On Monday, June 3, Ceres2030 convened at the Food and Agricultural Organization of the United Nations (FAO). Barely eight weeks had passed since the lead authors of our seven evidence syntheses had met, for the first time, in Seattle to settle on the most promising intervention questions that would—if answered—help us to get to the goal of SDG 2, “Zero Hunger.”

In that short time, the lead authors had grown into seven teams of 75 researchers in 23 countries; the questions had been refined to meet the demands of an evidence synthesis protocol; the first searches spanning more than 50 repositories of the world’s knowledge were executed. Sixty-one members of the research teams were able to make it to Rome. Their goal: after running hard, to run even harder and faster.

They had two stages to complete: Finalize their research protocols—an awareness tool to inform the rest of the world that this research is underway and that includes a detailed plan to assess the research base in advance of looking at it—and screen their search results of titles and abstracts into manageable and meaningful numbers of full text articles to analyze.

One of Ceres2030’s guiding principles is that we had to be as inclusive as possible in seeking out relevant research, especially from the gray literature; but now we had to test that relevance with highly precise searching. If it was important not to miss anything, it was equally important that our researchers not go blind from having to look at everything.

The meeting would put the Ceres 2030 model to the test: Could we really accelerate evidence synthesis without sacrificing quality?
Turning stakeholder feedback into evidence syntheses

One of the challenges of conducting evidence synthesis in agriculture is that almost everything works in some context. Unlike biomedicine, where evidence for “what works” can be synthesized in clinical isolation, evidence in agriculture is confounded by many contextual factors—environmental, institutional, political, and societal.

Our research questions had to be broad enough to account for these factors yet precise enough to reveal the kind of robust evidence for “what works” that donors can actually use to make investment decisions.

This tension can be resolved by illuminating a spectrum of evidence for each of the broad intervention questions, and illustrated by looking at one example of a research question from the Rome meeting:

“What are the market, non-market, regulatory and compliance incentives or compulsory/voluntary programs for farmers to adopt environmentally sustainable practices?”

“In our group, we are from very different countries. Our perspectives are so different. I am from Argentina, where Argentina is a net exporter of food exporters. So I need someone from Ghana with food security problems. The disciplines are different, too, so we can combine our experience to get more feasible proposals for sustainable food production.”

— Pablo Elverdin, economist, Group of Producing Countries of the Southern Cone (GRUPO)

The team engaged in this evidence synthesis expects to review a range of market and non-market based interventions composed of regulations, voluntary or compulsory measures, public and private standards, and cross compliance incentives for farmers to adopt sustainable environmental practices and integrated risk management systems (crop insurance, catastrophic insurance, price options, mitigation and adaptation programs, etc.).

Their evidence synthesis will conclude where the evidence is strong for these interventions, where gaps exist, and make appropriate recommendations.
Structuring the synthesis

Given the volume of research to support each question, each research team has two experienced research methods experts guiding them through the evidence generation process.

This team of research methods experts held their own retreat at Cornell in early May to work through the search strategies for each intervention question—a time-intensive process. (See illustration 1 for an example of a search strategy).

As a result of their efforts, the Rome meeting began with a cumulative total of 90,000 articles to digest, and critically, with 25 percent coming from grey literature. Illustration 2 shows each team’s question and the total number of results they started with at the beginning of the week and the progress immediately following the workshop.

1. What are the interventions small-scale producers and associated value chain actors in Sub-Saharan African and South Asian LMIC countries can adopt/adapt to reduce postharvest losses along food crop value chains?
   
   Initial search results: 12,786  Post workshop: 6,679

2. Under what conditions do actors in the midstream and downstream of food products and inputs supply chains improve the welfare of small-scale producers?
   
   Initial search results: 12,458  Post workshop: 3,779

3. What are determinants that lead small-scale producers in low- and middle-income countries to adopt climate resilient crops and crop-varieties?
   
   Initial search results: 5,651  Post workshop: 3,294

4. What is the impact of ruminant feed interventions on livelihoods among small scale and agro-pastoral livestock keepers in LMICs in Africa, Asia, and Latin America?
   
   Initial search results: 22,205  Post workshop: 7,770

5. What spectrum of farm-level interventions to alleviate water scarcity have been tested to increase small-scale farmers’ incomes, yields, and productivity?
   
   Initial search results: 17,681—Team is still experimenting with a machine model

6. What are the market, non-market, regulatory and compliance incentives or compulsory/voluntary programs for farmers to adopt environmentally sustainable practices?
   
   Initial search results: 17,936  Post workshop: 6,095

7. What are the effects of skills training interventions on educated and non-educated youth employment outcomes in agricultural value chains, agribusiness or contract farming in developing countries?
   
   Initial search results: 4,789  Post workshop: 4,020
How semantic searching transforms the evidence base

The keys to our “lean start up” approach to evidence synthesis is that we emphasize collaboration between researchers and information specialists and using machine learning to accelerate the screening work.

We needed the teams to speed through these 90,000 articles—and to do that, a small team of computer and information scientists led by Jaron Porciello at Cornell developed an open-source machine-learning model. The model was trained in advance of the meeting and the data team was on-hand during the meeting to discuss search issues in real time with the author teams.

They were able to help the teams analyze their individual datasets by using machine learning to generate new fields and metadata. This provided new datapoints to help the teams decide what was “in or out” during the title and abstract screening, the precursor stage to examining full research papers.

To say this approach is transformational is an understatement. The teams were able to quickly and intelligently eliminate thousands of irrelevant articles as defined by their inclusion and exclusion criteria. Initial inclusion searches that yielded eye-wateringly large datasets of 12,000 papers were refined to more manageable numbers (see illus. 2).

Even though semantic searching was an entirely new concept to most of the researchers, the teams rapidly began to use the technology, with assistance from our computer scientists, to design and customize the machine learning outputs. The teams all said that what they valued most about the Rome meeting was being able to work together in person for a week.

What we learned

In attempting something so large, complex, and unprecedented, it is inevitable that not everything goes quite to plan. As noted, we have one team that is behind schedule, and is wrestling with the complexity of the question they want to ask and answer. On our final day in Rome, they were the last to present, outlining their challenges. Exhausted though everyone was, researchers from the other teams stepped up to help.

The random allocation of team authors meant that some teams were quick to innovate and others less so. The global spread of team members meant that data sharing could have been better in advance of Rome. In hindsight, it would have been ideal to have all authors meet in April in Seattle to work on the questions and immerse themselves in the methodologies they would be using. Valuable time in Rome went on catch up.

“For me, this has been quite exciting. Two things I am excited about are the extent to which this process will prove superior to traditional approaches to reviews and second, the extent to which we will find other people studying these topics, which are not as evident as one would think in the general literature. Sometimes, when we have a few very important and powerful voices that we know, we tend to go to these voices for what they have said. Already we have seen some very interesting work that is not from the people that we would typically see—and it is work which has actually been done well.”

— Saweda Liverpool Tasie, Evidence Synthesis Lead Author, Associate Professor at Michigan State University, and Principal Investigator with the Feed the Future Nigeria Agricultural Policy Project.
Purposeful and forward-looking

Jaron Porciello, Cornell University primary investigator for Ceres2030, concluded the week in Rome by enunciating three guiding principles for Ceres2030 that described her vision for collaborative and coordinated evidence for agriculture:

“I believe that when we give experts better tools and support to generate evidence, and when we work on this endeavor in truly collaborative ways, then we stand a better chance of donors and governments implementing evidence-based solutions. I am so incredibly proud of the work we achieved in Rome and the passion shown by all the researchers. Given that our planetary pressures are at an all-time high, we need to be creative and we need to embrace risk to advance the generation of evidence.”

What we achieved in Rome

We challenged our author teams and they responded magnificently. We pushed them intellectually. We pushed them physiologically. Each day was a 10-14 hour marathon, with many giving up sleep too.

The result is one of the largest literature-based research projects ever conducted in agriculture—if not the largest. In approach, nothing like this has been attempted before.

The design, testing, and deployment of machine-learning tools in the service of evidence synthesis is an intellectual achievement that will become the norm in academic research. Technical papers on the process are being written.

In terms of our schedule, seven of the eight teams had finalized and uploaded their research protocols onto the Open Science Framework on June 4.

By the end of the Rome meeting, seven of the eight teams informally presented their research questions, protocols, and datasets at FAO. We have one team that is behind schedule.

Going forward, the teams are aiming to screen all their titles and abstracts by June 24. Full text analysis will be completed by the end of July. Manuscripts are due by the end of August.

— Jamie Morrison, Director and Strategic Programme Leader, Food and Agriculture Organization of the United Nations, and Ceres2030 expert advisory board member asks a question at the Evidence Synthesis presentation session on the closing day of the Rome meeting.