RFP 31/2025:
Appointment of a service provider for the acquisition of eDiscovery tool

Communication #1

Date of Issue: 16 January 2025

These questions and answers will be posted to all potential bidders.

Questions and Answers:

No	Question	Answer
Deployment Options		
1.	Your RFP document states in the Technical Specifications table, 6.4 "The tool must be able to be hosted On-Prem, running Windows Environment, and using MS SQL as a DB. For On-Prem should be able to be hosted on Cloud." This statement appears to be somewhat contradictory. Would a Software-as-a-Service (cloud based) offering be a viable solution for SARS?	The tool will be hosted on-premises by SARS. However, should SARS wish to host a component in the cloud from SARS's side, the tool must have this capability.

No	Question	Answer
2.	If so, could that cloud-based solution be hosted in a public cloud, or would it need to be in a private cloud?	It must be a private cloud – SARS must control the cloud environment.
3.	Furthermore, are there any data sovereignty issues that apply to SARS and if so, what are they and what is the applicable legislation or guideline?	Taxpayer data under SARS's control is proprietary to SARS, as directed by all legislation administered by the Commissioner of SARS. Additional personal and cybersecurity legislation, such as the Cybercrimes Act, will also apply. The Tax Administration Act and the SARS Act require that all taxpayer information is confidential and must remain so.
4.	The infrastructure to support a system to meet the order of volumes described in your RFP will be significant. If on-prem, have you made provision for such an infrastructure and is it / would it be physical or in a private cloud?	Provision has been and will be made for the volume of data and the required infrastructure.
Volumes		
5	Projected Volumes. We'd like to get more insights to the stated 100TB to 300TB of data per year mentioned in the RFP. Would that entire data volume need to be pushed into the eDiscovery system, or is that measure more an assessment of the total captured data that SARS receives in a year? In our experience, even 100TB would equate to a vast amount of documents, far more than the stated 50 users could ever hope to	The volume of data is increasing per annum from the current 30TB of data, hence the estimation of 100 – 300TB over the term of the contract. The 50 users will be focusing on specific data for the investigations from the volumes of data acquired from SARS enforcement actions.

No	Question	Answer
6	Are you able to give an indication of the type of data (e.g. documents, financial records, databases etc.) that compromises the 100TB to 300TB and the repositories/locations in which it is currently stored?	The data encompasses the types of records specified per the Technical evaluation criteria – Microsoft (Word, Excel, PowerPoint, Outlook), Adobe, accounting records, other e-mail formats, images, cellphone communications, etc. The data acquired is held by the Digital Forensic Investigations per servers that unit has in place and will then be imported into the eDiscovery tool. All the data is currently stored within SARS.
7	How many matters per year do you anticipate and what is the average duration a matter will reside live in the system? We are trying to assess what the maximum concurrent volume of live cases might be per year and in combination with the typical duration, how aggregate volumes might grow over the proposed 5-year term.	Investigation durations vary, but there are an anticipated 10 projects comprising of multiple taxpayers that will be part of the new take on. Historic projects (40) will not be migrated at this time, but data acquired on overlapping focus areas may be brought in to enhance current and new investigations. The duration is estimated to be a minimum of 5 years but may exceed 10 years. The aggregate volumes will result in the 100 – 300TB requirement over the term of the contract.
8.	Related to data volumes and case load, can SARS provide any insights on the typical anticipated use cases for the eDiscovery system?	The data will be utilized to determine the tax liability, address tax disputes (objections and appeals) and any other litigation actions.

No	Question	Answer
		It will also be used for SARS enforcement actions such as tax inquiries, search & seizure and preservation order applications.
Integration with SARS Processes and Systems		
9	The RFP talks of integration/compatibility with tooling such as Magnet AXIOM and Oxygen Forensics. Both are tools which can be used for pre-processing of imaged/captured content. Can we assume that the SARS DFI Unit will perform the pre-processing of collected data, with the resultant (likely significantly lower) volumes pushed into the eDiscovery process?	The DFI will perform the pre-processing utilising the mentioned tools confirmed.
10.	Can you provide any more specific description of the broader integration with SARS systems and processes stipulated in the RFP?	Future integration may include SARS Documentum and IBR systems. Integration with different data types from different sources such as Axiom and Oxygen.
11.	Self-sufficiency. To what extent does SARS anticipate being able to run the entire solution and process independently? Is there scope for a managed service (eDiscovery-as-a-Service) to remove the need for up-skilling within SARS?	Self-sufficiency with support from the service provider on issues such as (training and maintenance of the system).
12.	The RFP talks of real-time financial analytics. Can SARS please provide further insights as to your anticipated requirements. Financial analytics is	The financial data extracted through the tool will be analyzed to determine the tax liability or address the tax disputes phase of the investigation.

No	Question	Answer
	not traditionally considered an eDiscovery capability but is available through either dedicated third party solutions or metadata analysis using Business Intelligence tools such as Microsoft Power BI.	If the data from different documents needs to be compared and differences identified, this will be beneficial. SARS Power BI tools may be used once the data has been extracted.