Innovative Big Data Approaches for Capturing and Analyzing Data to Monitor and Achieve the SDGs

UNITED NATIONS
ESCAP
Economic and Social Commission for Asia and the Pacific
East and North-East Asia Office

ACKNOWLEDGEMENTS

The present paper was drafted by Dr. Soenke Ziesche, Data Science Consultant, ESCAP ENEA. The author whishes to thank Li Zhou, Associate Social Affairs Officer, ESCAP ENEA, Tianyi Gu, Intern, ESCAP ENEA and Charlotte Arribe, Intern, ESCAP ENEA for their inputs into this report.

This working paper was initially prepared as a discussion paper for an expert group meeting titled Forum on Innovative Data Approaches to SDGs held in Incheon, Republic of Korea, from 31 May – 2 June 2017, jointly organised by ESCAP East and North-East Asia office, and ESCAP Statistics Division. At this meeting, experts from over 12 countries discussed and reviewed this report, with feedback integrated subsequent to the meeting. The author is thankful for the inputs from the experts.

Disclaimer

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or its authorities, or concerning the delamination of its frontiers or boundaries. The opinions, figures and estimates set forth in this publication are the responsibility of the authors, and should not necessarily be considered as reflecting the views or carrying the endorsement of the United Nations.

Innovative Big Data Approaches for Capturing and Analyzing Data to Monitor and Achieve the SDGs

UNITED NATIONS
ESCAP
Economic and Social Commission for Asia and the Pacific
East and North-East Asia Office

Table of Contents

EXECUTIVE SUMMARY 008

INTRODUCTION: SDGS, TARGETS AND INDICATORS 011

Background 011

Tier classification of SDG indicators 012

INTRODUCTION: BIG DATA, INTERNET OF THINGS AND ARTIFICIAL INTELLIGENCE 015

Big data 015

Terminology and definitions 015
Background 017
Taxonomy of big data 018
Opportunities 019
Challenges 023

Internet of Things 028

Terminology and definitions 028
Background 029
Taxonomy of IoT sensors 029
Opportunities 030
Challenges 030

Artificial Intelligence 032

Terminology and definitions 032
Background 032
Taxonomy of AI methods 033
Opportunities 033
Challenges 034

Synergies for capturing and analysing data 035

OVERVIEW OF BIG DATA APPROACHES 036

Background 036

Structure of the overview and explanations 036

Exhaust data 037

Mobile phone data 037
Financial transactions 042
Online search and access logs 044
Administrative data / citizen cards 046
Postal data 048
Benefits, risks and recommendations 050

Sensing data 050

Satellite and UAV imagery 050
Sensors in cities, transport and homes 062
Sensors in nature, agriculture and water 066
Wearable technology 071
Biometric data 074
Benefits, risks and recommendations 076

Digital content 077

Social media data 077
Web scraping 082
Participatory sensing / crowdsourcing 085
Health records 089
Radio content 092
Benefits, risks and recommendations 094

AI applications towards SDGs without big data 094

Benefits, risks and recommendations 097

KEY FINDINGS AND ANALYSIS 098

Not suitable indicators for big data 099
Indicators with big data approaches 099
Targets with big data approaches 101

CONCLUSION 102

REFERENCES 103

ANNEX - SDG TARGETS, INDICATORS AND BIG DATA APPROACHES 108

List of Tables

Table 1 SDGs with numbers of targets, indicators and tiers **014**

Table 2 CDR example **016**

Table 3 Categories and sources of big data **018**

Table 4 Types of sensors, functionalities and examples **029**

Table 5 AI research topics **033**

Table 6 Big data approaches: mobile phone data **041**

Table 7 Big data approaches: financial transactions **043**

Table 8 Big data approaches: Online search and access logs **045**

Table 9 Big data approaches: Administrative data / citizen cards **047**

Table 10 Big data approaches: Postal data **049**

Table 11 Big data approaches: Satellite and UAV imagery **061**

Table 12 Big data approaches: Sensors in cities, transport and homes **065**

Table 13 Big data approaches: Sensors in nature, agriculture and water **070**

Table 14 Big data approaches: Wearable technology **073**

Table 15 Big data approaches: Biometric data **075**

Table 16 Big data approaches: Social media data **081**

Table 17 Big data approaches: Web scraping **084**

Table 18 Big data approaches: Participatory sensing / crowdsourcing **088**

Table 19 Big data approaches: Health records **091**

Table 20 Big data approaches: Radio content **093**

Table 21 Big data approaches: AI application towards SDGs without big data **096**

Table 22 Approaches for calculation of indicators by tier category **099**

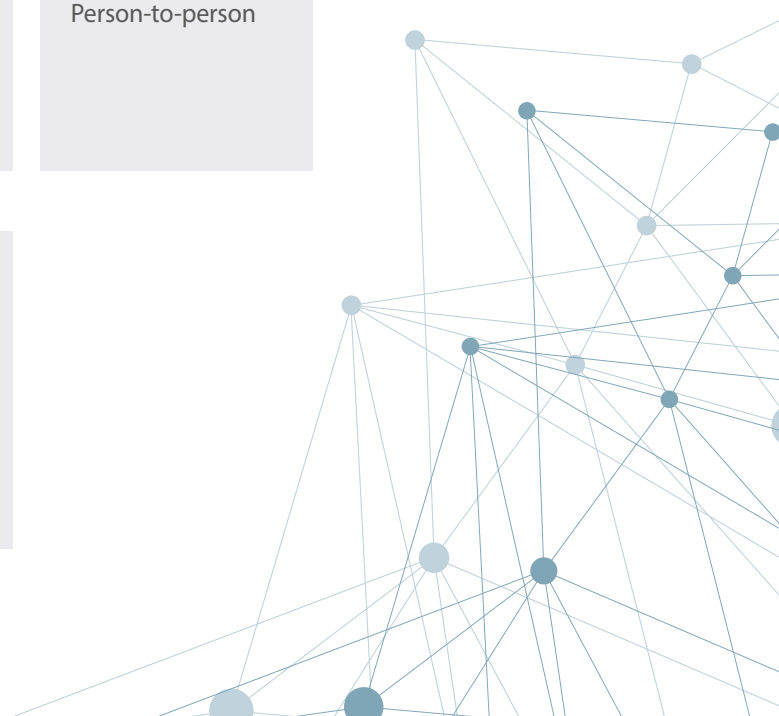
Table 23 Approaches for calculation of indicators by SDG **100**

Table 24 Approaches for achievement of targets by SDG **101**

Table 25 SDG targets, indicators and big data approaches **132**

List of Abbreviations

AI	Artificial intelligence
CDR	Call detail record
CSR	Corporate social responsibility
ENEA	East and North East Asia
ESCAP	Economic and Social Commission for Asia and the Pacific
IAEG	Inter-Agency and Expert Group
IoT	Internet of things
MDG	Millennium Development Goal
M2M	Machine-to-machine
M2P	Machine-to-person
P2P	Person-to-person
SDG	Sustainable Development Goal
UAV	Unmanned aerial vehicle



Executive summary

This report showcases around 140 big data approaches to potentially assist traditional statistics methods in capturing and analysing data to support the calculation of SDG indicators and the achievement of SDG targets. The presented approaches also aim to replace costly occasional surveys of traditional statistics with cheaper real-time information. The structure of the report is as follows: First, the SDGs are introduced with a focus on current challenges regarding lacking data as well as methodologies. Then an overview of big data, IoT and AI is given with a focus on categorization, opportunities and challenges. The main section is dedicated to describing, classifying and linking the aforementioned approaches to suitable SDG indicators and targets. Benefits, risks and potential recommendations for pilot projects are discussed per big data category. This is followed by a summary of the key findings, an analysis and the conclusion.

The 2030 Agenda for Sustainable Development for the time period from 2016 until 2030 comprises 17 Sustainable Development Goals (SDGs), subdivided into 169 targets and 232 indicators. In comparison, there were only 8 Millennium Development Goals (MDGs) with 21 targets and 60 indicators for the previous period from 2000 until 2015. Not only do the SDGs cover a much broader range of issues, the SDG indicators are also very different from and more complex than the MDG indicators, thus in many instances challenging for traditional statistics. Therefore, innovative approaches are required. The technological environment has continued to advance in recent years to a stage where it now appears promising to harness big data for both the achievement of SDG targets as well as the calculation of SDG indicators. Many of the new big data are passively emitted and collected as by-products of people's interactions with and uses of digital devices. Data coming from various sources provide unique insights about human behaviour and beliefs, which could be harnessed to increase the quality of life of these people, thereby contributing ultimately to the achievements of the SDGs.

In contrast to MDG data, which were mostly collected and owned by Governments, critical SDG data are produced passively by people, collected by machines and owned by corporations. Under the umbrella of corporate social responsibility, data philanthropy is a win-win opportunity for corporations to cost-efficiently improve their reputation while the UN or other organizations are receiving the data in order to use them. The concept of open data calls for Governments to provide free of charge, up to date, openly licensed and machine readable online data to enable data analysis by NGOs or other stakeholders.

Consequently, bigger quantities of data have come with low-costs and real-time availability, enabling the collection

预览已结束，完整报告链接和二维码如下：
https://www.yunbaogao.cn/report/index/report?reportId=5_1257