

Contribution of the CSM WG on Data to e-consultation on the scope of the HLPE Report on Data collection and analysis tools for food security and nutrition

21 March 2021

Introduction

The objectives and current scope of the HLPE Report on Data for Food Security and Nutrition (FSN) are insufficient because they lack a purpose and a normative dimension. The report should be clear about collecting data for what, how and for whose benefit. We propose, based on the CFS normative anchoring, that this report should identify how data collection and use can contribute to upholding the rights of Indigenous Peoples, women, peasants and family farmers, workers throughout food systems, fisherfolk, pastoralists, consumers; and how data collection and use should be governed and regulated in order to respect and protect human and peoples' rights.

The current scope refers to concepts of statistical data assuming that data are something neutral, objective and free from specific social, environmental, and political conditions. For peasants, small-scale food producers and rural workers, what has been called *data* is neither abstract nor neutral. The selective capture of data from the fields and its further transformation on information becomes a powerful economic commodity. This is not the approach that the actors in the above mentioned foods systems have when it comes to data collection and analysis.

This workstream offers an opportunity to address the wide variety of issues related to data collection and analysis for FSN that the actual proposed scope is not mentioning: digitization of food systems, unregulated uses of big data, ownership of data infrastructures, and methodological assumptions for collection and analysis. In approaching these issues, the HLPE needs to develop a clear definition of data for the public interest that addresses the concerns of Indigenous Peoples, women, peasants and family farmers, workers throughout food systems, fisherfolk, pastoralists and consumers.

The HLPE's definition and analysis of data must be guided and framed by the Human Right to Food and the entire international human rights framework, highlighting in particular, ICESCR, CEDAW, UNDRIP, UNDROP, ICCPR and relevant ILO Conventions. It should provide a clear analysis of the responsibilities and obligations of governments, the private sector, and rights holders and how data should be collected, analyzed, and governed.

The link between data recollection and analysis and the ongoing, aggressive digitalization of the food systems (from automation to robotization to the use of *artificial intelligence* for data processing and analysis) is taken for granted. The diverse aspects of digitalization in food and agriculture are being contested and widely debated among all sorts of civil society organizations. The environmental impact of data recollection and analysis, sustainability, and

the digitalization of food and agriculture should also be considered as part of the scope of this report.

Conceptual frameworks and methodological assumptions behind data collection and analysis

Every sector of the industrial economy, including agriculture, is amassing data and striving to make commercial use of it. At the same time, many states lack resources (human and structural) to conduct high quality data collection and update their statistics. Official pools, surveys and census, generally need “significant” numbers in order to produce the samples, percentiles, averages, probabilities, frames. The way in which statistics proceed may discard variabilities and differences that characterize the contexts within which small scale farmers, indigenous peoples, peasants and familiar agriculture units thrive.

The objectivity assumed by processes and tools that rely on artificial intelligence and machine learning should not be assumed. Behind the programming of the machines are human software designers and programmers, with interests, assumptions, and normative judgments, that construct the algorithms and formats through which data is collected and analyzed. Assumptions are also embedded within the algorithms that “teach” the machine learning systems to become “intelligent.” Data collection and analysis with new digital tools can translate into the discarding of variability, which in turn can impact the appreciation for diversity. The report should consider the persons, interests and tools at the foundation of any data-collection initiative. Aside from commercial interests that can permeate digitization and data related processes, their impossible neutrality makes this a political issue.

Statistics languages could preempt the use of methodologies that come from the most impacted for the lack of FSN. And in parallel, there are raising concerns about the new *legality* that is being established through digital tools for registering land use, land property, land use changes, establishment of protected areas, among other cases, where peasants and local communities have no say in these processes.¹

Data governance, ownership and privatization. Barriers to quality data collection and analysis

Focusing on questions of data collection is too limited in scope. The processes of digitalization, which include the entire infrastructure of data collection and its network effects, must be addressed by the HLPE. The construction of these data infrastructures as well as the governance of data must be addressed, given that data is increasingly understood as a key commodity and economic unit in the “Fourth Industrial Revolution.”

Private data infrastructures raise critical issues of power and control over data, who owns data, and who has access to it. The HLPE report must consider the current political economy of the

¹ See, for instance, GRAIN, 2020, “Digital fences. The financial enclosure of farmland in Latin America”, in <https://grain.org/en/article/6529-digital-fences-the-financial-enclosure-of-farmlands-in-south-america>

digital economy as a whole, which is characterized by the private nature of most digital infrastructure, a highly concentrated Information and Communication Technologies sector as well as geopolitical issues that are equivalent to a new, digital colonialism.²

Digitalization of agriculture can lead to de-skilling and loss of local knowledge, as more decisions are made without the most impacted by lack of FSN. Decisions farmers make are complex, and tied to contexts of place. The data farmers have access to is not the same as that of large data companies, as generally they don't participate in the recollection protocols or criteria, neither have access to / expertise in managing the platforms. Generally, small scale farmers, local communities and indigenous peoples are involved only as informants that give up ownership of their data.

The report should address the total privatization of data infrastructures. Data extractivism and wider digital and technological divide among all the actors in the food systems could only be avoided if data infrastructures are public. There is increasing vertical and horizontal integration and concentration in the management of data across agriculture and information technology sectors, threatening further cross-sector concentration for corporative profits.

Indigenous Peoples, women, peasants and family farmers, workers throughout food systems, fisherfolk, pastoralists and consumers need to be able to voice strong opinions in processes of digitalization and retain agency of their data. Loss of ownership and control of data can lead to a loss of decision-making capacity, loss of local knowledge, and loss of economic value. The value systems of local knowledge and agro-industry are different, and therefore require different understandings of data ownership. Farmer-to-farmer networks built with digital applications, such as the ones that emerged during COVID-19, serve as an example of alternative uses of data recollection, channeling and use of digital tools and applications.³

Without transparent and inclusive decision-making processes around data collection and analysis, global inequalities may be intensified. Data collection is a crucial element of digitalization across all stages of food production and consumption, and as far as it is seen, growing use of automation and robotics conducts to a loss of livelihoods while the benefits of data ownership are weighted to agribusiness.

Along issues related to lack of neutrality, lack of inclusiveness, loss of agency of food systems actors over information yielded from their contexts and environments, the report needs to acknowledge that data collection can be blocked by local political conflicts, natural disasters, violence, illiteracy and isolation.

² See United Nations Conference on Trade and Development. (2019). Digital Economy Report 2019. Value Creation and Capture: Implications for Developing Countries. Available at: https://unctad.org/system/files/official-document/der2019_en.pdf

³ More examples of alternatives to corporate-controlled digital agriculture include: FarmHack (an open-source, free online community for sharing tools and information, using the Commons Based Peer Production model), as well as IT companies using crowd-sourced information for applications like pest control techniques.

People-led data use and digitalization processes & What kind of data is needed

Data collection and digital tools can potentially be useful for local communities, indigenous peoples and small-scale farmers. But for farmers to lead their data and digitalization-related processes, particular training is needed, and this adds other layer of difficulty if really the initiatives led by the people are to be taken into account.

The possibilities of using data through digital tools in agriculture can positively impact livelihoods and quality of life of food producers, through the opportunity to “strengthen the centrality of peasant knowledge in the production process and in the organizational and social system”, as Schola Campesina affirms.

When Indigenous Peoples, women, peasants and family farmers, workers throughout food systems, fisherfolk, pastoralists and consumers have access and are trained to interpret the information derived from data, decision making processes for local sustainable development can be strengthened. As we saw during the hardest moments of the confinement in 2020, digital communication within local communities can be a valuable tool to facilitate collective solutions to common problems.

Alliances between food producers (including workers), open-source software engineers, lawyers, and experts can aid in maintaining the centrality of local knowledge in data uses. In any case, the collection and analysis of information should take place on a bottom-up principle (the ground) to the top (public power) with the legitimacy to make decisions.

The HLPE report on data should talk about what conditions are needed so the most impacted by lack of FSN can make use and benefit from data recollection and analysis and digitalization processes.

But better involving farmers within digitalization process is not enough. Maintaining the idea that there can be a *good digitalization* and a bad one is mistaken. The fact that we all use internet is not a reason to accept it, it's rather proof of how strong the constraint to use it is in order to survive economically and socially. As Indigenous Peoples, women, peasants and family farmers, workers throughout food systems, fisherfolk, pastoralists and consumers we need to strongly defend our capacity to perceive the world with a sensitivity that has nothing in common with data.

Policies needed to strengthen the capacity to make a good data collection and analysis

The draft scope rightly notes that a gap exists in the quality and quantity of data at the national level on food production and household FSN. Yet while supporting the public collection of statistical information for public policy is essential, the HLPE must consider the wider context in

which data is collected. Indeed, data is information and knowledge that is abstract from social context. Power relations and inequalities can often be reproduced through data, based on who is collecting data, from what sources it is collected, through what methods, and for what purpose. In addressing public data collection for political decision-making related to FSN, the HLPE should consider:

1. **Currently there is inadequate availability and accessibility of data.** Most data on FSN are provided by the industrial food web, but even the private sector has volunteered less data as company and industry analysts grow more secretive. Data developed by academics is also often published behind paywalls that limit public access. Public decision-making should be premised on statistical data that is available and accessible to the public.
2. **The biased conceptual framing against small-scale food producers and territorial food systems is also a major problem.** Data collection is limited to a small number of crops and food production activities that often benefit the industrial food system. Little accurate data exists about the peasant food web. Peasants grow around 7,000 crops, but most data collection centers on only 150 crops. The world does not have accurate information on the contributions of peasant producers to FSN. Similarly, registration of diversity is neglected. Through selective data collection, public policies about the importance of certain crops and decisions related to the concentration of land and water, water pollution or sexual harassment throughout the food system, will be totally biased.
3. **Data collection done from a distance without a relationship to small-scale producers and workers is imposed.** In the case of data related to land ownership, spatial planning, forest or fisheries registries, remote control of this information has always been the rule.⁴ Partnership with communities, especially those most marginalized—such as women, elderly, minorities, and LGBTQ+ people.—in data collection is essential.
4. **Census and surveys claimed to be for public policy making are not free from manipulation in hands of who decides what and how to conduct the research.**
5. **The value of both quantitative and qualitative data for addressing FSN should be addressed by the report.** A holistic understanding of food systems and nutrition requires multiple forms of knowledge to be drawn on, this includes quantitative and statistical data, but also ethnographic data and individual testimonies of rights-holders. Quantitative data and indicators have often been given greater authority in decision-making, but in the context of FSN not all forms of knowledge can be quantified. If the goal is the production of data for the public good, is imperative to understand what data and knowledge means for small-scale producers and workers. Translation of knowledge into commercial information or information for the public good, is a political process.
6. **Digitalization can improve the collection of statistical public data, but it is important to delineate different kinds of data and to assess the quality of different data.** Different forms of data (statistical public data, commercial data, microdata, microdata, and metadata) all have different uses and impacts.

⁴ See FIAN, January 2021, “Digital technologies cut access to land” in <https://www.fian.org/en/press-release/article/digital-technologies-cut-off-access-to-land-2699>.

7. **The use of big data for public policy raises significant concerns over the ownership of data, transparency of algorithms, and accountability.** As the United Nations Development Group notes, “Big data often contains personal data and sensitive data. The use of personal data should be based on one or more of the following legitimate and fair bases, subject to implementing UNDG member organizations’ regulations, rules and policies (including data privacy and data protection policies): (i) adequate consent of the individual whose data is used, (ii) in accordance with law, (iii) furtherance of international organizational mandates, (iv) other legitimate needs to protect the vital or best interest of an individual(s) or group(s) of individuals.”⁵ However, the use of big data for public policy raises additional concerns. Big data is the product of commercial technologies that shape the kind of data that is extracted. Moreover, for big data to be made intelligible, algorithms must be developed by scientists that in almost all cases are not transparent. The HLPE report should take into account the UN Special Rapporteur on the [Right to Privacy framework](#).
8. **The problems of data extraction through surveillance technologies (satellites, GPS, or radio frequencies) for public policy need to be addressed by the HLPE.** The dematerialisation of activities and objects or life forms creates value in information from digitized data. A new *resource* that can be used to produce profit from agricultural activities, better than planting potatoes. The report should challenge the de-facto right of commercial digital platforms to use and derive economic benefit of non-personal data.
9. **Structural inequalities and injustices are often reproduced in processes of data collection and analysis.** Lack of data on FSN for marginalized communities and populations can prevent adequate public policy responses. Moreover, the digital divide and gap in access to digital infrastructures, as well as algorithmic bias in data analysis, further effects the most marginalized populations.

About data analysis, the HLPE should consider:

1. **The role of public multilateral institutions for data collection and analysis related to FSN.** Private philanthropies have promoted the deinstitutionalization of data collection and analysis, moving them to private institutions and universities in power centers in the Global North and South. Multilateral institutions are rooted in human rights mandates. Deinstitutionalization can lead to distrust and significant concerns. Public-private partnerships between global governance institutions and tech giants would be as problematic as an entire privatization of data collection.
2. **Data modelling, algorithms for data aggregation, and machine-learning processes should be open, transparent, and accountable.** The recent use of artificial intelligence through processes such as machine learning to develop outcomes for FSN (including projects such as Ceres 2030) can lead to manipulation of outcomes and mistrust. AI poses significant problems for public policy-making in terms of transparency and

⁵ UN Development Group. “Data Privacy, Ethics and Protection: Guidance Note on Big Data for Achievement of the 2030 Agenda.” https://unsdg.un.org/sites/default/files/UNDG_BigData_final_web.pdf

accountability that should be taken into consideration.⁶

3. **More consistency and transparency is needed in counting the food insecure.** The counting of people that are over nourished, malnourished, and food insecure has been inconsistent in the last few issues of the SOFI report. These numbers are critical for achieving the CFS, yet the way that these numbers are counted has changed. There are a proliferation of “dashboards” and metrics that neither disaggregate data sufficiently, nor consistently.

Conclusion

Increasingly, many of the tools, methods, and platforms of data collection and analysis for Food Security and Nutrition and public policy-making more generally are in the hands of the corporate sector, including agribusiness. Governments, public institutions, and people all depend on cloud services and satellites that are developed and managed by big business. Any collection storage, and analysis of information for FSN should consider this context.

Indigenous Peoples, women, peasants and family farmers, workers throughout food systems, fisherfolk, pastoralists, consumers, defend their capacity to perceive the world with a sensitivity that has nothing in common with "data".

The problem of transferring total reliability to private data systems, of something as key as food security is something that needs to be discussed transparently, if not rejected since the beginning.

The extent of impacts of data collection and digitalization are not yet fully assessed, especially on food systems and local communities. Regulation at local, regional, and international levels is imperative to prevent asymmetries of power as well as the extraction of local knowledge to benefit global value chains. The concentration of data in private entities has unknown and potentially grave impact in food systems and should be approached with the precautionary principle. As noted earlier, the HLPE’s definition and analysis of data must be guided and framed by the Human Right to Food and the entire international human rights framework, highlighting in particular, ICESCR, CEDAW, UNDRIP, UNDROP, ICCPR and relevant ILO Conventions.

If not managed this way, data collection and analysis a form of extractivism. For the communities and collectives impacted, it means losing control of knowledge and threatens their food sovereignty.

Additionally, even if data collection were very thorough, data is only a static image of the facts or realities we want to analyze. Realities and facts are in movement and crosscut by a diversity of problems. Statistics give only a snapshot, even with very fine grain, of something in constant

⁶ Cannetaci, Joseph A. 2021. “Artificial intelligence and privacy, and children’s privacy: Report of the Special Rapporteur on the right to privacy.” <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G21/015/65/PDF/G2101565.pdf?OpenElement>

change. A computer will never be “able” to say to the farmer: I have the best solution to your problem. This is why agroecology—which is based on the centrality of the knowledge of food producers to guarantee their autonomy—must inform the analysis of data collection and analysis for FSN.

At the territorial level there are wide range of experiences around data usage that strengthen small scale farmers’, peasants’, food workers and indigenous peoples’ struggle agendas. But proprietary restrictions make it increasingly difficult to use data for the peoples’ needs, due to property rights and intellectual property systems. It is only with the participation of those impacted by the lack of FSN that any data collection and analysis can yield truthful and productive knowledge of the diversity of situations and dare to formulate any recommendations, either for recollection of new data or for better ways to analyze data collections in public institutions.

The HLPE report should emphasize the collective, inclusive, dialogic nature of data collection and analysis. It should move away from the idea that data collection and analysis are objective, technical exercises on which a successive political phase of policy decision-making can be drawn from.

Therefore it is important to widen the scope of this report as much as possible, encompassing issues related to methodologies, data definitions, due explanations about the figures on malnutrition, the narrow focus on variables not collectively decided with Indigenous Peoples, women, peasants and family farmers, workers throughout food systems, fisherfolk, pastoralists and consumers; transparency in the collection to avoid surveillance, reproduction of injustices and inequalities in data analysis, absence of approaches to data usage that come from territorial experiences, the production of narratives that translate wisdom, other types of knowledge, *qualitative experiences*, and diversity into data-coded language.

Due to the ownership of tools (from clouds to satellite to applications and platforms) and the biased methodological assumptions behind recollection and analysis, Data, after its transformation into information for public policies, has historically been used to selectively illuminate some aspects of reality while leaving others in shadow. The HLPE report should find ways to address that.