

# PRESS RELEASE

## COCO PROJECT

Service provider in the sense of §5 TMG adelphi research gemeinnützige GmbH

## MEDIA CONTACT

Paula Cros Marchena  
Alt-Moabit 91  
10559 Berlin

info@coco.eu

## 1,077 FARMS, 12 COUNTRIES, A HUGE DATASET

### The scale and complexity behind CoCo's pan-European farmer survey: A conversation with CITA researchers

25 March 2026

What does it actually take to understand how farmers across Europe live alongside wolves, bears and wolverine? For the CoCo project, it starts with a questionnaire and approximately 1,000 of them, collected from farms in 12 countries and 30 case study areas. Getting those questionnaires filled in was a feat in itself. But turning the resulting mountain of data into something researchers can actually analyse? That has been a different challenge entirely. We spoke to Ana Grau Valenciano and Eduardo Torres Martínez, researchers at CITA Spain, to find out what that process actually looks like.

Field workers collected farmers responses on paper during the face-to-face survey that lasted between 1 to 3 hours before transferring them into standardised Excel sheets, which were then uploaded to a shared drive. In theory, the built-in format constraints were designed to keep answers consistent. In practice, a dataset spanning a dozen countries, multiple languages and more than 40 interviewers and more than 1,000 farmers was always going to throw up surprises.

The job of cleaning and consolidating the data has fallen to a team of four researchers at CITA Aragon, the CoCo Spanish partner. Each researcher took responsibility for a set of countries, checking that formatting rules had been followed, that answers made sense and that nothing had slipped through the cracks. All of it had to be identified, flagged and either standardised or set aside, to ensure the final dataset is comparable across all 12 countries.

It sounds straightforward. It is anything but. They have been working on it since early February.

“Faced with such a complex and massive database, data entry can become a monumental task, where small errors naturally arise due to the overwhelming volume of information. However, finding and correcting these recording mistakes is essential to ensure that the final analysis is rigorous and reliable.” Eduardo Torres Martínez, Researcher, CITA Spain.

The challenges were as much about the complexity of European farming as about data processing. Across 12 countries, farming systems vary enormously, and fitting that diversity into the standardised, close-ended questions of a survey is no small task. Many respondents added written comments

alongside their answers, describing nuances and local realities that didn't quite map onto the questionnaire's categories. At the cleaning stage, the team had to make careful decisions about how to interpret and code that additional detail – judgement calls that could affect the entire dataset.

The technical challenges were equally demanding. Decimal separators alone –a comma in some countries, a full stop in others– created significant inconsistencies across the data. As Ana Grau explained to us, precision mattered enormously: a single digit wrong in the coordinates of a farm or grazing land could place a Spanish farm somewhere in Turkey, which actually happened during the review process. Meanwhile, the questionnaire's section on predator incidents required a separate row for each combination of predator species, livestock type and location, a structure that demanded painstaking attention to ensure every case was correctly recorded and complete.

Before the cleaning could even begin, the team spent around a week developing a precise protocol, a shared set of rules to ensure that every researcher made exactly the same decisions when faced with the same situation, accounting for the full range of exceptions and edge cases the data might throw up. Only then could the work itself start.

At the start of the process, working through a single country's dataset –around 100 questionnaires– took from a one full day to a couple of days for a first revision, then start the back-and-forth iterations with partners. By the end, the team had become so attuned to the patterns of each partner's data that the same task could be done in half a day. But speed in the cleaning phase was only part of the story: every query raised with a partner triggered a round of back-and-forth emails and online meetings that could stretch over weeks, particularly when field workers were still out in the field and unable to check the original paper questionnaires.

### **130,000 rows and counting**

The next challenge is combining all the cleaned national datasets into a single master database. That database –still being assembled– already contains more than 130,000 rows and is expected to reach at least 200,000. When the team attempted to merge it in Excel, the file exceeded the software's memory limits. The solution: converting everything to CSV format first, stripping out hidden characters and formatting quirks, before adding each country's data to the master file one by one. Plus, without an interface which doesn't need to render the data in a screen, they're saving a ton of memory in the process.

Once complete, the dataset will be divided among CoCo's research partners for analysis across several parallel workstreams: farming typologies and livestock management, the effectiveness of predator prevention measures, farmer attitudes and views towards large carnivores its management and other governance issues and a choice experiment exploring how farmers weigh up different policy options. It is, in short, the empirical foundation on which CoCo's eventual policy recommendations will rest.

The scale of the undertaking reflects the ambition of the project. Coexistence between people and large carnivores is one of the most contested issues in European conservation policy. Getting the data right –all 130,000 rows of it– is where that work begins.

Follow the project's progress and stay informed about new findings at <https://cocoproject.eu/>.

---

### **About the CoCo Project**

CoCo (Cocreating Coexistence) is a Horizon Europe-funded research project working to develop sustainable solutions for human-wildlife coexistence across European pastoral landscapes. The project brings together diverse stakeholders to co-create knowledge and tools that support both thriving pastoral communities and healthy carnivore populations.