Innovative Solutions Canada Program

Challenge EN578-170003/32: Automated redaction of video recordings for the purposes of Access to Information requests Challenge

Attachment 6
Questions and Answers #25 to #36

This document contains questions and answers related to this challenge.

Question #25:

Is there an expectation to cover video footage across "indoors/outdoors" as well as during "night & day" situation? If so, is it possible to get a screenshot of a night-time outdoor footage?

Response #25:

Yes there is an expectation to cover footage across indoors/outdoors as well as night and day. There are no samples of footage available during the bid submission phase.

Question #26:

Based on the previously answered Question 16 (3a), it seems the proposed solution would also need to understand the context of the audio. This would mean that the proposed solution would need to understand the audio context from multiple languages. Is our assumption correct? If yes, which languages would need to be covered besides French and English?

Response #26:

There is no requirement for foreign languages. Vetting outside the two Official Languages can be done manually.

Question #27:

Of the 2 meetings proposed as a part of Phase I (kick-off meeting & final review meeting), could you share:
a)   The agenda of the meetings (would it have mock videos, anticipated output)
b)   Tentative timelines (the gap between kick-off meeting & final review meeting, any mid-period check-ins etc)
This will help us submit a more detailed and accurate project plan.

Response #27:

Details regarding Phase 1 meetings will be discussed with selected bidders during the contract negotiations stage. The kick-off meeting would likely occur within the first couple weeks of contract award.

Question #28:

Given that AI technology is rapidly evolving, it is likely that the new advancements would make us want to upgrade our AI technology stack for the solution compared to our initial proposed plan. We would like to confirm with you if this is an acceptable approach on a supplier's part.

Response #28:

The total cost of the Contract must be the sum of the applicable direct and indirect costs which are, or must be reasonably and properly incurred and/or allocated, in the performance of the Contract, less any applicable credits. The Financial Proposal submitted will be negotiated before contract award and must be in accordance with the PWGSC Contract Cost Principles 1031-2 amended as follows:
At Section 7, subsection b:

DELETE: legal, accounting and consulting fees in connection with financial reorganization, security issues, capital stock issues, obtaining of patents and licenses and prosecution of claims against Canada;

INSERT: legal, accounting and consulting fees in connection with financial reorganization, security issues, capital stock issues, obtaining of licenses and prosecution of claims against Canada;

Question #29:

Building on top of the previously answered Question 9, our query is as follows:

a) Typically, algorithms for object recognition are evaluated over hundreds (if not thousands) of media files. As per the response to Q 9, there are only a dozen unclassified videos for testing. As such, the results may not be indicative of the true performance of competing solutions (because of increased standard deviation due to small sample size). Is there an idea to use CBSA classified videos to evaluate the solutions during Phase 1? If yes, how will the teams interact with these videos? We assume most teams would like to know about the actual failure modes to learn from them and improve their AI.

b) Are you planning to use open source datasets for evaluation? If so, which?

Response #29:

a) No, classified videos will not be used in Phase 1 or 2 of the resulting contracts.
b) The data sets being used for the testing phase are not yet identified. The intent is to work with the selected bidders to establish a testing protocol during the testing phase.

Question #30:

Of the 10,000 cameras CBSA has, can you please describe the position in which they are usually installed? For example, cameras could be installed on the ceiling (overhead cameras) or they could be installed at human head-level to record people in front of them (body cameras)? Are there usual camera position examples we should be aware of?

Response #30:

CBSA cannot describe our operational surveillance installations in detail to an open forum. Consider the systems being in the realm of traditional overt video surveillance. Cameras are installed on walls, ceilings, outdoor polls, etc like any other surveillance environment. Some cameras have a wide angle view that present a storytelling or overview view while others are focused on specific objects of interest such as faces and licence plates to provide operational detail views– again just like any traditional video surveillance system.

Question #31:

Are security cameras close enough to the people filmed to see and recognize their faces?

Response #31:

Yes.

Question #32

Are there special cameras installed with the sole purpose of filming people's faces, as opposed to filming crowds from a distance??

Response #32:

Yes.

Question #33:

a) Are scenes in the videos well-illuminated?
b) Should our system be able to deal with footage recorded in the darkness?

Response #33:

Yes to both, but the video in darkness is usually monochrome and may be illuminated with infrared lighting.

Question 34:

Is our system required to handle visually noisy, pixelated or damaged video recordings? Is this a priority?

Response #34:

Aspects of the video might be pixelated/ visually noisy. The redacting needs in the video decrease as the clarity of the video decreases. It is expected that as the quality/clarity of the video degrades, more manual processing will be needed. The priority would be to minimize the time and effort around this manual processing.

Question #35:

For the audio masking: will other than English and French languages be required? I.e. Middle Eastern person who speaks English says a name of a third party in Arabic in between English Sentences.

Response #35:

No – only English and French will be handled systematically. All other languages would require manual processing.

Question #36:

a) Regarding essential outcome #1 when it says "allow the processing of digital audio and video recordings, by redacting all instances of all personal identifiable content", it is understood that the proposed solution is required to partly automate the redaction of personal identifiable content in a video.
b) Is this also required for audio?
c) Is the goal only to provide software that will purely allow for the manual redaction of audio segments?
d) Or to also partly automate this process (ex. to use speech recognition techniques to help the user identify personal identifiable content in the audio recordings)?

Response #36:

a) Yes. However the focus of a solution is to minimize manual intervention as much as possible
b) Yes
c) No
d) Yes with the ultimate solution having little to no user intervention