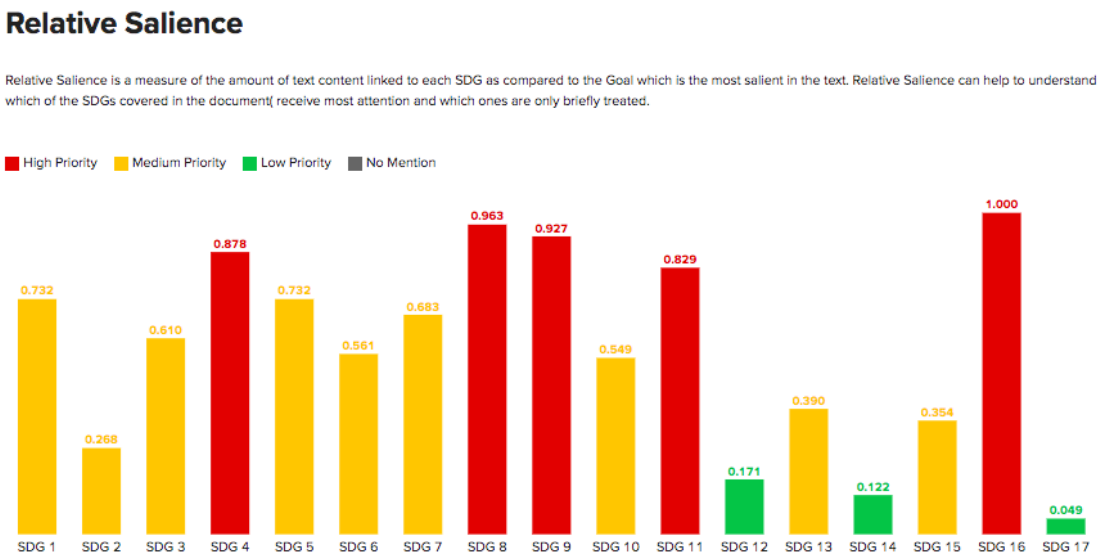
The second tab, Current Priorities, sets out to provide an in-depth analysis of text documents with respect to the Sustainable Development Goals (SDGs) using natural language processing technologies. In so doing, the tab serves a two-fold purpose.

First, it provides users with results of the analysis of the latest Voluntary National Report (VNR) submitted by the selected country. The VNRs are taken from the website of the High-Level Political Forum on Sustainable Development. While most VNRs are written in English, some countries submit their reports in a different official UN language such as Arabic, Chinese, French, Russian or Spanish. In such cases, the analysis is performed on a machine-translated English version of the original document. For VNRs that are available in multiple languages, the English version is always used for the purpose of the analysis.

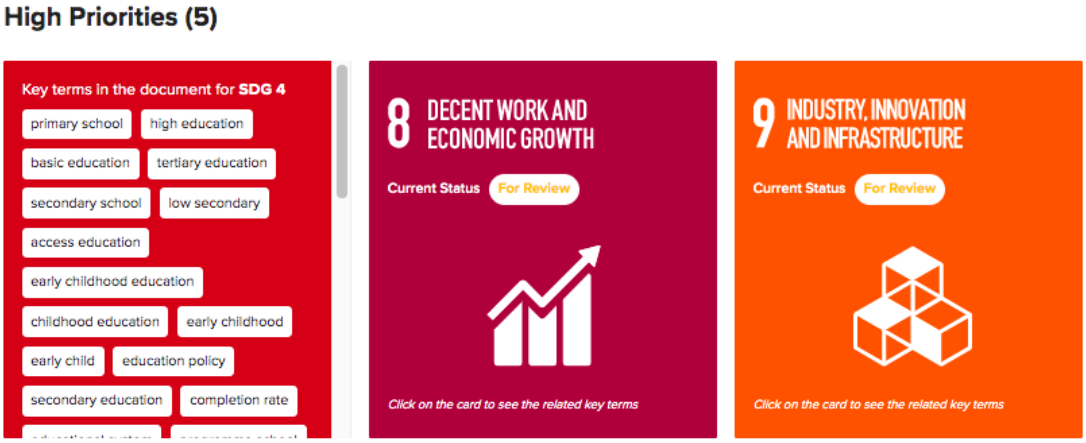
Second, the Current Priorities tab enables users to go beyond VNRs and analyze a report, policy document or any other text document of interest. This can be done by uploading an arbitrary PDF file. Note that this must be a searchable PDF and not a scanned copy of a document. Also note that unlike VNRs, in the October 2022 version, such documents must be in English as non-English texts will not be translated.

The analysis of VNRs and user-provided documents is performed using a **custom-built machine learning model for SDG classification.** The model analyses the document by linking its constituent pieces to SDGs and then aggregates the statistics before presenting the results to the user. The underlying assumption of the analysis is that the amount of text linked to each SDG defines how important each SDG is. In this context, we refer to the "importance" of an SDG as salience. Thus, the most salient SDG is the one to which most of the text pieces, e.g., paragraphs, could be linked.

There are two main UI elements that present the results of the analysis. The Relative Salience bar chart depicts the salience of all 17 SDGs. The most salient SDG, by definition, obtains a relative salience of 1.0 and other SDGs can have any of the values from 0.0, i.e., this SDG is not covered in a text at all, to 1.0, i.e., this SDG is the most prominent/salient in the text. This measure of relative salience is used to categorize SDGs into High-, Medium- and Low-priority tiers. This is done based on thresholds. The thresholds were determined such that, on average, approximately 33.3% SDGs would fall into the Low-priority group across all VNRs.



Below the chart, there is a SDGs by Priorities section that shows all SDGs grouped by priority. By clicking on any SDGs, users can also see key topics that are discussed in the document in relation to the selected SDGs.



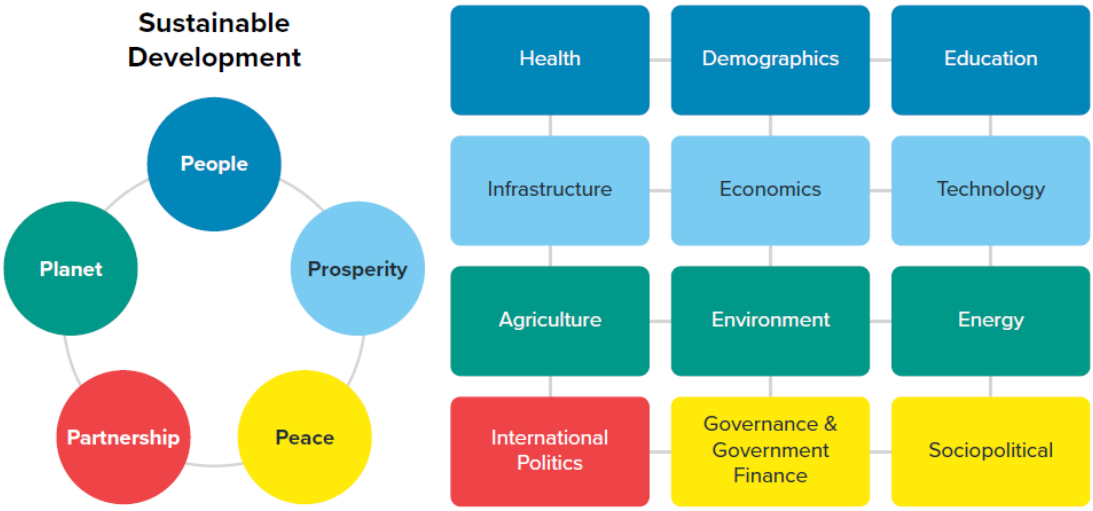
It is also possible to compare two documents by uploading two files. In such case, the users are presented with a table that lists each SDG and its priority tier in both documents side by side. This may be useful to compare VNRs from different years for the same country, two VNRs from two different countries or two completely different policy documents.

SDG Push

The final tab of the main page is labelled SDG Push. The results in this section is building upon UNDP's 'SDG Push' scenario – which features an ambitious yet feasible set of SDG accelerators across governance, social protection, green recovery, and digitalization – the diagnostic supports governments to mainstream and disseminate accelerators to recover better from the COVID-19 pandemic.

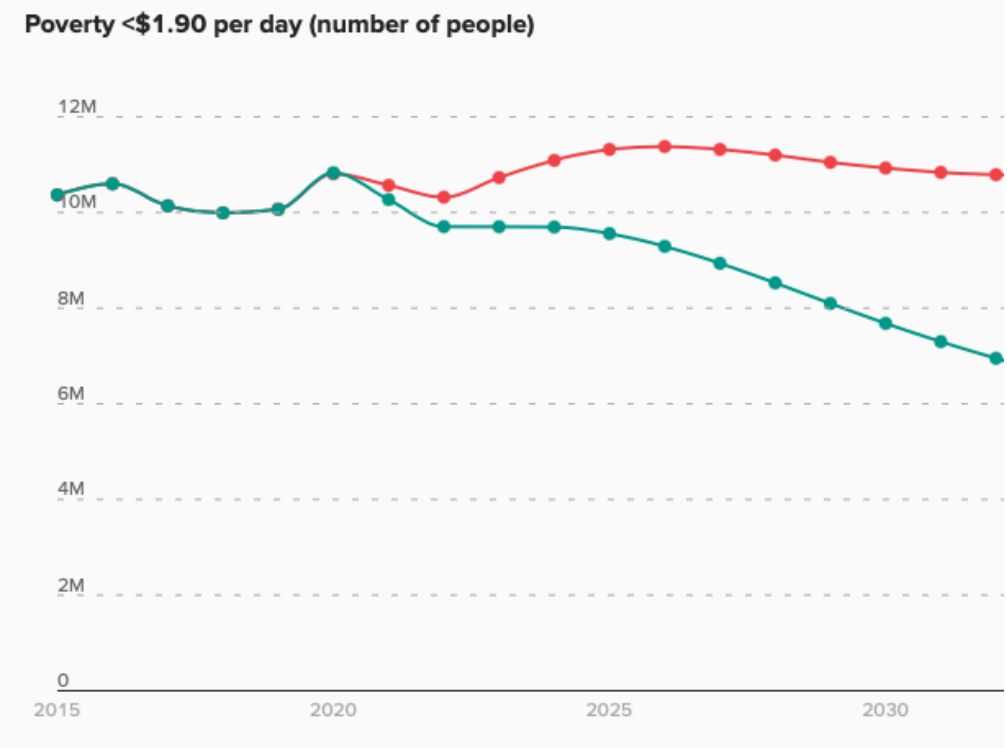
It has two parts: the first part describes Future Scenarios which is based on International Futures (IFs) model used to forecast the development path and the effects of combined set of interventions in the areas of Governance, Social Protection, Green Economy and Digital Disruption across different and predefined set of scenarios. IFs is an open-source integrated assessment modeling tool that simulates interactions within and across 186 countries and 12 core systems: agriculture, demographics, economics, education, energy, environment, finance, governance, health, infrastructure, international politics, and technology. These systems are dynamically connected, so that changes in one system may lead to changes across all others, which can illuminate spillover effects as well as tradeoffs and synergies with different policy choices.

The representation below shows how the models within IFs correspond to the 5 "Ps" and associated SDGs.



The second part illustrates the impact of Future Scenario investments in five SDGs and its targets and compares it with Baseline scenario where such investments are not undertaken. The time series graphs show progress until 2050 on a yearly basis for every target related to specific SDG tab shown in the subheading. Each of these graphs provide two scenarios for each country in terms of what the gaps would look like with and without the "SDG push" scenarios, color-coded in green and red, respectively.

For example, for the default country South Africa, users can see that SGD 1 tab produces three graphs, one each for poverty under $1.90 per day (by number of people and percent of population) and female poverty headcount by number of people. The top line in each of these graphs is given in red and is a scenario when "SDG Push scenario" is not implemented and the bottom line is when the "SDG push scenario" is implemented and shown as a green line. Therefore, by having these graph lines right on top of each other, users can easily identify the gaps with and without push scenarios See example below:


Poverty <$1.90 per day (number of people)
12M
10M
8M
6M
4M
2M
0
2015
2020
2025
2030