

# Sentinel

Social and Environmental Trade-Offs  
in African Agriculture

## Horizon scanning draft technical report – April 2020

**Adam Devenish<sup>1</sup> & Joe Tobias<sup>1</sup>**

<sup>1</sup> Department of Life Sciences, Imperial College London

# Background

## What is the problem?

Continued population and consumption growth means that global food demand is likely to be one of the major future challenges facing humanity<sup>1</sup>. With regions such as Sub-Saharan Africa (SSA) facing an expected doubling of human population and tripling of food demand over the next half century, posing a range of severe environmental, political and socio-economic challenges<sup>2–5</sup>. It is therefore important to realise that decisions made today, can have profound effects on the future.

In some cases, key United Nation's Sustainable Development Goals (SDGs) are in direct conflict, raising difficult policy-related decisions particularly in relation to trade-offs between food production and ecosystem health. Moving forward, it is therefore vital that we ensure decision makers have access to the most up to date and relevant knowledge to aid them in the decision-making process. The first step towards achieving this is through the identification of the most critical research questions that, if answered, would have the greatest positive impact on addressing these trade-offs and ensuring future productivity and resilience of the African food system.

**“Achieving food and nutrition security, reducing inequality, and preserving terrestrial ecosystems: 100 critical research questions for decision makers in sub-Saharan Africa.”**

### *Objectives:*

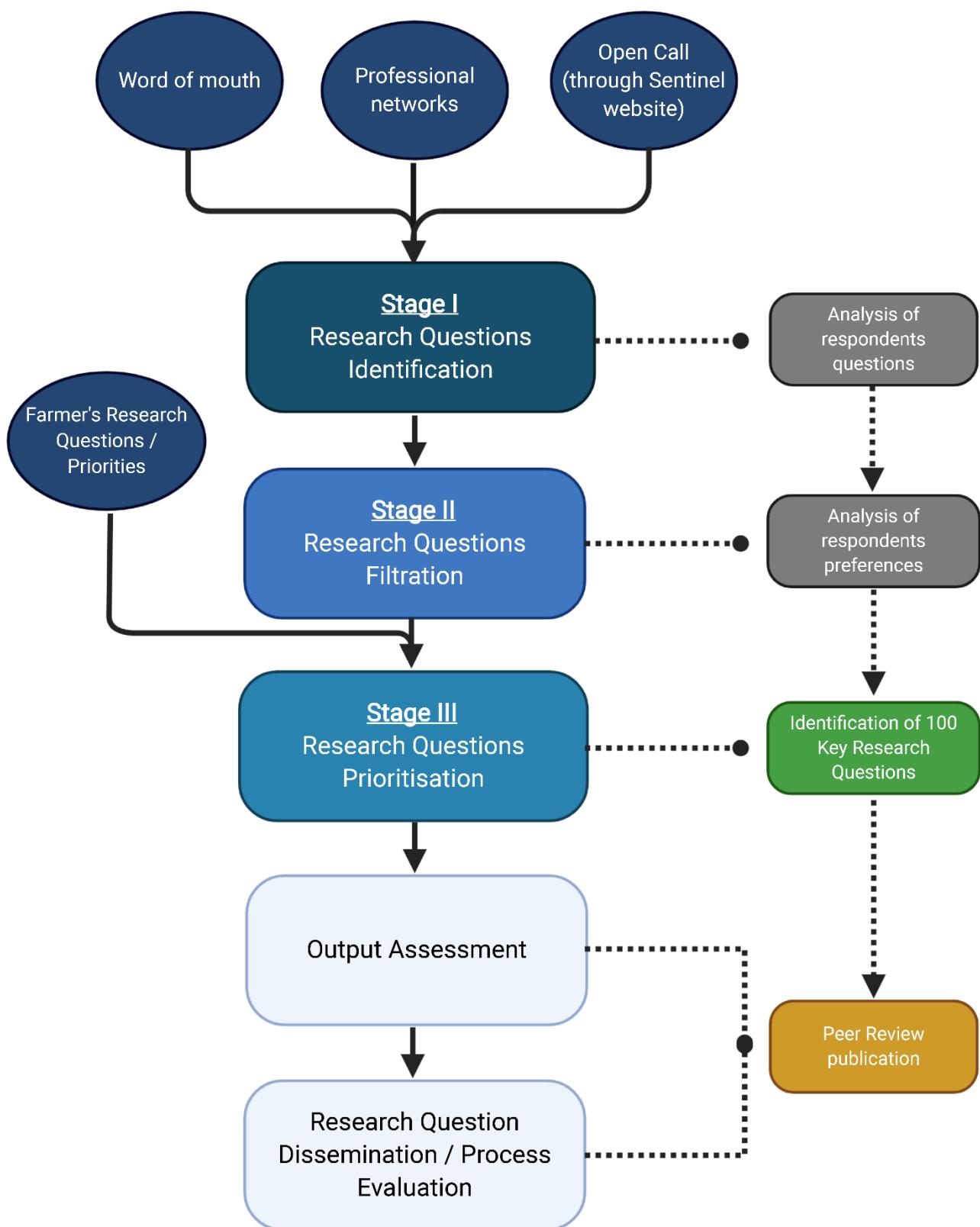
- 1.** To identify the 100 most critical scientific questions that can be used to steer research and aid policy makers in meeting the challenges of the three key SDGs (by 2030).
- 2.** Therein also strengthening and building capacity of UK and African partners to carry out interdisciplinary research on trade-offs that are relevant to research users.

# Method

## What is the approach?

Despite the urgent need for evidence-based policy, there is often very little information exchange between decision-makers and researchers. This is driven by the fact that there is a potential mismatching between the priorities of the producers (i.e. academic researchers) and consumers (i.e. the policy makers and local practitioners) of research. In order to narrow this gap, an interdisciplinary and inclusive approach is required, that engages and involves representative from a wide range of organisations and professions across SSA.

With this objective in mind, this paper uses a horizon scanning exercise to co-create research questions, in doing so build capacity. This approach has been adapted from other similar horizon scans<sup>6–8</sup>, and follows a three-stage process (Fig 1).



**Figure 1:** Sentinel horizon scanning process

Background.....	2
Method.....	2
Stage I: Research question identification.....	5

## Stage 1: formulation and gathering of research questions

Contributors were contacted either directly (via email/telephone/workshop meetings) or indirectly (via an online call for participants) and were asked to put forward research questions relating to topics of food and nutrition security, reducing inequality, and ecosystem health in Sub-Saharan Africa. Each participant was informed that while there are no restrictions on the type or number of questions that could be submitted, however, to ensure that the questions address the call, and make their way to the final stage, they are required to meet four key requirements:

1. Must address a knowledge gap (or gaps) that can be filled within a reasonable time frame (e.g. <5 years), using a realistic research design.
2. Cannot depend on the outcome of another question.
3. Should not be able to be answered with a simple ‘yes’ or ‘no’ response.
4. Should define a subject, an intervention, and a measurable outcome, if related to an impact and intervention.

In addition to submitting research questions, participants were asked to provide further background information on themselves (incl. their occupation, affiliation, area of work, country or work, whether they are consumers or producers of research), and to score seven questions aimed at determining their potential research priorities. It is important to note that while this information collected had no bearing on the validity or relevance of the participants research question, it was instead used as a relatively simple means by which to gauge the demographic and research priorities of the participants. Furthermore, the participant’s questions (and priority scores) were clustered and analysed using a variety of thematic methods, including text-mining and topic models.

Prior to transitioning to stage two, all questions will be screened against the above criteria. Those that do not meet the criteria may either be omitted, or wherever possible modified sufficiently to ensure that they are eligible for the exercise.

## Stage 2: voting and short-listing

All respondents involved in the stage one shall be invited to participate in the anonymous short-listing of questions. To avoid potential voter fatigue, each participant will be asked to score a random subset of thirty questions. While at this point it would be desirable to get the participants to rank the questions, previous studies have noted that ranking can lead participants to feeling pressurised to select broad questions<sup>7</sup>. To avoid this potential problem, each participant instead will be asked to assign each question to one of three pre-defined categories: “essential”, “possible” or “irrelevant”. This selection will be based on the context, that this question if answered, should have the greatest positive impact on addressing the challenges of achieving food and nutrition security, reducing inequality, and/or preserving terrestrial ecosystems. Importantly, no limit shall be given to how many questions can be assigned to each category, however, each respondent will be given a fourth category (labelled “additional”), in which they can record any other questions that

they think should be put forward to Stage three. All questions, bar those that only ever get placed in the “irrelevant” category will be carried forward to the next stage. Additional questions may be carried forward based on the activities of other Sentinel research partners, specifically those which engage with local practitioners on the ground (incl. farmers) to identify what their research interests and priorities are.

Finally, similarly to stage one, the question assignments will be analysed in order to determine which topics are considered of greater importance, as well as what effect (if any) participant demographics or research priority ranking has on research question selection.

### **Stage 3: research question prioritisation**

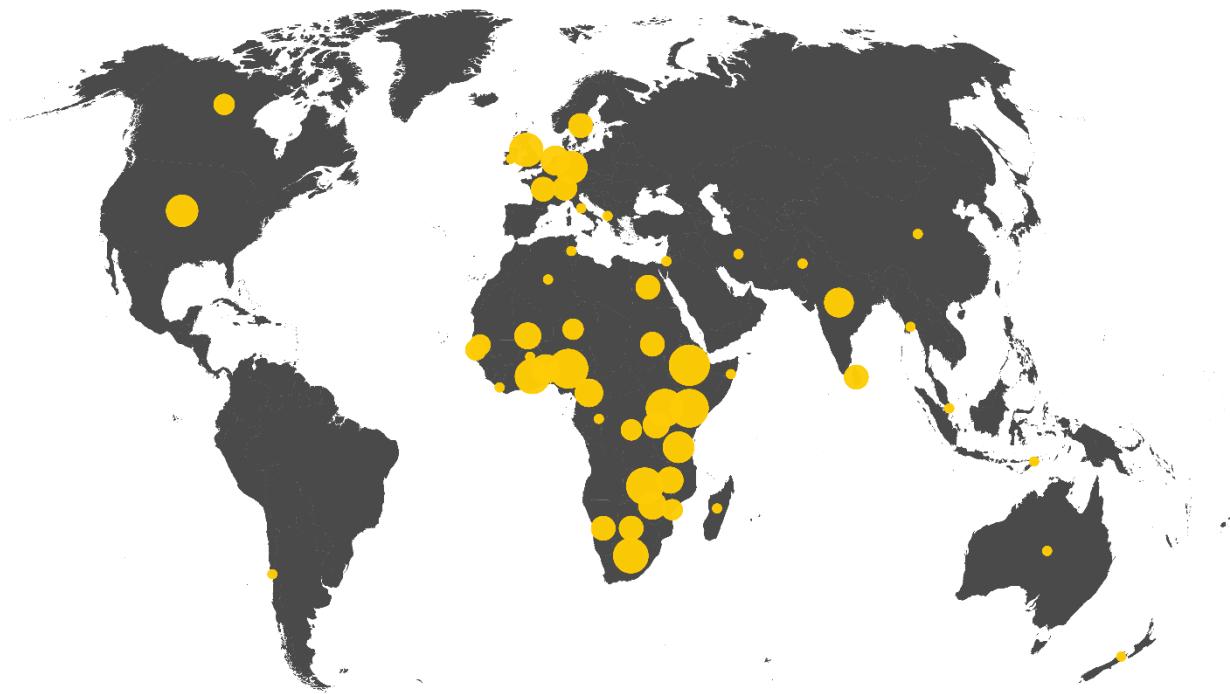
In this last stage, a subset of 30 respondents will be selected and invited to attend a two-day workshop, aimed at identifying and selecting the final 100 research questions. Selection of these participants will be done so as to best reflect the diversity of participants involved in the study and cover a broad array of stakeholder groups, expertise and demographics. Given that this exercise is aimed at a capacity building between UK and SSA stakeholders, an equal number of respondents will be selected from both UK and SSA.

Furthermore, as the objectives of this exercise align closely with the priorities of the RUFORUM project partners, this workshop will most likely be hosted in Uganda. However, given the ongoing global situation in relation to COVID-19 the venue and timing is still up for discussion.

## **Stage I: Research question identification**

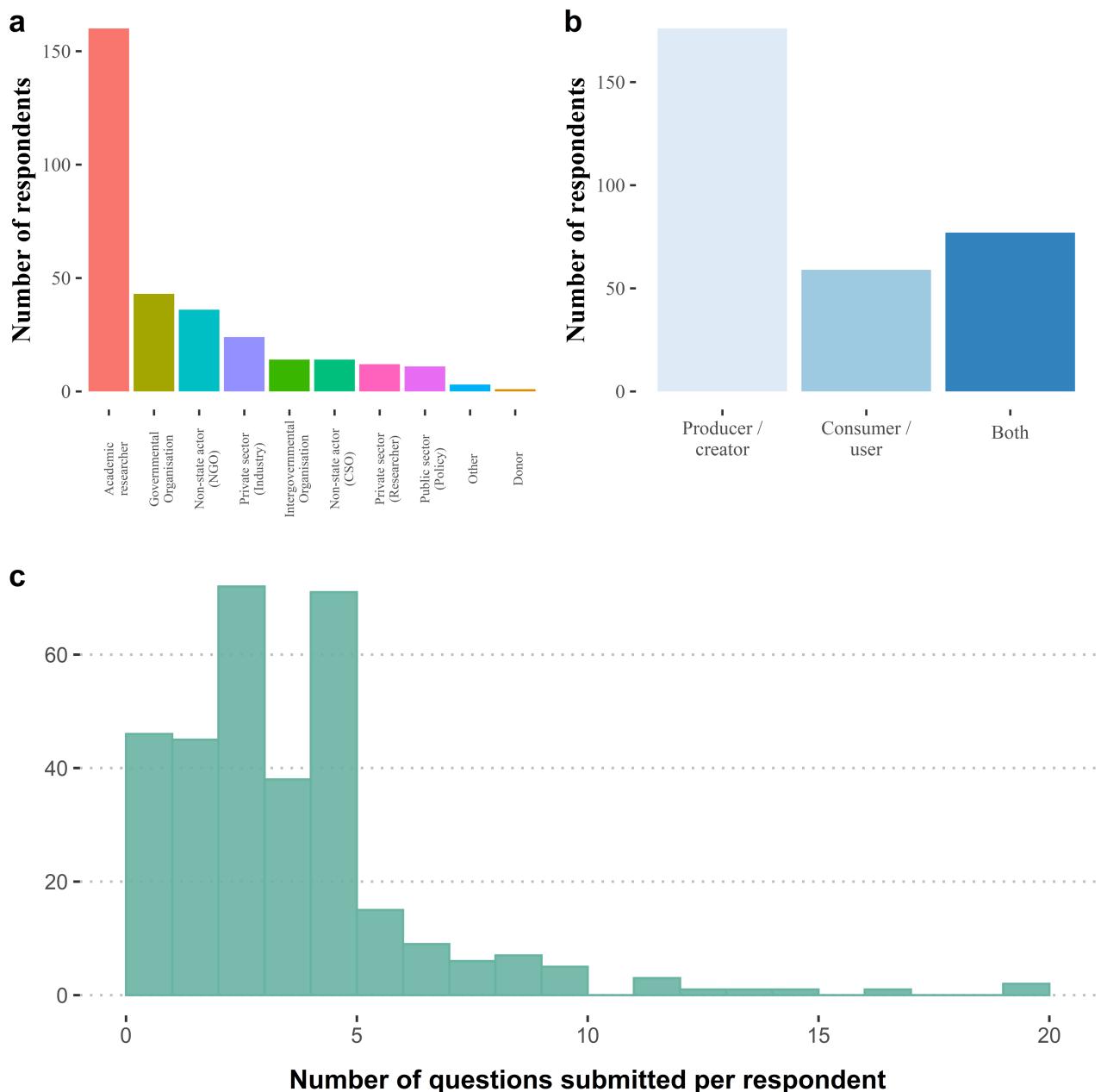
**Where are we so far...**

In total 318 complete responses have been received so far, with over half of the responses being received through the Sentinel website (open call). With respondents from 55 countries submitting questions; 75.5% of which work or undertake research in sub-Saharan Africa (Fig 2).



**Figure 2:** Global distribution of horizon scanning respondents. Circle size reflects number of respondents (log transformed).

Unsurprisingly, based on the responses received so far, the majority of respondents self-identified as academics (Fig 3a) and/or producers of research (Fig 3b). However, it's important to note that many of the respondents from SSA reported additional professions, in addition to their primary profession. As such, primary professions alone may not be a true reflection of the distribution of stakeholders involved in this exercise.

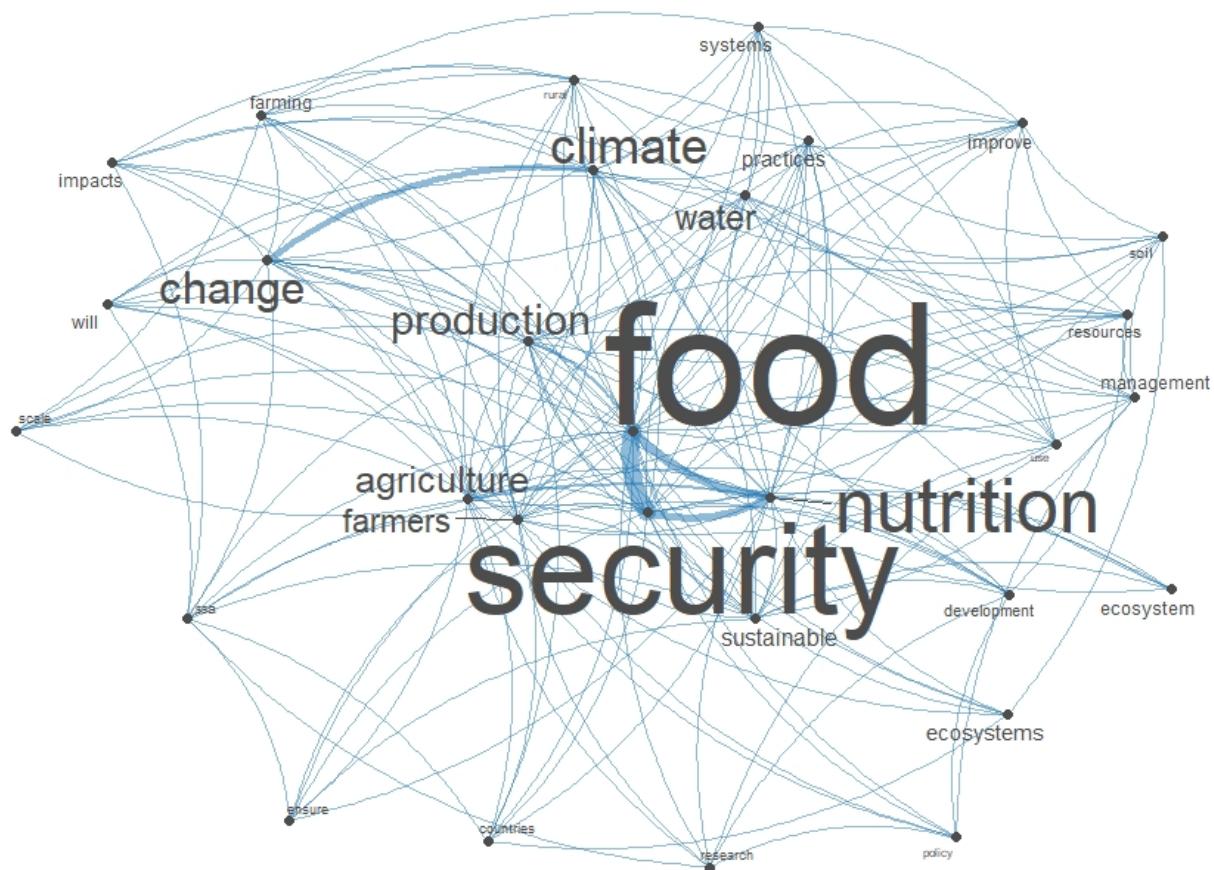


**Figure 3:** Horizon scanning respondents demographics

A combined total of 1339 questions were submitted (in both English and French), and while on average each respondent submitted four questions, there was some variation in the total number of questions received from each participant (Fig 3c).

### Preliminary analysis 1: key research topics

To identify the key topic areas within the pool of received research questions, the raw (unedited) responses were tokenized and sorted according to frequency. These text fragments were then arranged and networked according to their frequency and co-occurrences (Fig 4).

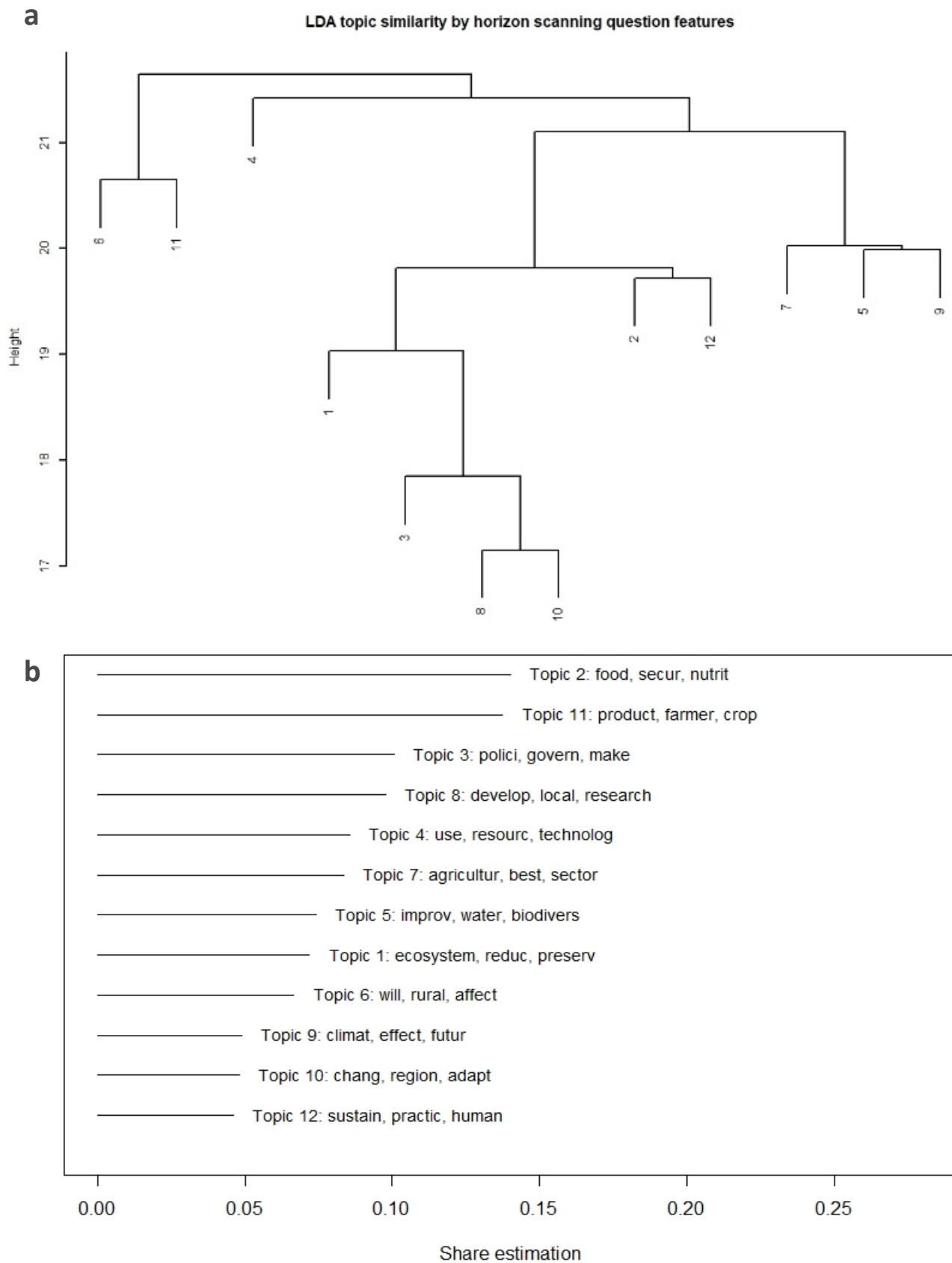


**Figure 4:** Network of research question feature co-occurrences. Node (word) size reflects term frequency.

Based on this information alone, it's clear that food security/nutrition is by far the major topic reflected in our research questions, followed by climate change. To further explore and identify the major topics or concepts present within the pool of research questions, the tokenised text fragments ("bag of words") were analysed using a series of unsupervised topic modelling methods<sup>1</sup>. In this instance, an arbitrary number (12) of topics was selected and the text fragments were clustered accordingly (Fig 5a). These text clusters were then passed back through the pool of research questions to see what proportion of topics are found in the overall corpus (pool of research questions). Using this approach, we identified 12 defined research topics or thematic research clusters (Fig 5b).

At this resolution, it's clear that at least two of major topic areas (food and nutrition security, and ecosystem health) are clearly reflected within our pool of thematic clusters. With both topics on food security/nutrition (topic 2) and farmer/crop production (Topic 11) forming the largest share of research questions (Fig 5b).

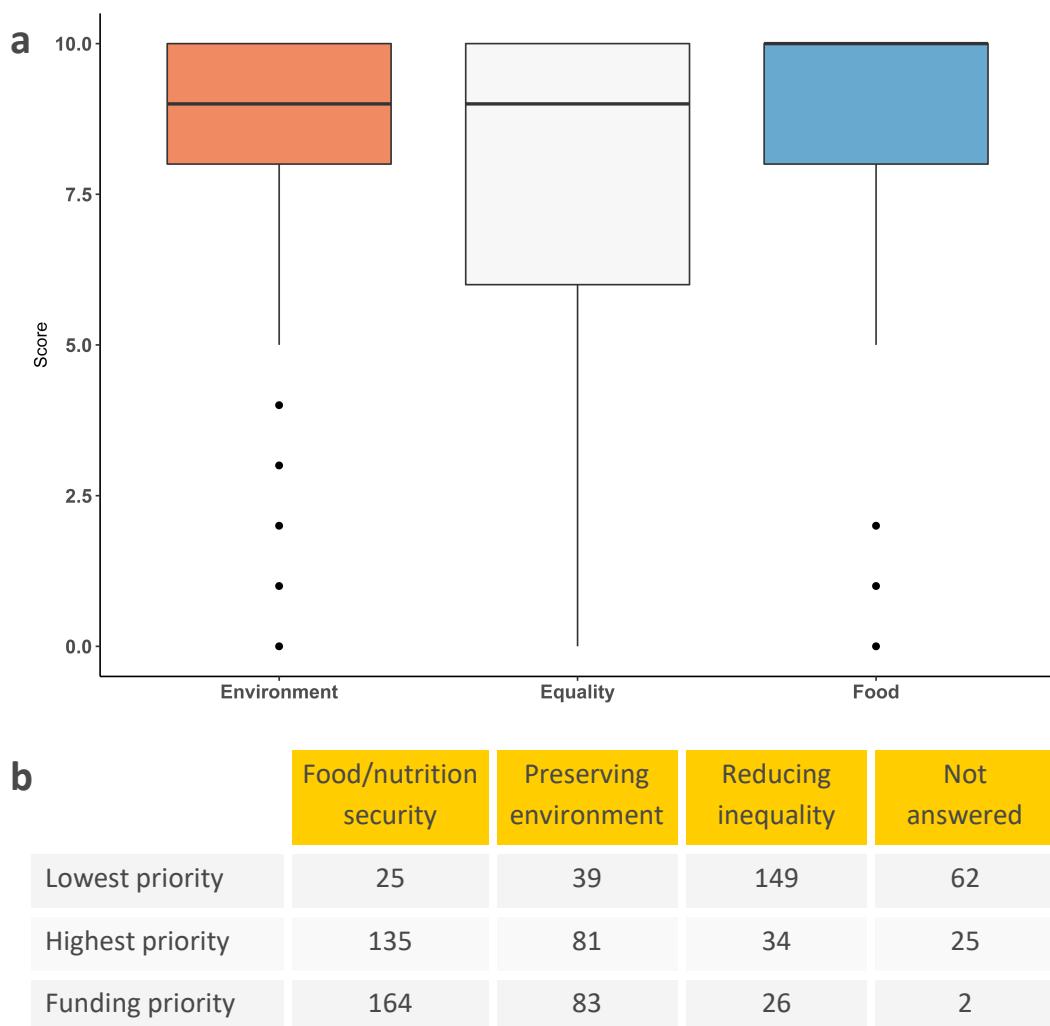
<sup>1</sup> Latent Dirichlet Allocation (LDA) and Structural Topic Modelling (STM)



**Figure 5:** Topic modelling results of horizon scanning questions

## Preliminary analysis 2: research priority

All respondents were asked to score (on a scale of 1 [lowest] to 10 [highest]) the importance of preserving the environment, food/nutrition security and reducing inequality when making decisions concerning agricultural development. Despite all three topics scoring relatively highly, on average reducing inequality scored consistently lower across all respondents<sup>2</sup> (Fig 6a). A finding that was further reflected in the aggregate choices of research/funding priorities (Fig 6b).



**Figure 6:** Boxplot and table of respondents' research and funding priorities

Interestingly, respondent's location (in terms of whether they were based in/outside SSA) and/or profession had no discernible effect on scoring of research priorities<sup>3</sup>. While this priority scoring is a relatively crude approach to determining respondent research priorities, these results do line up with the semantic networks/topics identified from the research questions (Fig 4 and 5).

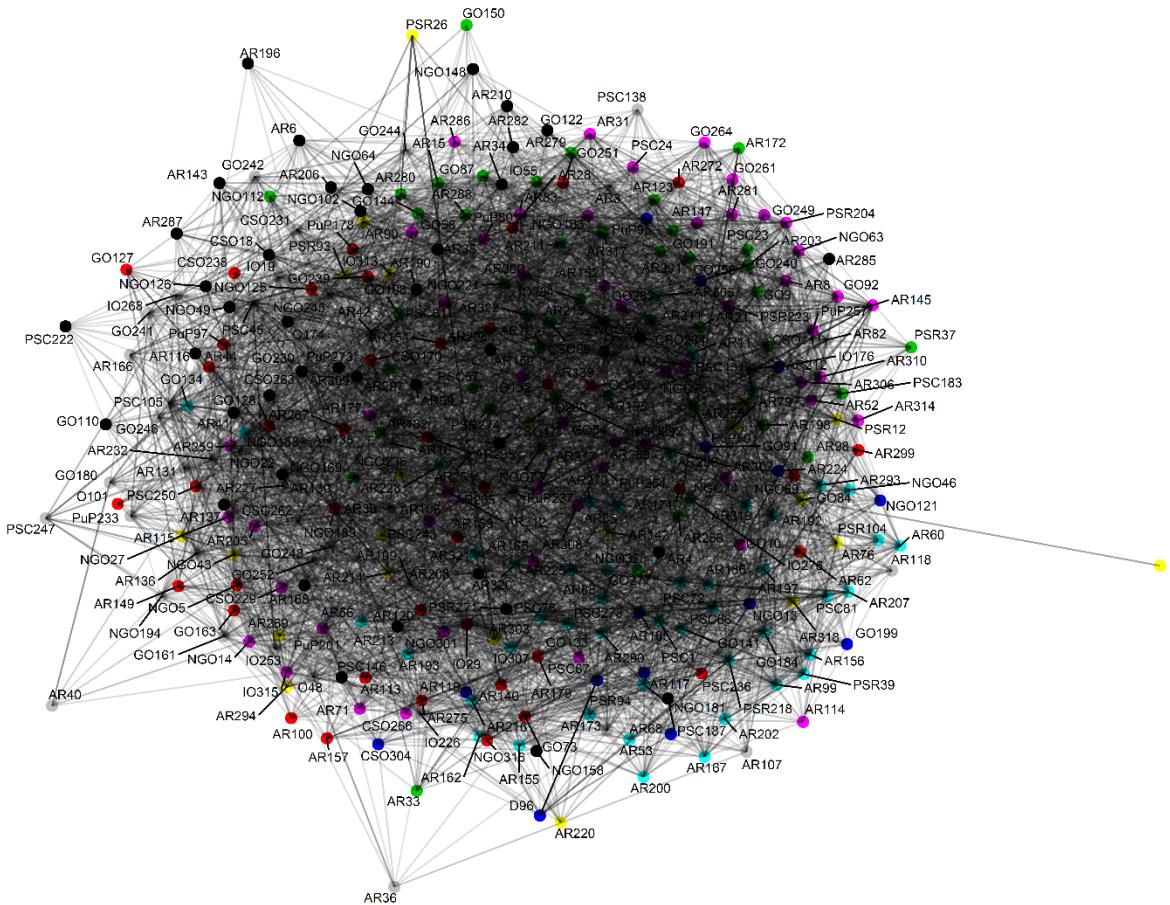
<sup>2</sup> Significantly lower than both food/nutrition security (*GLM*: SE = 0.29, z = 4.15, p < 0.0001) and preserving the environment (*GLM*: SE = 0.30, z = -2.30, p < 0.05)

<sup>3</sup> Not a significant predictor (*GLM*: SE = 0.27, z = 0.98, p = 0.326) of research priority scoring

## Preliminary analysis 3: respondent clustering

Based on the assumption that a knowledge gap exists between producers and consumers of research, we revisited the respondents research questions to see whether there was a relationship between word use (frequency and distribution of words across research questions) and primary profession. With the assumption being that this relationship would manifest in clear clustering according to profession.

Using noun phrases<sup>4</sup> extracted across our corpus (pool of research questions) we found no evidence of clear clustering according to primary profession (Fig 7).



**Figure 7:** Network of respondents (nodes) based on text fragments found in research question.

It's worth noting that this first pass was done using an unsupervised method and is likely unsatisfactory due to the low repetition of noun phases needed to create clear clustering. Therefore, a further in-depth supervised analysis is planned for the future, in which similar noun phases are grouped across the corpus, prior to networking.

Nonetheless, in light of these limitations and caveats this approach did identify seven distinct clusters based on a unique set of noun phases (Table 1). With academics forming a significant portion of each group or research question cluster (Table 1).

<sup>4</sup> Approach recommended to overcome the poor semantics of traditional bag-of-word methods

**Table 1: Groups identified using textmineR and their associated elements**

Group	Proportion (%) of academics	Key text elements unique to each group
1	62	food security, program, control, gm foods, water availability, climate predictions, food productivity, influence, agriculture, fishery sector, pollination enhance, nutrition security
2	52	composition, groundwater, insecurity, protein, conflict zone, sensing, theory
3	47	wastewater, cultivation, plant, crops, regeneration, forest, ecosystem, value chain, smallholder, irrigation
4	45	disease prevention, contribute, food self-sufficiency, women empowerment, invasion, mosquito, soil increase, crop production
5	53	encouragement, greenhouse, farming, willingness, fisherfolk, rate, nature, cropping, degradation, drought
6	44	community forest, equity, disaster, consumption, equality, inequality, process, price commodity, IoT technology, development strategy, agribusiness model
7	47	plant biology, scale, farm produce, pesticide safety, food quality, soil management, fish, frontier, food quality, soil management, environment, accessibility, progress

## Next steps: Stage II and onwards

Now that funding has been secured to host the workshop (stage three) we will be progressing on to Stage two. However, we still have a rather large skew in the demographics of the respondents. It may therefore be advisable that in the process of organising Stage two that we reopen the call for participants, focusing on stakeholders based outside of SSA, and/or from non-academic professions/organisations. In line with this, the Sentinel website pages on the Sentinel project will be redevelop to reflect the progress of the exercise (using selected infographics, including attached sunburst chart) and to fill in stakeholder gaps wherever possible.