Public Works and Government Services Canada | Travaux publics et Services gouvernementaux Canada

RETURN BIDS TO:
RETOURNER LES SOUMISSIONS À:

Bid Receiving - PWGSC / Réception des soumissions - TPSGC
11 Laurier St. / 11, rue Laurier
Place du Portage, Phase III
Core 0B2 / Noyau 0B2
Gatineau
Quebec
K1A 0S5
Bid Fax: (819) 997-9776

SOLICITATION AMENDMENT
MODIFICATION DE L'INVITATION

The referenced document is hereby revised; unless otherwise indicated, all other terms and conditions of the Solicitation remain the same.

Ce document est par la présente révisé; sauf indication contraire, les modalités de l'invitation demeurent les mêmes.

Comments - Commentaires

Vendor/Firm Name and Address
Raison sociale et adresse du fournisseur/de l'entrepreneur

Issuing Office - Bureau de distribution

Systems Software Procurement Division / Division des achats des logiciels d'exploitation
Terrasses de la Chaudière
4th Floor, 10 Wellington Street
4th etage, 10, rue Wellington
Gatineau
Quebec
K1A 0S5

Title - Sujet Automated GEOSCAN Mineral Mapping	
Solicitation No. - N° de l'invitation 23240-200239/A	**Amendment No. - N° modif.** 002
Client Reference No. - N° de référence du client 23240-200239	**Date** 2019-11-14
GETS Reference No. - N° de référence de SEAG PW-$$EE-017-36880	
File No. - N° de dossier 017ee.23240-200239	**CCC No./N° CCC - FMS No./N° VME**
Solicitation Closes - L'invitation prend fin **at - à 02:00 PM** **on - le 2019-12-23**	**Time Zone** **Fuseau horaire** Eastern Daylight Saving Time EDT
F.O.B. - F.A.B. **Plant-Usine:** ☐ **Destination:** ☐ **Other-Autre:** ☐	
Address Enquiries to: - Adresser toutes questions à: Lessard, Peter	**Buyer Id - Id de l'acheteur** 017ee
Telephone No. - N° de téléphone (613) 850-7602 ( )	**FAX No. - N° de FAX** ( ) -
Destination - of Goods, Services, and Construction: **Destination - des biens, services et construction:**	

Instructions: See Herein

Instructions: Voir aux présentes

Delivery Required - Livraison exigée	**Delivery Offered - Livraison proposée**

Vendor/Firm Name and Address
Raison sociale et adresse du fournisseur/de l'entrepreneur

Telephone No. - N° de téléphone
Facsimile No. - N° de télécopieur

Name and title of person authorized to sign on behalf of Vendor/Firm (type or print)
Nom et titre de la personne autorisée à signer au nom du fournisseur/ de l'entrepreneur (taper ou écrire en caractères d'imprimerie)

Signature **Date**

Canada

Amendment 002

This amendment 002 is raised to modify the Request for Proposal and answer questions.

1) At article 1.0 Requirement of the Artificial Intelligence Contract

Delete article 1.0 Requirement in its entirety

Insert the following:

1.0 Requirement

1.1 The Contractor agrees to supply the services and deliverables to the Client the as described in Stage I of the Statement of Work and the Contractor's technical bid entitled _____, dated _____in accordance with, and at the prices set out in, the Contract.

1.2 Optional Goods and Services for Stage II and III. The Contractor grants to Canada the irrevocable option to acquire the goods, services or both described at:

a) Optional Stage II, of the Statement of Work in Annex ‘’B’’; and

b) Optional Stage III, of the Statement of Work in Annex ‘’B’’;

under the same conditions and at the prices and/or rates stated in the Method and Basis of Payment in Annex C. The option may only be exercised by the Contracting Authority and will be evidenced, for administrative purposes only, through a contract amendment.

The Contracting Authority may exercise the option at any time before the expiry of the Contract by sending a written notice to the Contractor.

2) At Annex B – Statement of Work

No modifications

3) Respond to questions from suppliers:

Q1. Could you elaborate on the answer to question 11 in the presentation, related to the criteria to select bidders?

R1. The selection criteria is determined by the client department, not by the Government of Canada. For example, supplier’s responses to the NPP, information gathered from the engagement session and previous knowledge of the suppliers and the work that they do may all be considered.

Q2. Will you have additional data provided, like a training set?

R2. Yes, a training set will be provided with 234 publications. Predictive models in the first stage is based on key words and titles in the metadata.

Q3. Do you require a visualization solution for the data extraction?

R3. No, we are not considering a visualization solution.

Q4. Do you have any constraints with a cloud-based hosted solution?

R4. No, we are already working in the cloud with various cloud suppliers. A cloud-based solution is a preferred solution. The digital solution will be developed outside NRCan’s cloud and NRCan premise, but once completed, it will be hosted in our cloud, and this cloud-hosted solution does not have security constraints.

Q5. How many examples are in the training set and what is the total number of publications to be examined?

R5. Currently, there is a total of 234 records in the training set and a total of approximately 83,000 records in the GEOSCAN database. More examples for the training set can be provided (e.g., weaker matches), if needed.

Q6. Do you have bilingual records in your database?

R6. Yes, records are in English, French, and bilingual. The majority of the metadata contains English and French language. The proposed digital solution must be compatible with both official languages.

Q7. How was the training set assembled?

R7. As we are interested in nickel ore systems, we searched GEOSCAN using key words, such as “nickel,” to select candidate training files. Research scientists then went through the titles and abstracts of these files to examine which ones truly focused on nickel ore systems. This subset of files became the training set. Less obvious key words, such as “ultra-mafic,” should presumably be identified by the AI solution.

Q8. How do you determine if the result is positive or negative (i.e., whether a publication is relevant to nickel ore systems)?

R8. When we developed the training set, we focused on fields such as titles containing words that clearly indicate relevance to nickel ore systems. However, the AI solution will hopefully draw on all metadata fields and potentially less obvious phrases to find “hidden” but relevant publications.

Q9. The metadata sounds somewhat limited. Is it within the scope of work to also use PDF scans?

R9. PDF scans are not in the scope of work for stage 1 (mandatory stage of the contract), but will be a key component of the optional stage 3 work. The metadata contains abstracts that are often quite detailed.

Q10. AI contractors are generally not geoscience experts. Will there be an expert in geoscience field (i.e., with domain knowledge) available to support the contractor?

R10. Yes. We expect industries expertise in AI, weekly meetings with the client are part of the scope of work (minimum frequency) to ensure that the project is on track.

Q11. Are you expecting a multi-class algorithm to classify the relevant constructs?

R11. It will be multi-class due to the words that predict nickel ore systems are quite varied. To clarify, the scope is only related to nickel ore systems.

Q12. Do you want to know the type of algorithm we will use and why our algorithm is more performant than alternatives?

R12. No, we are not interested in why your algorithm performs better than others; we are only interested in the desired results.

Q13. Are you providing sensitive data?

R13. All the data is open data, and there is no sensitive data for this project. An algorithmic impact assessment was conducted with Treasury Board Secretariat at the beginning of this project and it did not identify any risk associated with this project. The assessment also indicated that this project would result in a responsible and ethical artificial intelligence output.

Q14. Will this project focus only on nickel ore systems or will others be incorporated?

R14. This project will focus on nickel ore systems.

Q15. In training data sets, are there examples of false positives?

R15. No, but some can be added if needed.

Q16. The hyperlink given in section SW 5.2 is not working. Could you send the updated link to the GEOSCAN public database?
https://geoscan.nrcan.gc.ca/starweb/GEOSCAN/servlet.starweb

R16. See the following link, https://geoscan.nrcan.gc.ca

Q17. As the RFP indicates: Stage 1 scope only requires the accuracy of the search engine to work with the 80,000 entries that exist in XML document format. Can you confirm if PDF's are in scope or out of scope?

R17. PDFs are not in the scope of work for Stage 1, but form part of optional Stage 3 work.

Q18. On the Webinar we heard that the key words and abstracts have been translated, if in French originally. Can you confirm that all the key words and abstracts in GEOSCAN are all in English?

R18. Entries in GEOSCAN may occur in English, French or both. The digital solution must take both languages into consideration.

Q19. Can you confirm if the output from the predictive model is required in both English & French?

R19. The raw output of the predictive model for stage 1 can be an electronic data table containing the GEOSCAN ID number, spatial coordinates and the machine learning results (entirely numeric). However, the entire digital solution must be able to accommodate and build predictive models using English and French language.

Q20. Are there page limitations to the proposal? Please indicate if there are page limitations by proposal sections or for the entire proposal.

R20. There are no page limitations for the entire proposal or individual sections.

Q21. At any stage, is there a possibility of utilizing supplemental datasets whether provided by NRCan or publicly available, to improve the model? For example, there may be ways to bridge gaps between data sets to improve the models.

R21. The focus of the current proposal is on natural language processing tools of the NRCan metadata. Means of improvement and supplemental data integration could be considered in the future depending on the outcomes of the current proposed project.

Q22. In later stages, would NRCan be open to changing how data is stored within the publication database to better improve how the process works? For example, it might make sense to update the metadata format to better predict very large tags in keeping with best practices.

R22. There are currently no plans to change how data is store in the publication database.

Q23. Is there currently, or has there been prior to the issuance of the NPP, an external organization working with the Government of Canada in the development of requirements or potential solutions. If yes, can the Government of Canada provide the name of the consultant or company? Will this consultant or company be permitted to bid on this procurement?

R23. There is no other project of this kind at NRCan. This Statement of Work has only been developed by NRCan staff, with support of PSPC guidance on procurement. No external firms were involved or consulted in the conception of this project.

Q24. As stated in the SOW on page 3, the last deliverable for Stage 1 (Stage 1 Reporting) is due on March 31, 2020. The SOW also states on page 4 that the Delivery of Predictive Model Report is due on March 18, 2020. Could the Government please indicate if these dates are correct, and if not, please provide revised due dates?

R24. The different dates are correct. The March 31, 2020 date is the contract end date for Stage 1 deliverable and the March 18, 2020 is the deliverable date.

Q25. Does the Government of Canada have a preferred format for the draft project plans requested for this proposal i.e. Excel, Microsoft Project, or Word tables? Would you prefer the images of the draft project plans to be included in the Word document proposal or submitted as attachments to supplement the Word document?

R25. A single Word or pdf with images included in the document would be preferable.

Q26. Does the $499,999 ceiling price include license costs?

R26. Yes, any licensing cost must be included in the price ceiling.

Q27. Availability of Named Individuals. The Bidder certifies that, if it is awarded a contract as a result of the bid solicitation, every individual proposed in its bid will be available to perform the Work as required by and at the time specified in the bid solicitation. If for reasons beyond its control, the Bidder is unable to provide the services of an individual named in its bid, the Bidder may propose a substitute with similar qualifications and experience. The Bidder must advise the Contracting Authority of the reason for the substitution and provide the name, qualifications, and experience of the proposed replacement. For the purposes of this clause, Canada will consider only the following reasons as being beyond the control of the Bidder: death, sickness, maternity and parental leave, retirement, resignation, dismissal for cause, and termination of an agreement for default.

Would Canada consider amending the language to the following: "The bidder must make their best efforts to ensure that the proposed resource is available for this project. Canada will accept an equally qualified replacement if the proposed resource is no longer available at time of contract award.

R27. We are not going to amend the language. We are requesting the same resource expertise as the proposed resources in the proposal.

Q28. You mention accuracy in the description but is there any preference for reducing false positives or false negatives? What I mean by this is should the work focus on reducing either false positives or false negatives.

R28. The binary classes of the entire dataset will be heavily unbalanced (e.g., 90% of publications unrelated to ore systems; 10% related to ore system). For that reason, a good model won't be based entirely on accuracy. In your proposal, consider approaches/tools that will have good model precision. Because the results of this proposal will be fed into our mineral potential model, false positives (i.e., predicting the presence of a mineral deposit when that area is in fact barren, which put mineral exploration dollars in the wrong place) tend to be more costly.

Q29. Can the Government please explain the advantages of selecting the 7 additional vendors randomly based on expression of interest rather than being based on the merits of the solutions potentially available from the marketplace?

R29. Reducing the bidding pool enables PSPC and client departments to maintain procurement timelines and project deliverable deadlines. Once the bidder's pool is reduce, suppliers will be equally evaluated based on the criteria in the solicitation.

Q30. What is the criteria being used by the Government of Canada to select the 3 vendors?

R30. The client will select three pre-qualified bidders at its discretion from among those who have self-identified their interest to bid on this Request for Proposal. A specific timeline (November 12, 2019 at 2PM EST) is specified for the suppliers to express interest to compete.

Q31. Why choose accuracy and precision as performance measures (rather than the usual measures for recommendation systems - accuracy, recall and, if only one measurement is to be used, F_1 measurement)?

R31. The training data is imbalanced and so we are interested in precision in addition to accuracy and recall. The more validation procedures the better.

Q32. Can we have more details on "cross-validation using learning data provided by NRCan and manual validation performed by NRCan with a smaller subset of GEOSCAN inputs"?

R32. As an extra measure of validation we will select a subset of documents highlighted by the predictive model as a match and qualitatively assess their fit. This will involve just reading the abstracts of that subset of matches with subject matter experts.

Q33. Would it be possible to consult the data used for this cross-validation?

R33. Not now. The idea is to have the bidder and NRCan follow independent validation procedures.

Q34. With respect to learning data, would it be conceivable to also provide a set of entries that are not related to the targeted mineralized system, i.e., could the learning data include both positive and negative examples (not to be confused with metadata for which no relevance judgment has been made)?

R34. Yes, in principal. We could select publication on rocks that are not know to host Ni or publications on other ore systems.

Q35. For a given mineralized system, how many examples can we expect to find in the learning data?

R35. The Ni training dataset contains about 200 publications.

Q36. The mandatory technical criteria speak of a cloud solution, what type of minimal interaction is expected? In particular, how many user groups are to be created and what would be their role?

R36. Once completed, the digital solution will be hosted in NRCan's Cloud. The interactions and use of the digital solution with NRCan users are not a determinant of the proposed solution.

Q37. Speaking of minimal solution, is it sufficient to add a new metadata listing to the mineral systems associated with a publication?

R37. No. The GEOSCAN metadata will remain untouched. The expected output of stage 1 is an electronic data table containing the GEOSCAN ID, spatial coordinates and predictive model results.

Q38. In order to avoid the multiplication of cloud computing solutions within the organization, would it be conceivable to add the proposed digital solution to an existing cloud computing solution?

R38. You are encouraged to offer the most powerful machine learning solution.

Q39. For the third stage, is it required to be able to extract spatial coordinates, including topographic data from a mineralized system from a map rather than text?

R39. The predictive model should be based on the content of the PDF, however, the spatial coordinates can still come from the metadata.

Q40. Given this, how many examples can we expect to find in the learning dataset?

R40. See response 35.

All other terms and conditions of the Request for Solicitation remain the same