ENGAGING CITIZENS FOR SUSTAINABLE DEVELOPMENT: A DATA PERSPECTIVE

MAKING THE GLOBAL AGENDA THE CITIZENS' AGENDA

REPORT

UNITED NATIONS UNIVERSITY

UNU-CS

Computing and Society

About This Report

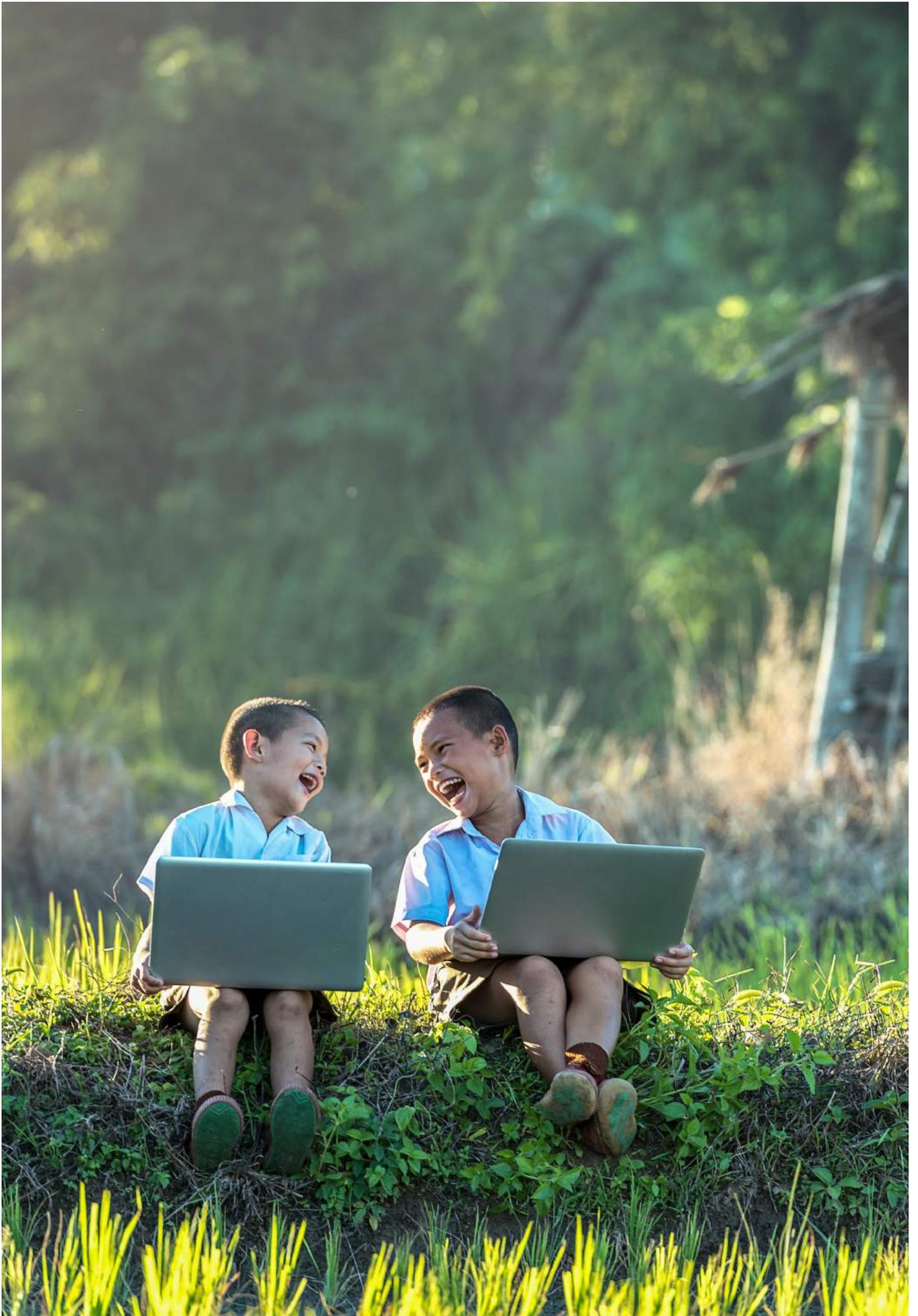Engaging Citizens for Sustainable Development: A Data Perspective – Making the Global Agenda the Citizens' Agenda considers the case of the engagement of citizens and community-level actors within the United Nations sustainable development agenda, also known as the Global Agenda. As primary agents of their development and the ultimate beneficiaries of the Sustainable Development Goals, citizens have a pivotal role to play not only in terms of the effort and action towards the achievement of the goals but also in terms of the associated monitoring of the progress towards these goals. This report presents findings of case study research exploring the dynamics of participation and engagement of citizens within the Global Agenda, highlighting issues of data marginalization and exclusion, the case for community-driven indicators and community-based monitoring systems, data ownership and sharing, as well as citizen-centric data valorization.

This report is produced from research that is undertaken at the United Nations University Institute on Computing and Society (UNU-CS), within the Small Data Lab. It provides a policy-relevant digest and an overview of the research and frames the findings for practitioners and stakeholders working on programs that advance the Sustainable Development Goals agenda, especially, at sub-national and community levels.

Author
Mamello Thinyane, Principal Research Fellow, UNU-CS

Design
Geneva Damayanti, Communications Officer, UNU-CS

Acknowledgments

Table of Contents

Executive Summary

The United Nations 2030 Agenda for Sustainable Development is a global aspirational roadmap towards achieving a sustainable future for all – through reducing poverty, protecting the environment, and building inclusive societies. At its core, it is guided by principles of inclusivity and participation, and its success is predicated on establishing strategic cross-cutting and multi-stakeholder partnerships towards the achievement of the goals.

The engagement, inclusion, and participation of citizens, and perhaps more significantly the ownership of the Global Agenda by empowered citizens, is pivotal towards realizing the broad goals of sustainable development in such a way that no one is left behind. From the findings of case study research, this report makes note of and discusses the following key considerations that must be taken into account for meaningful engagement of citizens for sustainable development.

- With the centrality of the "Leave no one behind" principle within the Global Agenda, it is necessary to formulate and develop frameworks and instruments that operationalize this principle for a broad range of stakeholders. Further, the myriad of ways through which individuals and population groups get left behind, need to be further explored, expounded on, and mitigated against. Towards this end, a typology of data marginalization and exclusion, that highlights forms of marginalization in indicators data, is presented in this report. The typology identifies the following five marginalized populations groups, presented through the "voices" metaphor: unknown voices, silent voices, muted voices, unheard voices, and ignored voices.
- Awareness of the agenda for sustainable development remains a critical precursor to an active engagement of citizens, communities and other grassroots-level stakeholders. Evidence presented in this report suggests that individuals' lack of awareness of the sustainable development goals remains very high. However, despite the lack of awareness of the SDGs, there is a high level of significance and relevance that individuals generally ascribe to each of the 17 goals.
- Engaging citizens for sustainable development demands that the sustainable development goals be given local and contextual relevance to the realities of citizens and different communities. Community-driven indicators and the associated community-based monitoring systems are a crucial mechanism for making the Global Agenda relevant to the hyperlocal contexts and for garnering community ownership and buy-in.
- Meaningful participation of citizens within the Global Agenda is enabled and achieved when the citizens' agency is amplified and when the benefits of participation, especially in indicators data collection, primarily accrue to the citizens.
- The role of civil society organizations, and community-based organizations, for brokering and intermediating the participation of citizens within the Global Agenda is critical – even more so for the marginalized and vulnerable populations groups.

INTRODUCTION

INTRODUCTION

Meeting the goals of the 2030 Agenda for Sustainable Development requires the engagement of stakeholders from all sectors of society, including the civil society organizations (CSO) and individual citizens. The different stakeholders have specific roles to play and responsibilities towards the achievement of the 17 developmental goals. Beyond the individual stakeholder effort, however, enhancing the engagement between the stakeholders is also crucial, as suggested by the definition of the specific goal around strengthening partnerships for the goals – SDG 17.

Catalyzing the engagement of the diverse stakeholders within the sustainable development goals (SDGs) agenda requires effort to give the SDGs a contextual relevance to the diverse realities of the different regions and countries, and to the varied interests of the stakeholders. The need for this "localizing of the SDGs" [1] has been given significant recognition through efforts within the United Nations towards engaging with local actors to translate and "channel the global goals into local actions" and through the adoption of the Mainstreaming, Acceleration, and Policy Support approach for engaging national stakeholders.

There are efforts around the world that seek to enhance the engagement of the various stakeholders within the SDGs program. For example, the UN Global Compact (UNGC), through the SDG Compass [2], is developing tools and guidelines to support the private sector on aligning business strategies and operations towards the achievement of the SDGs. The UNGC is also working on developing frameworks and tools to support reporting on the SDGs by the private sector companies. At the intergovernmental level, the High-Level Political Forum (HLPF) on Sustainable Development facilitates broad engagement and participation through the avenue of the Major Groups and other Stakeholders (MGoS) which currently has representation from the following nine stakeholders [3]: business and industry, children and youth, farmers, indigenous people, local authorities, non-governmental organizations, scientific and technological community, women, and workers and trade unions.

Engaging citizens for sustainable development and within the Global Agenda is not only motivated from the "localizing the SDGs" perspective, but it is also given impetus primarily from the following core principles of the 2030 Agenda for sustainable development:

- Human-rights based approach
- Leaving no one behind
- Preserving the planet and inter-generational responsibility
- Integrated nature and equal importance of economic, social, and environmental pillars
- Inclusivity, solidarity, and participation
- Transparency and accountability

These principles place individual citizens squarely at the center of the 2030 agenda for sustainable development, not only as beneficiaries of the development outcomes, but also as responsible agents for the achievement of the goals, as partners within the SDG programs, and as custodians of the developmental aspirations. Further, in the context of SDG indicators monitoring, these principles emphasize the position and role of citizens as more than just subjects that provide indicators data, but more importantly as partners actively engaged in monitoring progress on the goals and holding their governments accountable towards the achievement and the enjoyment of the goals.

This report considers the case of the engagement of citizens within the SDG agenda. However, it does so specifically, by highlighting the considerations from the citizens' perspective and discussing the key issues associated with their engagement and participation. The first section "Leaving them behind – data marginalization and exclusion" starts by focusing on the principle of leave no one behind (LNOB) - considering how individuals get left behind in social indicators data. This section expounds on the notion of data marginalization and exclusion and presents a typology that provides support for the operationalization of the LNOB principle in indicators data.

The report then presents findings from case study research that further highlights the need for concerted awareness-raising around the SDGs; the importance of community-driven indicators framework to facilitate increased localization and buy-in at the community and individual level; and the importance of citizen-centric indicators data collection, processing, and sharing. Lastly, the role of civil society organizations, in particular, community-based organizations (CBOs), for brokering and intermediating the meaningful engagement of citizens and vulnerable population groups is explored and discussed through the case study of one such CBO.

LEAVING THEM BEHIND

Data marginalization and exclusion

LEAVING THEM BEHIND

One of the explicit beneficiaries of the development imperatives identified in the 2030 Agenda for Sustainable Development are individuals, and this focus on the individuals is further emphasized through the pledge to 'leave no one behind.' This principle of broad inclusion and participation is motivated from the findings of the assessment of the Millennium Development Goals (MDGs) program, where despite major achievements being made across the eight MDGs, recognition was also given to the fact that various population groups and individuals across the world did not enjoy the benefits of the development achievements. These groups and individuals are generally clustered as the 'marginalized' or the 'excluded' and they not only represent groups that usually experience marginalization and exclusion as part of their everyday societal experience, they also face marginalization from accessing the very resources and developmental opportunities that can expand their capabilities towards development and improved livelihood.

Marginalization occurs across the full spectrum of international and national development activities - from planning and implementation to monitoring and evaluation. There has been a general tendency to cluster the different types of marginalization as a monolith, without unpacking the diversity and variety of the forms of marginalization that individuals experience. This not only constrains the discourse on marginalization, but it also limits the effectiveness of interventions for addressing marginalization and for mitigating the effects of exclusion. However, more pertinently, it hampers the operationalization of the "leave no one behind" principle.

Marginalization (n):
the processes that lead to individuals being excluded from participation within the wider societal life.

Exclusion (n):
the outcome of marginalization and the state of being disadvantaged in societal engagements.

Data marginalization occurs in cases where, through the use of data and the associated data analytics, individuals' meaningful participation in the broader societal life is limited and constrained. In the context of the SDG program, data marginalization can occur across the full length of the indicators data value chain, from data collection to impact (Figure 1). Addressing marginalization in social indicators is important because "as a society, we care about what we measure, we use what we measure, and what we measure drives policies and society in a particular direction"[4]

COLLECTION   PUBLICATION   UPTAKE   IMPACT

Figure 1 - Social Indicators Data Value Chain [16]

In this section, the notion of data marginalization and exclusion is expanded to highlight five distinct types of marginalization - to show how individuals within these categories experience marginalization and exclusion. The metaphor adopted for this classification is to consider the marginalized and excluded populations as "voices" that within normative development processes would be listened to, heard, and engaged. These different types provide a typology that should enrich the discourse on data marginalization, and it should also facilitate effective and targeted interventions to address the specific forms of marginalization. The development of this typology followed a multidimensional clustering approach – which is presented in Annex A. The five types of data marginalization formulated in this typology are unknown voices, silent voices, muted voices, unheard voices, and ignored voices.

THE UNKNOWN VOICES

The unknown voices are the population groups that are unknown to the social indicators data collecting entities. Around the world, various peoples fall into this category, whose social wellbeing realities are thus excluded from the social indicators data. Two broad categories within this population are the unknown unknowns and the known unknowns. The existence and characteristics of the latter group are more readily defined than the former group, and this includes the isolated and untouched communities, concealed individuals, and in some cases, victims of modern-day slavery.

ISOLATED AND UNTOUCHED COMMUNITIES
– there remains population groups that are completely isolated and who have no contact with the wider and mainstream societies. The so called "untouched" tribes fit in this category.

MODERN DAY SLAVES – there are
approximately 40.3 million people who are in different situations of modern-day slavery, including various forms of forced labor, trafficked individuals, and sex slaves. These groups of individuals face extreme levels of abuse and marginalization, and they usually lead lives that are removed from mainstream society. While some of these groups are engaged and actively included in social development programs, there are forms of enslavement that completely exclude and isolate individuals.

40.3 million
people in modern-day slavery around the world

Guilherme Gnipper Trevisan/Hutukara

CONCEALED INDIVIDUALS – one of the groups that get uncounted are individuals who are concealed. The reasons and motivations for concealment include: wanting to protect resources (e.g., illicit income), membership to violent gangs, protecting recipients of social welfare and assistance, drug dealing, and undocumented immigrants.

CASE #1

An ethnographic study of the behavioral causes of census undercounting among migrant Hispanic workers in a small rural farming town in Oregon – USA, found numerous factors that led to concealment and being uncounted [16]. The fear of government officials, and the immigration authorities specifically, was identified as one of the key reasons in the community for not participating and not being engaged in censuses. Due to the constant fear of arrests, widespread secrecy surrounded many of the households. Further, the complex and ad hoc living arrangements, also meant that none of the usual assumptions about households held true for this community. These factors led to many individuals being uncounted and to the reality of the living situation within this community being unknown and misunderstood.

GROUPS ACTING IN RESISTANCE TO GOVERNMENT – these are groups who either actively or passively resist government (and external) interventions. This resistance usually is due to mistrust of the government and the associated social development programs, or due to ideological commitments that are hostile to governments and external entities.

THE SILENT VOICES

The silent voices are the individuals who, due to personal factors, lack the capacity for vocalizing - so that while their objective wellbeing could be observed and noted, their subjective wellbeing is difficult to ascertain. These population groups include the destitute, the weak and vulnerable, children, and the persons with Intellectual and Developmental Disability (I/DD), who are generally isolated and restricted from participating in public societal domains or from being engaged politically.

It is worth noting that while this population group would be silent primarily to conventional methods of social indicators data collection and monitoring, it is possible to encourage engagement through targeted and specialized interventions. As an example, the social wellbeing of the I/DD individuals might not be easy to establish through standard instruments of data collection, however, through the use of assistive technologies, these individuals could overcome the barriers and limitations to their full participation and engagement.

WEAK AND VULNERABLE – vulnerable people, such as the homeless living on the streets or the aged living in institutions, are at the risk of being left behind when they are limited from expressing and contributing to social indicators data.

PERSONS WITH DISABILITIES – the World Health Organization has estimated that approximately 15% of the global population are people with disabilities. Various forms of disabilities place functional and cognitive limitations on individuals' ability to express and articulate their wellbeing – especially through conventional tools of data collection (e.g., census, questionnaires, interviews, self-reporting).

15%
of the world population are people with disabilities
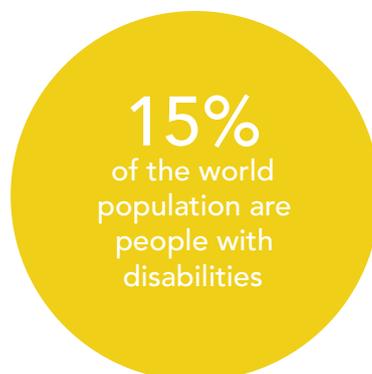
CHILDREN – at different stages of their lives and in different national contexts, children face constraints that limit their visibility in social indicators data. In their report on achieving the SDGs for children [17], the UNICEF notes that over half a billion children around the world are at risk of being unheard and unseen.

THE MUTED VOICES

The muted voices are the population groups that within specific societal contexts are marginalized and excluded. In this case, the marginalization is attributable to factors such as social norms, societal values, and practices. Examples of these population groups include LGBTQIA communities, stigmatized individuals, individuals who are considered at the bottom of the social ladder - such as the Dalits and untouchables, low-skilled migrant workers, and refugees. The nature of the marginalization of these individuals is that they are suppressed, and 'muted' through various structural instruments, resulting in them not being able to participate fully in the political processes and social development programs.

The aspect of being 'muted' is very context specific - in different socio-cultural and political contexts the same individuals would experience different levels of marginalization. For example, the extent to which domestic violence against women is noted and addressed will differ significantly based on the societal context. In one society, it could be normalized as part of the social fabric of that society, and in a different society, it could be considered as a violation of basic human rights and therefore addressed accordingly.

LGBTQIA COMMUNITY – the levels of acceptance of the LGBTQIA community in mainstream society varies significantly around the world – from extreme state-sanctioned persecution to full legal recognition and integration. This affects and determines the level to which they are given a voice and visibility in social indicators.

1 in 3
victims of family violence is male, according to 1IN3, an Australian campaign to address domestic violence against males

WOMEN – despite the general awareness of the injustices of sexism and gender discrimination, in many parts of the world and in many sectors of society, women still face levels of marginalization that hamper their full and meaningful participation in the life of a society. By extension, they are also limited from expressing their social wellbeing and from being counted.

STIGMATIZED GROUPS – around the world individuals are stigmatized for a variety of reasons. The stereotypes, scorn, ostracization, and prejudice that these individuals face limits their engagement and participation in societal life. Examples of some of the stigmatized and therefore muted voices include people living with HIV and male victims of domestic violence.

LOW-SKILLED MIGRANT WORKERS AND REFUGEES – migrant workers, in particular, low-skilled migrants, and refugees are often relegated to the periphery in host countries where their position is considered demeaning and less desirable. The International Labor Organization (ILO) estimates that there are 150.3 million migrant workers in the world, of which 11.5 million are migrant domestic workers [5].

John Ferguson/Oxfam

THE UNHEARD VOICES

These are the population groups that are excluded in the sampling or the data collection phase of social indicators monitoring. This exclusion can also be attributed to, among other factors, the capabilities of the data collecting agencies. Due to constrained resources, monitoring social indicators may be limited to 'low-hanging fruit,' for example, based on convenience sampling. Examples of population groups that face this marginalization include:

DIGITALLY UNCONNECTED – increasingly the advantages of using electronic and online survey tools mean that more data collection is done online or using electronic instruments. The convenience and efficiency of these instruments means that the people and groups who are digitally unconnected get excluded.

ILLITERATE – the population groups who are not able to read and write get excluded from engaging with any written data collection instruments – such as self-reported surveys.

LANGUAGE EXCLUDED – closely related to the exclusion of illiterate populations includes all forms of marginalization pertaining to language limitations.

ECONOMICALLY EXCLUDED – social indicators, including proxy indicators, are increasingly being derived and determined from individuals' economic activity. For example, using banking transaction records to determine income levels, or using expenditure on mobile phone services as a proxy for income. For cases where this electronic transactional data is utilized for social indicators monitoring, individuals who are not part of the mainstream economy become excluded.

4 BILLION
According to the World Economic Forum, 4 billion people are still unconnected due to lack of infrastructure, lack of affodability, poor skills, awareness and cultural experience, and low local adoption and use.

THE IGNORED VOICES

The ignored voices are characteristically excluded and marginalized during the analysis of the collected data. The traditional methods of statistical analysis and the focus of reporting on social indicators have always had as the focus the aggregate phenomena typically at the meso and macro levels. This has meant that in the reporting, detailed information and individual specifics are lost due to aggregate statistical analyses. This loss of detail and aggregation of individual data points is an example of traditional statistical challenges which are well understood and have been documented. There is, however, another form of exclusion and marginalization that is linked to new approaches of analysis associated with what is generally termed Big Data. Social indicators monitoring is set to increasingly make use of the various forms of Big Data and the associated analytics to derive relevant insights about the various social phenomena.

TRADITIONAL STATISTICS - one of the mechanisms through which marginalization can occur in statistical analyses is through the well-known effects of aggregation bias and ecological fallacy. Ecological fallacy occurs when ecological correlations, which are correlations at an aggregate or group level, are assumed to be valid substitutions for individual correlations. Through the use of ecological level analyses, the social wellbeing aspects of individuals can be misunderstood, misrepresented and ignored.

CASE #2

The implication and impact of aggregation bias is illustrated in case of countries where notwithstanding the improvements in the recorded development indicators, many individuals do not experience the developmental benefits. For example, Geiser notes situations where despite the improving development indicators, vast and increasing differences were identified between minority populations living in urban centers and the vast majority living in rural areas [18]. In this case the aggregation to arrive at the national level indicators 'ignores' the minority populations whose wellbeing and development would require attention and intervention. The underlying distributional realities which need to be highlighted and emphasized especially in development activities, instead get 'ignored' and concealed through this 'tyranny of averages' [19].

NEW DATA APPROACHES - the advances made in the broad field of Big Data analytics and Machine Learning represent an opportunity and a potential for social indicators monitoring - by deriving insights from data that is characterized by its high volume, variety, velocity, value, and veracity. This data is collected passively from individuals' online activities, social media information, crowd-sourced data, and from Internet of Things (IoT) data. In considering the impact of Big Data on social indicators monitoring, it is essential to understand the marginalization aspects that are attributable to the underlying data, and also to the process of data analysis.

Fundamentally Big Data exhibits a phenomenon of selection bias, which is well known from large datasets. The impact of the increased selection bias extends to the subsequent analysis methods including data models that rely on correlation.

Secondly, the processing undertaken in data analysis makes use of advanced computational algorithms. The increasing autonomy and opaqueness of algorithms and growing proliferation of their use not only in social indicators monitoring but also in other societal activities means that they have increased potential for unchecked marginalization and exclusion. This is mainly linked to the fact that these algorithms are susceptible to inadvertent bias, advancing and reinforcing historical discrimination, and favoring specific views and perspectives.

The decisions that are made by algorithms have consequences on societies and individuals that are not fully understood, and that are not always transparent. The individuals who are adversely affected by the outcomes of these algorithms and those who get unrepresented or misrepresented by the data models, face a subtle and yet increasingly prevalent form of data marginalization.
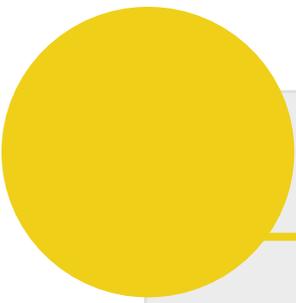
SO WHAT?

These different forms of marginalization demand concerted effort to guarantee the pledge of the UN General Assembly to leave no one behind. Clear strategies, along with the collection of 'high-quality data disaggregated by key dimensions that are relevant to the specific national context', need to be further formulated, agreed to, and put into action.

The following are policy-relevant recommendations on mitigating the forms of marginalization noted in this section:

THE VILLAGE PERSPECTIVE – from the old African adage, 'it takes a village to raise a child', the corollary 'it takes a village to leave no one behind' follows. The nuanced understanding of social phenomena is lost when the focus on social indicators monitoring is only to inform policy and action for national-level and multilateral stakeholders. The attention that is required for engaging some of the most vulnerable and marginalized population groups is sometimes only accessible to community level actors, such as non-governmental organizations, community-based organizations and civil society organizations. Therefore, explicit partnerships, frameworks and an intention for social indicators to support grassroots action are needed to ensure that community actors are sufficiently capacitated and therefore enabled to contribute effectively to the sustainable development agenda.

ALTERNATIVE DATA APPROACHES – alternative data approaches, including the use of perception data, thick data, small data, and micro data, stand to mitigate the limitations and weaknesses of current data approaches and to highlight situations of data marginalization experienced by the unknown, silent, and ignored voices.

FAT ALGORITHMS AND DATA MODELS – with the increasing incorporation of sophisticated data modeling algorithms into social indicator monitoring processes, it is paramount that checks and balances are put in place to ensure that the algorithms are FAT - Fair, Accountable and Transparent. In doing so this would ensure that they don't inadvertently perpetuate data marginalization of the ignored voices.

AWARENESS AND ADVOCACY – the plight of the marginalized population groups needs to further be highlighted and brought to the fore through effective awareness campaigns and advocacy action. This strategy can effectively be used for challenging and engaging the context-specific socio-cultural norms and values which perpetuate marginalization of the silent voices and the muted voices.
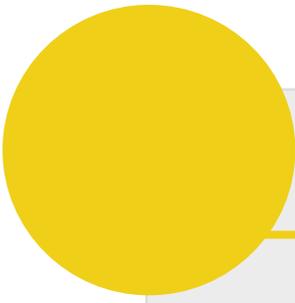
MULTI-LEVEL ANALYSIS – allows for the analysis of data to be undertaken not only at the macro aggregate level, but also at the micro and meso levels. This strategy potentially addresses the marginalization and exclusion that is experienced by the ignored voices.

PROXY INDICATORS – proxy indicators give insight on phenomena to which indicator metrics are lacking. In cases where individuals' social wellbeing phenomenon would be ignored due to lack to accurate indicators, relevant proxy indicators can be utilized, thus ensuring

CITIZEN AWARENESS

Precursor to engagement

CITIZEN AWARENESS

One of the reasons and indicators of the missed opportunity for engaging individuals within the Millennium Development Goals (MDGs) program was the widespread lack of awareness of the MDGs and the associated program activities by the average citizen. Without the awareness of these global development goals, collective action, community mobilizing and civic engagement towards the achievement of these goals was hampered. Raising awareness and sensitizing citizens about the sustainable development goals remains a critical precursor to their engagement and collective participation.

- I don't know the SDGs - 62.8%
- I have heard about the SDGs - 19.2%
- I am generally aware of the SDGs - 11.1%
- I am very aware of the SDGs - 4.5%
- I know all the 17 SDGs in detail - 2.5%
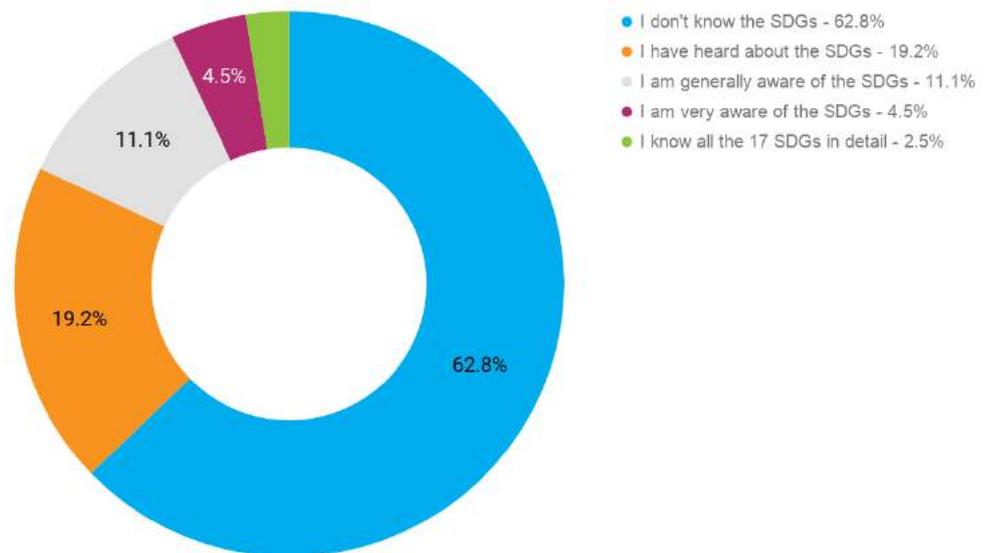
Figure 2 - Awareness of the Sustainable Development Goals

62.8%
of respondents do not know and have no awareness of the Sustainable Development Goals

A survey (further details in Annex B) was undertaken to explore the awareness levels and attitudes of individuals towards the SDGs. A staggering 62.8% (n = 1170) of respondents indicated that they did not know and were not aware of the SDGs. Only 19.2% (n = 357) had heard about the SDGs, and a further 11.1% (n = 206) expressed a general level of awareness of the SDGs (Figure 2).

Notwithstanding the high lack of awareness of the SDGs, in their official UN formulation, the goals that have been articulated generally hold strong significance and relevance for individuals.

Figure 3 shows the evaluation, on a 5-star rating scale, of the importance that individuals ascribe to the different goals. For this evaluation the short narratives of the SDGs were utilized, for example, for SDG1 the question was framed as "Rate the importance of the goal to 'end poverty in all its forms everywhere'."
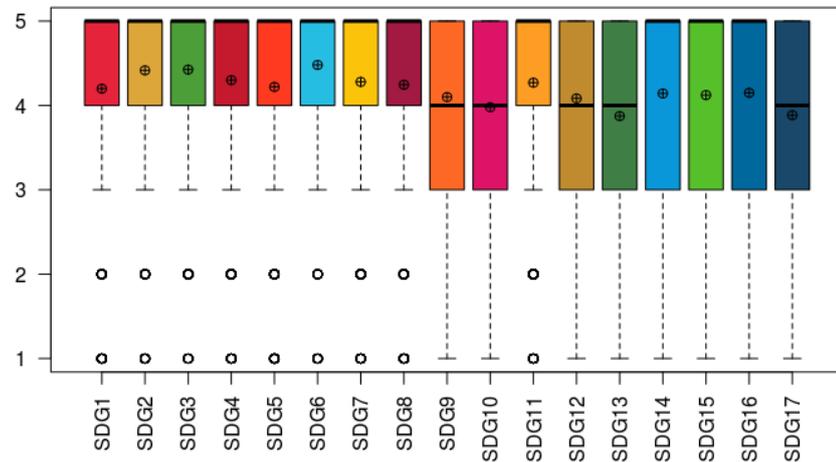
Figure 3 - Relevance and Significance of the Sustainable Development Goals

The findings from the survey illustrate the clear potential and opportunity of engaging citizens towards the achievement of the SDG imperatives, due to the strong sense of importance and relevance that individuals ascribe to the goals.

Beyond identifying the strong relevance of the SDGs, the survey also undertook an inquiry into the extent to which various stakeholders (i.e., the government, community-based organizations, non-governmental organizations, educational institutions, citizens/individuals, private industry and corporations, and international community) are ascribed with responsibility for the achievement of the SDGs. On a 7-point scale, from 1 (low importance) to 7 (high importance), overall the respondents expressed the broad responsibility and the vital role to be played by all the stakeholders – with the highest responsibility (at the mean rating of 5.07) for the achievement of the SDGs being ascribed to Citizens/Individuals (Figure 4).

Figure 4 - The Role of Various Stakeholders within the Global Agenda

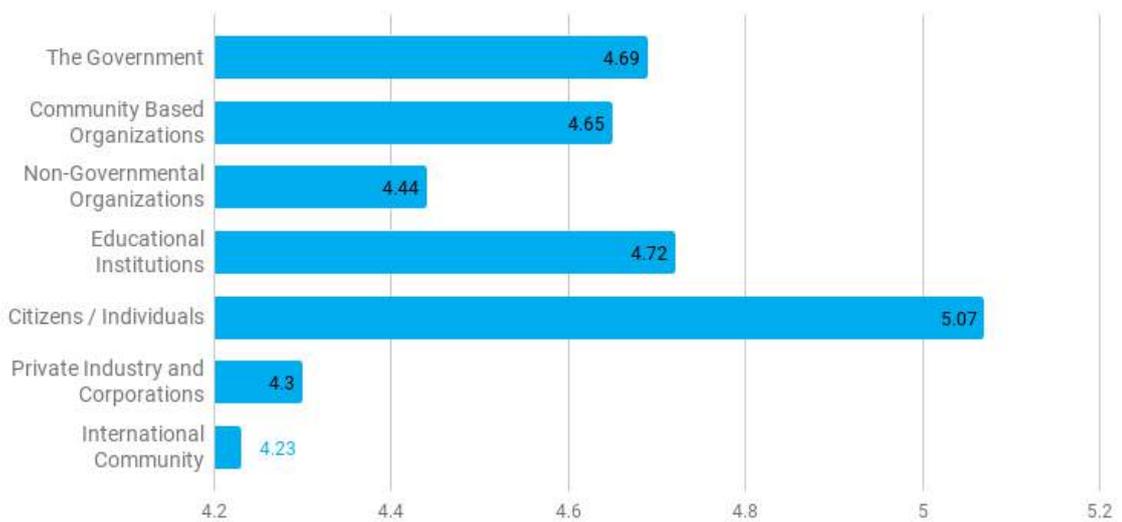The engagement and participation of individuals within the Global Agenda is grounded in the principles upon which the agenda is conceptualized – such as, human rights approach, inclusivity and participation, and leave no one behind. However, beyond these underlying principles, individual citizens' agency drives and motivates their participation and engagement towards the achievement of the kind of the lives and the kind of future that they have reason to value.

SO WHAT?

As primary agents of their own development and wellbeing, the engagement and participation of citizens and community-level actors within the Global Agenda is invaluable and can be leveraged for greater action, monitoring, and impact.

There is, however, currently a high lack of awareness of the agenda for Sustainable Development Goals, and consequently of the programs and actions around the SDGs. Without the awareness of the SDGs, participation by citizens and community-level actors will remain hampered and limited. The need for extensive awareness-raising around the 2030 agenda for sustainable development remains as a precursor to participation and engagement.

The citizens' agency and recognition of the importance and relevance of the Global Agenda goals stands as further motivation for facilitating their active engagement and participation towards the achievement of the SDGs.

SOCIAL
INDICATORS

The case for community-driven
indicators framework and community-
based monitoring

"Because economic, social, and cultural right are rights of the individual, the general scope of many indicators may preclude their direct application, other than for obtaining a general overview of a particular situation" [20]

SOCIAL INDICATORS

Social indicators monitoring, which traditionally has custodianship within the National Statistics Offices (NSOs), serves to inform the public and policymakers by providing statistics regarding social well-being phenomena. The failure of social indicators to have an impact on developmental imperatives has been attributed to [6]: having a descriptive approach that isn't situated within a clear developmental theory or conceptual framework; confining social indicators monitoring to the production of statistics without a larger plan of action that links to policy and action impacts; not having a democratic indicators program with a clear public participation process. Further, it has been noted that the scoping of the social indicators, both at the macro level and from the obligation perspective precludes their direct application to individuals' development.

The focus of human development monitoring has traditionally been for informing strategy and action for national level and multilateral stakeholders, with a minimal emphasis on enabling and supporting local community-level and grassroots action. This invariably means that the lenses of data analysis and reporting are typically skewed towards identifying the macro phenomenon as opposed to the small-scale micro details in the data. The framing and articulation of the SDG indicators, while relevant and necessary for government-level Planning, Monitoring, and Evaluation (PM&E) processes, limits the potential for direct contribution by citizens and community-level actors to the social indicators monitoring.

Obligation perspective

Indicators framed for determining the extent to which governments are complying with their obligations to the citizens under human rights law.

This limitation is because, despite the alignment of the SDGs with the developmental aspirations and goals of the individuals, individuals and communities employ a different set of indicators in their everyday life for monitoring their sustainable livelihood and wellbeing. Thus, the macro-level aggregate posturing represents a potential overall risk and limitation to the SDGs indicators monitoring effort. It is a risk that directly affects the very principle of "leave no one behind" which is core and foundational in the articulation of the SDGs

The interplay between the SDG indicators and the individuals' indicators data towards the goals has been explored through a case study of SDG3 - to "ensure healthy lives and promote well-being for all at all ages", which is made up of 13 targets with 26 accompanying indicators (Table 1). Refer to Annex B for further details on this case study
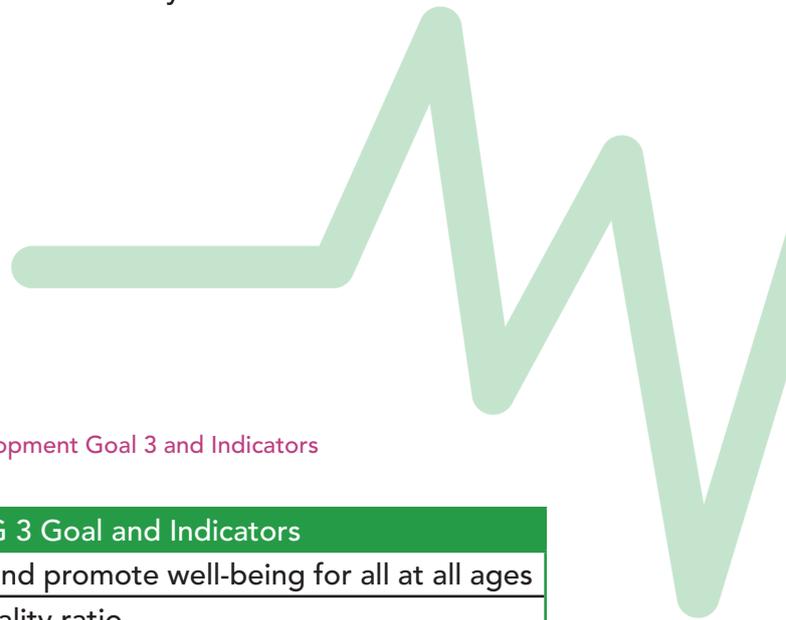
Table 1 - Sustainable Development Goal 3 and Indicators

SDG 3 Goal and Indicators	
Ensure healthy lives and promote well-being for all at all ages	
3.1.1	Maternal mortality ratio
3.1.2	Proportion of births attended by skilled health personnel
3.2.1	Under-five mortality rate
3.2.2	Neonatal mortality rate
3.3.1	Number of new HIV infections per 1,000 uninfected population, by sex, age and key populations
3.3.2	Tuberculosis incidence per 1,000 population
3.3.3	Malaria incidence per 1,000 population
3.3.4	Hepatitis B incidence per 1,000 population
3.3.5	Number of people requiring interventions against neglected tropical diseases
3.4.1	Mortality rate attributed to cardiovascular disease, cancer, diabetes or chronic respiratory disease
3.4.2	Suicide mortality rate

SDG 3 Goal and Indicators (continued)	
Ensure healthy lives and promote well-being for all at all ages	
3.5.1	Coverage of treatment interventions (pharmacological, psychosocial and rehabilitation and aftercare services) for substance use disorders
3.5.2	Harmful use of alcohol, defined according to the national context as alcohol per capita consumption (aged 15 years and older) within a calendar year in litres of pure alcohol
3.6.1	Death rate due to road traffic injuries
3.7.1	Proportion of women of reproductive age (aged 15-49 years) who have their need for family planning satisfied with modern methods
3.7.2	Adolescent birth rate (aged 10-14 years; aged 15-19 years) per 1,000 women in that age group
3.8.1	Coverage of essential health services (defined as the average coverage of essential services based on tracer interventions that include reproductive, maternal, newborn and child health, infectious diseases, non-communicable diseases and service capacity and access, among the general and the most disadvantaged population)
3.8.2	Number of people covered by health insurance or a public health system per 1,000 population
3.9.1	Mortality rate attributed to household and ambient air pollution
3.9.2	Mortality rate attributed to unsafe water, unsafe sanitation and lack of hygiene (exposure to unsafe Water, Sanitation and Hygiene for All (WASH) services)
3.9.3	Mortality rate attributed to unintentional poisoning
3.a.1	Age-standardized prevalence of current tobacco use among persons aged 15 years and older
3.b.1	Proportion of the population with access to affordable medicines and vaccines on a sustainable basis
3.b.2	Total net official development assistance to medical research and basic health sectors
3.c.1	Health worker density and distribution
3.d.1	International Health Regulations (IHR) capacity and health emergency preparedness

Social indicators can typically be framed from two perspectives, the obligation perspective and the enjoyment perspective. From the obligation perspective the indicators are framed to establish the extent to which the government is fulfilling its obligation towards its citizens, and as such, these are typically articulated in terms of the provisioning of the structures and the supporting mechanisms for service provisioning to the citizens.

Enjoyment perspective

Indicators framed to establish the extent to which individuals are enjoying the benefits of development.

From the enjoyment perspective, the indicators are framed to assess the extent to which individuals are directly enjoying the benefits of development. Most of the SDG 3 indicators are framed from the enjoyment perspective, as they largely assess the health and wellbeing experience of the individuals, for example – illness prevalence (3.3.1, 3.3.1, 3.3.3) and alcohol use (3.5.2). The obligation based indicators that have been defined for SDG 3 are 3.1.2, 3.5.1, 3.7.1, 3.8.1, 3.8.2, 3.b.1, 3.b.2, 3.c.1, and 3.d.1. The following are further general observations regarding the SDG3 indicators (with the corresponding text formatting to highlight the associations in Table 1):

1. Of the 26 indicators, 9 of them are largely framed from the obligation perspective
2. Of the 17 indicators framed from the enjoyment perspective, 10 of them are for monitoring mortality or birth rates. While mortality and birth rates can give insights on the health and wellbeing of a community of people at an aggregate level, at an individual level, mortality or birth is not a relevant indicator for health and wellbeing.
3. Five of the indicators are associated with monitoring disease infections (i.e. HIV, TB, Malaria, Hepatitis B and tropical diseases)
4. Two of the indicators are associated with monitoring health related lifestyle choices i.e. alcohol use and tobacco use.

To further explore the interplay dynamic with the SDG3 indicators, the macro-level indicators were juxtaposed against micro individual-level indicators. A survey instrument was used to identify the data that individuals regularly use for monitoring their health and wellbeing. This was explored through an open-ended question that was framed as:

"What information and data do you use in your everyday life that you find relevant for your wellbeing?"

This question primarily sought to discover the data, metrics, and information that individuals are currently utilizing for their health and wellbeing indicators monitoring. However, beyond just highlighting the specific data around SDG3, the analysis also shed light on: the motivations and reasons why participants collect and monitor health and wellbeing data; the tools and techniques that they utilize for monitoring; as well as the dynamics between the motivations for monitoring and the type of data that is monitored.
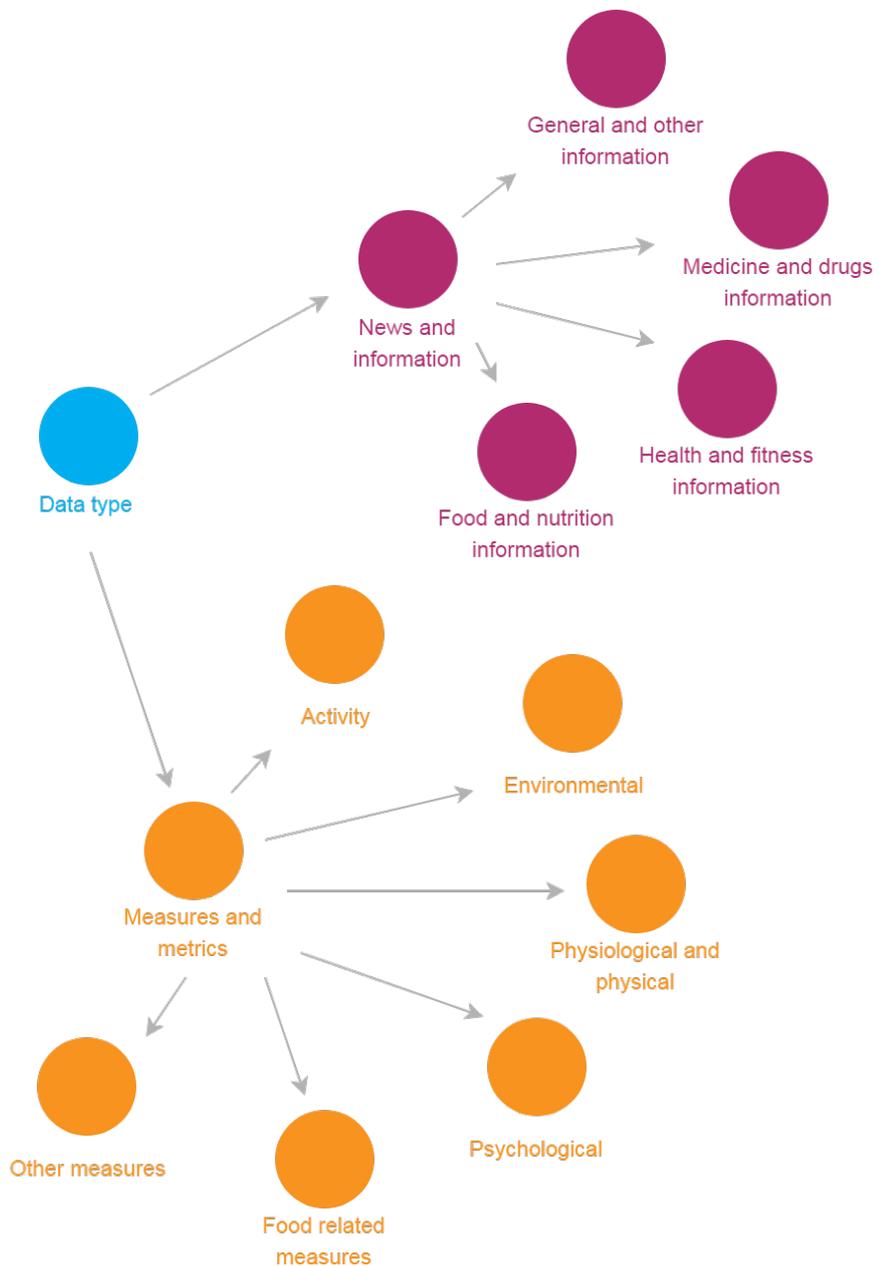
Figure 5 - Individuals Health and Wellbeing Indictaors Data and Information

From the analysis of the responses, the different types of data and information that individuals make use of to monitor their health and wellbeing were noted as highlighted in Figure 5. The main clusters of indicators that individuals make use of for their health and well-being are: activity – data related to individuals' physical activity, e.g., walking and exercising; environmental – external environmental phenomenon data, e.g., the weather; physiological and physical – measures associated with the biological functioning of the body as well to aspects of the physical body, e.g. weight, blood pressure; psychological – measures related to the emotional and mental wellbeing of the individuals; food-related measures – are associated with the food and drink intakes of individuals, e.g., sugar intake; and lastly other measures – which categorizes other measurement data that individuals noted and highlighted in their responses, e.g., economic indicators, work times and loads.

The following key observations have been made regarding the health and wellbeing indicators for that individuals use:

1. The data and information that individuals make use of mostly aligns with the enjoyment perspective evaluation of social indicators, and therefore regards the assessment and monitoring of the extent to which individuals experience and enjoy the specific health and wellbeing targets.

2. Individuals undertake the monitoring of their health and wellbeing at a finer granularity and much more nuanced level. For example, under the psychosocial measures and metrics, the following further indicators have been noted: Stress level, Mood, Happiness level, Emotions, and Levels of socializing. Some of these different and detailed metrics can be observed from the thematic analysis of the responses as captured in the associated psychological measures and metrics word cloud below (Figure 6).

Figure 6 - Psychological Measures and Metrics Wordcloud

3. Individuals make use of not only quantitative observational throughput data, but also qualitative perception data and data that is associated with small incremental and individual level changes.

4. The use of indicators data by individuals is nuanced and varied. Beyond using the health and wellbeing data for the basic planning, monitoring and evaluation associated activities, individuals also use data (among other things) to inform behavior change and for reaching new goals and improving their health and wellbeing (Figure 7).
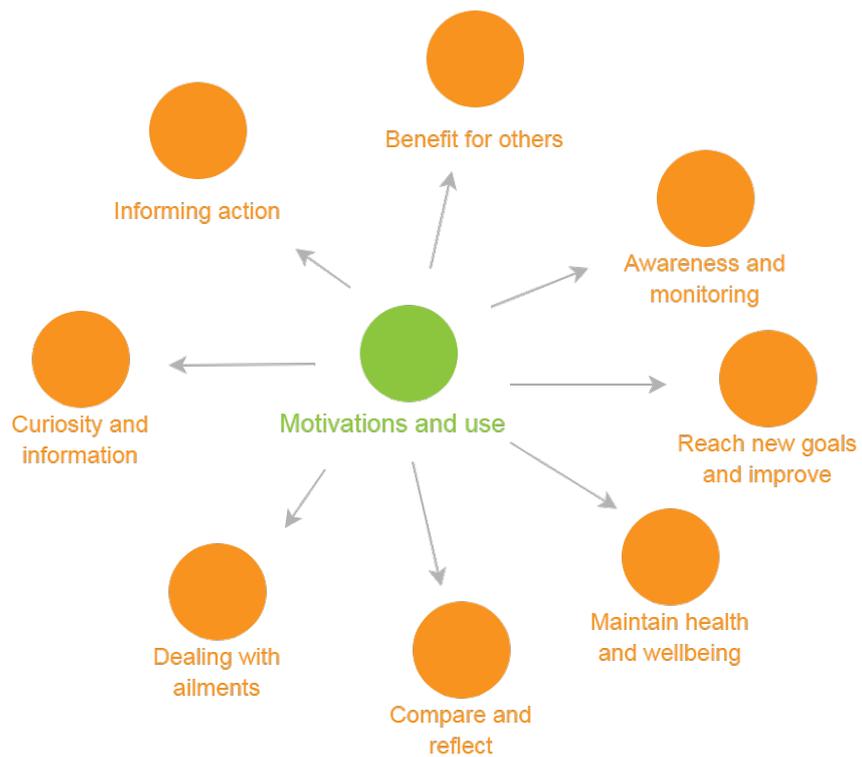
Figure 7 - Health and Wellbeing Indicators Uses by Individuals

The 2030 agenda for sustainable development, as an aspirational global vision of a possible future, requires and participation from the national level and multilateral stakeholders, as much as it does from non-governmental organizations, civil society organizations, communities, and citizens. Further, the engagement of these stakeholders, at their respective macro, meso and micro levels, requires an accompanying indicators framework that is relevant, meaningful and actionable to each of their realities. Having juxtaposed the SDG3 indicators with the indicators that individuals use for their health and wellbeing, the need for harmonizing interaction within the indicators ecosystems becomes apparent. Since the current SDG indicators framework mainly caters to the macro-level interests and perspectives of national level and multilateral stakeholders, community-driven indicators frameworks would cater to the micro-level, hyperlocal perspectives of the community-level actors and citizens and elevate their role in monitoring progress on the SDGs.

MobiSAM is a Mobile Social Accountability Monitoring project implemented in South Africa, that provides platforms, including technology-based, to enhance citizens' ability to voice their concerns with regards to service delivery and to hold their local government accountable to service delivery through citizen monitoring [21]. The platform allows for direct reporting on lack of service delivery (through issue tickets) and on locally-relevant indicators (delivered as polls) on issues such as water (SDG6), sanitation (SDG6), roads (SDG9), and electricity (SDG7).

Since its inception in 2013, the project has demonstrated the opportunities and challenges of engaging grassroots action, community organizing, and citizen monitoring towards the achievement of the locally relevant developmental aspirations and imperatives. The key lessons and insights from MobiSAM include [21]:

- Effective engagement in citizen monitoring needs to be grounded in the context of empowered citizens who understand their rights to services and the government responsibilities and obligations that those rights entail.
- Citizens' primary motivation for engagement in monitoring is for the direct utility that accrues to them – e.g. a solution to a service delivery issue.
- In the absence of strong political will and leadership, the role of 'trusted messengers' and intermediaries is important for supporting citizen engagement and participation.
- The embedding structural and political context directly influences the success and effectiveness of citizen monitoring and generally of citizen participation efforts.

The data revolution for sustainable development presents an opportunity for a revolution in social indicators monitoring which would amplify the role of citizens within the SDG program. In doing so, this would support social metrics and indicators to be acquired, analyzed and packaged to be actionable by citizens and community-level actors. There is an opportunity to establish frameworks and mechanisms for the active and explicit incorporation of citizens within the SDG agenda, thereby: leveraging grassroots action and collective community action towards the SDG goals and employing citizen-generated data and community-based monitoring to augment and supplement national social indicators datasets.
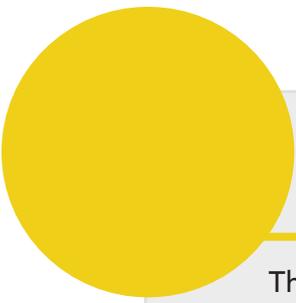
SO WHAT?

There is currently effort towards the localization of the SDGs, which fundamentally recognizes the importance of giving the SDGs a local relevance and sensitivity to contextual factors. Localizing the SDGs facilitates meaningful engagement of countries within the SDG program.

The current SDG framework defines 232 indicators across the 17 goals. These indicators are mostly geared towards identifying the aggregate metrics that capture the broad global dimensions and features of international development. As such the engagement of local and non-governmental stakeholders in the monitoring of progress towards the SDGs can be hampered by the disconnection between these 232 indicators and the "grassroots" community-level developmental metrics. Defining locally relevant indicators allows for a meaningful engagement in the SDG monitoring process, in a manner that is relevant, adds value, and creates local ownership.

Engaging citizens and facilitating their participation within the sustainable development agenda necessarily demand contextualizing the SDGs to their realities and situations. One of the critical steps towards aligning the SDGs with the contextual realities of individual and their communities, is through the formulation of community-driven indicators and the implementation of the associated community-based monitoring of indicators.

Elevation the role of the citizens within the social indicators data ecosystem, is an opportunity to:

1. Balance the current largely top-down and ecological (i.e. focusing on macro-level aggregate phenomena) approach to the SDG indicators monitoring.
2. Actively and explicitly engage citizens and community-level actors in action towards the achievement and the monitoring of the 2030 agenda for sustainable development.
3. Leverage and contribute to the larger goal of more empowered and engaged citizens.
4. Contribute novel and locally relevant data towards the SDG indicators framework.
5. Give a voice to individuals and population groups who are currently excluded and marginalized in social indicators monitoring.

It is an opportunity to embrace the Data Revolution for Development with potential to leverage the citizens' agency towards their own development and the development of their own communities.

CITIZEN GENERATED DATA

Data sharing and utility

CITIZEN GENERATED DATA

Individuals collect and make use of data for a myriad of purposes including awareness and monitoring, maintaining health and wellbeing, informing action, comparing and reflecting, and reaching new goals and improving their situation. Considered from the Personal Informatics perspective, the use of data is typically for gaining awareness and providing insights to the individuals about a phenomenon or a situation of interest. While deriving relevant insights from the collected data primarily ensues through the individuals' engagement with their data, research has found that individuals also engage in data sharing with others for sense making purposes. The collection and use of data by individuals therefore naturally comprises the social dimension.

In the context of the sustainable development data ecosystem or that of data-driven societies, sharing of personal data needs to be considered not only within individuals' own social circles but also with other stakeholders within the broader data ecosystem. The individuals' attitudes towards sharing their personal data within the sustainable development data ecosystem was explored, in terms of both the willingness of the participants to share their data with specific stakeholders in the health and social indicators monitoring sectors; as well as the factors that would inform their willingness to share (or to not share) their personal data.

On a continuous scale between 1 ("low willingness to share") and 7 ("high willingness to share"), the participants were most (mean 5.996) willing to share their personal health data with their doctors, followed by the willingness to share health and wellbeing data with their families (Figure 8). In the context of the SDGs, where the National Statistic Offices currently have custodianship of the collection of social indicators data, the slightly below average (mean 3.661) willingness of the individuals to share their health indicators data with the NSOs can be observed.
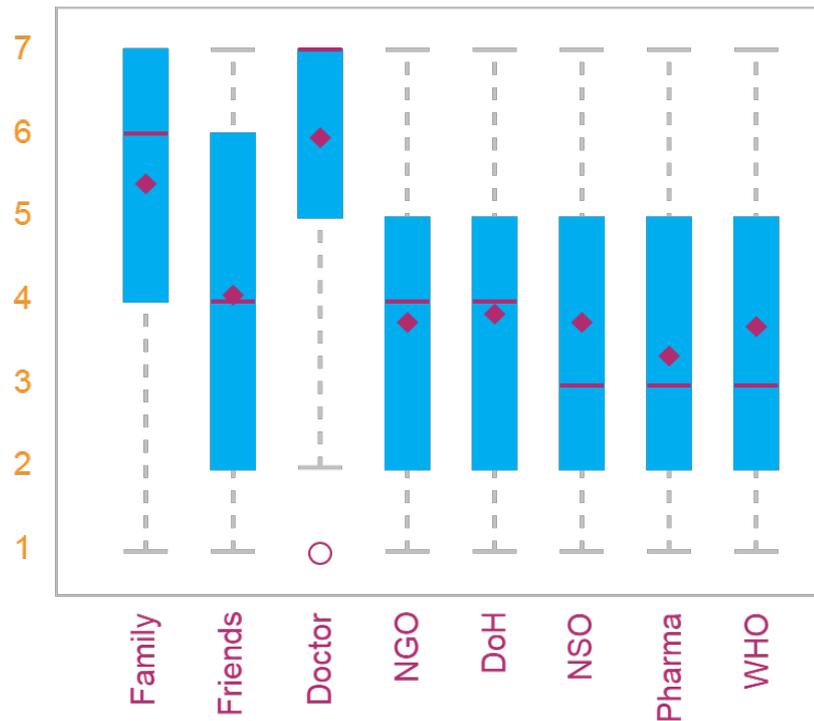
Figure 8 - Preferences for Data Sharing with Various Stakeholders

Further analysis was undertaken to understand how the participants' attitudes towards sharing their personal data correlate across the different stakeholders.
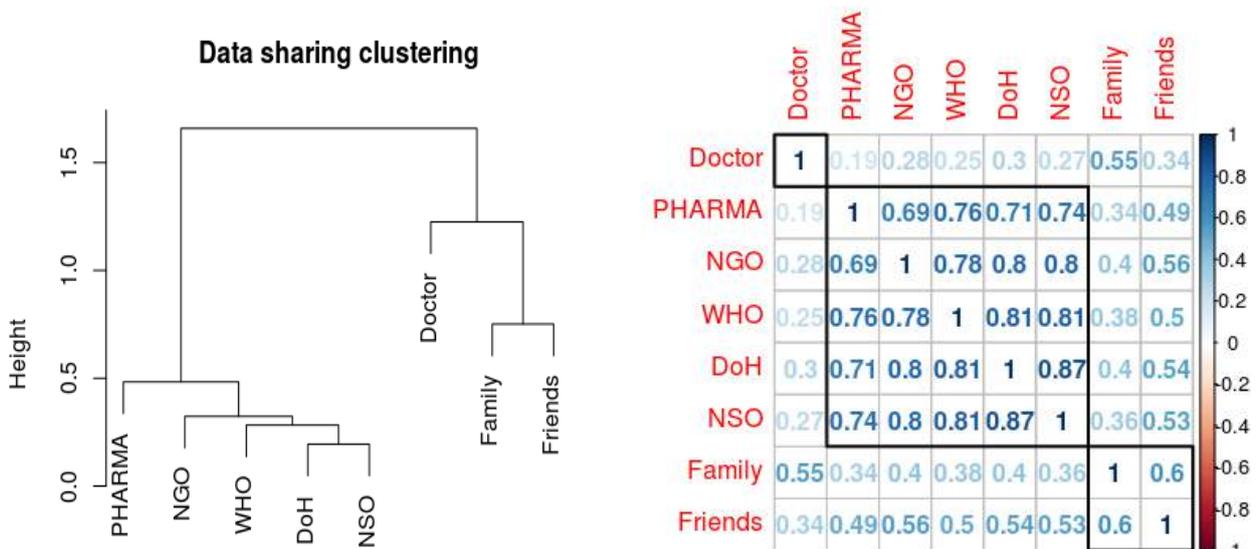

Figure 9 - Hierarchichal Clustering and Correlations for Data Sharing with Different Stakeholders

The correlation analysis and the subsequent clustering identified three groups of entities and stakeholders for data sharing with individuals (Figure 9). Table 2 lists the entities within each of the three clusters that have emanated from the analysis.

Table 2 - Data Sharing Stakeholders Clusters

Cluster #	Entities (mean)
Cluster 1	Doctor - Your doctor (5.996)
Cluster 2	PHARMA – A pharmaceutical company (3.38) NGO – A local NGO working on health issues (3.732) WHO – The World Health Organization (3.658) DoH – The Department of Health (3.86) NSO – The National Statistics Department (3.661)
Cluster 3	FAMILY – Your family members (5.354) FRIENDS – Your friends (4.102)

The first cluster consists of a single stakeholder, the individuals' doctor or medical provider. While not established explicitly in our research, it is evident that this cluster represents the stakeholder who is able to provide a health service for the individuals with the benefit accruing directly to the individual. The sharing of health and wellbeing data with this specific stakeholder, therefore, contributes to the direct enjoyment and achievement of the individuals' health and wellbeing goals.

The second cluster from the analysis comprises pharmaceutical companies, NGOs working on health-related issues, the World Health Organization, national department of health, as well as the national statistics department. The stakeholders in this cluster are characterized by being organizations within the wider health sector including governmental, non-governmental as well as international/multinational organizations. As far as sharing data with these organizations are concerned, it can be noted that the benefit that accrues to the individuals would mostly be indirect and not immediate. For example, sharing data with the national statistics department, could inform the reporting on national health and wellbeing and subsequently the development of relevant health policies, which would in time be of benefit to the individual. In the case of sharing data with the local NGO working on health-related issues, there are possible scenarios where the resultant benefit could accrue more directly to the individuals.

The third cluster identified from the analysis above comprise entities who share high social proximity and relational coupling with the individual. The sharing of data with these stakeholders is typically for sense making and social-support.
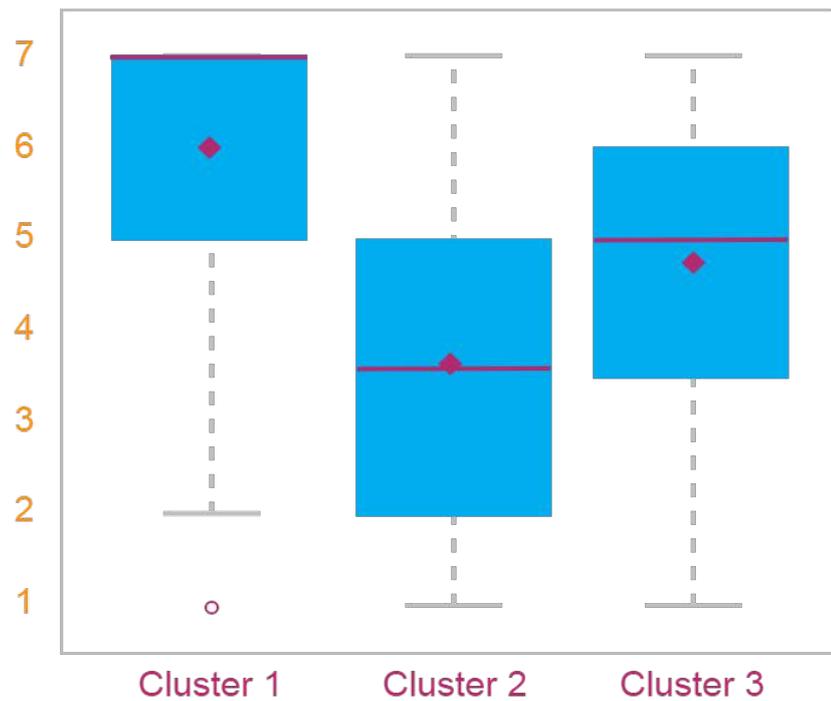
Figure 10 - Stakeholder Clusters for Data Sharing

The three clusters observed in this analysis are statistically significantly different, and from this, it can be concluded that the individuals' attitudes and perceptions regarding data sharing with various stakeholders within the data ecosystem vary significantly per cluster (Figure 10). Furthermore, sharing data with stakeholders that provide direct benefit and that support the direct enjoyment of the health and wellbeing goals (i.e., the doctor) has the highest ranking, followed by sharing data with friends and family typically towards social sense making and social support. Lastly, sharing data with other organizations and stakeholders within the data ecosystem (i.e., cluster 2) has the lowest ranking – this finding has implications for further consideration on engaging individuals and community level actors in the monitoring of personal social indicators for official purposes (i.e., Planning Monitoring and Evaluation). These results give empirical evidence to what would be the expectation with regards to personal data sharing considerations.
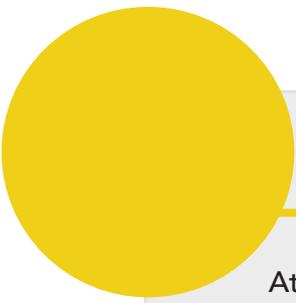
SO WHAT?

At the center of data-driven societies are individuals and citizens who not only generate data, but who should also benefit from the outcomes of data-driven development. The motivation from the Personal Informatics perspective, for people to collect information about themselves is usually to advance self-knowledge, self-insight and to promote positive attitudes and behaviors. Notwithstanding the various motivations for the collection of data and tracking of indicators, the following are the key observations.

DATA UTILITY

There is an indication of a willingness for individuals to share their indicators data (i.e. generated and collected by them) primarily in cases where the utility and benefit of sharing the data accrues to them. As such in the case of health and wellbeing data, individuals are happy to share their data primarily with their health service providers, presumably because this data be used towards their direct and immediate health benefit. Thus, the various engagements (i.e. policies, programs, projects) around citizen-generated data need to primarily prioritize the direct benefit and impact on the citizens (i.e. citizen-centric data valorization).

SOCIAL SENSEMAKING AND COLLECTIVE CAPABILITIES

The role of individuals' social network is an important component of the utilization of social indicators data for the purposes of social sensemaking and for social support. Leveraging these social and community networks also has the potential to amplify the collective capabilities of communities around sustainable development. Along with considering individually generated citizen data, there is potential to leverage the insights from collective and community data.

DIFFERENTIATED DATA SHARING

At one level the findings from the research point to the need for differentiated data sharing arrangements with entities within the data ecosystem, but also point to the willingness of the participants to consider sharing their indicators data across the ecosystem. The use of data collected from individuals however needs to be grounded in the understanding of individuals' preferences. Thus, the various data collection, aggregation, curation, and processing entities within the data ecosystem have unique, non-transferable data sharing requirements from and commitments to citizens.
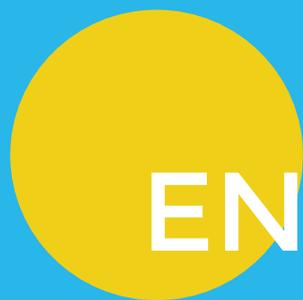
ENGAGING CITIZENS

On their own terms

ENGAGING CITIZENS

The evolution of the social indicator data ecosystem will see participation and engagement between increasingly diverse stakeholders, in various roles including data collection, aggregation, curation, and processing – with the NSOs playing a facilitation and coordination role. The position of individuals within this ecosystem remains central, as the primary subjects of indicators data and the primary beneficiaries. Mortier et al. have suggested that, with the growing ubiquity and pervasiveness of data, there is a need to explicitly consider Human-Data Interaction (HDI), which they define as being constituted of three key themes and domains [7]: legibility, agency, and negotiability. Legibility regards ensuring the comprehensibility of data and the associated algorithms, so that the individuals are aware of their data, and the implications of its use; agency in this context is about allowing individuals the freedom and capacity to act within the data ecosystems; negotiability concerns the dynamic relationships that emanate from the individuals' interaction with the data. These three themes provide a suitable initial framing of the critical concerns for consideration in the role and engagement of citizens in indicators data ecosystems.

Beyond understanding the individuals' attitudes towards sharing data with specific stakeholders (as discussed in the previous section), it is therefore also important to understand the terms on which individuals can be engaged, as well as the factors that affect the individuals' willingness to be engaged towards data sharing. This inquiry was framed in the research survey as "How much would the following factors influence your willingness or lack of willingness to share your personal health information?" and the participants' rated ten (10) factors on a continuous scale from 1 to 7, at "low influence" and "high influence" respectively.

The ten factors for this question were informed by the following considerations: factors associated with privacy (n = 5), factors related to the personal benefit of sharing data (n = 1), and factors associated with external use within the wider data ecosystem (n = 4). Privacy is a pertinent issue in the discussions around the collection and sharing of personal information and in wider discourses on freedom and democracy. Solove, in his publication on "conceptualizing privacy" offers a bottom-up perspective to thinking about privacy, taking into consideration the contextual factors and the dynamic nature of privacy, instead of defining privacy in singular, universal and abstract terms [8]. He notes the following six conceptualizations of privacy, which have been incorporated in this inquiry:

1. The right to be left alone
2. The ability to shield oneself from unwanted access by others
3. Secrecy and the concealment of certain aspects from others
4. Control over personal information
5. Personhood and the protection of one's personality, individuality and dignity
6. Control over one's intimate relationships or aspects of life

These six concepts of privacy have informed five of the ten factors explored in the survey, with "control over personal information" and "control over one's intimate relationships or aspect of life" having been merged and condensed into the "need to control access to personal information" (i.e., control) option in the survey.
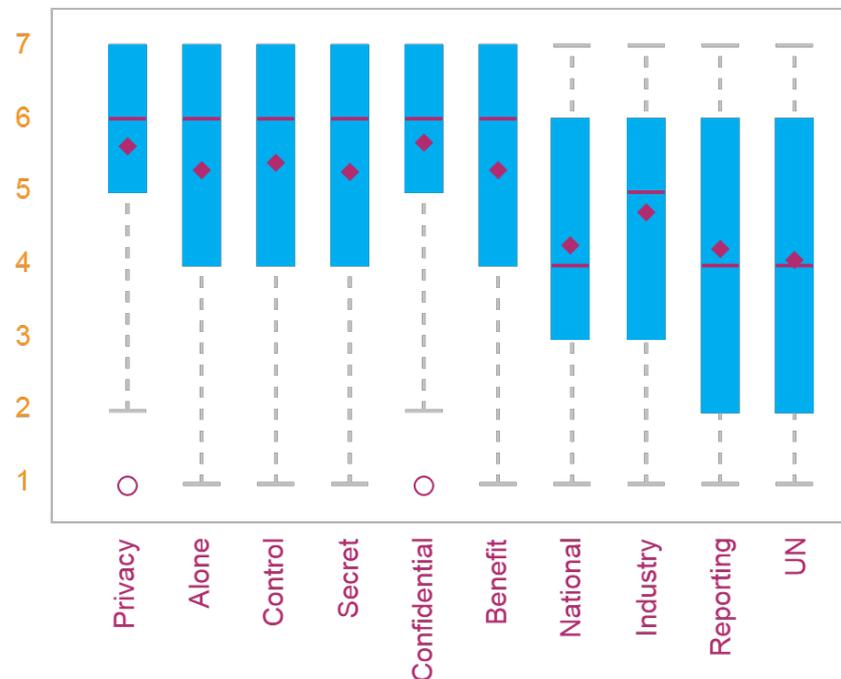

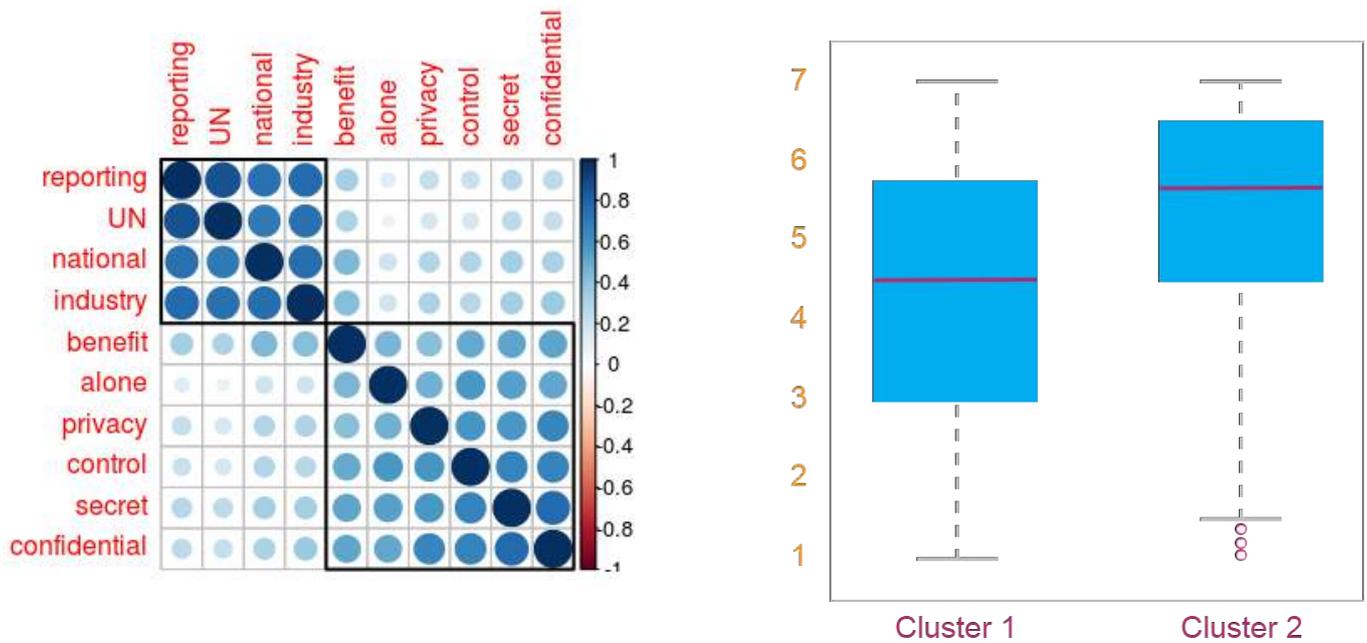Figure 11 - Factors Affecting Indicators Data Sharing

Figure 12 - Correlation and Clustering of Factors Affecting Indicators Data Sharing

From the data, "privacy" and "confidentiality" are the highest rated factors and concerns (Figure 11). Further, the general observation is that the factors associated with the various notions of privacy (i.e., Cluster 2 factors from Table 3) have more of an influence on the individuals' decision to share their personal data and information. This is also revealed by the correlation analysis and the hierarchical clustering of these factors (Figure 12).

Table 3 - Sharing Influence Factors Clusters

CLUSTERS 1 FACTORS
• reporting: Assisting the government to report on the quality of health in the country • UN: Assisting the United Nations to compare the quality of health across different countries • national: That the information is used for national health policies • industry: Contributing to improving the health industry through better medicines

CLUSTERS 2 FACTORS
• benefit: That the information is used for my direct benefit and wellbeing • alone: My right to be left alone • privacy: The privacy of the information • control: The need to control access to personal information • secret: The secrecy of the information • confidential: The confidentiality of the information
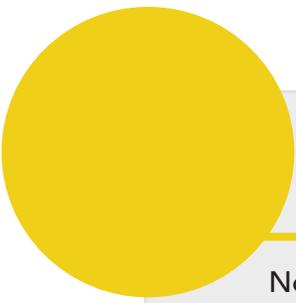
SO WHAT?

Notwithstanding the expected observation of the importance of privacy considerations in the sharing of personal information and indicators data, the findings further highlight the centrality of this over the other factors associated with data sharing within the indicators data ecosystem. Thus, the engagement of individuals for social indicators monitoring, by leveraging citizen-generated data, crowd-sourced data and specifically personal data, needs to be undertaken on terms that amplify not only the individuals' privacy concerns but also that generally places the individuals' interests at the center of the negotiated data sharing engagements, in a manner that amplifies the Human Data Interaction principles of legibility, agency, and negotiability.

DATA INTERMEDIATION

The role of community-based organizations

DATA INTERMEDIATION

Community-based organizations are organizations whose mission and vision is serving people in a local community and mobilizing individuals around specific, local issues within the community. These organizations typically have as their goals the undertaking of projects towards improving the wellbeing of communities. Due to their embeddedness in and proximity to communities, these organizations are in contact with some of the most vulnerable, marginalized, and isolated population groups. As such beyond providing services to these specific groups, they also play a role of amplifying and giving voice to individuals who would otherwise not have strong societal and political representation.

The work of CBOs holds a strong alignment with the developmental imperatives articulated in the 2030 Agenda for Sustainable Development, and so, they are a significant ally not only towards the achievement of the development goals but also for the monitoring of progress towards the achievement of these goals.

A case study of a community-based organization that provides a broad range of social services was undertaken to understand the potential role and positioning of this CBO towards brokering and intermediating the engagement of marginalized individuals [9]. The study focused on one of the centers of this CBO, the homeless center, which provides shelter and housing for homeless individuals, undertakes outreach to investigate reported cases of homelessness, and engages with the government through reporting homelessness data. This CBO holds a unique position in being the only entity that has the mandate and responsibility for providing housing services for the homeless. This unique positioning affords them the potential for playing an intermediation role for their service clients.

The following observations were made on the role and potential engagement of the CBO for advancing the Global Agenda imperatives [9]:
• The CBO center has a unique and nuanced awareness of the social wellbeing situation of the population group that they serve, i.e., the homeless people.
• In their daily operations, the CBO processes a lot of data rich with social indicators. This includes health data, socioeconomic well-being data, dwelling conditions and safety data, and homelessness prevalence data.
• The standard workflow is that data gets collected on paper forms, a subset of that gets digitized into spreadsheet documents, and a further subset is analyzed and used in reports, internally for management and externally for partners, funders and to report national homelessness indicators.

Community-based organizations are able to facilitate increased participation of vulnerable and marginalized groups and to broker increased collaboration for social indicators monitoring in the context of the Sustainable Development Goals.

Notwithstanding the above factors, which invariably place the CBO in a position where they can contribute to the understanding of the situation of homelessness in a nuanced and detailed way, the following further considerations remain:

- The bulk of the data that the CBO collects and processes is sensitive, private, and confidential data. Thus, the most immediate use of this data is within the confines of the operations of the CBO where institutional mechanisms guarantee privacy and confidentiality preservation.
- The CBOs, like most NGOs, operate in the context of limited financial resources. While there is some use of data for supporting internal organizational efficiencies, in general, external requirements for collecting and processing data can place an extra burden and can be disempowering for the CBOs [10].
- The use of technology within the CBO, which would support active engagement within the indicator data ecosystem, is unplanned and somewhat haphazard. This is linked to the factor of limited financial resources, which means that the CBO is not able to have dedicated IT personnel nor an extensive IT strategy.
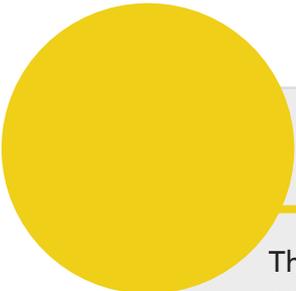
SO WHAT?

The important role that CBO's and NGO's stand to play towards the achievement of the Global Agenda is well recognized and acknowledged. The High-Level Political Forum on Sustainable Development, through the avenue of Major Groups and other Stakeholders (MGoS) supports the participation of NGOs within the global deliberations on the SDGs. Around the world various other interventions take cognizance of the unique role of NGOs towards the achievement of the Global Agenda. Overall, much can be learned about creating a more inclusive climate for stakeholders in the SDG monitoring and reporting processes by considering the promising practices implemented in other areas of the UN ecosystem. For example, participation and accountability are key principles in the human rights framework and one mechanism, which could provide a model for increasing citizen and CBO/NGO participation in data reporting at the national level, is the Universal Periodic Review (UPR) model.

Further, while there has been recognition given for the need for financing capacity building for NSOs to enhance their ability to collect indicators data, there is also a need to explore mechanisms for resourcing and capacitating NGO's and CBO's for active participation within the indicators data ecosystem – in line with the specification of SDG 17 – target 17.17

"Encourage and promote effective public, public-private and civil society partnerships, building on the experience and resourcing strategies of partnerships"

CONCLUSION

CONCLUSION

The importance of the engagement and participation of citizens and the civil society towards the sustainable development goals is espoused and enshrined in the core principles that have guided the formulation of the Global Agenda. Further, the sustainable development goal 17 on "strengthening the means of implementation and revitalizing the global partnership for sustainable development" has a specific target on "Encouraging and promoting effective public, public-private and civil society partnership…" This report has presented key considerations, from a data perspective, that must be taken into account for the meaningful engagement and participation of citizens in the agenda for sustainable development.

The data revolution for sustainable development is an opportunity for:

1. Ensuring that citizens are given visibility and a voice in indicators data, by accounting for and mitigating against the various forms of data marginalization and exclusion in social indicators.

2. Developing frameworks and tools that operationalize the 'leave no one behind' principle for various stakeholders within the social indicators data ecosystem.

3. Engaging citizens not simply as data subjects for SDG indicators nor simply as beneficiaries of the sustainable development goals, but as primary agents responsible for their own development and the development of their communities.

4. Encouraging buy-in, ownership and hyperlocal contextualization through community-driven indicators and community-based monitoring systems.

5. Connecting citizens with actionable insights for their everyday living through citizen-centric indicators data valorization.

6. Empowering citizens to hold their governments accountable to their responsibilities and obligations to citizens.

7. Leveraging the role of community-based organizations towards intermediating and brokering the participation of marginalized citizens

In the end, the engagement, inclusion, and participation of citizens, and perhaps more significantly the ownership of the Global Agenda by empowered citizens, remains one of the most critical factors towards realizing the broad goals of sustainable development in such a way that no one is left behind.

ANNEXES

ANNEX A
Developing the Typology of Data Marginalization

Clustering all the individuals and population groups who suffer marginalization and exclusion into a monolith misses out on the nuances of the phenomenon of marginalization. While it might make the discursive engagement with the matter simpler, it does not accurately account for the diversity and the heterogeneity in the characteristics of these groups of individuals. A typology of data marginalization that has been presented in this report has been developed through a multidimensional clustering approach which is depicted in Figure 13 below.
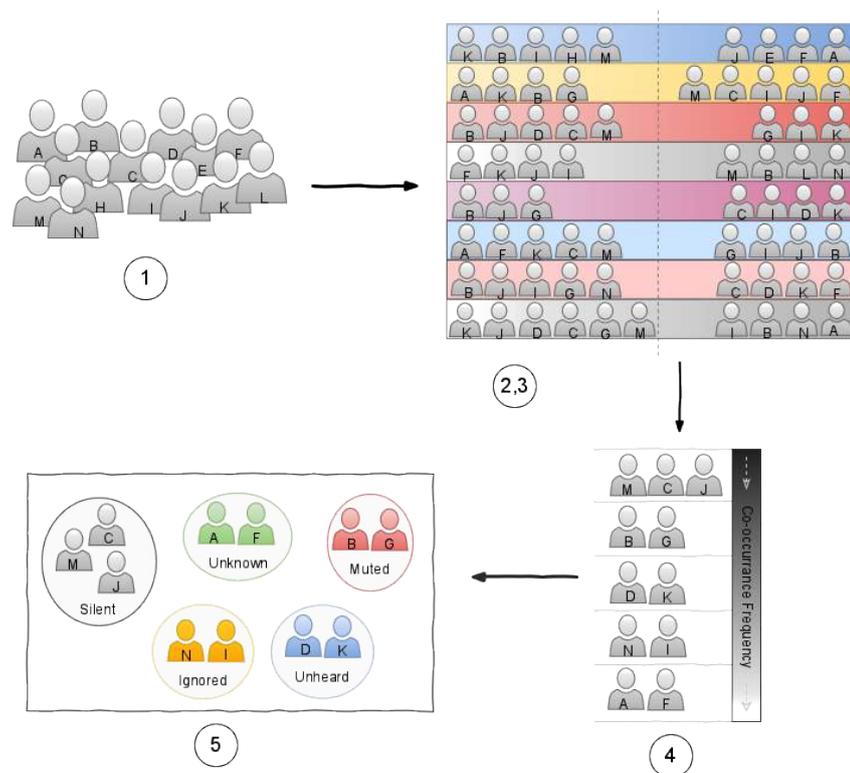

Figure 13 - Multi-Dimensional Clustering Approach for Typology Development

The approach consisted of the following steps:
1. A list of personas and scenarios of marginalization and exclusion was drawn up. Examples of the personas included:
   a) A destitute, homeless person living in a high-income country
   b) A homosexual person in a homophobic social context
   c) A family living in a remote and isolated village in one of the low-income countries
   d) A victim of domestic violence in a country with very high gender equality metrics

ANNEX A
Developing the Typology of Data Marginalization

2. A list of binary aspects (e.g., activities, attributes, dimensions) of social indicators monitoring was drawn up. These aspects provide a non-exhaustive nuanced list of different ways of marginalization and exclusion that can occur in the process of social indicators monitoring. Some of the aspects identified in this section are:

    a) Digitally accessible vs. Not digitally accessible. This aspect indicates whether the individuals have access to digital and electronic services and therefore if they generate data from which social indicators can be derived. This was further broken down into the four levels (i.e., Data exhaust, online information, physical sensors, and crowd-sourced data) as defined in the Global Pulse taxonomy of digital data [11]

    b) Societal outlier vs. Societal average. This aspect identifies the extent of integration, cohesion, and homogeneity of the individuals within the social context that they exist. A victim of domestic violence in a country with high gender equality would perhaps represent an exception and an outlier in their community. However, for a woman who lives in highly patriarchal society, despite the fact of her situation being common and therefore a 'societal average,' such a woman would still face forms of marginalization in social indicators monitoring linked to the underlying phenomenon of marginalization due to the patriarchal context.

    c) Vocalizing and expressing vs. Silent. This refers to the extent to which the individual enjoys the capabilities of expressing their social situation (whether positive or negative). In this case, the notion of capabilities is as adopted by Nussbaum and Sen's capabilities approach [12]. Therefore a destitute and homeless person on the streets of one of the low-income countries, would most likely not have a platform for expressing their social woes, but beyond that, due to the phenomenon of adaptive preferences might not have the agency to express their social woes [13]. Within this aspect, that individual would be classified as not vocalizing and not expressing. On the other hand, an unemployed, hungry person in a country with high development metrics would likely have avenues through which they can express and make their situation known.

3. The personas that have been defined were then characterized according to the list of binary aspects identified. This step is the mapping of the aspect of marginalization and exclusion to each of the personas. For example, it is within this step that a homeless, destitute person in a low-income country, as well as a person living in an isolated and remote village, would both be characterized under being "Not digitally accessible." After this step, all the personas were mapped and given attribution linked to the dimensions of marginalization and exclusion.

4. The defined personas were then clustered based on the strength of the association from the aspectual characterization step (the previous step). The association is established in terms of the frequency of co-occurrence of personas across the various aspects of exclusion and marginalization. For example, after this step and based on high co-occurrence frequency, the women in patriarchal societies, and the homosexual living in a homophobic society end up being clustered together.

5. The final step was the labeling and description of the major clusters of personas that have been derived from this process. This step provided the final formulation of the key forms of exclusion and marginalization.

This process of distilling the significant clusters of marginalization and exclusion resulted in the identification of five key forms of marginalization. Invariably the five forms are neither an exhaustive list nor the most nuanced articulation of the various forms of marginalization in social indicators monitoring. The five dimensions represent a usable typology that recognizes on one end that the singular notion of marginalization and exclusion falls short of accounting for the inherent and observed heterogeneity and diversity within these population groups, and on the other end that there is a need to aggregate the individual instances of exclusion and marginalization based on actionable commonalities.

ANNEX A
Developing the Typology of Data Marginalization

The categories identified largely define the core characteristics of the various forms of marginalization and exclusion, without strict enforcement of the boundaries between these categories. As such the categories are not mutually exclusive and they also share some overlaps, wherein, for example, individuals within the 'inaudible/silent voices' could also be in 'muted voices.' One such example would be individuals with intellectual and developmental disability who due to their internal mental capability would be limited from full participation in societal and political activities (i.e. the 'inaudible/silent voice' aspect) and who would in a certain socio-cultural setting be stigmatized and socially marginalized (i.e. the 'muted voice' aspect).

Invariably these categories are tightly coupled to the socio-cultural context in which they are considered. Population groups that find themselves in one category in one country could be in a different category or not even experience any form of marginalization and exclusion altogether in a different country. For example, LGBTQIA communities who would be 'muted voices' in a country that discriminates against them, could be fully and actively engaged in a different context. The use of these categories to identify the various groups that are marginalized therefore always needs to be tied to a well-defined socio-cultural context. As a corollary to that, the mitigation strategies that are implemented to address the different categories of marginalization will have a high specificity to the socio-cultural context.

ANNEX B
Methodological Design for SDG3 and CBO Case Study

SDG3 CASE STUDY
The findings presented in this report are based on research that was undertaken to investigate the potential for engagement of individuals within the sustainable development agenda. The case study scenario for this investigation is on SDG3 "Ensuring healthy lives and promoting well-being for all at all ages." This is motivated by the fact that in recent years there has been an increase in the use of data for monitoring individuals' health and wellbeing, which has been fueled in part by the developments in sensors technologies. The proliferation of mobile devices, smart wearable devices, and health-monitoring consumer devices that bundle these sensors means that more people are increasingly using data for personal health informatics [14].

The methodological design of this case study is framed around the following lines of inquiry with regards to individuals' use of personal health data - to understand: the reasons and motivations for data collection and monitoring; the current practice around personal health informatics; and the attitudes and values around data sharing, and social sense making. Towards this end, a semi-structured survey instrument was developed and has been used for data collection. The survey consisted of 14 questions: two closed questions on demographics, five closed questions on current personal health informatics practice, two closed questions on sustainable development goals, three closed questions on data sharing, and one open-ended question to discover individuals' current practice around health and well-being monitoring. This survey instrument was administered online, and the recruitment of the participants was done through email invitations, social media, virtual snowballing techniques, and paid participant recruitment through a survey portal. Thus, the survey was framed to inform a non-probabilistic descriptive understanding of these issues for the specific individual participants in the survey, without seeking to make wider population group generalizations.

The data collection through the survey is ongoing with further participant recruitment being undertaken through virtual snowballing and social media channels. Two sets of analyses were conducted - the first was conducted with responses from 981 participants, and the second was conducted with responses from 1863 participants. The latter analysis informed the section on "Citizen Awareness – Precursor to engagement" and the former analysis informed the rest of the sections on the SDG3 case study.

ANNEX B
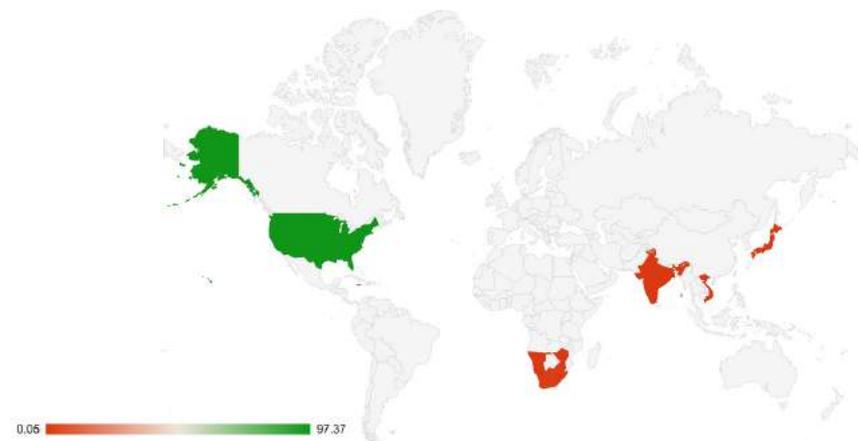## Methodological Design for SDG3 and CBO Case Study

Figure 14 - Respondents Geographic Location

Figure 14 highlights the geographic distribution of the participants in the survey. The majority (97.37%, n = 1815) of the respondents are from the United States of America, and these are mainly the participants who were recruited through QuestionPro.com - the online platform that was utilized for the survey.

ANALYSIS OF THE QUANTITATIVE DATA
- Basic descriptive statistical analysis was undertaken for the quantitative data from the survey. Non-parametric methods were utilized for the analysis due to the non-probabilistic sampling in the survey.
- The Spearman rank correlation measure was used for the correlation analyses undertaken.
- The cluster analysis conducted employed Agglomerative Hierarchical clustering method, using the complete linkages method, and the Euclidian distance between the correlation scores.

ANALYSIS OF THE QUALITATIVE DATA
- The analysis of the qualitative data from the open-ended question in the survey was undertaken by a single researcher through a Template Analysis (TA) approach. In TA an initial a priori set of codes is usually developed, then a subset of the data is analyzed - throughout which process further emergent codes are added to the coding template. This emergent coding template is then applied to the rest of the data, allowing for the flexibility for new codes that emerge from the rest of the data to be added to the coding template. The a priori codes in this analysis were informed by literature from the field of personal informatics, in particular, the stage-based model of personal informatics [14] and the lived informatics model of personal informatics [15].

ANNEX B
Methodological Design for SDG3 and CBO Case Study

CBO CASE STUDY
The case study on the role and positioning of CBOs for contributing towards the Global Agenda was undertaken through a contextual inquiry engagement with a partnering community-based organization. The engagement, which occurred over a period of five months, consisted of nine site visits involving interaction with the staff of the CBO center as well as with the service clients through participant observation and informal, unstructured interviews. The visits were framed around:
- relationship building with the CBO
- gaining a broad understanding of the operations of the CBO
- exploring the current use of information and communication technologies (ICT) within the organization
- investigating the use of social indicators data by the organization

RESEARCH LIMITATIONS
The evidence presented in this report is based on case study research focusing on SDG3 and on the specific partnering CBO. In the case of the SDG3 case study, the findings presented from the survey must be understood with consideration of the following limitations:
- The use of convenience sampling through the online survey platform means that the responses are non-random and therefore non-probabilistic. Non-parametric statistical analyses have thus been undertaken with the quantitative data
- The majority (97.37%) of the respondents in the survey are from a single country, the United States of America. This means that the views and perspectives derived from the analysis are largely representative of that specific context
- The template analysis methodology utilized for the qualitative data has limitations associated with researcher bias.

Overall the findings from the case study research are specific to the respondents and to the situation of the CBO and thus cannot necessarily be generalized to other population groups and other CBOs.

Despite these limitations, the case studies have enabled a deep and focused study of the issue of citizen participation in social indicator data, and of the intermediation role of CBOs within the broader indicators ecosystem. The case studies have also sufficiently highlighted issues that should be of key consideration in interventions that's seek to encourage citizen engagement and participation within the Global Agenda programs around the world.

REFERENCES

[1]   "Localizing the SDGs." [Online]. Available: http://localizingthesdgs.org/.

[2]   "SDG Compass." [Online]. Available: https://sdgcompass.org/.

[3]   "About Major Groups and other stakeholders." [Online]. Available: https://sustainabledevelopment.un.org/aboutmajorgroups.html.

[4]   A. Braun, "Gross National Happiness in Bhutan: A Living Example of an Alternative Approach to Progress," Soc. Impact Res. Exp., 2009.

[5]   ILO, "ILO Global estimates on migrant workers: Results and Methodology," Geneva, Switzerland.

[6]   C. W. Cobb and C. Rixford, Lessons learned from the history of social indicators, vol. 1. San Francisco: Redefining Progress, 1998.

[7]   R. Mortier, H. Haddadi, T. Henderson, D. McAuley, and J. Crowcroft, "Human-Data Interaction: The Human Face of the Data-Driven Society," SSRN Electron. J., Oct. 2014.

[8]   D. J. Solove, "Conceptualizing Privacy," Calif. Law Rev., vol. 90, no. 4, p. 1087, Jul. 2002.

[9]   M. Thinyane, L. Goldkind, and H. I. Lam, "Data Collaboration and Participation for Sustainable Development Goals—a Case for Engaging Community-Based Organizations," J. Hum. Rights Soc. Work, pp. 1–8, Feb. 2018.

[10]   C. Bopp, E. Harmon, and A. Voida, "Disempowered by Data," in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17, 2017, pp. 3608–3619.

[11]   E. Letouzé, Big Data for Development: Challenges & Opportunities. United Nations Global Pulse, 2012.

[12]   M. C. Nussbaum, Creating capabilities : the human development approach. Belknap Press of Harvard University Press, 2011.

[13]   M. Nussbaum and A. Sen, The Quality of Life. Oxford University Press, 1993.

[14]   I. Li, A. Dey, and J. Forlizzi, "A stage-based model of personal informatics systems," in Proceedings of the 28th international conference on Human factors in computing systems - CHI '10, 2010, p. 557.

[15]   D. A. Epstein, A. Ping, J. Fogarty, and S. A. Munson, "A lived informatics model of personal informatics," in Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '15, 2015, pp. 731–742.

[16]   M. D. Montoya, "Ethnographic evaluation of the behavioral causes of undercount: Woodburn, Oregon," Ethnogr. Eval. 1990 Decenn. Census Rep., vol. 25, pp. 6–90, 1992.

REFERENCES

[17]   UNICEF, "Progress for every child in the SDG Era," New York, 2018.

[18]   A. Geiser, "Social Exclusion and Conflict Transformation in Nepal : Women, Dalit and Ethnic Groups FAST Country Risk Profile Nepal," http://edoc.vifapol. de/opus/volltexte/2011/2436/pdf/WP_5_2005_nepal. pdf, 2005.

[19]   D. Nayyar, "The MDGs after 2015: Some reflections on the possibilities," 2012.

[20]   M. Green, "What we talk about when we talk about indicators: Current approaches to human rights measurement," Hum. Rights Q., vol. 23, no. 4, pp. 1062–1097, 2001.

[21]   H. Thinyane and I. Sieborger, "MobiSAM: Reflections from a four year case study using technology to increase public participation in local government in South Africa," J. Community Informatics, vol. 13, 2017.

FURTHER READINGS

This report is based on the research undertaken at the United Nations University Institute on Computing and Society within the Small Data Lab. The following academic publications provide further discussion of the research.

- Thinyane, M., Goldkind, L., & Lam, H. I. (2018). Data Collaboration and Participation for Sustainable Development Goals—a Case for Engaging Community-Based Organizations. Journal of Human Rights and Social Work, 3(1), 44-51.
- Thinyane, M., Bhat, K., Goldkind, L., & Cannanure, V. (2018). Critical Participatory Design: Reflections on engagement and empowerment in a case of a community-based organization. In Participatory Design Conference 2018, Hasselt & Genk – Belgium
- Thinyane, M. (2017). Small data and sustainable development—Individuals at the center of data-driven societies. In ITU Kaleidoscope: Challenges for a Data-Driven Society (ITU K), 2017 (pp. 1-8). IEEE.
- Moorosi, N., Thinyane, M., & Marivate, V. (2017). A Critical and Systemic Consideration of Data for Sustainable Development in Africa. In IFIP WG9.4 International Conference on Social Implications of Computers in Developing Countries (pp. 232-241). Springer, Cham.
- Thinyane, M., (2017) A Typology Classification of Exclusion and Marginalization in Data-driven Digital Societies, European Symposium series on societal challenges in Computational Social Science – Inequality and imbalance, 15-17 November, London – UK.
- Thinyane, M., (forthcoming), Towards informing human-centric ICT standardization for data-driven societies, Journal of ICT Standardization, River publishers.

IN SUPPORT OF
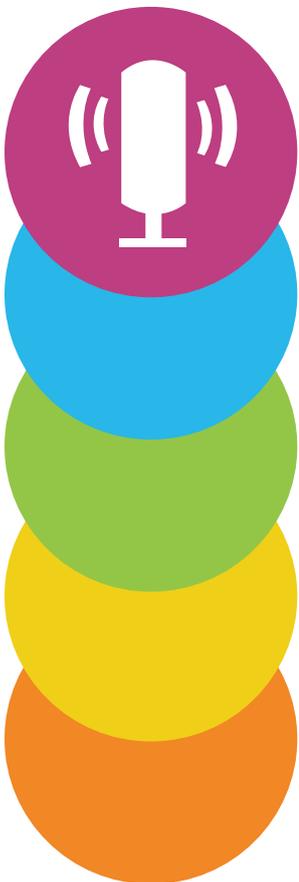
ENGAGING CITIZENS FOR SUSTAINABLE DEVELOPMENT: A DATA PERSPECTIVE

MAKING THE GLOBAL AGENDA THE CITIZENS' AGENDA

United Nations University Institute on Computing and Society
Casa Silva Mendes
Estrada do Engenheiro Trigo No. 4
Macau SAR, China
Tel: +853 2871-2930
https://cs.unu.edu

@UNU_CS

/company/UNU-CS

/ComputingandSociety