GLOBSEC
IDEAS SHAPING THE WORLD

[TO ALL PROSPECTIVE TENDERERS]

In Bratislava
On 28.06.2019

RE: <u>Clarification of Tender Document No. 2 – Request for Proposals - HADES</u>

On 28th May 2019 the Purchaser, GLOBSEC, with its registered seat at Kuzmányho 4031/3, Banská Bystrica, 97401 have issued the Tender Document for procurement of „Human-AI Disinformation Extraction and analysis System (HADES)", Tender No. 01-2019.

The Purchaser received a request for clarification containing questions below and in accordance with clause 9 Clarification of Tender Document of Part A. Instructions to Tenderers of the Tender Document, the Purchaser hereby issues this Clarification of Tender Document No. 2 as follows:

Clarification to Tender Document		
Question No.	Request for Clarification	Clarification
Request for Clarification dated on 26.09.2019		
1	Should a price for using cloud and cloud services during the development and during the first two years of production us be included in the price proposal	Yes, it should be included in the price proposal.
2	Who needs to guarantee that we can scrape Facebook, Instagram, and other services from a legal standpoint? a) Who is responsible if a service closes its API and the content will not be accessible anymore. b) We cannot contact Facebook and other services prior to the contract (because of the NDA) and agree on the API usage. If this will not be provided by Globsec, access to API cannot be guaranteed in the proposal.	a) Responsibility over changed API takes the supplier, however in the case the API will be deleted or the access to data will be legally or otherwise denied, nor it is not possible to replace the deleted API with a new API, this will be taken into consideration and the supplier will not be responsible in this case. b) The Purchaser hereby grants the tenderers as the "Receiving Party" under the signed NDAs the approval to disclose necessary scope of Confidential Information to social services providers in order to receive a commercial calculation necessary to prepare and submit a tender to this tendering procedure. However, the scope of the approval is limited only to (i) information necessarily required to

		prepare and submit the bid and (ii) is limited only to social services providers whose web content is not publicly accessible (to the extent and for the purposes of the calculation of tender price proposal) and a particular contract or other approval for the access to such data is required from the social services provider.
3	Similar problem arises with storing content. Quoting CTK for an example: "CTK - Všechna práva vyhrazena. Určeno jen pro osobní využití. Bez předchozího písemného souhlasu ČTK je zakázána jakákoli další publikace, přetištění nebo distribuce jakéhokoli materiálu nebo části materiálu zveřejněného na těchto stránkách a to včetně šíření prostřednictvím elektronické pošty, SMS zpráv, MMS zpráv a včetně zahrnutí těchto materiálů nebo jejich části do rámců či překopírování do vnitropodnikové či jiné privátní sítě a včetně uchovávání v jakýchkoli databázích." Have you investigated the legal aspects of scraping the sources you envision will likely qualify as our data sources?	The aim of the Purchaser is to find an agreement with some of the mainstream media protected by law or paywall. However, the core of the sources will be disinformation sources. In the case of disinformation or conspiracy outlets, they are normally easily accessible, because they want to spread the disinformation across the society as much as possible. However, similarly as per request no. 3above, if scraping will be legally or otherwise denied, nor it won't be possible to secure the access otherwise, this will be taken into consideration and the supplier will not be responsible for unavailability of such source.
4	Could you elaborate "Constantly updated list of disinformation/information operation sites based on a predefined list of criteria;" z "ANNEX B.1"? How exactly will this list look like? Should a generic scraper working for each news website be developed? Or if users liked to monitor different sites, a new scraper would need to be developed prior to this? Is the system responsible for updating the list of operation sites automatically, or will an admin of the system do this e.g. based on some suggestions from the system?	The Purchaser aims to monitor pages/groups based on their content. A generic scraper is acceptable, if it can scrape the content from the sources. The sources will be various webpages and social media accounts spreading disinformation rather than news websites. Therefore, the crawler should be able to scrape the content also from different websites (some use for instance rss, some different technologies). The list will consist of the websites and the social media accounts (Facebook, Twitter, Instagram, Youtube) based on pre-defined criteria about the content. The system should download the content automatically on pre-defined interval (for example every 15 minutes). However, taking into account the nature of disinformation web structure, the list of web information services cannot be definite as old sites are still disappearing and new ones are emerging. Ideally the HADES tool should be able to update the list automatically based on automatic qualification and evaluation of the web content. If the tenderer's technical proposal of the tool will be able to detect

		potential new disinformation web information services automatically the Purchaser will consider this in evaluation of the tenderer's technical proposal. However, such automatic qualification of the web content is not essential requirement on the tool. If it won't be able to automatically evaluate the content, the list of web information services will be updated by the Purchaser periodically. At the same time the Purchaser does not limit the number of sources to be processed as the tool should be able to include new sources as they will emerge in time. Just for informational and not binding purposes actual list of sources contains approximately 200 webs per country.
5	Could you elaborate on "Mapping functionality" present in "Form F.5"?	Mapping should show, where are which disinformation sources (geographical location), which platform (Facebook, Twitter, Youtube, etc.), how are the sources connected (how do they share documents amongst each other, where does given information origin, etc.), identifying and adding to the system new acteurs.
6	Will the volume of data to download be specified in advance? And related to question 4, is there going to be a final list of sources which are to be monitored and scraped? Will there be quotas for the number of videos of news articles?	The volume of data is the subject to the number of sources and their activity on the internet. Therefore, the Purchaser can not specify the volume of data in advance. For the list of sources please refer to answer no. 4 of this clarification.
7	Do you require the system to link single data instances (messages, posts, news articles) to individual authors? This is -- to our understanding -- not compliant with GDPR	The Purchaser requires the system to be GDPR compliant, therefore the Purchaser aims for the publicly accessible data like for example public Facebook pages. If linking of any data to individual authors should not be compliant with GDPR, such linking shall not be allowed and shall not be a requirement for the tool.
8	Are deadlines, payment phases, handovers negotiable? According to the contract proposal, the first payment is due after six months of work with a risk that Globsec can still terminate the contract after that time.	Some details are negotiable, however there are also hard deadlines like the delivery for instance, which are not negotiable. However, concerning the termination, financial compensation and settlement conditions in case of termination of the contract from the Purchaser shall be negotiable.

9	What is the expected amount of users? What is the expected amount of data scientists? Will the users have a formal contract with Globsec?	More or less 30 users are expected, whereas the number of users can be later higher so this number is not binding or limiting. The users will have formal contracts with Globsec.
10	In "Form F.5", it seems that scoring rationale "1 point for at least one in all projects" in cloud technologies, it seems this formula does not sum to 2 points maximum as expected.	It does, if the mentioned cloud providers were used, you can receive one point for all project. If you have experience even with some other (which was not mentioned as example in the upper) you will receive one additional point.

Please take into account this Clarification of Tender Document.