

Exploring the Potential of Customized AI Chatbots in Food Policy Research: Capabilities and Constraints in Comparative Perspective

Todd Benson^{a*}

^aIndependent Consultant, Washington, DC, USA

INFO

<i>Submitted</i>	10 December 2023
<i>Keywords</i>	Generative Pre-trained Transformer (GPT), Artificial Intelligence (AI), Chatbot
<i>Flagship</i>	Co-LAB
<i>Work Package</i>	Enabling environment

ABSTRACT

Artificial Intelligence (AI) tools are rapidly transforming the way people work, providing opportunities for improving efficiency, accelerating innovations, and supporting human decision making. The area of policymaking is no exception. To explore the advantages and disadvantages of customized chatbots for supporting food policy research, we created a virtual food policy expert using a customized AI system using training literature drawn from about two dozen food policy research reports or food policy documents. A range of questions seeking food policy advice were then posed to the chatbot to better understand the limits to its expertise, the accuracy of its responses, and the value it might offer users seeking specific advice to guide the formulation and implementation of food policy related strategies and programs. A relatively simple comparative assessment was made of the results obtained from asking the same food policy-related questions to different AI sources, including the customized AI virtual food policy expert. Overall, we found that AI-based chatbots can be useful tools for summarizing research and providing “first cut” information only. We also conclude that, while customized chatbots have the advantage of offering the possibility of uploading pre-selected papers, unless a large amount of information is uploaded or the chatbot is also allowed by design to draw on information from other sources including the internet, the responses the virtual coach provides may be limited to only a small and specific set of questions. To be more credible and useful for policymakers, chatbots need to be enabled to provide more specific information. For the foreseeable future, providing such detailed and often country-specific information will still be in the hands of human experts.



1. Introduction

The significant expansion over the past two years in the availability of artificial intelligence (AI) tools based on generative pre-trained transformer (GPT) large language models to mimic informed human responses on a wide range of topics suggests that thematic expertise might be as effectively communicated virtually as in-person. While not infallible, chatbots like ChatGPT and Bard readily provide quite sophisticated, clearly stated responses to relatively complex queries. The International Food Policy Research Institute (IFPRI) is a global center of excellence for advising national governments and other stakeholders on food policy issues, including on agricultural development, food security, and nutrition. Might a virtual food poli-

cy expert that draws on the high-quality research output from its researchers be a valuable tool for IFPRI to use in its efforts to communicate evidence-based solutions for addressing food policy problems around the globe?

IFPRI researchers are exploring the impact that AI-powered chatbots might have in strengthening and improving the quality of agrifood system policymaking. This work is being done to support the Research Initiative on Digital Innovation of the CGIAR, the consortium of international agricultural research centers of which IFPRI is a part. There are several strands to this research initiative. All include examining the effectiveness of communication products or channels to provide evidence-based food policy support, comparing those produced by or involving traditional methods with those that draw upon AI tools. AI-driven virtual coaches or experts are among the AI-powered tools whose value for policy communication the IFPRI researchers seek to assess.

*Corresponding author. Previously worked at IFPRI (1999-2021). Email address: todd.d.benson@gmail.com. The Virtual Food Policy Expert logo credit: DALL-E by OpenAI.

This document reports on an exercise undertaken by an experienced food policy researcher contracted by IFPRI to explore the opportunities that such a virtual food policy expert might offer IFPRI, while also identifying any challenges or risks the Institute might face in doing so. A virtual food policy expert was created using the Coachvox AI system, an online proprietary virtual coach creation tool, using training literature drawn from about two dozen food policy research reports or food policy documents. The themes of these documents included agricultural and broader economic development, food security, and nutrition, particularly public health nutrition. They were drawn from country-specific or more generally applicable research and policy support work that involved the consultant. Once the virtual food policy expert was created, a range of queries were then used to better understand the limits to its expertise, the accuracy of its responses, and the value it might offer users seeking specific advice to guide the formulation and implementation of food policy-related strategies and programs.

This report describes the Coachvox AI system that was used to create the virtual food policy expert and the challenges that had to be overcome in processing the food policy documents so that the virtual expert could use their content. Once the virtual food policy expert was operational, a relatively simple comparative assessment was made of the results obtained from

asking the same food policy-related questions to different AI sources, including the Coachvox AI virtual food policy expert. The report includes some discussion of this analysis, and the output from the various AI tools used is presented in an annex. The report concludes with some recommendations and cautions that IFPRI might consider in using virtual experts that rely on AI to offer policy solutions.

2. Methods

2.1. The Coachvox AI System

Coachvox AI (<https://coachvox.ai>) is a recently established company based in the United Kingdom that, on a commercial basis, uses artificial intelligence tools to create virtual coaches¹. Users can interact with such a virtual coach around the clock through a chat screen to obtain AI-assisted guidance and advice immediately. The principal advantage of the virtual coach is the immediacy of interaction. This information is presented in a style consistent with a personal interaction with the individual being represented by the virtual coach.

The principal market targeted by Coachvox AI comprises lifestyle coaches who provide personalized services for individuals to achieve their personal goals, whether around employment, personal finances, social relationships, health and fit-

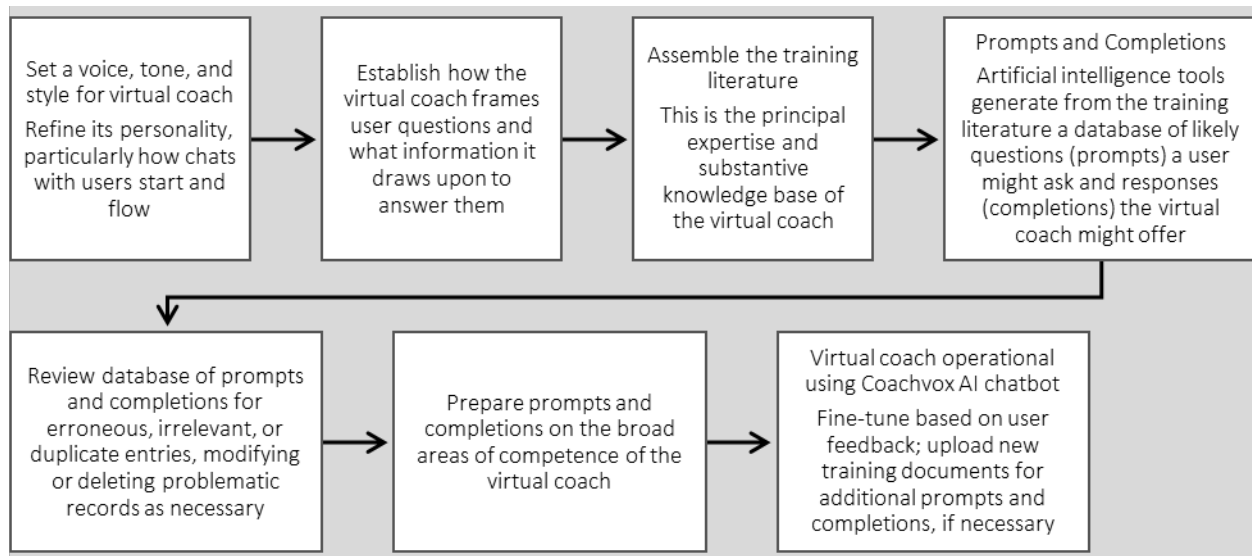


Figure 1 Steps to creating a virtual coach using the Coachvox AI system (Source: Adapted from Coachvox AI online documentation)

¹ Currently, there is considerable competition in the business of creating virtual coaches—a basic web search on ‘creating chat-based virtual coaches’ demonstrates this. The choice of Coachvox AI for the exercise to create a virtual food policy expert was made by the IFPRI researchers involved based on their assessment of Coachvox AI as a relatively accessible and cost-effective means of creating a virtual coach. Selecting the virtual coach creation platform to use was not part of the consultant’s scope of work.

ness, avocations, or otherwise. Being available to their clients, even if in virtual form, 24 hours a day is advantageous to such service providers and a useful and potentially profitable addition to their businesses. Despite being designed for a market unlikely to involve the provision of food policy advice, Coachvox AI can also be used to create a virtual food policy expert, as for the exercise here².

Coachvox AI uses a stepwise approach to create a virtual coach (Figure 1). This process is described in considerable detail here to highlight how certain of these steps used by Coachvox AI and its competitors may account for some of the strengths and weaknesses of virtual coaches, including those seen in the virtual food policy expert developed in this exercise. As virtual coach technologies evolve, some of the weaknesses now seen may be addressed.

First, the voice, tone, and style of the virtual coach are set. These characteristics should reflect the objectives for which it is being created. Depending on their area of expertise, lifestyle coaches will likely choose their virtual representation to be friendly; chatty, with considerable give-and-take with users; quite empathetic; and supportive. However, for the virtual food policy expert, an authoritative, relatively verbose and detailed, objective, and somewhat distant voice, tone, and style was adopted. An extensive set of phrases consistent with these traits is then developed to handle how conversations will start and then flow between the virtual coach and users.

Thereafter, how the virtual coach works with new users to better understand their questions is established. This is done by deciding the depth to which the virtual coach questions the user to obtain further specifics on their issue of concern. The better the virtual coach understands the user's needs, likely the more usefully it will frame its responses. However, a balance must be achieved to avoid the virtual coach being overbearing.

The creator also specifies what should be done if a user asks questions for which no information has been included in the expertise and knowledge base created for the virtual coach. Two options follow:

The first is to have the virtual coach rely on the internet to a specified degree to generate an answer. Coachvox AI provides a 7-level slider that runs from "Exclusively use my content" (1/7) to "Use the internet only when necessary" (2/7) to "A preference for my content" (4/7) to "Use my content and that of the internet equally" (7/7). For preference levels 2 to 7, the virtual expert will always provide an answer, drawing on the internet if necessary.

If the virtual coach is designed to use its own content exclusively (preference level 1/7), then the creator develops a set of phrases the virtual coach will employ to inform the user that the question posed is outside of its knowledge bank, encouraging the user to continue the chat by rephrasing their questions or by asking about a different dimension of the issue of concern.

Where this internet content slider is set will be an important design consideration in developing a virtual food policy expert by IFPRI, since the validity of the evidence base for any information drawn from the internet may be unclear.

The next step is to create the expertise and knowledge base upon which the virtual coach draws in its chat-based interactions with users. This is developed from training literature that the creator uploads for the AI tools of Coachvox AI to analyze and extract relatively extensive sets of unique prompts and completions (P&C) that reflect the information in the training literature. Each P&C is a possible user query ("prompt") and a virtual coach response ("completion") to that query, both of which are developed by the AI tools from the uploaded documents.

In selecting documents for the training literature, Coachvox AI suggests that a tree-based conceptual framework be used to identify what inputs are needed for the knowledge bank for the virtual coach. The 'trunk' consists of the broad but well-defined area of expertise the coach possesses. 'Branches' of the tree are specific thematic sub-topics within that broad area of expertise on which the coach will have particular knowledge and experience. The 'leaves' in this conceptual framework are the specific details within each sub-topic that are extracted as P&Cs from the documents in the training literature.

For personal services coaches creating a virtual coach, Coachvox AI suggests that in creating the training literature they consider articles, blogs, certain types of social media postings, presentation files or slide decks, training course materials, book chapters, and emails that they have created which contain relatively clear and straightforward articulations of the content they would want their virtual coach to convey in response to users seeking their guidance and support. Documents that are reasonably simple in structure and content and well-focused are better than those that are complex and nuanced. If complex documents are included in the training literature for the virtual coach, Coachvox AI recommends that they be run through an AI chatbot like ChatGPT to extract their major points. The synopses would then be uploaded as

² To date, Coachvox AI has only created virtual coaches that work in the English language. However, as AI tools are readily adaptable for use with other languages, virtual food policy experts could be created using similar virtual coach creation platforms optimized for other languages.

part of the training literature, rather than the detailed documents.

The training literature for a virtual food policy expert will be quite different from that used to create a virtual personal coach. Complexity, context-specificity and detail, empirical results, and nuance are inherent to much applied food policy research. As will be discussed, in using Coachvox AI to create a virtual food policy expert, the creation of the P&C database upon which the virtual expert would run using typical policy research outputs for the training literature proved to be a significant challenge.

Each document selected to make up the training literature for the virtual coach is uploaded in turn to Coachvox AI as *.pdf files. Coachvox AI sets no specific length limitations on the documents. When uploading, the system asks the creator to give each document a unique name for inclusion in the database. The AI tools of Coachvox AI then analyze and extract a relatively extensive set of unique P&Cs from each. This is done within a black box—the creator does not provide any parameters to guide the AI tools of Coachvox AI as to the types of P&Cs that should be developed from each document.

The processing of each document by Coachvox AI results in a flat P&C database with five fields for each P&C record:

- an *id* field (usually hidden);
- a *prompt* field containing the text of the AI-generated query;
- a *completion* field containing the text of the AI-generated response to the prompt/query;
- a *labels* field that initially is empty, but to which one or more subject keywords for the P&C can be added by the creator later; and
- a *source* field containing the document name supplied by the creator.

The number of P&C records the AI tools of Coachvox AI will generate for a document will depend on its length and nature. For example, for a working paper of about 50,000 words reviewing in technical detail a range of soil fertility management issues in Malawi, Coachvox AI generated over 850 P&Cs. In comparison, a policy brief of about 500 words resulted in about two dozen P&Cs.

The creator then reviews within Coachvox AI online each P&C one-by-one, accepting, modifying the prompt or the completion if either needs additional information added or errors removed, or deleting the P&C altogether if it is inaccurate or is not useful for the specific purposes for which the virtual coach is being created³. With the training literature for the virtual food policy expert, many of the P&Cs generated by Coachvox AI referred to empirical or bibliographic details contained in a document; information that generally would be irrelevant to a user seeking broader guidance on how to deal with a specific food policy challenge. Moreover, as will be discussed, the AI tools created some wrong or nonsensical P&Cs from the food policy documents, although fewer than expected. All irrelevant or erroneous P&Cs need to be modified or deleted from the database.

However, the step of reviewing the P&Cs generated in Coachvox AI was very tedious. Only a single record could be reviewed and modified at a time. After reviewing one by one the P&Cs created by Coachvox AI for several lengthy food policy documents, deleting in some cases more than half, the assistance of the technical staff at Coachvox AI was sought to identify a better way to generate and review the P&Cs to optimize the P&C database for use by the virtual food policy expert. They suggested going outside of Coachvox AI and using ChatGPT 3.5 to do so. The P&C database created by ChatGPT from the training literature could then be reviewed and modified much more efficiently using spreadsheet software and uploaded as a *.csv file for Coachvox AI to use to create the virtual food policy expert. This detour to the normal Coachvox AI virtual coach creation process is described later.

Once the reviewed and modified P&C databases for all of the documents in the training literature have been uploaded to Coachvox AI, the substantive databank on which the virtual coach will draw is in place. However, some final tweaks are needed. These include preparing a set of P&Cs to specifically provide answers to questions from a user on the broad areas of competence of the virtual coach, drawing on the tree-based conceptual framework for the virtual coach described earlier. These prompts can be created by asking ChatGPT to list the most important questions someone might ask a coach with specific expertise. The completions to these prompts can then be created by the creator of the virtual coach or by asking ChatGPT or another AI tool to provide relatively generic answers to those questions. These P&Cs, together with the voice, tone, and style characteristics of the virtual coach estab-

³ At this stage in creating the virtual coach, subject keywords for each P&C can also be added, inserting them in the labels field of the database. While these keywords are not used extensively in how virtual coaches under Coachvox AI currently operate, they are expected to become more important in future versions of the system. Similarly, while the P&C databases created within Coachvox AI contain a source field identifying the document from which each P&C was generated, this also is not currently used in how virtual coaches developed through the Coachvox AI system operate. Moreover, if P&Cs are uploaded to Coachvox AI as a *.csv file, data in a

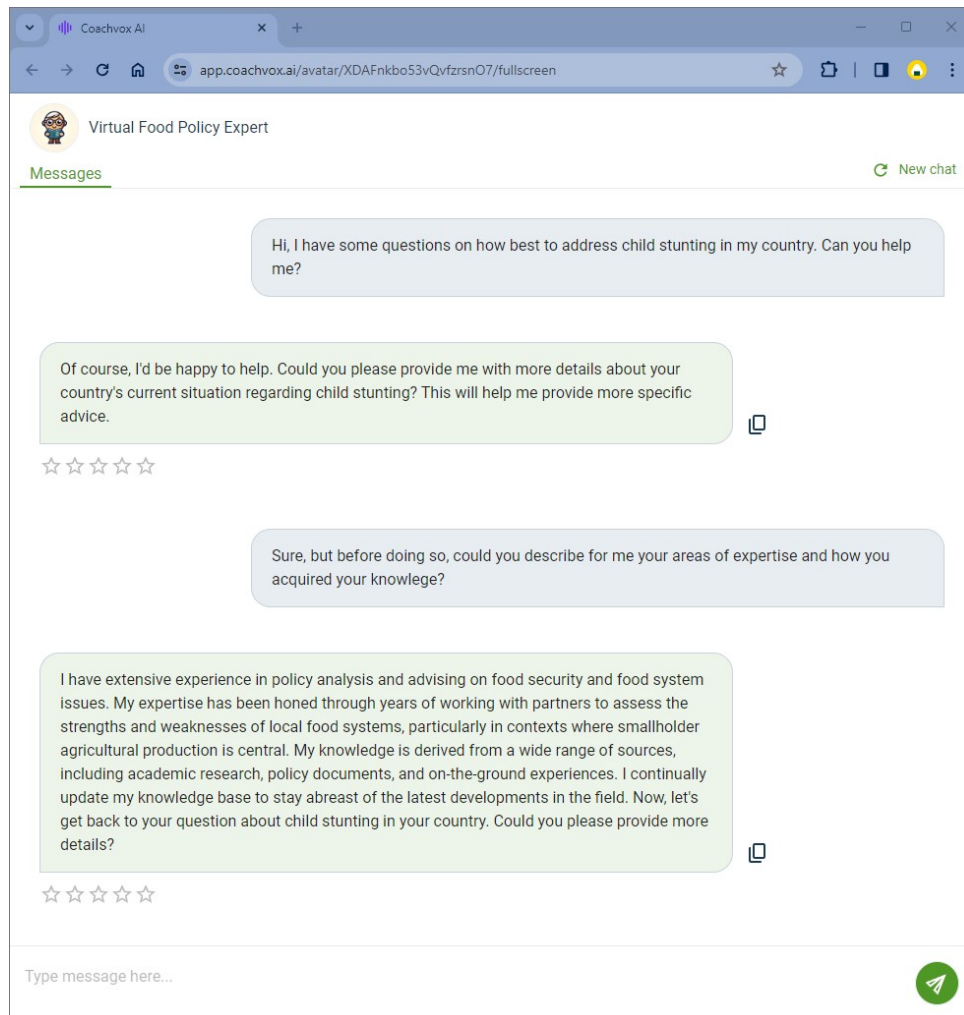


Figure 2 The Coachvox AI chat-based interface (Source: Author interaction online with the Virtual Food Policy Expert at <https://app.coachvox.ai/avatar/XDAFnkbo53vQvfzrsnO7/embed>)

lished earlier in the process, should enable a smooth start to a chat-based conversation with the virtual coach.

Once the virtual coach has been developed to a point where it can be tested more broadly, a web link for the virtual coach is provided to the creator by Coachvox AI. Clicking on the web link will open a relatively simple chat interface in a browser (Figure 2). The user can first be asked to provide a name to personalize the conversation and an email address in case of

the need for any follow-up, but this is optional. The chat with the virtual coach can then begin.

Based on initial interactions with users, the characteristics of the virtual coach can be modified to streamline the chat conversations better. Equally important, it may become apparent that not all of the areas of expertise of the coach were included in the original training literature or users frequently seek guidance on issues for which the coach has some insights, but no information on those issues was included in the training litera-

³ At this stage in creating the virtual coach, subject keywords for each P&C can also be added, inserting them in the labels field of the database. While these keywords are not used extensively in how virtual coaches under Coachvox AI currently operate, they are expected to become more important in future versions of the system. Similarly, while the P&C databases created within Coachvox AI contain a source field identifying the document from which each P&C was generated, this also is not currently used in how virtual coaches developed through the Coachvox AI system operate. Moreover, if P&Cs are uploaded to Coachvox AI as a *.csv file, data in a

ture originally. In such cases, additional P&Cs for the virtual coach can be created either by uploading documents for Coachvox AI's own AI tools to process or by using other AI tools independently of Coachvox AI to create P&Cs from new documents and uploading them to Coachvox AI as *.csv files.

2.2. Training AI

The training literature data

In considering how best to go about creating and evaluating the strengths and weaknesses of a virtual food policy expert for food policy communication purposes, the IFPRI researchers working under the CGIAR Research Initiative on Digital Innovation thought a relatively straightforward approach to doing so would be, in a sense, to replicate an actual person with considerable food policy research experience. Consequently, the consultant IFPRI hired for this exercise was Todd Benson, the author of this report. From 1999 to 2021, I was a researcher with IFPRI, primarily working on country-level development policy issues around agricultural and economic development, food security, and nutrition. Most of this work was done in Africa, although not all. For five years before joining IFPRI, I conducted agricultural research for the government of Malawi. To be clear, I have considerable experience on how best to address a relatively wide range of food policy challenges at country level, but I have no expertise in AI, in the use of AI tools, or in how such AI tools are designed and their consequent strengths and weaknesses. What I know now on these AI topics primarily comes from the work on creating the virtual food policy expert documented here.

The training literature proposed to be used to create the P&C database for the virtual food policy expert was some of the research output and country-level policy documents I had a hand in creating over the years, either as the principal author or as a co-author. In selecting the documents for this literature, diversity was sought in:

- thematic content—agricultural and economic development along multiple dimensions, food security, or public health nutrition;
- scope of analysis—technical or more strategic policy oriented;
- methods—quantitative analyses, model simulations, qual-

itative analyses, and conceptual work; and

- context within which the work was situated—country-level work in Africa (Malawi, particularly) and elsewhere, multi-country or continental-scale analyses, or focused on methods or concepts.

The documents included national policy documents on food security or nutrition, strategic reviews of food policy issues, policy briefs, research working papers, journal articles, and book chapters. The documents processed by AI tools to create a P&C database from each are listed in Annex 1. Twenty-one documents or compilations made up the initial training literature for the virtual food policy expert⁴.

Creating the Prompt and Completion (P&C) database from the training literature

To generate P&Cs from this training literature, an initial attempt used the tools in Coachvox AI. This was not successful on two counts. The initial upload was of all the documents in the training literature as a set. When the P&C databases created for each document were examined, only the first documents uploaded from the set were found to have accurate information in their records. The records of documents updated later were corrupted in various ways—uploading the complete set of documents appeared to have overwhelmed the Coachvox AI processing system. Upon request, the Coachvox AI technical staff deleted these initial databases from the system. The documents were then uploaded again, but one at a time, with the P&Cs for each being checked before uploading the next.

The second problem resulted from the unguided way Coachvox AI created the P&C database for each document. As noted, the system is designed to work with relatively simple, short, unnuanced, and quite narrowly focused documents. The reports, briefs, and book chapters drawn from applied policy research compiled for the training literature for the virtual food policy expert were considerably more complex and wordy than those for which Coachvox AI was designed. Consequently, when processed, several hundred P&Cs were often generated for each food policy research document, with many of those P&Cs being overly detailed or empirical in content, rather than providing more general guidance on food policy

⁴ Copyrights should be considered when selecting documents for the training literature. For this proof-of-concept exercise, this was not consistently done. While most of the documents used are working papers of various sorts or IFPRI publications published under a Creative Commons license, some copyrighted journal articles were also used. Most were earlier published as working papers. To respect copyrights, in retrospect, the better choice would have been to use the earlier working

priorities or strategies that countries might adopt.

After not very satisfactorily and tediously processing six documents from the training literature for the virtual food policy expert through Coachvox AI, the Coachvox AI designers were approached for a solution to the apparent mismatch between the sorts of documents Coachvox AI was designed to process and the food policy research documents being used to develop the virtual food policy expert. They quickly acknowledged that their tool was not very well suited for processing the sorts of documents being used to create the virtual food policy expert. Instead, they suggested ChatGPT 3.5 be used to generate the P&Cs from those documents⁵. The ChatGPT-generated P&Cs could be compiled and reviewed using basic spreadsheet tools and, when we had compiled a set of acceptable P&Cs for each document from all of the P&Cs ChatGPT generated, they could be appended into a single comma-delimited text file (*.csv) and uploaded to Coachvox AI. When uploaded, the P&Cs in the *.csv file are added to the P&C database already created in the Coachvox AI system for the virtual food policy expert.

Using ChatGPT 3.5, sets of P&Cs were generated for the remaining documents in the training literature. The following query was used in ChatGPT 3.5 to do so.

```
The following text has been copied from a [brief / book chapter / research report / etc.] on [summary of the focus of the text, e.g., how to make sure that reducing malnutrition is a national development priority]. I want to make the recommendations contained within it to be accessible to those who have not read the entire document, so I require as many of the key [recommendations / findings] to be distilled from it as possible. I would like this presented as a two-column table of questions and answers.
```

After providing this query to ChatGPT, it returned a request

for the text. A section of each document from the training literature not exceeding about 4,000 words was then copied into the ChatGPT interface. (The word limit is due to the token limitations built into ChatGPT for each interaction⁶.) Depending on the nature of the document, between a half dozen and three dozen P&Cs would then be generated by ChatGPT from the block of text it was provided.

In reviewing and, if necessary, modifying the ChatGPT-generated P&Cs manually, two issues related to how Coachvox AI uses P&Cs in actuating the virtual coach need to be taken into consideration. First is that the P&Cs are unique data points—P&Cs drawn from the same document will not contain information linking them, but must stand alone in terms of the topics which each addresses or the contexts to which each applies. As a result, information on the topic and context of applicability must be presented in the prompt and possibly also in the completion of each P&C pair⁷.

The P&Cs generated by ChatGPT for each block of text were reviewed to ensure this information was included in the prompt at least so that the P&C could be used in isolation from other P&Cs in the database. Most commonly, it was found that the country or regional context of a training literature document needed to be added to some of the prompts that ChatGPT had generated. Also, some prompts were generated that had no topical content, so that too was added. This information was added manually to the prompts when the P&Cs were reviewed in a spreadsheet.

More important for the logical operation of the virtual coach is the second concern of duplicate prompts in the P&C database. When faced with a question from a user that the Coachvox AI chatbot matches to more than a single prompt in the P&C database, Coachvox will be unable to logically choose which of the completions from the duplicate prompts to provide the user in response. This can lead to erroneous or nonsensical responses from Coachvox. Consequently, after using an AI tool, like ChatGPT, to generate a set of P&Cs on key findings or recommendations from a document, the dataset of P&Cs

⁵ At the time of this writing, 4.0 is the latest version of ChatGPT (<https://chat.openai.com>). However, the Coachvox AI designers noted that the earlier 3.5 version is stronger for generating chat-like responses, of which the P&Cs used by Coachvox AI are a form. All work using ChatGPT for creating the Coachvox AI virtual food policy expert was done using version 3.5.

⁶ Another limitation is a 30 query-per-hour limit for basic (non-fee) users of ChatGPT. This limit was reached several times in processing blocks of text from the training literature to create P&Cs for the virtual food policy expert.

⁷ However, this is only a characteristic of the records in the P&C database used to create the virtual coach. When engaging in a chat with the virtual coach on the

	A	B	C
1	Prompt	Completion	Labels
966	What are the main areas where governments can work to enhance fertilizer use by smallholder farmers?	Governments should improve information dissemination on fertilizer use, develop crop-specific fertilizer recommendations, and ensure clarity in regulations for importers and traders.	Uganda, Tanzania, Mozambique, agriculture
967	How should African governments regulate high-analysis fertilizers compared to other agrochemicals?	Governments should adopt less restrictive regulations for high-analysis fertilizers to lower prices for farmers while ensuring quality through self-regulation processes.	Uganda, Tanzania, Mozambique, agriculture
968	What should African governments consider regarding direct interventions, like input subsidies, for fertilizer use?	Governments should be cautious with direct interventions, as fertilizer demand depends on profitable crop sales. Focus on building efficient and reliable output markets instead.	Uganda, Tanzania, Mozambique, agriculture
969	Who are the nutritionally vulnerable groups of interest in the study of chronic malnutrition in northern Nigeria?	The study focused on infants and young children under two years of age, their mothers, women of childbearing age, and adolescent girls, as undernutrition has the most adverse consequences for them and society.	Nigeria, nutrition
970	What were the findings related to stunting in young children in northern Nigeria?	Stunting rates in northern Nigeria were more than twice the national average. There was no significant improvement in stunting levels between 2008 and 2013, with the rest of Nigeria making more progress.	Nigeria, nutrition
971	What is the significance of the generalized decline in the nutritional status of infants immediately after birth in northern Nigeria?	Newborns in northern Nigeria experience an immediate decline in nutritional status from birth due to maternal undernutrition, which is then compounded by poor post-birth nutrition or health challenges. This contributes to high rates of stunting.	Nigeria, nutrition

Figure 3 Random extract of records from the prompt and completions database file uploaded to Coachvox AI for creating a virtual food policy expert (Source: Author)

produced needs to be reviewed to identify any duplicate prompts, either removing a P&C pair for one of those duplicate prompts or modifying one of the duplicate prompts so that it is unique and better tailored to eliciting the completion with which the prompt is paired.

To avoid duplicate P&Cs, it is generally advantageous to use only a single ChatGPT query to extract P&Cs from a text block. If more than one ChatGPT query is used to generate P&Cs from a single text block, the chances of duplicate P&Cs being created are high. What is called ‘prompt engineering’ in using AI tools is important here. A virtual coach creator should experiment with different ChatGPT queries to obtain as output the sorts of P&Cs that are most useful for a virtual coach to use. In terms of content, how a ChatGPT query is formulated can result in the generation of considerably more P&Cs that reference substantive recommendations or research findings from the text block that are valuable for guiding responses to food policy challenges and far fewer P&Cs that are concerned with detailed context-specific empirical results or bibliographical content. Similarly, the formatting of the output from ChatGPT in a prompt and completion (question and an-

swer) format can be adjusted by changing the wording of the query.

However, even if only a single well-tested ChatGPT query is used to generate P&Cs from a block of text, two reviews to find duplicate prompts will still be required before the P&C database can be uploaded to Coachvox AI. The first review will be done at the document level. Because a longer document on a single topic will be made up of several blocks of text of up to 4,000 words from which P&Cs are generated by ChatGPT separately, it is quite likely that duplicate prompts will be found across all of the sets of P&Cs created from the text blocks making up the entire document. The second review will be necessary when creating a single *.csv file of the P&Cs created by ChatGPT from all training literature documents for uploading to Coachvox AI. As the training documents are related in their content, whether at the broad food policy level or by sub-themes within food policy, all P&Cs generated from the training literature should be jointly reviewed for duplicate prompts and modifications made so each prompt in the entire P&C database is unique.

Identifying duplicate prompts takes effort. One approach is to place all the P&Cs being examined in a spreadsheet and sort them by prompt to find duplicates. ChatGPT can also be used by asking it to find duplicates in a complete list of the prompts being examined. The following ChatGPT query generates reasonable results: Show me the pairs of duplicate statements in the following list. The duplicate prompts identified by ChatGPT are then located in the spreadsheet and modifications are made to them.

After using ChatGPT to create the P&Cs and reviewing and modifying as necessary the P&Cs for each document in a spreadsheet, labels were added to each P&C. While they are not yet a core element of how the virtual coach operates, the Coachvox AI designers say that these labels help “join the dots” between the question a user poses and the response provided by the virtual coach. More than one label can be attached to a single P&C. For the P&C database for the virtual food policy expert, eight thematic labels were used—‘food policy’, ‘agriculture’, ‘food security’, ‘nutrition’, ‘economic development’, ‘social protection’, ‘soil fertility’, and ‘agricultural markets’. Country name labels were also attached to P&Cs that were derived from country-specific research or policy engagements. The same labels were applied to all P&Cs drawn from a single document in the training literature. These are shown in the table describing the training literature for the virtual food policy expert in Annex 1.

In its current form, Coachvox AI is not conducive to experimentation on the content of the P&C database. Beyond the removal of individual P&C records, the creator of the virtual coach cannot backtrack on any decisions around what will be used to constitute the knowledge base for the virtual coach. Once any uploads are made to the P&C database on the Coachvox AI platform, it is difficult on one’s own to reverse those changes—the assistance of the Coachvox AI technical team would be required. Consequently, the creator should think hard about what they would like the virtual coach to do before uploading any new P&Cs to effect changes. Having learned lessons from earlier mistakes in the process, caution was taken before uploading to Coachvox AI the *.csv file that contained about 1,350 of the P&Cs that had been generated by ChatGPT, reviewed, and accepted or modified (Figure 3). With the P&Cs from the half-dozen documents from the training literature for the virtual food policy expert that Coachvox AI had earlier processed, the virtual food policy expert drew on about 1,650 P&Cs developed from documents in the training literature (Annex 1).

However, once operational, if it becomes clear through use

that the P&C database for the virtual food policy expert is insufficient or otherwise incomplete, additional documents can be processed using ChatGPT to create additional P&Cs. However, before these are uploaded, they will need to be compared to the existing P&Cs to ensure that no duplicate prompts are introduced with the new P&Cs. While Coachvox AI does not permit extensive editing of the P&C database within the platform, it does enable the creator to easily download the database that is currently being used for the virtual food policy expert. Comparisons between the existing P&Cs and the possible new P&Cs can then be made.

Using ChatGPT to create P&Cs to specify the domains of expertise of the virtual food policy expert

The virtual food policy expert will be operable once the P&C database developed from the documents in the training literature is uploaded to Coachvox AI. There was no delay between providing the database to the platform and the virtual coach being operational.

However, Coachvox AI suggests that anyone creating a virtual coach should create several dozen P&Cs that will introduce a potential user to the broad topics the coach covers. This can be done using AI tools, like ChatGPT, and employing the tree-based conceptual framework for the virtual coach sketched out earlier. For the virtual food policy expert, the ‘trunk’ is the broad field of expertise of the expert, ‘food policy’. The ‘branches’ here are ‘agricultural development’, ‘food security’, and ‘nutrition’—we also added ‘smallholder agricultural development’ to further refine the broader agricultural development theme. The following five ChatGPT queries were used to generate prompts that a client might use with the virtual coach:

```
What are 20 questions a policymaker from a developing country might ask [a food policy expert / an agricultural policy expert / an expert on smallholder agricultural development / an expert on household food security / an expert on reducing malnutrition in the population]?
```

ChatGPT drew the 20 prompts produced for each query from the internet and the rest of the large set of training literature used to develop ChatGPT. However, the creator may also manually create the prompts.

Initially, the consultant created completions (responses) to these prompts manually. However, laziness set in, and it was found that ChatGPT could create shorter and more focused completions for each prompt than the consultant could. So the

100 prompts that ChatGPT generated from the five queries were then used to generate a completion for each. The ChatGPT query used was:

Tell me in 30 words [the content of the Chat GPT-generated prompt]?”

An example of one of the ChatGPT-generated prompts is “How as a developing country can we address the issue of malnutrition and improve the nutritional quality of our food supply?”. So this text was simply copied into the ChatGPT prompt bar and the prefix statement of “Tell me in 30 words” was added to create the ChatGPT query:

Tell me in 30 words how as a developing country can we address the issue of malnutrition and improve the nutritional quality of our food supply?

The completions generated by ChatGPT from this and the other prompts were then edited and included in a *.csv file with their respective ChatGPT-generated prompts and uploaded to the Coachvox AI for appending to the virtual food policy expert P&C database.

As noted, the content of these ChatGPT-generated prompts and completions on the domains of expertise of the virtual food policy expert was drawn from content on the internet and elsewhere, and not from the documents in the training literature used to create the more specific P&Cs for the virtual coach. Consequently, some of the P&Cs ChatGPT generated on the domains of expertise for the virtual food policy expert were outside of the expertise of the training literature used in creating the virtual expert. For example, some prompts were generated by ChatGPT on fisheries management or on food sovereignty issues, both of which are not covered in the training literature used. Such P&Cs were not included in the *.csv file uploaded to Coachvox AI. In total, 95 P&Cs were generated using ChatGPT to introduce a client to the broad topics on which the virtual food policy expert might have expertise.

3. Results

3.1. Strengths and weaknesses of the AI tools in creating a virtual food policy expert

In the exercise to create the virtual food policy expert, AI tools were used twice to generate P&Cs for the virtual expert. The first effort used the Coachvox AI tools that are part of its online coach creation process with six of the food policy research documents in the training literature created for the virtual food policy expert. Being somewhat dissatisfied with how Coachvox AI processed the documents, the second effort to

use AI tools to generate P&Cs involved using ChatGPT 3.5 with text blocks drawn from all remaining documents in the training literature. Here we describe and assess the quality of the output of the AI tools used in the two efforts.

Using the Coachvox AI tools, the creator of the virtual coach has no control over how those tools process the documents in creating the P&Cs. The prompt engineering that controls how the AI tools operate was done by Coachvox AI technical staff in creating the system. A user creating a virtual coach using Coachvox AI cannot change those prompts. The control the user has over the P&Cs that will be output by the Coachvox AI tools is achieved through modifying the inputs—the documents in the training literature for the virtual coach. As noted, Coachvox AI recommends that creators simplify any complex documents in the training literature for a virtual coach to focus on key concepts and information, removing secondary details. This can be done by pre-processing them through an AI tool like ChatGPT, asking the AI to summarize the document and extract the key points that would be beneficial to those seeking advice related to the expertise intended for the virtual coach. If this is done, one can expect quite helpful P&Cs to emerge from the analysis of the pre-processed documents by the Coachvox AI tools. If the documents provided for analysis are simple and direct and the information they present can be used without consideration of context or other nuances, the black box approach Coachvox AI adopts in using its AI tools to generate P&Cs from such documents will likely work satisfactorily.

Moreover, the P&C generation process of Coachvox AI has some valuable functions not found in generic AI tools. These include the ability to accept as inputs and process text blocks longer than 4,000 words and to ensure that duplicate prompts are not created in generating P&Cs from a document. As managing the text length constraints and dealing with duplicate prompts were a large part of the effort to use generic AI tools to generate P&Cs for the exercise here, the value of these Coachvox AI functions is appreciated.

However, for complex documents, such as those in the training literature for the virtual food policy expert, the Coachvox AI tools are unsatisfactory. When using Coachvox AI to create the P&Cs for the exercise here, no pre-processing of the food policy research documents was done to simplify and to only present the higher-level findings and recommendations from each. Doing so was judged incompatible with the breadth of insight that a user consulting a food policy expert would expect to be provided, including context-specificity and nuance. Moreover, some experimentation

would have been necessary to determine how best to simplify the food policy research documents so that Coachvox AI could process them satisfactorily. However, in its current form, the platform does not facilitate much experimentation in creating the P&C database for a virtual coach. There is no easy, user-invoked way on the Coachvox AI platform to undo and remove unsatisfactory results from the P&C database.

Using ChatGPT 3.5 to generate P&Cs from the documents in the training literature provided more control over the output obtained than the Coachvox AI tools did. By experimenting with the text of the queries used with ChatGPT in relatively small ways, such as by changing one or two words, outputs were obtained that primarily focused on each document's broader research findings and policy recommendations. Generally, the ChatGPT queries that included the terms "key recommendations" or "key findings" produced higher-level qualitative results rather than detailed content. Relative to what the Coachvox AI tools generated, far fewer P&Cs emerged from the ChatGPT analyses that focused on specific empirical details of the research described in the documents, details that are unlikely to be of interest to a user seeking strategic advice for dealing with a food policy challenge. Information drawn from tables and literature citations in the research documents were also less common in the P&Cs output by ChatGPT relative to that produced by the Coachvox AI tools.

The text block length limitations and the issue of P&Cs with duplicate prompts associated with ChatGPT have been noted, as these required considerable effort to address to produce a useful P&C database for the virtual food policy expert. Although not linked to the content of the P&Cs, another problem encountered in using ChatGPT was that the formatting requested for the output "as a two-column table of questions and answers" was not consistently produced. This format is necessary to create the P&Cs for the virtual coach efficiently. Similarly, in using the AI tool to find prompts that may not be duplicates based on their specific wording but were duplicates in content and so would pose a problem for a virtual coach, the output would not always consist of pairs of such duplicate prompts. Instead, ChatGPT would do something other with the pairs than simply present them, such as by rewriting and consolidating them into a single prompt. These formatting issues generally were solved by emphasizing in one or two additional queries to ChatGPT the specific formatting required for the output. However, sometimes it was necessary to restart the ChatGPT chat to have it reliably provide output in question-and-answer format.

However and more broadly, this inconsistency in performance

on this secondary but important issue of the format of the output points to the general challenge of replicability in using these AI tools. The average user cannot be sure that the same type of output from the AI tool—whether in terms of content or format—will always be obtained using the same query on the same document.

3.2. Accuracy of AI analysis in extracting information from food policy research documents

In designing the exercise to create a virtual food policy expert, a significant concern of those involved was the accuracy with which the AI tools would extract and present information from food policy research-related documents. A standard critique of these tools is that inherent to their current design is that they may "hallucinate" and produce output that does not accurately reflect reality. In consequence, ChatGPT now includes at the bottom of all of its web pages the disclaimer "ChatGPT can make mistakes. Consider checking important information."

This problem was of considerably less significance for the exercise than feared. In reviewing the P&Cs that Coachvox AI and ChatGPT generated, identifying erroneous prompts or completions was critical. While no count was made of such P&Cs, retrospectively, a rough estimate is that only one or two out of 100 P&Cs were found to reflect the content of the source document wrongly. The accuracy of these AI tools is impressive.

The wrong P&Cs identified seemed primarily to result from sentences or paragraphs in the source material that involved contrasts, such as between the characteristics of or analytical results for different sub-populations. Two examples:

- Sentences or paragraphs with the logical structure of while X implies Y, not-X implies (or does not imply) not-Y, repeatedly were found to result in erroneous P&Cs. Errors typically seemed to result from ChatGPT omitting in the completion statement one of the four components of the information—X, Y, not-X, and not-Y. By ChatGPT dropping one of them in its processing the information into a P&C, the contrast was lost, and Y or not-Y was assigned to both X and not-X.
- Similarly, for a report that contrasted three agricultural marketing laws, ChatGPT created a completion that assigned the same characteristic to all three laws. In the original document, the point being made was that laws A and B had that characteristic, while law C notably did not.

While these were the primary patterns seen in relating the source materials to the erroneous P&Cs identified, other sen-

tence and paragraph forms may pose challenges for AI tools to process them into accurate P&Cs.

Identifying such erroneous P&C requires that the creator of the virtual coach knows the training literature quite well. The AI tools are verbally adept, being large language models. So, their erroneous interpretation of such text does not result in gibberish and nonsense, but in a well-crafted P&C that in form could seem plausible and correct. Consequently, in reviewing the P&Cs, the creator needs a reasonably close understanding of the source document and its conceptual foundations to identify nicely written P&Cs that are just plain wrong. Unfortunately, identifying such problematic P&Cs needs to be done by manually examining each in turn—ChatGPT or other AI tools are the source of the problem, so it is unlikely that any AI-enabled methods for doing so efficiently can be developed.

The level of detail from the source material the AI tools present is a second concern related to their accuracy. It has been noted repeatedly above that it is expected that any person consulting a virtual food policy expert will primarily be seeking strategic but somewhat generalized guidance on what policy reforms or public investment programs might effectively help overcome specific agricultural development, food security, or nutrition challenges. So in developing the P&C database for the coach, there will be some preference for generalized statements over specific ones. However, what if nuance and context-dependent considerations are essential to identifying appropriate solutions to those challenges?

In the initial testing of the virtual food policy expert after it became operational within Coachvox AI, the food policy researchers involved were immediately impressed with how intelligently and eloquently it engaged in a chat conversation. However, a typical second reaction was that the virtual food policy expert seemed to be too general in its responses. This may reflect the somewhat general nature of the questions (prompts) posed to it. However, it may also reflect a bias towards generalized statements in the answers (completions) it offers, which is linked to the creation of the P&C database by the AI tools.

Testing whether such a bias exists would be challenging to do partly because the provision of robust evidence-based guidance suitable for application by a wide range of users is among the reasons the virtual food policy expert was created in this exercise. Meeting this objective requires that the information the virtual expert provides is somewhat generalizable. Consequently, in engineering the ChatGPT queries used to extract P&Cs from the training literature, the creator of the virtual food policy expert will see value in more generalized,

rather than detailed, P&Cs, and engineer the queries accordingly. However, that the generalized knowledge base on which the virtual food policy expert runs may be a somewhat inaccurate presentation of the source material for that knowledge may also reflect how ChatGPT processes that source material and not only due to how the queries used with the AI tool were written.

3.3. Value of the virtual food policy expert in use—a comparative assessment

After the virtual food policy expert was developed in Coachvox-AI, to assess its performance and its likely value to users, the same six questions or question pairs that a user might ask were posed to three versions of the Coachvox-AI virtual food policy expert as well as to ChatGPT 3.5 (<http://chat.openai.com/>) and Bard (<http://bard.google.com>). The three versions of the virtual food policy expert were created by adjusting in the Coachvox AI platform the degree to which the virtual coach could rely on information other than that in the training literature used in its creation—none (1 of 7 possible degrees of reliance on other sources); partial (4/7); and equal reliance on other sources as on the training literature (7/7). ChatGPT and Bard were chosen for use in this assessment because, among the chatbots available for public use, they currently have the largest numbers of users and remain in active development.

The first question asked of each AI tool concerned the sources it draws upon in providing guidance on food policy issues and how it decides which sources to favor. The other questions posed reflect in various combinations the food policy thematic areas of agricultural development, food security, and nutrition. The following topics were the focus of the six questions:

- The sources the AI tools used to obtain food policy guidance and how they determined which of possibly competing insights obtained to offer the user;
- Smallholder agriculture development and poverty;
- How soil fertility management contributes to young child nutritional status;
- Agricultural production for national food security;
- How best to build political will to improve nutrition; and
- How to achieve food security for all.

The specific queries submitted to the AI tools for each question or question pair (two topics involved a pair) are presented in Annex 2 at the start of the sub-sections in which the re-

sponses obtained for each from the AI tools are presented.

For all but the first, the questions were asked in three forms that varied by geographic specificity. Each question was asked first with reference to all developing countries to reflect general insights the tools might provide a user. The other two question forms were country-specific—to Malawi and to Burkina Faso. Malawi featured strongly in the research literature used to create the virtual food policy expert. Consequently, one might expect that the three Coachvox AI virtual food policy experts would be able to provide considerable detail on how best to address food policy challenges there, likely more than the detail offered by ChatGPT or Bard. In contrast, none of the documents in the training literature for the virtual food policy expert drew from work in Burkina Faso. Consequently, it was expected that no answer would be offered for questions on the country posed particularly to the Coachvox AI virtual food policy expert that only relied on its training literature to formulate responses. Instead, the virtual expert would respond that the question was outside of its areas of expertise.

The AI tools were asked to limit their responses to 50 words or less to keep the responses manageable for review. When asking a new question, a new chat was started with each. This was done to avoid any previous chat interactions affecting the content of the response made to the new question. However, for the two question pairs, the follow-up questions were asked in the same chat.

4. Discussion

The first question asked of each AI tool, a pair of questions, concerned the sources each used to advise on food policy-related issues and how the chatbot chose between contrasting evidence-based recommendations in those sources. All three versions of the Coachvox AI virtual food policy expert provide very similar responses to both questions—not word for word, but pretty much point for point. The answers to the first question were well-stated, if general and not very insightful. Those from the three virtual food policy experts for the second question made reference to “empirical validity”, “credibility of their sources”, or robustness, but without stating how that validity or robustness was determined. (However, recall that responses were limited to 50 words.)

ChatGPT provided responses that were more technical in nature, noting that the most recent sources it used were from January 2022, so they may be dated, and providing a basic description of how AI techniques were used on those sources. The last sentence in each response from Chat GPT to both questions in the set were disclaimers, instructing the user to

cross-verify its output with “domain experts” and more recent information, as well as to consider whether context-specific factors may be important.

In contrast, Bard showed a larger personality than the other chatbots from its first response, saying, “Trust me because ...”. Bard seemed to be making a robust and enthusiastic pitch for the user to partner with it in finding solutions, rather than simply viewing Bard as a source. The enthusiasm is quite compelling, while possibly being overdramatic for users seeking guidance couched in more strongly objective and empirical terms. However, in responding to the second question of the pair, Bard provided a list of the sorts of evidence it draws upon and how it uses them.

The responses to the other five questions posed of the five different chatbots were reviewed closely to determine how performance across the chatbots differs. The questions were asked with different geographic specificity to assess how much the significant information on Malawi and the paucity of information on Burkina Faso in the training literature for the virtual food policy expert was reflected in the answers the three versions of that expert provided. Here, several dimensions of the performance of the five chatbots are discussed:

Their answers are quite generalized without as much detail as expected. More context-specific detail was expected from the virtual food policy experts than was provided, since these were developed from over 1,500 P&Cs drawn from about two dozen food policy documents or research reports. This perception of the chatbots providing only general advice is partly due to imposing the 50-word limit on responses. Additionally, the AI tools behind these chatbots are based on probabilistic models that guide their word choices for phrase and sentence construction. Consequently, these tools can be expected to preferentially provide conceptually general information, rather than specific or, particularly, atypical information, given how much more likely it is that general information appeared in the documents used to train the AI tools. This preference for providing generalities rather than specifics or exceptions to general cases would seem to be a bias associated with how the tools were constructed and operate.

Considerable commonality in the responses that were provided was observed. All the chatbots seem to have drawn from much the same training literature. The addition of the training literature for creating the Coachvox AI virtual food policy expert did not alter or expand the content of the responses much.

There were exceptions, however. It was expected that Malawi-

specific questions posed to the virtual food policy expert that only relied on the training literature for its knowledge base would provide responses that differed in some ways from those of the other chatbots. This was not the case for most questions—all three virtual food policy experts provided much the same response, even if not word for word.

However, for the last question asking the chatbot to offer its blue-sky thinking on how to achieve food security for all, the version of the virtual food policy expert that only relied on training literature information quite clearly drew on key points in some documents in that literature to generate its response. These points were absent or more softly stated in the responses of the other two versions of the virtual food policy expert that also relied on internet information. For the Malawi-specific blue-sky thinking question, ChatGPT and Bart provided poor responses, drawing from sources that seem not to offer any insights on achieving food security in Malawi. For this context-specific question, the reliance of ChatGPT and Bart solely on internet sources to obtain context-appropriate guidance did not work well.

For the Burkina Faso-specific blue-sky thinking question, however, the three versions of the Coachvox AI virtual food policy expert provided only slightly better responses than did ChatGPT and Bart. The advantages they offered likely stemmed from common food policy challenges both Malawi and Burkina Faso face, so some appropriate guidance for Burkina Faso could be obtained from Malawi, as well as from other countries from which the training literature was drawn. However, it also was apparent that the Coachvox AI virtual food policy experts were adlibbing somewhat, inserting some guidance for addressing food insecurity in Burkina Faso that seemed inappropriate or only of secondary value to the country. It was evident in the responses that the virtual food policy experts had less context-specific knowledge on Burkina Faso than they did for Malawi.

Given the commonality in responses across the five chatbots, their broad perspectives on strategies to address food policy challenges generally were aligned. Particularly for the generic developing country questions, those of ChatGPT and the Coachvox AI virtual food policy experts often matched quite closely—again, not word-for-word, but in the points covered by the responses of each. (This suggests that ChatGPT is likely one of the AI tools used to create and operate virtual coaches created using the Coachvox AI system.) The responses of Bard differed from those of the others in style—in keeping with the more expansive personality it has relative to the other chatbots. However, the substance of the points Bard made in

response to each question overlapped substantially with those made by the other chatbots.

However, there was one notable exception to this pattern. The fourth question on agricultural production for national food security asked for guidance on whether pursuing self-sufficiency in the production of staple food crops was one of the best ways to assure food security. The three versions of the Coachvox AI virtual food policy expert and Bard all provided recommendations suggesting that self-sufficiency strategies were not always the best choice to achieve food security. In contrast, ChatGPT strongly supported such a strategy without much equivocation—“self-sufficiency in staple food crops is crucial for a developing country with a large smallholder farming sector”. ChatGPT asserted this perspective for developing countries in general and for Malawi and Burkina Faso. While the wisdom of self-sufficient production to achieve national food security still is a point of debate, particularly within governments having to manage food crises, the development economics literature over the past several decades certainly is more aligned with the perspectives offered by the three versions of the virtual food policy expert and Bard, than with those of ChatGPT. To better understand the risks of using AI tools as a source of policy recommendations, a case study on what accounts for the different recommendations the AI tools offered on national self-sufficient crop production as a food security strategy would be valuable.

It was noted earlier in this report that to ensure that the information and recommendations provided have been properly vetted and validated, restricting the information base upon which the virtual food policy expert draws to that of the training literature provided will be important. Consequently, perhaps the most troubling observation made of the results from this assessment of AI tools was that some information not included in the training literature appeared in the responses of the Coachvox AI virtual food policy expert that only relied on the training literature for its knowledge base and not the internet and other sources. This included content on post-harvest losses and food waste, which was not mentioned in any of the almost two dozen documents in the training literature, as well as specific references to topics that were only referred to in passing in one or two documents, such as on climate change and employing climate-smart agricultural techniques, agricultural credit, or the Sustainable Development Goals. Why this inclusion of information from unspecified sources occurs is unclear, but likely is linked to the much broader training literature used to create the AI tools, some of which may have included documents offering perspectives and conceptual insights on food policy issues, including some not considered in

the training literature for the virtual food policy expert.

This leakage of potentially unvetted information into the knowledge base of the virtual coach makes their use as a channel to communicate information and recommendations that are purported to be evidence-based and sound quite risky. For IFPRI, an institution that must maintain its reputation as an objective provider of evidence-based guidance on food policy issues around the world, that reputation may be placed in jeopardy if an IFPRI-branded virtual food policy expert promoted recommendations to users for which there was minimal supportive evidence. Before virtual coaches can be endorsed as a new way to convey evidence-based information to users, a better understanding is needed of how internet-derived content might come to feature in the responses of virtual food policy expert that was designed only to draw upon high-quality, vetted evidence for the source of those responses.

It was noted that Bard is a chatbot with a bigger personality than the other chatbots used in this assessment, providing in its output some statements that could be critiqued for being somewhat normative or otherwise subjective. For example, it included in its response to the second of the pair of questions on building political will to improve nutrition in Malawi, “We need strong evidence, local champions, and clear links to economic growth to convince politicians that investing in tiny tummies reaps big rewards for Malawi’s future.” While one would hope that that the evidence underlying Bard’s statement is as strong as that used by ChatGPT in its much more staid and somber response to the question, this contrast highlights that style will matter to how users assess the reliability and value of these virtual coaches, including virtual food policy experts.

In designing the assessment, Burkina Faso was included in the questions in part to see if the virtual food policy expert that relied only on the training literature supplied to answer the questions would apologetically state that it could not answer questions posed on Burkina Faso. One step in specifying the general characteristics of the virtual coach within Coachvox AI is to provide “let down information”, a dozen or more responses the virtual coach should give if it chooses not to answer questions outside the knowledge bank established for it through the training literature. However, efforts to get the most restricted Coachvox AI virtual food policy expert to decline to offer questions on Burkina Faso did not succeed. Apparently, the training literature used contained sufficient generalized information for the virtual food policy expert to develop reasonable answers concerning food policy strategies for Burkina Faso, even though no specific information on the country was included in the training literature. Its ability to

respond also may reflect the contents of the much more extensive training literature used to build the AI tools employed by Coachvox AI. Chatbots seemingly are not humble—they would rather provide a wrong answer than admit to not having the expertise to offer a correct answer. Users of a virtual food policy expert must keep this in mind.

Although producing considerable AI-based output, the assessment here is simple and restricted. The range of questions asked to each AI tool could be expanded to better understand how each responds and to identify where some caution is needed in drawing inferences from those responses to guide the design of specific food policies or agricultural development, food security, or nutrition strategies or programs. These might include both general and context or country-specific questions that:

- Examine broad concepts to guide the design of strategies to accelerate agricultural growth, particularly for the smallholder sector; to achieve national and household food security; and to reduce malnutrition, particularly for young children and mothers.
- Request detailed guidance on specific dimensions or components of such strategies.
- Contrast the recommendations made by the virtual food policy expert for a developed country context to those made for a developing country context. Are different recommendations offered? If so, is it clear what factors drive those differences?
- Probe the limits of the guidance the virtual coach can provide. Does it provide insights on:
 - Costs of public investments to address food policy challenges;
 - Political economy considerations in adopting the strategies or public investments recommended; or
 - Technical choices in program design.

The assessment here primarily sought to identify weaknesses in how virtual coaches operate and use the information in their knowledge base in providing responses. Flaws were found. However, many of these are likely to be overcome, possibly quite soon, given the recent pace of development of these technologies. The assessment here should be recognized as reflecting a specific point in their development. The potential for virtual food policy experts to be an effective channel through which appropriate and validated guidance can be pro-

vided to those who require it remains great, even if more work is needed on their design to ensure that all the guidance such virtual experts offer is appropriate and valid.

Virtual food policy experts for IFPRI communication—cautious exploration

For certain users of IFPRI's research, virtual food policy experts may better or more efficiently meet their needs than relying on IFPRI working papers, policy briefs, or blogs or on peer-reviewed journal articles written by IFPRI researchers. The chat format of questions and answers may enable the user to obtain the results or policy recommendations being sought more quickly, while also pointing the user to related food policy issues to explore.

However, the exercise described in this report of creating a virtual food policy expert with relatively wide-ranging expertise across agricultural development, food security, and public health nutrition cannot be recommended for replication by IFPRI. There are too many uncertainties around the provenance of the information and recommendations such a general virtual expert would provide. Moreover, instilling sufficient nuance and context specificity into its responses and determining how the chatbot would handle conflicting evidence-based recommendations on specific food policy issues would be complex challenges to overcome in designing a reliable virtual expert.

Instead, what merits some consideration for inclusion in the strategies IFPRI uses to communicate its research is creating a virtual expert on a relatively narrowly defined area of food policy expertise, such as might emerge from a research paper or program. Such a virtual expert could be added to the other channels that IFPRI uses to communicate its research results. The Coachvox AI system of creating a virtual coach may work relatively well for doing so. Detailed working papers from the specific research could be used as the training literature to keep the area of expertise of the virtual expert narrowly focused.

It would be important to design the virtual expert in such a way that it will not try to improvise a response to any questions posed outside of its narrow area of expertise. How this can be done is not clear. We were unable to get virtual food policy expert created under the exercise here to decline to answer questions that we thought would be outside of its areas of expertise. As was evident in the assessment made of the virtual food policy expert, even when the virtual expert is set up to rely only on information from documents in its training literature, it will draw upon other information from outside of those

documents. We assume this is due to the much more extensive training literature used to create the AI tools on which the virtual food policy expert operates, but further investigation needs to be done.

This challenge in designing a virtual food policy expert to provide results and guidance on a narrow sub-set of food policy issues is a manifestation of the general problem that IFPRI will face in seeking to use AI tools for policy communication: How can IFPRI ensure that insights and recommendations the AI tools offer users is valid and constitutes the best evidence-based information available? To maintain its reputation as an objective and impartial evidence-based food policy research institution, IFPRI must be transparent with users on the sources of the information and recommendations it offers. There is much that can go wrong were IFPRI to incorporate virtual food policy experts and other AI tools into the communication channels it uses to disseminate guidance on food policy issues without understanding how and why such tools may improvise responses based on invalid or irrelevant information. The current state of the technologies driving virtual experts requires that there be relatively close oversight by experienced researchers of the information and recommendations they might provide.

If IFPRI can be confident that any virtual food policy experts used to disseminate knowledge and recommendations from its research only use validated information, several new challenges then may arise. The AI tools used to create the P&Cs from the training literature supplied for the virtual food policy expert in the exercise here worked relatively well, although detailed reviews of the P&Cs generated from each document were still required. However, there may be some food policy sub-issues or research methods on which the AI tools perform unsatisfactorily in creating useful P&Cs. These may include processing research documents focused on food systems performance in all their complexity or on political economy considerations in food policy or which are based on simulations run through economic models. Can the AI tools process these types of documents into useful P&Cs for a virtual food policy expert? P&C creation efforts using AI tools should be made on a variety of food policy research outputs to better understand the limitations or weaknesses each type of output may reveal in the capabilities of the AI tools.

Creating and operating virtual food policy experts is only one of several ways that AI tools can be used by IFPRI researchers. The tools are almost certainly already being used to assist with assembling literature reviews on specific topics, designing experimental protocols, drafting survey questionnaires,

conducting statistical analyses, and helping with visualization of the results obtained, among other tasks. These trends will continue. However, in all these uses of the tools, some caution is merited for two reasons.

First, concerns about replicability are inherent to the use of AI tools, including virtual experts, and the seemingly black-box manner in which they operate. In the exercise here, it was evident that rerunning the same query or prompt to the virtual food policy expert or to the more general chatbots used in creating the P&Cs for the virtual expert would not produce the same result as previously obtained. It seems unwise to simply view this variation in outputs from the same input as random and of little significance to what inferences are drawn from the use of the AI tools.

Second, as discussed earlier, being probabilistic tools developed on existing information, AI tools may introduce nuanced biases that work against innovation. A cautious approach is necessary to balance reliance on these tools for innovative thinking against the risk of perpetuating the status quo. While some aspects of the research process may greatly benefit from reliance on these tools, others may be harmed, potentially leading to erroneous conclusions. Researchers at IFPRI must critically assess the value of AI tools in all dimensions of their work, recognizing potential limitations in how they might fos-

ter innovative research and thinking for addressing food policy challenges.

Finally, in the operations of IFPRI, the virtual food policy experts may be valuable entry points for the Institute to begin directly engaging with potential clients or donors on complex food policy research questions. However, user interactions with the virtual experts will need to be monitored to improve IFPRI's understanding of the limits of virtual experts for such communication purposes. As these limits become clearer, efforts can be made to strengthen the virtual experts, such as by adding P&Cs to expand their areas of expertise or to enable users to engage more productively with the tool. However, becoming better cognizant of these limitations may result in IFPRI deciding that virtual food policy experts pose a risk to IFPRI's reputation as an impartial provider of evidence-based food policy advice that cannot be managed appropriately. It would then make sense for IFPRI to pull the plug on its virtual food policy experts.

In conclusion, IFPRI's research and communication efforts may be significantly advanced by integrating and refining virtual food policy experts in its work. However, this is not certain. More assessments of the benefits and risks of using these and other AI tools are needed. The cautious approach recommended here will ensure that advancements in the capa-

This publication has been prepared as an output of CGIAR Research Initiatives on [Digital Innovation](#), which researches pathways to accelerate the transformation towards sustainable and inclusive agrifood systems by generating research-based evidence and innovative digital solutions, and [National Policies and Strategies](#), which co-creates demand-driven policy solutions with national institutions, supporting countries to transform food, land and water systems for development and sustainable futures. This publication has not been independently peer-reviewed. Any opinions expressed here belong to the author(s) and are not necessarily representative of or endorsed by CGIAR. In line with principles defined in [CGIAR's Open and FAIR Data Assets Policy](#), this publication is available under a [CC BY 4.0](#) license. © The copyright of this publication is held by [IFPRI](#), in which the Initiative leads reside. We thank all funders who supported this research through their contributions to [CGIAR Trust Fund](#).

Annexes

Annex 1. Training literature used to create Coachvox-AI virtual food policy expert

no.	Document	P&Cs, est.	Process	Labels
1	Coote, C., K. Giller, A. Sutherland, T. Benson. S. Mughogho, N. Oudwater, & J. Coulter. 1998. <i>Soil fertility in Malawi: A review of policies, productivity, and perceptions</i> . London: Natural Resources Institute & University of Greenwich.	158	Coachvox AI	Malawi; soil fertility; agriculture
2	Benson, T, with others. 2005. <i>Framework document for a National Nutrition Strategy for Ethiopia</i> . Washington, DC: IFPRI. http://www.ifpri.org/PUBS/wpapers/ethionutrition.asp	49	Coachvox AI	Ethiopia; nutrition
3	Wiebelt, M., K. Pauw, J.M. Matovu, E. Twimukye, & T. Benson. 2011. <i>Managing future oil revenue in Uganda for agricultural development and poverty reduction: A CGE analysis of challenges and options</i> . IFPRI Discussion Paper no. 1122. Washington, DC: IFPRI	34	Coachvox AI	Uganda; agriculture
4	Benson, T., E. Schmidt, H. Namusoke, V.J. Temple, B. Holtemeyer, K. Codling, and C. Rudert. 2020. "Limits to commercially iodized salt to address dietary iodine deficiency in rural Papua New Guinea" <i>Asia-Pacific Journal of Clinical Nutrition</i> , 29 (2): 414-422.	13	Coachvox AI	Papua New Guinea; nutrition
5	Malawi ZHMSR Research Team. 2019. <i>Malawi Zero Hunger and Malnutrition Strategic Review (ZHMSR)</i> . Full report	110	Coachvox AI	Malawi; food security; nutrition; agriculture; social protection
6	Set of 14 three to five page briefs developed from: 2019 <i>Malawi Zero Hunger and Malnutrition Strategic Review (ZHMSR)</i> .			
	ZHMSR Brief 1 — <i>The Malawi ZHMSR Process</i>	6	Coachvox AI	Malawi; food security; nutrition; agriculture; social protection
	ZHMSR Brief 2 — <i>Situation Analysis around Hunger and Malnutrition in Malawi</i>	9	ChatGPT; Coachvox AI	Malawi; food security; nutrition
	ZHMSR Brief 3 — <i>Sustaining Recent Positive Nutrition Trends in Malawi</i>	9	ChatGPT; Coachvox AI	Malawi; nutrition
	ZHMSR Brief 4 — <i>Institutional Coordination for Action to Reduce Chronic Malnutrition in Malawi</i>	4	ChatGPT; Coachvox AI	Malawi; nutrition
	ZHMSR Brief 5 — <i>Food Security in Malawi – Going Beyond Agricultural Production</i>	11	ChatGPT; Coachvox AI	Malawi; food security; agriculture; agricultural markets
	ZHMSR Brief 6 — <i>Social Protection – Meeting the Nutritional Needs of Those Not Economically Productive or Whose Livelihoods Are Adversely Affected by Shocks</i>	6	ChatGPT; Coachvox AI	Malawi; food security; social protection
	ZHMSR Brief 7 — <i>Agriculture in Malawi – Strategic Considerations</i>	13	ChatGPT; Coachvox AI	Malawi; food security; agriculture
	ZHMSR Brief 8 — <i>Structural Transformation of Malawi’s Economy– Strategic Considerations for Well-nourished Citizens</i>	8	ChatGPT; Coachvox AI	Malawi; food security; nutrition, agricultural markets
	ZHMSR Brief 9 — <i>Achieving Zero Hunger and Zero Malnutrition in Malawi by 2030 – A Policy Research Agenda</i>	14	ChatGPT; Coachvox AI	Malawi; food security; nutrition; agriculture; social protection
	ZHMSR Brief 10 — <i>Actions to Ensure Food Security for All Malawians – Beyond Food Production</i>	8	ChatGPT; Coachvox AI	Malawi; food security; agriculture; social protection
	ZHMSR Brief 11 — <i>Actions to Accelerate Improved Nutrition for All Malawians</i>	7	ChatGPT; Coachvox AI	Malawi; nutrition; social protection
	ZHMSR Brief 12 — <i>Actions to Expand Employment Options beyond Agriculture for All Malawians</i>	9	ChatGPT; Coachvox AI	Malawi; food security; nutrition; social protection
	ZHMSR Brief 13 — <i>Actions to Enable Agriculture to Be a Principal Driver of Economy-wide Transformation in Malawi</i>	8	ChatGPT; Coachvox AI	Malawi; agriculture
	ZHMSR Brief 14 — <i>Actions to Meet the Nutritional Needs of the Chronically Poor and Disaster-stricken in Malawi</i>	4	ChatGPT; Coachvox AI	Malawi; food security; social protection

no.	Document	P&Cs, est.	Process	Labels
7	<u>Seven book chapters</u> from: Benson, T. 2021. <i>Disentangling Food Security from Subsistence Agriculture in Malawi</i> . Washington, DC: IFPRI			
	Chapter 1, Introduction	14	ChatGPT; Coachvox AI	Malawi; food security
	Chapter 2, Food Security in Malawi	39	ChatGPT; Coachvox AI	Malawi; food security
	Chapter 3, Agriculture in Malawi	84	ChatGPT; Coachvox AI	Malawi; agriculture
	Chapter 4, Agricultural Markets in Malawi	48	ChatGPT; Coachvox AI	Malawi; agricultural markets
	Chapter 5, Recognizing Rural Differentiation in Malawi	53	ChatGPT; Coachvox AI	Malawi; agriculture
	Chapter 6, Structural Transformation of Malawi's Economy to End Hunger	17	ChatGPT; Coachvox AI	Malawi; economic development; agriculture
	Chapter 7 Actions to Advance Toward a Reliably Food-secure Malawi	25	ChatGPT; Coachvox AI	Malawi; food security
8	<u>Six chapters</u> from: Benson, T., and J. De Weerd. 2023. <i>Employment options and challenges for rural households in Malawi</i> . MaSSP Working Paper 40. Lilongwe: IFPRI.			
	Chapter 1—Introduction	22	ChatGPT; Coachvox AI	Malawi; economic development; agriculture
	Chapter 2— Economic differentiation and change among Malawian households, with a focus on agriculture	37	ChatGPT; Coachvox AI	Malawi; economic development; agriculture
	Chapter 3—Household farming to keep out of poverty—an income analysis	42	ChatGPT; Coachvox AI	Malawi; economic development; agriculture
	Chapter 4—Household farming to keep out of poverty—a crop productivity analysis	23	ChatGPT; Coachvox AI	Malawi; economic development; agriculture
	Chapter 5—Off-farm employment in Malawi	42	ChatGPT; Coachvox AI	Malawi; economic development
	Chapter 6—Discussion and policy recommendations	39	ChatGPT; Coachvox AI	Malawi; economic development; agriculture
9	Benny, D., T. Benson, M. Ivekolia, M. Kedir Jemal, and R. Ovah. 2022. <i>Improving agricultural productivity in Papua New Guinea: Strategic and policy considerations</i> . Working Paper. Washington, DC: IFPRI. https://doi.org/10.2499/p15738coll2.134987	96	ChatGPT; Coachvox AI	Papua New Guinea; agriculture
10	Benson, T. 2021. "Making the achievement of zero hunger and malnutrition a policy priority: A critical assessment of recent national Zero Hunger Strategic Reviews from Asia and the Pacific" <i>Asia-Pacific Journal of Clinical Nutrition</i> . 30 (2):316-328.	36	ChatGPT; Coachvox AI	food security
11	Benson, T., M. Amare, and A.I. Ogunniyi. 2020. <i>The relative commercial orientation of smallholder farmers in Nigeria: Household and crop value-chain analyses</i> . Nigeria Strategy Support Program Working Paper 66. Abuja: IFPRI	37	ChatGPT; Coachvox AI	Nigeria; agriculture; agricultural markets
12	Comstock, A., T. Benson, F. Nankhuni, J. Kanyamuka, Z. Nyirenda, & C. Nyondo. 2019. <i>A critical review of Malawi's Special Crops Act and Agriculture (General Purposes) Act</i> . IFPRI Discussion Paper 1792. Washington, DC: IFPRI.	80	ChatGPT; Coachvox AI	Malawi; agriculture; agricultural markets
13	Benson, T., and T. Mogue. 2018. "Constraints in the fertilizer supply chain: evidence for fertilizer policy development from three African countries." <i>Food Security</i> , 10 (6): 1479-1500.	44	ChatGPT; Coachvox AI	Uganda; Tanzania; Mozambique; agriculture
14	Benson, T, M. Amare, M. Oyeyemi, and O. Fadare. 2017. <i>Study of the Determinants of Chronic Malnutrition in northern Nigeria: Qualitative Evidence from Kebbi and Bauchi states</i> . Nigeria Strategy Support Program Working Paper 44. Abuja: IFPRI	53	ChatGPT; Coachvox AI	Nigeria; nutrition

no.	Document	P&Cs, est.	Process	Labels
15	Benson, T., & S. Mugarura. 2013. "Livestock development planning in Uganda: Identification of areas of opportunity and challenge." <i>Land Use Policy</i> . 35: 131-139	19	ChatGPT; Coachvox AI	Uganda; agriculture
16	Benson, T.; N. Minot; J. Pender; M. Robles; & J. von Braun. 2013. Information to guide policy responses to higher global food prices: The data and analyses required. <i>Food Policy</i> . 38: 47-58.	48	ChatGPT; Coachvox AI	Food security; agricultural markets
17	Benson, T. 2008. <i>Improving nutrition as a development priority - Addressing undernutrition within national policy processes in sub-Saharan Africa</i> . IFPRI Research Report 156. Washington, DC: IFPRI.	69	ChatGPT; Coachvox AI	Nutrition
18	Benson, T.; S. Mugarura; & K. Wanda. 2008. Impacts in Uganda of rising global food prices: The role of diversified staples and limited price transmission. <i>Agricultural Economics</i> . 39 (suppl.): 513-524.	28	ChatGPT; Coachvox AI	Uganda; food security; agricultural markets
19	Benson, T, with others. 2008. <i>National Food Security and Nutrition Strategy for Liberia</i> . Monrovia: Ministry of Agriculture, Republic of Liberia.	82	ChatGPT; Coachvox AI	Liberia; food security; nutrition
20	Benson, T. 2007. "Cross-sectoral coordination failure – How significant a constraint in national efforts to tackle malnutrition in Africa?" <i>Food and Nutrition Bulletin</i> . 28 (2, supplement): S323-S330.	23	ChatGPT; Coachvox AI	nutrition
21	Benson, T. 2004. <i>Africa's food and nutrition security situation – Where are we and how did we get here?</i> 2020 Discussion Paper no. 37. Washington, DC: IFPRI.	64	ChatGPT; Coachvox AI	food security; nutrition
	Prompts and Completions (P&C) to introduce a user to the broad topics on which the virtual food policy expert has expertise. The prompts were developed using ChatGPT and internet content, not the training literature for the virtual food policy expert.	95	ChatGPT (based on internet content); Coachvox AI	ONE OF: food policy, agriculture, food security, or nutrition
	TOTAL Prompts and Completions, estimate	1,739		

Annexes

Annex 2. Responses obtained from asking the same food policy-related questions to different AI sources—a comparative assessment

After the virtual food policy expert was developed in Coachvox-AI, the same set of questions that a user might ask were posed to three versions of the Coachvox-AI food policy expert as well as to ChatGPT 3.5 (<https://chat.openai.com>) and Bard (<https://bard.google.com>). The three versions of the virtual food policy expert were created by adjusting in the Coachvox AI platform the degree to which the virtual coach relies on information other than that contained in the training literature used in its creation—none (1 of 7 possible degrees of reliance on other sources); partial (4/7); and equal (7/7).

Most of the questions asked were varied by geographic specificity, being asked with reference to developing countries in general; specifically for Malawi, which was the research location for many of the documents in the training literature for the virtual food policy expert; and specifically for Burkina Faso, which did not feature in the training literature. The questions reflect in combinations the food policy thematic areas of agricultural development, food security, and nutrition.

To keep the responses manageable for review, the chatbots were asked to limit their responses to 50 words or less. In asking a new question, a new chat was started with each chatbot. This was done to avoid any previous chat interactions affecting the content of the response to the new question. However, for two questions, follow-up questions were asked.

Q1) The sources AI tools used to obtain food policy guidance and how they choose among them

- Question one of two: I am a senior civil servant in the government of a developing country. The issues I deal with include smallholder agricultural development, improving household and national food security, and ensuring that our citizens, particularly young children and their mothers, are adequately nourished. I understand you may be able to guide me in how best to address these issues. However, before I ask you questions on specific food policy challenges, please tell me in 50 words or less why I should trust what you tell me on these issues. What is the evidence you draw upon in responding to my questions on these issues and formulating food policy or program recommendations?
- Question two of two: If there are different opinions expressed in the evidence you consult on food policy issues, how do you as an AI tool determine which of these opinions is likely the best one to follow? Please explain in less than 50 words.

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

(1) I appreciate your query. My guidance is based on extensive research, analysis of successful food policy programs, and collaboration with experts in the field. I draw upon evidence from diverse contexts, considering local nuances. My aim is to provide informed, context-specific advice to enhance food

security and nutrition.

(2) I analyze diverse opinions based on their empirical validity, relevance to local context, and alignment with sustainable development goals. The best opinion is typically one that is evidence-based, context-specific, and contributes to long-term food security and nutrition improvement.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

(1) I draw upon a wealth of experience in food policy analysis and a deep understanding of food systems, particularly in contexts of smallholder agriculture. My recommendations are based on rigorous research, best practices, and collaboration with a wide range of stakeholders in the food security sector.

(2) I synthesize diverse opinions based on their alignment with proven best practices, the credibility of their sources, and their relevance to the specific context. This approach ensures that the most appropriate and effective recommendations are made.

Coachvox AI virtual expert using training literature equally with internet (7/7):

(1) I draw upon extensive research, best practices, and global case studies in food policy. My recommendations are based on evidence and experience in diverse contexts, particularly where smallholder agriculture is central. I aim to provide balanced, context-specific advice to enhance food security and nutrition.

(2) I analyze diverse opinions based on their empirical evidence, relevance to the context, and alignment with best practices. The most robust, context-specific, and globally successful solutions are prioritized to ensure effective food policy recommendations.

ChatGPT 3.5:

(1) I provide information based on a diverse range of publicly available sources up until my last update in January 2022. My responses are generated from patterns learned during training on a mixture of licensed data, data created by human trainers, and publicly available data. Always cross-verify with domain experts and updated resources for the latest information.

(2) I don't have personal opinions. Instead, I analyze patterns in data and present information based on learned associations. Evaluating conflicting opinions requires considering diverse sources, expert consensus, and recent evidence. Always consult with domain experts and consider context for nuanced decision-making.

Bard:

(1) Trust me because I learn from vast data on successful food security policies, agricultural development, and nutrition initiatives. I analyze studies, reports, and expert opinions, then identify common factors in positive outcomes. I'll share evidence with you, not just tell you what to do. Let's tailor solutions to your context, drawing on global best practices that fit your local needs.

(2) I weigh opinions based on research quality, data robustness, alignment with your context, and consensus among experts. No single opinion reigns, but strong evidence, local relevance, and diverse perspectives guide me.

Q2) Smallholder agriculture development and poverty—global, Malawi, Burkina Faso

- Globally, we see that smallholder households are more likely to be poor than households pursuing other livelihoods. What policies should be prioritized by developing countries around the world so that the welfare of smallholder farming households will substantially increase and they can meet their basic needs? Please explain in less than 50 words.

DEVELOPING COUNTRY:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

Governments should prioritize policies that enhance access to

affordable credit, improve rural infrastructure, and promote education and training in modern farming techniques. These measures can increase productivity, income, and overall welfare of smallholder farming households.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

Developing countries should prioritize policies that enhance access to affordable credit, improve agricultural extension services, and promote market access. These measures will boost productivity, increase income, and ensure food security for smallholder farming households.

Coachvox AI virtual expert using training literature equally with internet (7/7):

Developing countries should prioritize policies that enhance access to affordable credit, improve agricultural extension services, and promote market access. These measures can increase productivity, incomes, and food security for smallholder farming households.

ChatGPT 3.5:

Developing countries should prioritize policies that enhance access to credit, provide agricultural training, and improve market access for smallholder farmers. Additionally, investing in rural infrastructure and social safety nets can boost their productivity and income, addressing poverty and meeting basic needs.

Bard:

Prioritize policies that: 1. Boost productivity: Invest in education, infrastructure, and climate-resilient farming techniques. 2. Connect farmers to markets: Facilitate fair trade, storage solutions, and value addition. 3. Empower women farmers: Secure land rights, provide financial access, and promote leadership roles. 4. Strengthen safety nets: Implement social protection schemes and disaster risk reduction plans. By tackling these key areas, developing countries can empower smallholder households to thrive, not just survive.

MALAWI:

- In Malawi, we see that smallholder households are more likely to be poor than households pursuing other livelihoods. What specific policies appropriate to the Malawian context should be prioritized be pursued by the government of Malawi so that the welfare of smallholder farming households in the country will substantially increase and they can meet their basic needs? Please explain in less than 50 words.

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

The government should prioritize investments in agricultural research, transport, and communication services. Strengthening markets and investing in cooperative development can also help improve the welfare of smallholder farming households.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

The government should prioritize investments in agricultural research, infrastructure, and market development. Additionally, strengthening cooperatives and providing targeted support for commercially-oriented smallholders can drive economic transformation and improve welfare in rural areas.

Coachvox AI virtual expert using training literature equally with internet (7/7):

The Malawian government should prioritize policies that enhance access to agricultural inputs, improve market infrastructure, and provide training in modern farming techniques. These measures will increase productivity, income, and ultimately improve the welfare of smallholder farming households.

ChatGPT 3.5:

The Malawian government should prioritize policies that enhance smallholder farmers' access to education, technology, and markets. Investing in agricultural extension services, providing credit facilities, and promoting sustainable farming practices will empower smallholders, reduce poverty, and improve overall welfare.

Bard:

Prioritize policies that address interconnected challenges: access to credit, land ownership & inheritance rights for women, drought-resistant seeds & irrigation, and post-harvest storage & market access. Empowering women farmers and mitigating climate risks will be crucial for long-term success.

BURKINA FASO:

- In Burkina Faso, we see that smallholder households are more likely to be poor than households pursuing other livelihoods. What specific policies appropriate to the Burkinabé context should be prioritized by the government of Burkina Faso so that the welfare of smallholder farming households in the country will substantially increase and they can meet their basic needs? Please explain in less than 50 words.

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

The government should prioritize policies that enhance access to quality inputs, improve agricultural extension services, and facilitate access to markets. Additionally, social safety nets should be strengthened to protect these households from shocks and stresses.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

The government should prioritize policies that enhance access to affordable credit, improve agricultural extension services, and facilitate market access. Additionally, social safety nets for vulnerable households and investment in rural infrastructure can significantly improve the welfare of smallholder farming households.

Coachvox AI virtual expert using training literature equally with internet (7/7):

The government should prioritize policies that enhance access to credit, improve agricultural extension services, and promote market access. These measures will boost productivity, increase income, and improve food security among smallholder farming households.

ChatGPT 3.5:

Implement policies promoting access to credit, improved agricultural techniques, and market linkages for smallholder farmers in Burkina Faso. Strengthen social safety nets, invest in rural infrastructure, and provide education to enhance their productivity, income, and overall well-being.

Bard:

Prioritize policies that diversify income beyond rain-fed agriculture:

- Microfinance and training for non-farm activities: Weave baskets, sew clothes, sell produce.
- Invest in drought-resistant crops and irrigation: Reduce climate vulnerability and boost yields.
- Connect farmers to markets and fair prices: Cut out middlemen, empower cooperatives.

These steps can break the cycle of poverty and empower Burkinabé farmers to thrive.

Q3) How soil fertility management contributes to young child nutritional status—global, Malawi, Burkina Faso

- I am a senior policymaker in [a developing country / Malawi / Burkina Faso]. Recently I have received many ad-

vocacy pitches to support public investments to improve the soils on which our farmers grow their crops. One advocate said that among the benefits of doing so would be fewer undernourished children. I don't understand their thinking on this. Please explain the relationship between soil fertility management and reducing the prevalence of young children that are stunted in their growth [in a developing country / specifically in Malawi / specifically in Burkina Faso]. Do so in 50 words or less.

DEVELOPING COUNTRY:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

Enhanced soil fertility boosts crop yield and diversity, leading to improved dietary diversity and nutrition. This can reduce child stunting by providing essential nutrients for growth and development.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

Soil fertility management enhances crop yield and diversity, leading to improved household food availability and nutritional quality. This, in turn, can reduce child stunting by ensuring adequate nutrient intake for optimal growth.

Coachvox AI virtual expert using training literature equally with internet (7/7):

Enhanced soil fertility boosts crop yield and diversity, leading to improved household food availability and nutritional quality. This can contribute to better child nutrition, reducing the prevalence of stunting.

ChatGPT 3.5:

Enhancing soil fertility improves crop nutrient content, combating malnutrition. Nutrient-rich crops foster healthier diets, reducing stunting in young children. Supporting soil investments promotes sustainable agriculture, directly impacting child nutrition and well-being in developing countries.

Bard:

Healthier soil yields more nutritious crops, giving children the building blocks they need to grow. Better harvests also boost family income, enabling them to afford diverse, healthy diets. That breaks the cycle of malnutrition and stunting for generations.

MALAWI:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

Improved soil fertility can boost crop yields, leading to increased food availability and diversity. This can enhance the nutritional intake of children, potentially reducing stunting rates in Malawi.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

Enhanced soil fertility can increase agricultural productivity, leading to better food availability and diversity. This can improve children's nutritional intake, reducing the prevalence of stunting.

Coachvox AI virtual expert using training literature equally with internet (7/7):

Improved soil fertility can enhance crop yields and diversity, leading to better household food availability and nutrition, which can reduce child stunting in Malawi.

ChatGPT 3.5:

Improving soil fertility enhances crop nutrient content, addressing micronutrient deficiencies in diets. This promotes healthier food consumption, reducing child stunting in Malawi. Improved agricultural practices contribute to enhanced nutrition, positively impacting child growth and development, aligning with the advocated public investments for long-term health benefits.

Bard:

When crops grow in fertile soil, they pack more nutrients. This translates to more nutritious food for families, especially young children. In Malawi, where stunting is high, better soil means better diets and a chance to break the cycle of malnutrition. It's like giving every child a fighting chance to thrive. Invest in soil, invest in Malawi's future.

BURKINA FASO:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

Improved soil fertility boosts crop yields, leading to increased food availability and diversity. This enhances children's nutrition intake, reducing the risk of stunting. It's a direct link between soil health and child health.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

Improving soil fertility can enhance crop yields and diversity, leading to better household food availability and diet diversity. This can improve child nutrition, reducing stunting rates in Burkina Faso.

Coachvox AI virtual expert using training literature equally with internet (7/7):

Improved soil fertility can boost crop yields, enhancing food availability and diversity. This can improve children's nutrition, reducing stunting. In Burkina Faso, where many rely on agriculture, soil management can significantly impact child health.

ChatGPT 3.5:

Enhancing soil fertility in Burkina Faso boosts crop yields and nutritional content, addressing malnutrition. Improved harvests provide diverse, nutrient-rich foods, crucial for child development. This, in turn, reduces stunting, promoting healthier growth and lowering the prevalence of undernourished children.

Bard:

Healthier soil nourishes healthier crops, leading to more nutritious food for families. This directly impacts children, reducing stunting caused by micronutrient deficiencies and insufficient calorie intake, a major issue in Burkina Faso. Investing in soil fertility empowers families to grow diverse, nutrient-rich food, giving their children a strong start in life.

Q4) Agricultural production for national food security—global, Malawi, Burkina Faso

- Is pursuing self-sufficiency in the production of staple food crops one of the best ways for [a developing country with a large smallholder farming sector / Malawi specifically / Burkina Faso specifically] to assure the food security of its citizens? Provide me guidance in 50 words or less.

DEVELOPING COUNTRY:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

Self-sufficiency can enhance food security by increasing local production. However, it's not always the best choice, especially where food market systems are good. Diversification, non-farm employment, and resilience to shocks are also crucial for food security.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

While self-sufficiency can enhance food security, it's not the only solution. Diversification, market integration, and non-farm income sources can also contribute significantly to food security in developing countries with large smallholder farm-

ing sectors.

Coachvox AI virtual expert using training literature equally with internet (7/7):

Self-sufficiency can enhance food security by increasing local food production. However, it's not always the best choice, particularly where food market systems are good. Diversification, non-farm employment, and resilience to external shocks are also crucial for food security.

ChatGPT 3.5:

Yes, pursuing self-sufficiency in staple food crops is crucial for a developing country with a large smallholder farming sector. It enhances food security by reducing dependence on imports, stabilizing prices, and empowering local farmers. Diversifying crops and implementing sustainable agricultural practices further strengthen resilience against external shocks.

Bard:

While self-sufficiency can be tempting, it's not always the best bet. In 50 words:

- Focus on increasing productivity and diversifying crops. This boosts food availability and resilience to shocks.
 - Trade strategically. Import staples where inefficient to grow, and export high-value crops for income.
 - Invest in infrastructure and storage. Reduce losses and connect farmers to markets.
 - Empower smallholders. Provide access to land, credit, and knowledge.
- Remember, food security isn't just about production, it's about access and affordability too.

MALAWI:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

Self-sufficiency in staple crops can contribute to Malawi's food security, but it's not the only solution. A comprehensive approach, including diversification, market access, and institutional support, is crucial for sustainable food security.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

Self-sufficiency in staple food crop production is a component, but not the sole solution for Malawi's food security. It must be complemented by other strategies like diversifying food sources, improving storage and distribution, and strengthening institutional frameworks.

Coachvox AI virtual expert using training literature equally

with internet (7/7):

Self-sufficiency in staple food crop production is a component, but not the sole solution for Malawi's food security. A comprehensive approach, including diversification, market access, and improved infrastructure, is crucial for sustainable food security.

ChatGPT 3.5:

Yes, pursuing self-sufficiency in staple food production is crucial for Malawi's food security. Investing in agriculture, promoting sustainable farming practices, and ensuring efficient distribution systems can enhance domestic food production, reduce dependency on imports, and safeguard citizens against external supply disruptions, fostering long-term food security.

Bard:

While self-sufficiency is tempting, it's not the be-all for Malawi's food security. Focusing solely on staple crops like maize can miss opportunities for diversifying diets and boosting incomes. Investing in climate-resilient crops, market access, and storage could be more sustainable and empowering for Malawians in the long run.

BURKINA FASO:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

Self-sufficiency in staple crops can contribute to food security in Burkina Faso, but it's not the only solution. Diversification, market access, and income generation are equally important. A comprehensive approach is needed for sustainable food security.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

Self-sufficiency in staple crops can contribute to Burkina Faso's food security. However, it's not the sole solution. Diversified farming, improved market systems, and non-farm income sources are also crucial for sustainable food security.

Coachvox AI virtual expert using training literature equally with internet (7/7):

Self-sufficiency in staple crops can contribute to food security in Burkina Faso, but it's not the only solution. Diversification, market access, and income generation are also crucial. It's important to consider the broader food system and socio-economic context.

ChatGPT 3.5:

Yes, pursuing self-sufficiency in staple food crops is crucial for Burkina Faso's food security. By prioritizing domestic production, the country can reduce dependency on imports, enhance resilience to external shocks, and ensure a stable and affordable food supply for its citizens.

Bard:

While pursuing self-sufficiency in staples like millet and sorghum seems ideal for Burkina Faso's food security, it's not the only solution. Here's why:

—Climate change: Burkina Faso's erratic rainfall makes consistent harvests difficult. Diversifying crops and promoting drought-resistant varieties is crucial.

—Conflict and displacement: Violence disrupts agricultural activities and displaces people, making self-sufficiency harder. Investing in conflict resolution and supporting displaced farmers is key.

—Trade and regional cooperation: Burkina Faso can't produce everything it needs. Trade with neighboring countries and regional food reserves can fill gaps and stabilize prices. Focus solely on self-sufficiency might be short-sighted. A multi-pronged approach that tackles climate challenges, conflict, and trade will be more effective in assuring long-term food security for Burkina Faso.

Q5) How best to build political will to improve nutrition—global, Malawi, Burkina Faso

- Question one of two: I am a senior policymaker in [a developing country / Malawi / Burkina Faso]. Recently I have been pressed by advocates to support public investments to improve the nutrition of our young children. Why should I care? Isn't this the responsibility of their parents and not government? Explain in 50 words or less why my reaction may be wrong?
- Question two of two: What are the challenges to building political will to increase public investments to improve the nutrition of young children [in developing countries / specifically in Malawi / specifically in Burkina Faso], and how can they be overcome? Please explain in 50 words or less.

DEVELOPING COUNTRY:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

(1) Public investment in child nutrition is crucial as it shapes a nation's future. Adequate nutrition in early years leads to improved health, cognitive development, and productivity in adulthood. While parents play a role, government intervention

ensures equal opportunity for all children, irrespective of their socio-economic background.

(2) The main challenges include competing priorities, lack of awareness, and limited resources. Overcoming these requires raising awareness about the long-term benefits of child nutrition, showcasing successful interventions, and advocating for resource allocation towards nutrition in national development strategies.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

(1) Public investment in child nutrition is crucial as it shapes a nation's future. Good nutrition in early years leads to better health, cognitive development, and productivity. While parents play a role, government intervention ensures equitable access, especially for disadvantaged families.

(2) Building political will for child nutrition investments can be challenging due to competing priorities and limited resources. Overcoming this requires raising awareness about the long-term benefits of such investments, and advocating for policy changes that prioritize nutrition in national development strategies.

Coachvox AI virtual expert using training literature equally with internet (7/7):

(1) Nutritional status of children is a public concern, not just parental. It shapes a nation's future workforce and productivity. Government investment in child nutrition can yield high returns in health, education, and economic growth. It's a strategic move for national development.

(2) Challenges include competing priorities, lack of awareness, and resource constraints. Overcoming these requires advocacy, evidence-based arguments showing high returns on investment in child nutrition, and mobilizing diverse stakeholders for a common cause: a healthier, more productive future generation.

ChatGPT 3.5:

(1) Investing in children's nutrition is crucial for national development. Malnutrition hampers cognitive development, reducing future productivity and perpetuating poverty. Government intervention ensures equitable access to nutrition, mitigating long-term societal costs. Prioritizing child nutrition aligns with broader economic and social goals, fostering a healthier and more prosperous nation.

(2) Challenges include competing priorities, budget constraints, and a lack of awareness. Overcoming requires effec-

tive advocacy highlighting economic benefits, framing nutrition as a strategic investment, engaging stakeholders, and demonstrating successful interventions' impact, fostering political commitment for sustained public investments in children's nutrition.

Bard:

(1) You're right, parents are primarily responsible for their children's well-being. But, investing in early nutrition goes beyond individual families. It's an investment in your country's future. Healthy, well-nourished children become productive adults, boosting your economy, healthcare system, and social stability. It's like planting seeds for long-term prosperity.

Think of it this way: would you wait for each family to build their own road before investing in national infrastructure? Early nutrition is the same - a foundational investment that unlocks a brighter future for all.

(2) Building political will for early childhood nutrition is tough. Benefits take years to show, competing needs grab attention, and data on returns is weak. Many leaders don't grasp the long-term costs of malnutrition. Plus, industries selling unhealthy food may lobby against change.

To overcome this, strong advocacy, compelling data, and strategic communication are key. Build coalitions with diverse stakeholders, and find champion leaders who prioritize child well-being. It's not just the right thing to do, it's smart politics for a healthier, more prosperous future.

MALAWI:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

(1) Public investment in child nutrition is crucial. Malnutrition in early life can lead to lifelong health and cognitive issues, impacting national productivity. While parents play a role, government intervention can ensure equitable access to nutrition, benefiting society as a whole.

(2) Challenges include competing priorities, limited resources, and lack of awareness on nutrition's long-term impact. Overcoming these requires evidence-based advocacy, highlighting nutrition's role in national development, and demonstrating the cost-effectiveness of investing in child nutrition.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

(1) Public investment in child nutrition is crucial. Malnutrition in early life can lead to lifelong cognitive and physical deficits, impacting productivity. Government intervention can

ensure equitable access to nutrition, fostering a healthier, more productive future generation, beyond what individual parents can provide.

(2) Challenges include competing priorities, limited resources, and lack of awareness about the long-term benefits of child nutrition. Overcoming these requires strong advocacy, evidence-based policy dialogue, and demonstrating the economic benefits of investing in child nutrition for Malawi's future productivity and growth.

Coachvox AI virtual expert using training literature equally with internet (7/7):

(1) Public investment in child nutrition is crucial as it shapes the nation's future. Malnutrition hampers cognitive development, affecting education and productivity. While parents play a role, government intervention ensures equitable access to nutrition, especially for vulnerable groups. It's an investment in human capital.

(2) Challenges include competing priorities, limited resources, and lack of awareness about nutrition's long-term benefits. Overcoming these requires evidence-based advocacy, demonstrating the economic returns of nutrition investment, and building alliances with influential stakeholders to prioritize child nutrition.

ChatGPT 3.5:

(1) Investing in children's nutrition is crucial for Malawi's future. Adequate nutrition enhances cognitive development, improves school performance, and boosts productivity. Government intervention ensures vulnerable families have access to essential resources, addressing systemic issues and fostering a healthier, more capable generation, ultimately benefiting the nation's socio-economic development.

(2) Challenges in Malawi include competing priorities, limited resources, and awareness gaps. Overcoming these requires effective communication to highlight long-term benefits, engaging key stakeholders, aligning nutrition goals with broader development objectives, and implementing targeted policies to demonstrate tangible impacts, fostering political will for sustained public investments in children's nutrition.

Bard:

(1) Malnutrition stunts children's physical and cognitive growth, impacting future earning potential and draining the economy. Investing now saves lives, boosts lifelong potential, and pays back handsomely. It's not just about parents; it's about building a stronger Malawi.

(2) In Malawi, short-term pressures (e.g., elections) and competing priorities (e.g., infrastructure) eclipse long-term gains of child nutrition. We need strong evidence, local champions, and clear links to economic growth to convince politicians that investing in tiny tummies reaps big rewards for Malawi's future.

BURKINA FASO:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

(1) Child nutrition is a public concern. Malnutrition affects cognitive development, impacting future productivity. Government investment can ensure all children, regardless of their parents' resources, have access to nutritious food, promoting national development and economic growth.

(2) Challenges include limited awareness of malnutrition's impact and competing priorities for public funds. Overcoming these requires raising awareness among policymakers about the long-term benefits of child nutrition and advocating for increased budget allocation for nutrition programs.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

(1) Public investments in child nutrition are crucial. Malnutrition can hinder cognitive development, affecting future productivity. While parents play a role, government intervention can ensure equitable access to nutrition, benefiting the entire society in the long run.

(2) Challenges include limited awareness of malnutrition's societal costs and competing priorities. Overcoming these requires raising awareness among policymakers about the long-term economic benefits of improved child nutrition and advocating for increased budget allocations for nutrition interventions.

Coachvox AI virtual expert using training literature equally with internet (7/7):

(1) While parents play a crucial role, government investment in child nutrition is vital. It ensures equitable access to nutritious food, fosters healthy growth, and reduces future healthcare costs. This investment ultimately contributes to a stronger, healthier, and more productive society.

(2) Challenges include limited awareness of nutrition's societal impact, competing priorities, and resource constraints. Overcoming these requires raising awareness among policymakers about the long-term benefits of child nutrition, prioritizing nutrition in policy agendas, and exploring sustainable funding

sources for nutrition programs.

ChatGPT 3.5:

(1) Investing in child nutrition is crucial for Burkina Faso's future. Malnutrition hinders cognitive development, affecting the nation's human capital. Government intervention ensures access to essential resources, benefiting societal well-being and economic productivity. Prioritizing child nutrition is a strategic investment in the nation's health, education, and long-term prosperity.

(2) Challenges in Burkina Faso may include competing priorities, limited resources, and awareness gaps. Overcoming these requires targeted advocacy, highlighting the economic benefits of improved nutrition, engaging with influential stakeholders, and implementing evidence-based policies to demonstrate positive outcomes, fostering political will for sustained investments in child nutrition.

Bard:

(1) Your reaction is understandable, but consider this: brain development peaks in early childhood. Investing in their nutrition now isn't just about filling bellies, it's about building a brighter future. Healthy kids learn better, earn more, and contribute more. It's an investment, not a handout, with returns for generations.

(2) In Burkina Faso, convincing folks to invest in young children's nutrition is tough. Some see it as a family's duty, not the government's. Others doubt the returns, unaware of the life-long benefits. We need to show, not just tell. Simple, local data on how healthy kids do better in school and later life can be powerful. Share stories of transformed lives, not just statistics. Partner with trusted local leaders and parents to champion the cause. Make it clear: investing in our children is investing in our nation's future.

Q6) Blue sky thinking on achieving food security for all—global, Malawi, Burkina Faso

- What is the most cost-effective way that [a developing country / Malawi / Burkina Faso] can achieve sustainable food security for all of its citizens? Please explain in 50 words or less.

DEVELOPING COUNTRY:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

The most cost-effective way is to invest in small-scale farming, promote sustainable agriculture, reduce food waste, and

improve market access. Additionally, strengthening safety nets and prioritizing nutritional education are crucial. Collaboration among all stakeholders is key to success.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

A cost-effective strategy involves enhancing local agricultural productivity, reducing food waste, improving market access, strengthening social safety nets, and prioritizing nutritional education. Collaboration among government, communities, and organizations is key to ensure sustainable food security.

Coachvox AI virtual expert using training literature equally with internet (7/7):

A cost-effective approach involves promoting sustainable agriculture, reducing food waste, strengthening safety nets, and improving market access. Collaboration among governments, communities, and organizations is crucial. Nutritional education and support for farmers are also key components.

ChatGPT 3.5:

Investing in small-scale, diversified, and environmentally friendly agriculture, coupled with improved infrastructure, education, and access to markets, is the most cost-effective way for a developing country to achieve sustainable food security, ensuring long-term resilience and equitable access to nutritious food for all citizens.

Bard:

Empower small farmers with climate-resilient seeds, training, and local markets. Invest in infrastructure and post-harvest storage to minimize food loss. Encourage diverse, nutritious crops and connect rural areas to urban demand. It's a multi-pronged approach, but it empowers communities, builds resilience, and maximizes resources.

MALAWI:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

Malawi can achieve sustainable food security cost-effectively by increasing agricultural productivity, strengthening food markets, targeting investments towards smallholder farming, encouraging non-agricultural work, and reforming the food security policy framework. Social protection measures are also crucial for vulnerable households.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

Malawi can achieve sustainable food security cost-effectively

by enhancing agricultural productivity, strengthening food markets, investing in smallholder farming, creating non-agricultural job opportunities, and reforming food security policy to include social protection and broader economic development.

Coachvox AI virtual expert using training literature equally with internet (7/7):

Malawi can achieve cost-effective sustainable food security by investing in agricultural productivity, strengthening food markets, diversifying the economy beyond agriculture, and implementing social protection policies for the most vulnerable households.

ChatGPT 3.5:

Implementing diversified, climate-resilient agricultural practices, investing in irrigation infrastructure, promoting efficient water management, enhancing access to modern farming technologies, and providing comprehensive education on sustainable farming methods can help Malawi achieve cost-effective and long-term food security for all its citizens.

Bard:

Empowering smallholder farmers with climate-smart techniques, like agroforestry & drought-resistant crops, is Malawi's best bet. It leverages local knowledge & resources, boosts yields, and builds resilience to climate shocks, all at a fraction of traditional interventions.

BURKINA FASO:

Coachvox AI virtual food policy expert using only training literature as knowledge base (1/7):

Burkina Faso can achieve sustainable food security cost-effectively by investing in smallholder farmers, improving

agricultural practices, enhancing market access, and strengthening social safety nets. These strategies can boost production, reduce food waste, and ensure food access for all.

Coachvox AI virtual expert mainly using training literature, but using internet if necessary (4/7):

Burkina Faso can achieve sustainable food security cost-effectively by enhancing local agricultural productivity, improving market access for farmers, reducing post-harvest losses, and strengthening social safety nets. Collaboration among stakeholders is crucial for success.

Coachvox AI virtual expert using training literature equally with internet (7/7):

Burkina Faso can achieve sustainable food security by investing in smallholder farming, improving market access, enhancing agricultural practices, and prioritizing nutrition education. Collaboration among stakeholders is key to ensure cost-effectiveness and sustainability.

ChatGPT 3.5:

Burkina Faso can achieve cost-effective food security through investments in smallholder farmer education, improved irrigation, and resilient crop varieties. Strengthening local markets, promoting sustainable farming practices, and leveraging technology for efficient resource management can enhance productivity, reduce food waste, and ensure long-term food security for all citizens.

Bard:

One highly cost-effective approach for Burkina Faso: empower local farmers with climate-smart techniques like "zai" pits (capturing rainwater), drought-resistant crops, and farmer-led seed banks. This boosts local yields, resilience, and food security at minimal cost, while empowering communities.

This publication has been prepared as an output of **CGIAR Research Initiatives on Digital Innovation**, which researches pathways to accelerate the transformation towards sustainable and inclusive agrifood systems by generating research-based evidence and innovative digital solutions, and **National Policies and Strategies**, which co-creates demand-driven policy solutions with national institutions, supporting countries to transform food, land and water systems for development and sustainable futures. This publication has not been independently peer-reviewed. Any opinions expressed here belong to the author(s) and are not necessarily representative of or endorsed by CGIAR. In line with principles defined in [CGIAR's Open and FAIR Data Assets Policy](#), this publication is available under a [CC BY 4.0](#) license. © The copyright of this publication is held by [IFPRI](#), in which the Initiative leads reside. We thank all funders who supported this research through their contributions to [CGIAR Trust Fund](#).