

RESPONSE TO BIDDER QUERIES - RFP/RBE/001/2024

MODELLING RETURN SCENARIOS FOR REFUGEES FROM UKRAINE

Question 1: In relation to the in-person workshop, could you provide details about the venue? We need this information to calculate travel and accommodation expenses accordingly.

Answer 1: As indicated in Annex A – Terms of Reference, the in-person workshop/seminar is expected to take place in a European country, preferably in Geneva –Switzerland. This said, the precise geographical location will be shared with the winning bidder once determined.

Question 2: Could you outline the expected level of accuracy for the model?

Answer 2: As indicated in Annex A – Terms of Reference, the expected outputs of the model are to simulate potential flows and profiles of refugees returns to Ukraine under different future scenarios or key contextual factors (i.e. legal avenues for refugees in host countries, socio-economic opportunities and potential evolution of security situation in Ukraine). Nonetheless, it is expected that the proposed model has a solid theoretical framework and that the parameters and assumptions can represent accurately agent’s interactions and decision-making process as reflected in available evidence and the model associated uncertainty. The prospective bidder is expected to explain what type of sensitivity analysis and parameter calibration considerations will be considered during development and testing.

We refer also to the criteria indicated in Annex B – Technical Offer Form, Part II (Project Proposal Quality), which will be used to assess the appropriateness of the methodology proposed, the modality for model’s development and the feasibility of the model to be expanded or adapted as new evidence becomes available or as future changes in the environment become clearer.

Question 3: Regarding all the database, can you please let us know who is responsible for extracting and cleaning the data for the model?

Answer 3: As indicated in Annex A – Terms of Reference, it is expected that one of main data sources from the model will be surveys led or co-led by UNHCR. All microdata from UNHCR surveys is cleaned, anonymized and published for external use in UNHCR’s microdata library, but the selected contractor will be expected to extract and prepare the relevant data points or indicators needed for the model. Another main source of data will correspond to population statistics of refugees from Ukraine in different host countries, which UNHCR publishes in an open data portal and can also facilitate access to time series data. Finally, official administrative datasets or sources from host countries’ authorities that are publicly available are also expected to be key sources for the model. UNHCR will contribute to identify known sources that can be relevant for

the model, but the selected contractor will be expected to extract and prepare the relevant data points or indicators as well as to identify on their own relevant sources, including if relevant big data sources.

Question 4: What is the size and format of the data intended for use in the model?

Answer 4: Regarding microdata from UNHCR surveys, the size varies by survey, ranging from country surveys of around 800 to 1,000 households to regional surveys of 4,000 to 5,000 households. UNHCR usually publishes microdata in CSV format, along with its corresponding metadata. For publicly available data from official sources size and format will depend on type and nature of the data (i.e. whether it is time series data on population statistics, or sectorial administrative records of refugees hosted in the country, or subnational socio-economic indicators, among others).

Question 5: Are the indicators/factors/variables predetermined within the datasets, or is our assistance required in determining them? Please verify.

Answer 5: As explained above, the selected contractor will be expected to extract and prepare the relevant data points or indicators from available sources. Such selection should be based on the theoretical and methodological framework that will be proposed for the model (which will be later finetuned and agreed by UNHCR with the selected contractor). Furthermore, as indicated in Annex B – Technical Offer Form, one of the evaluation criteria of the proposals will be the *“Appropriateness of the methodology proposed, including the theoretical background and hypothesis to be used in the model’s development, the variables that can be explicitly included in the simulations, and the thematic and geographical disaggregation”*.

Question 6: Do we need to build APIs from the scratch to integrate it with envisaged model or APIs are available and client will provide us.

Answer 6: With regards to microdata from UNHCR surveys, data can be downloaded / shared in CSV format for each different survey. For some surveys UNHCR has developed and published interactive online dashboards using Powe BI where key results/indicators can also be accessed. Population statistics data compiled by UNHCR can be accessed in its open data portal as well as through API access provided by UNHCR. For data from official sources, some might include the option to access datasets via APIs, though this is not likely for most possible sources.

Question 7: Could you confirm the total number of systems/databases that need integration with the data model?

Answer 7: As mentioned above, one main source of data will correspond to UNHCR surveys. Several country and regional surveys have been implemented so far and it is expected that the selected contractor will identify and select the most relevant surveys and indicators in consultation with UNHCR. For population statistics compiled by UNHCR, the agency can facilitate access – including via API- to the historic time series data. Finally, for publicly available official sources it will depend on which are considered relevant for the proposed model.

Question 8: Can you confirm if a web interface is necessary for the data model? If so, kindly provide an overview of the modules/features and your technology stack preferences?

Answer 8: Having a web interface for the data model is not a minimum requirement for the project, but prospective bidders are welcome to propose it as an optional or added feature if they wish to do so. Regarding technological preferences, prospective bidders should specify and justify in their technical proposal the programming language they will use to develop and run the proposed model in order to satisfy the technical requirements. Open-source solutions are a preferred option. Currently, UNHR uses R Studio - Posit Connect to deploy data science applications developed using R and Python languages. While the possibility of using this platform to deploy any web-based interface will be preferred, prospective bidders can propose and justify the use of an alternative technology.

It is imperative, however, that any proposed solution must adhere to all relevant legal and regulatory standards, including those pertaining to data protection and intellectual property rights, as well as those established under UNHCR's [General Policy on Personal Data Protection and Privacy](#) and [UNHCR's Policy on Information Security](#).

Question 9: Regarding the web-interface, can you please let us know the user roles associated with it?

Answer 9: Please see answer to question 8. While a web interface for the model is not a minimum requirement for the project, prospective bidders are welcome to propose it as an optional or added feature if they wish to do so. However, access to any such web interface should be limited for internal use by specific UNHCR staff, and the parties should agree if it could be made public after the finalization of the project (see Annex A, Part III).

Question 10: Can you please let us know your preference regarding the hosting of the web-interface?

Answer 10: Please see answer to question 8. While a web interface for the model is not a minimum requirement for the project, prospective bidders are welcome to propose it as an optional or added feature if they wish to do so. Currently, UNHCR uses R Studio - Posit Connect to deploy and host data science applications developed using R and Python languages. While the possibility of using this platform to deploy any web-based interface will be preferred, prospective bidders can propose and justify the use of an alternative technology. As mentioned in answer 9, access to any such web interface should be limited for internal use by specific UNHCR staff.

Question 11: Is model training included in the scope, or is it scheduled for later phases? Please clarify.

Answer 11: The minimum expected results and deliverable are indicated in Annex A – Terms of Reference, Part III. While model training is not a minimum deliverable, prospective bidders are welcome to propose it as an optional or added component if they wish to do so considering that the expected completion date of the project is end-of June 2024. Please note also that one of the evaluation criteria specified in Annex B – Technical Offer Form, refers to “*Feasibility of the proposed model to be adapted or expanded to include additional factors and data sources, and/or replicate results under different future scenarios*”.

Question 12: Do we need to create the model from scratch, or can we leverage existing open-source models available in the market?

Answer 12: Please refer to Annex A for the technical requirements. Any product/model offered should meet these requirements. Given the tender is Request for Proposal, prospective bidders are expected to propose an approach which satisfies the requirement.

Prospective bidders can indeed propose to use or leverage existing open-source models or codes, though they will be expected to justify in their technical proposal the relevance of such previous models or codes to satisfy the technical requirements of the project. It is imperative, however, that any proposed approach, solution, or model must adhere to all relevant legal and regulatory standards, including those pertaining to data protection and intellectual property rights.

End of the document