

Introduction

The following reflects the comments developed by the Civil Society and Indigenous Peoples' Mechanism (CSM) of the CFS on the v0 draft of the *Data collection and analysis tools for food security and nutrition*. The topic is an emerging area of development and concern for all peoples, especially as the digitalization of agriculture expands in the private sector and in some public/private partnerships with little to no analysis of its contributions to food security. However, the report was only published in English, and the original deadline for comments provided only nine working days to analyze forty pages of text. This is unacceptable given the time it takes to consult with and develop comments from across the CSM's diverse constituencies and regions.

Despite these concerns, the CSM raises significant concerns with the draft report. The scope of the topic was never defined, which is reflected in this report. What types, sources and uses of data that the report is concerned with is never explicitly addressed. Indeed, "data" itself is not defined and as a result there is a worrying conflation between "data", "evidence" and "knowledge" for policymaking in a way that erases the latter two categories. Instead, diverse forms of data (digital sequence information, statistical census data, big data from the internet of things, etc.) are blurred together without elaborating the risks and governance issues raised by distinct forms of data.

A notion of "data-driven decision making" is presented without being examined or unpacked. The CSM supports knowledge-based and values-based decision making and sees evidence-gathering (including data-gathering) as essential to good decision making—but no case is made or explained for what "data-driven" decision making means or why data should be placed in the "driving" seat of decision making on food security and nutrition. In effect, the report promotes data as a vague and unalloyed good without acknowledging the role of bias in data or adequately addressing the risks to the privacy, autonomy, and sovereignty of small-scale producers, food chain workers, and indigenous peoples that data collection, extraction and commercial mobilization pose.

Moreover, by focusing only on *data*, without regard to the *infrastructures* of data collection, distribution, processing and ownership the report evades key issues about the political economy and physical impacts of data. Given that data is a commodity of growing importance in the food and agricultural sectors (indeed in the global economy overall) and is subject to ever growing controversies about monopoly power and distortion of governance and oversight, this is unacceptable. The voracious competition for processing and mobilisation of data by the private sector both within and beyond the agrifood sector has engendered increasing corporate mergers

CSM Comments on V0 *Data collection and analysis tools for food security and nutrition*

and consolidation, financialization of food systems, and inequality and asymmetrical control over information, land, trade and production.

The widening extraction of data through digital technologies is generating increasing surveillance as a core of new business strategies that we see across many different sectors and social spheres including at every step of the food value chain. Those benefiting most from this datafication of food chains also control major philanthropic and policy initiatives around agricultural and food related data. As private actors accumulate and process troves of data from the genomics of seed and soil to the mapped behaviours of food producers and consumers, they hold the possibility of exercising greater and unequal control over food and agriculture systems. Data when aggregated and processed, especially by modern AI and machine learning platforms, is neither neutral, nor technical, but rather is about access to and control over information, knowledge and automated means of production and distribution.

The emphasis on data along with new “data-driven” technologies, including robotics, machine learning, genomics and the internet of things are already replacing the work and the knowledge of farmers and food chain workers, creating conditions for the displacement of Indigenous peoples, peasants, and the landless, and radically changing the need for labor in the food system. Moreover, digitized technologies are putting ever more pressure on workers through new surveillance technologies, forcing them to perform at break-neck speeds and endangering their health. Deep ramifications for our food system are evident; a shift in this trajectory is necessary to end hunger and food insecurity and to ensure human rights (including the right to food and water and the rights of Indigenous peoples and small-holder/peasant food producers).

We have framed our comments through five sections below that respond to the questions posed by the High Level Panel of Experts in Food Security and Nutrition (HLPE) in the e-consultation on *Data collection and analysis tools for food security and nutrition*.

https://www.fao.org/fsnforum/consultations/HLPE_data_collection_analysis_tools_V0

Report framing and conceptual framework

The report starts from several problematic assumptions that should be addressed in the report’s framing and conceptual framework.

- 1) The paper suggests from its opening remarks that food systems have failed because of a lack of existing data (without clarifying which food systems have failed or what missing data has created this situation). There are many possible reasons why food systems are

CSM Comments on V0 *Data collection and analysis tools
for food security and nutrition*

failing at this time but this particular framing hides those many factors. It erases from view historical and acute issues of colonialism, trade, and corporate control all of which have been driven by unequal power and control, not simply ineffective data-gathering. It suggests that simply having more data in the aggregate will reverse the “failures” of food systems, rather than dealing with entrenched interests and inequalities. This is naive at best - a dangerous attempt to side-step or sweep away difficult political debates at worst. This technocratic framing is carried through in the conceptual framework, which also provides a model of “data-driven decision making” that suggests that more data alone will facilitate more effective, equitable, and sustainable governance - but assumes (and promotes) this without presenting evidence.

- 2) The report conceptualizes data as neutral and representative, rather than as decontextualized information that results from the choices made about what information to collect and how to classify, measure, and commensurate information. In other words, there are already biases and assumptions embedded within data that will be heightened as new data-processing and automated decision making tools rise in importance. The assumptions can make data appear as neutral and uncontested, further entrenching the assumptions embedded in seemingly neutral data. These assumptions can marginalize particular constituencies and groups. For example, data may be collected from primarily commercial farmers thereby erasing small-holders, subsistence growers, pastoralists, hunters and other peasants. We have seen a concerning example of this recently in FAO’s 2021 statistical paper by Lowder S.K., *et al.*, (2021) “Which farms feed the world and has farmland become more concentrated?”, (*World Development*, 142) which using FAO stat data systematically erased from view many small food production systems from an FAO accounting of who feeds the world.

This example has been highlighted by civil society groups. See:

<https://www.etcgroup.org/content/peasants-still-feed-world-even-if-fao-claims-otherwise>

Alternatively, data may not be collected on marginalized populations, which can drive policy choices that further marginalize them.

- 3) The report provides a linear conceptualization of data-driven decision making, overlooking that all policy making is based on particular values, interests, and power imbalances and that while evidence is important for policy making, ‘data’ is only one limited kind of evidence. Decision-making must be inclusive and rights-based, ensuring the participation of those most affected. It is more accurate and proportional to describe the need for “data informed” decision making, rather than data-driven. In fact, “data driven” decision making has been highly contested - especially as automatic decision-

CSM Comments on V0 *Data collection and analysis tools
for food security and nutrition*

making through machine-learning creeps into governance. For example predictive algorithms in policing have shown deepened policing of already highly-policed communities, reproducing racial, class and cultural prejudices.

- 4) The report fails to define the terms and concepts it is using. Data meaning sometimes information, and sometimes evidence are used interchangeably throughout the report and must be defined and carefully distinguished. Data is generally defined to specifically mean a type of information that is machine readable or can be processed by a computer but the word also has an occasional informal meaning that is synonymous with “evidence”. This report is seemingly primarily about data of the machine-readable type but sometimes it conflates or slips into the informal meaning of the word - especially in advocating around ‘data-driven decision making’ (does it mean ‘evidence-driven’?). Different forms of data are not distinguished—not only between qualitative and quantitative data—but also different forms of quantitative data that take different forms based on level of abstraction and construction. For example, composites and indicators are very different from statistical counts. They require different kinds of classifications and assumptions. All data is not representative, nor is it free from bias - indeed as current scholarship on algorithmic bias shows data and “data-driven decision making” can not only reflect but significantly deepen harmful biases around race, gender, culture etc, hard coding them into policy and practice. This is especially important in food and agriculture systems where land and other resources as well as political power may be concentrated in one economic, cultural or racial group. The report needs greater precision about the terms it is using if it is to be useful to establishing effective governance over data.

Relatedly, the terms “data value chain” and “data flows” are used in the report without defining them. The use of the term “data value chain” particularly should be reconsidered. It is an economic concept borrowed from commercial business models that use data-processing tools to upgrade the economic value of data to leverage bigger profits in the marketplace. Those new business models are based on surveillance and extractive data relationships. Just as the notion of ‘value chains’ in food systems poorly captures and distorts the multifunctional and complex nature and impacts of food systems, so describing data-handling and data-relations by this narrow commercial metaphor may be wholly inappropriate and have distorting implications. The term “user” is also used throughout the report, without distinguishing the different stakes of “users” nor reflecting on the power and social implications of data enterprises defining farmers, eaters, indigenous peoples and other human beings by this transactional term. As has been noted elsewhere, besides data-processing the other industry that terms its clients

CSM Comments on V0 *Data collection and analysis tools
for food security and nutrition*

as “users” are in drugs and pharmaceuticals - where there is a rich literature on the stigmatizing and social assumptions built into such nomenclature.

- 5) While the report focuses on data it should also be made clear that there are clear biases in terms of what is recognized as “data” and thereby admitted to consideration in policy making processes. Here data is primarily discussed as quantified information, but that excludes other forms of knowledge and information that may be more relevant to social contexts. More hard computable data is not always better, especially for agriculture which requires softer relational knowledge of land, seeds, and climate. The focus on quantified data risks alienating the real-world knowledge of peasants, Indigenous peoples, farmers, and food producers, shared through generations and between communities.
- 6) Not all forms of data *should* equally inform policy making, given the assumptions inherent within data. For example, the sources of much big data are private infrastructures generated from sensors, platforms, and the internet of things that uniquely measure commercial processes. When choices are made by private actors (whose interests are in profits rather than the public good) about what information to collect, their assumptions and biases are integrated into the decision-making process without adequately scrutinizing them and putting them to democratic debate. The conceptual framework thus requires greater clarity and specificity about the *risks* of drawing on data indiscriminately in policymaking. It should also elaborate the limits of data, and the necessity of other knowledge types for which ‘data’ is not available.
- 7) This report and workstream is funded by the Bill and Melinda Gates Foundation (BMGF), whose foundation has worked directly with its Co-Chairman’s former corporation, Microsoft, on the digitalization of agriculture. (See <https://cacm.acm.org/magazines/2021/12/256930-digital-agriculture-for-small-scale-producers/fulltext>). In line with the foundation’s source of wealth and worldview BMGF boast an explicit aim to move over 50% of smallholder farmers in its key focus areas onto data platforms within the decade - a move from which Microsoft’s Azure cloud services and “farm beats” platform is poised to profit most richly. The promotion of digitalization of agriculture and data governance by actors that also seek to directly profit from these new technologies are not only concerning, it represents a key problem of using private data infrastructure for public purposes, namely, conflicts of interest. The promotion of data collection through private infrastructure not only legitimates private surveillance of individuals and communities, it offers private actors the opportunity to shape information and knowledge of the food systems in ways that can further entrench their power by both

shaping individual and formal decision-making and skewing policy outcomes. Given this fundamental conflict of interest baked into the funding of this workstream it is essential that the HLPE properly examine and grapple with the more challenging complexities of this topic - especially regarding power, control and private interest.

Data for food security: agency and sustainability

Data collection and analysis feeds a model of society. Neutral data does not exist - it is always curated and shaped by the questions and aims of data gatherers. Collection of data for food security and nutrition (FSN) has therefore to be determined by the model of FSN that we seek to advance. Human Rights are acknowledged to be fundamental to reach FSN for all, data collection and generated knowledge have therefore to be designed to benefit them.

To broaden the definition for food security for data collection, analysis and use, and in particular around the dimensions of agency and sustainability, the report needs to further elaborate both the process-elements of data collections, analysis, and use as well as describe some substantive areas in which data for agency and sustainability should be developed.

In terms of the process-areas of data sustainability and use, the report should provide a discussion on what kind of knowledge and data analysis is necessary for food security policy-making. The HLPE and the CFS already provide a model for evidenced-based, inclusive multilateral decision making that could serve as a model for data informed decision-making. The HLPE offers a model of scientific, transparent and inclusive processes that incorporates different forms of knowledge across various disciplines and professional backgrounds. Moreover, it is rooted in an institution based on the right to food, where those most affected by food insecurity are given a voice in decision-making. Rights-based, data-informed decision making must include those on whom data is collected. Indeed, there are some populations that may refuse data collection and extraction including from their territories. Indigenous peoples have articulated the concept of indigenous data sovereignty to demand collective control over the kinds of data that is collected on them and to have a voice in what data is collected and how it is used. These rights must be respected.

There are also several substantive areas of data collection that could be highlighted in the report to support agency and sustainability. First, disaggregated data that address inequalities based on gender, race, nationality, language, and sexual orientation is necessary to promote equity and agency. An example to highlight the importance of asking the right questions in data collection is this study by Canadian data group PROOF in collaboration with civil Society Food Share that found that race in Canada is highly correlated with food insecurity - especially for black

Canadians: <https://proof.utoronto.ca/anti-black-racism/> - offering evidence for the need for racial justice policies at the heart of food security interventions. Second, data on market structures including market consolidation and monopoly control is also necessary to enable agency and sustainability. Finally, data that is collected by humans rather than by sensors or satellites and reported to the population within communities, builds agency into the process.

Treatment of existing data sources, agricultural digitalization, and other new technologies

Chapter 4 of the report identifies “new and emerging data driven technologies” and elaborates the kind of data which they might be used to collect. This chapter presents an anemic analysis of these technologies by focusing solely on their implications in terms of data, rather than the larger risks and benefits these technologies pose to food systems. The CSM raises several concerns with this chapter in its current form.

- 1) The report approaches data recollection for FSN as a purely technical problem, avoiding the distributive issues that are at the heart of the technology debate. This includes questions related to the ownership of technologies and infrastructures, the distributive effects of technologies, the effects of technology on employment, the effects on existing knowledge infrastructures and networks, and the socio-ecological effects of technologies.
- 2) Technologies cannot address structural and historical inequalities that pervade food systems that include access to land, inequality, climate change, and unequal distribution of resources. As noted earlier, the report is premised on the naive assumption that food insecurity can simply be addressed through more and better data.
- 3) The current draft of the report has an entire section that lists multiple new technologies that may be used to collect data. However, it has no explanation of the criteria used to identify these technologies. Are these technologies listed because they allow for the collection of data or do they contribute to food security and nutrition? It is not enough to present a list of “technologies” (see 4.3.). The report must develop a stronger set of criteria about their contribution to food systems based on the risks and benefits they pose.
- 4) Several key technologies are omitted from this report, even though they are components of the “datafication” of food systems. These include:
 - a) Digital Sequence Information (DSI) on genetic resources are one of the economically fundamental parts of the future development of the seed market

CSM Comments on V0 *Data collection and analysis tools
for food security and nutrition*

and thus of the control of agricultural and food production. The digitization or datification of DNA is undergoing contentious negotiations over access and benefit sharing (ABS) in the Biodiversity Convention and the FAO Seed Treaty. In a completely unregulated manner, vast quantities of agricultural production and consumption are being collected and stored – from fisheries, farms and forests to retail shops and homes. Terabytes of genomic information are also being extracted, sequenced and stored via several initiatives including the WEF’s Earth Bank of Codes, its related Amazon Bank of Codes, the International Barcode of Life, gene bankers “DivSeek,” and an unknowable number of private corporations.

- b) Digital land registries, which have been described as the new “birth certificates” for rural properties. Locating and measuring land using georeferencing technology (i.e. GPS) is increasingly becoming a mandatory requirement for gaining access to public policies and credit, and for complying with environmental regulations. Digital land registries are becoming a condition for defining property rights in land regularization processes in regions from Colombia, Brasil, Paraguay, Bolivia and Argentina, among many other examples. However, these land registries invisibilize centuries-old collective land rights that are impossible to measure in terms of data. See: Grain (2020). “Digital Fences.” <https://grain.org/en/article/6529-digital-fences-the-financial-enclosure-of-farmlands-in-south-america>

- c) Smart contracts - these are self-executing programs stored on a blockchain that run when predetermined conditions are met. They typically are used to automate the execution of an agreement such as the sale of commodities or collection of rent and fees on a service. Blockchains executing smart contracts (and exchanging cryptocurrencies and tokens) are becoming a common infrastructure for managing data-driven food value chains and also for data collection, management and automation. As such, smart contracts and their conditions will assume an ever greater opaque role in governance of digitized food systems with privately inserted ‘code’ acting as artificially responsible proxy for ‘law’.
See for example:
https://www-cdn.law.stanford.edu/wp-content/uploads/2018/09/Kolber_LL_20180910.pdf
Vulnerable to cyber-attack, coded in ways that are not easily readable or intelligible to lay people in the food chain or to policymakers, capable of locking food chain actors out of access, rights and assets - smart contracts need strong accountability to be enforced on their use or could hollow out governance of digitized food systems..

- 5) The issue of data misinterpretation needs to be further elaborated. The current draft acknowledges that “...Analyses that make use of sophisticated quantitative models – such as based on regressions, computable general equilibrium models, artificial intelligence, machine learning, etc., – are particularly prone to misinterpretation, given that a full comprehension of the nature and implications of the assumptions made to build the models is likely to remain beyond grasping for most of the intended readers and users of such data and modeled results, including policy makers...” We need to question why more data is equated with more accuracy, when it is clear that data can easily be misinterpreted especially when it is decontextualized from the complex human and ecological processes of food systems. It is a dangerous assumption and ignores the biases in huge data sets. Research on such biases show that the results of big data recollection processes give the researcher a vast sense of precision. “While big data have the connotation of being exhaustive in scope, aiming to capture entire populations or systems (n = all), in reality it is both a representation and a sample, shaped by the technology and platform used, the data ontology employed and the regulatory environment and (...) subject to sampling bias.”

See “Imprecision farming. Examining the (in)accuracy and risks of digital agriculture”

Journal of Rural Studies 86 (2021) 623–632, in

<https://www.sciencedirect.com/science/article/abs/pii/S0743016721002217?via%3DiHub>

Capacity and governance of data

As with other tools or processes that are intended to be used for the public benefit, the questions about data recollection and analysis are the same as always: Who controls what? Who decides who controls? What is the process to determine the data that is needed? How will these tools or processes benefit food producers, workers, and consumers, rather than benefit other actors such as agri-food businesses? Given that this is the first time that the HLPE and the CFS are entering into the complex world of data, it is vital that we begin with these questions and that we do not just slip into frames and language that will limit and determine what and how can be discussed later.

Data collection and analysis (including in their digital forms) are powerful tools to guide policies and development plans, under a rights-based approach (building on the implementation of UNDROP and UNDRIP). A rights-based approach in data collection and analysis—where participation and knowledge sharing are fundamental elements—is essential to keep these processes under their primary objective: effectively advance food security and nutrition for all. It is the only framework that will allow to determine the needed data based on the knowledge gaps

identified at local level. Local communities have to be involved in the design of data to be collected and the control on the data should be kept in public entities.

Although chapter five deals with institutions and governance of data, it is woefully underdeveloped. Several issues need to be addressed:

- 1) This section does not address the institutions and initiatives that are already developing from an array of different actors to govern data. There is no analysis of the current instruments of governance of data, of the impact of intellectual property rights on data and information, no analysis of the current control processes at the global level by the large digitalisation monopolies. These include intentional initiatives such as the International Platform for Digital Agriculture in the FAO as well as private codes of conduct. It also does not address the implications of existing treaties (Convention on Biological Diversity, FAO Seed Treaty, Agreement on Trade Related Aspects of Intellectual Property Rights), supranational regulations like the European Union's General Data Protection Regulation, or legal frameworks (from database rights to plant breeders rights). The existing regulatory and legal landscape needs to be mapped in order to understand the gaps and issues within data governance.
- 2) Open data is not necessarily the best solution to address access and capacity issues. Indigenous peoples, for example, have demanded *data sovereignty*, control over their data - for example in the matter of digital genetic sequences. Since data as a good confers more economic power when privately aggregated, establishing means of collective and community control over data governance is essential to fair and just data relations. The report needs to address the critiques and concerns about open data.
- 3) As noted above in discussing data governance, the report uncritically adopts language of data drawn from the private sector such as "data value chain" and "data life cycle." However such terms frame governance through a market or technocratic lens, rather than alternative terms including knowledge, information, and evidence. The report must address how these different frames are embedded with particular values and priorities.
- 4) Given that data from many new technologies come from private data infrastructures, data is primarily privately governed as a commodity. Indeed, Big Data handling is presently one of the most powerful industries. Something as complex and sensitive as food systems should not be taken as just another one of the new business niches of the big data industries. A huge blindspot of this report is any engagement or acknowledgement with the way that data is power in the world of 2022. It will be

important to examine how powerful data operators are seeking to institutionalize their services in the collection and processing of data for FSN with an eye towards leveraging that same data for commercial or geopolitical ways. Privately-held for-profit data analytics firms such as Palantir and Premise Data are already muscling into the food data arena. Palantir, which for example has data collection and processing agreements with the World Food Programme and US Food and Drug Administration is a consistent subject of concerns raised by privacy and digital justice advocates for its military, surveillance and political activities.

Other comments

Overall, the report paints an overly rosy picture of data as the solution to food insecurity and the problems caused by industrial food systems. Because the report neither delineates its scope, nor defines "data" it ends up promoting new tools of digital agriculture as opportunities for better data collection and policymaking. However, this framing not only legitimizes these technologies without adequately evaluating the broader risks they pose, it also profoundly reshapes the way evidence and policy-making are understood by the HLPE itself.

The process of reassembling data into knowledge is not transparent, not participatory and very often strengthens power imbalance and inequalities, as mentioned before. The digitization of data is accelerating these trends that are intrinsically in opposition with the SDGs 2030. Clear firewalls and blocks to prevent undemocratic data extraction and undue private influence and power over data governance should be explored and proposed to prevent private data actors from exploiting international, regional or national food, security and nutrition data infrastructure. This includes investigating the implications of deals such as the WFP-Palantir agreement and placing conflict of interest prohibitions ensuring that private actors involved in data activities for commercial food and agribusiness (eg major cloud firms such as Microsoft, Amazon, Alibaba) are excluded from managing, processing or influencing public data gathering for food security and nutrition.

More data itself will not address the structural and historical inequalities that are responsible for food insecurity. For example, Chapter 3 of v0 describes in detail the enormous difficulties involved in collecting data for food security and nutrition. These difficulties are concentrated in countries or within regions of countries that have mostly small-scale agriculture. Small scale farmers suffer from the digital gap, poor previous recollections, lack of coordination, financing and digital illiteracy. Data-based decision making is likely to reproduce existing inequalities and render them less visible.

CSM Comments on V0 *Data collection and analysis tools
for food security and nutrition*

Finally, the funding of this workstream by the Bill and Melinda Gates Foundation, a philanthropy linked with a corporation that seeks to profit through the development of private high tech infrastructures in the field of agriculture, reflects the profound conflicts of interest that are entangled in data collection through new tools of digital agriculture. Extreme caution is needed as we look to these actors and the technologies they develop to address food insecurity and other common goods.