

Request for Information: Development of a Customized Large Language Model for AQABA - "ASK AQABA"

Question and Answer sheet PoC-4

Q1	How can we participate in the tender? Is the link for the payment not working?
A1	<p>You can access the tender documents directly through the Ministry of Digital Economy and Entrepreneurship's website Tender Link (which has been already included in the announcement article on SNS but it seems that some applicants did not notice it). You can download detailed information such as TOR from this page. Participation is free of charge.</p> <p>If you are interested in participating, please submit your information through the provided Request for Information (RFI) link no later than 2nd Feb 2025. Qualified firms will subsequently receive the Request for Proposal (RFP) for further steps some time in February 2025, and you will usually have 30 days to prepare for the proposal.</p>
Q2	Can we ask for an extension?
A2	No, you cannot.
Q3	Can we submit it as an Emirati company based in UAE?
A3	No, the company must be a local company in Jordan because this PoC is intended to promote local AI industry in Jordan. Please refer to the requirements described in the TOR document.
Q4	Success Metrics: Beyond the outlined KPIs, are there additional metrics JICA considers critical for evaluating this project's impact?
A4	No additional metrics now but might add some at the time of RFP.
Q5	Future Scaling: Are there plans to expand the "Ask Aqaba" prototype to other sectors or regions? Should scalability considerations be detailed in the proposal?
A5	There is no immediate plan for the expansion now, but future scalability must be considered in the design of PoC system.
Q6	Consortium Composition: <ul style="list-style-type: none">• Can a Jordanian company apply as part of a consortium with international companies?• Are there specific requirements for foreign entities, such as percentage of work allocation or representation?
A6	Please refer to "5. Requirements for private IT company" in the TOR.
Q7	Academic AI Researcher: What are the eligibility criteria for the academic AI researcher regarding qualifications and experience?

A7	Refer to “6. Requirements for academic AI researcher” in the TOR.
Q8	Data Security: Are there specific guidelines for data security, storage, and privacy during the PoC phase?
A8	Data security must follow ASEZA’s guidelines as well as Jordanian regulations. In this PoC, we will use publicly available data for the training of AI model, and no confidential data will be used. This does not mean that the raw data used for the training can be exposed to outside of ASEZA. The original data (especially ASEZA’s internal data) should be kept secured, but the information contained in those data can be used for the training (and thus can be a part of answer for user query).
Q9	Hardware and Infrastructure: Should the proposal include hardware specifications and infrastructure requirements, or will ASEZA or JICA provide these?
A9	As outlined in the “Infrastructure Compatibility” section of TOR, the applicant should propose required hardware that is enough for PoC scope and not for full implementation. Based on similar use cases in Japan and other countries, “fine-tuning” of existing LLM requires significantly small GPU resources in both training and inference and that’s the advantage of creating domain-specific LLM by fine-tuning the existing general-purpose LLM. Similar PoCs in Japan normally requires at most only one NVIDIA Tesla A100 and it is more than enough for the purpose of this PoC. Based on the result of this RFI, we may modify the ToR to specify a particular GPU to be used. Please also refer to the answer of Q20.
Q10	Pre-trained Models: Are there preferences for specific pre-trained models (e.g., LLaMA, BLOOM), or should we recommend and justify the choice?
A10	The applicant should propose a pre-trained model with supporting facts and criteria.
Q11	Real-Time Updates: What level of real-time updates and scalability is expected during the PoC?
A11	It is required to implement an “automatic” periodical refinement mechanism for AI model. The frequency of update would be weekly or daily at most. In case the applicant proposes to use RAG method for instance, the mechanism for on-demand retrieval of data from data sources must also be implemented. "Scalability" in this PoC means that the system must be designed to be scalable in the future (after the PoC) in terms of covering other pillars of ASEZA as well as to easily distribute query loads.
Q12	Additional Funding: If the PoC demonstrates exceptional results, are there additional funds available for further expansion?
A12	No additional funding is available within the scope of this PoC.
Q13	Evaluation Criteria: What is the evaluation weightage for the technical proposal, cost proposal, and team qualifications?

A13	<p>Evaluation weight of technical proposal and financial proposal is currently planned as 70% for technical and 30% for financial in upcoming RFP process.</p> <p>Please note that this is the RFI (Request for Information) and you don't need to submit any proposal.</p>
Q14	Submission Format: Are there specific templates or formatting guidelines for the proposal, or is the structure at the company's discretion?
A14	<p>This is RFI and not RFP. All you need to do is to submit information through a designated online form. The format for proposals will be provided when we send RFP to selected applicants based on the result of this RFI.</p>
Q15	We would like to confirm if it is possible to attach documents while filling out the submission form?
A15	<p>At this stage you need to submit your information to the RFI online form only, and no additional attachments will be accepted.</p>
Q16	What is the domain of the requested LLM model? A data sample would also be helpful.
A16	<p>It is written in the TOR document. Tourism and investment. There is no data sample at this stage.</p>
Q17	Which languages should the LLM support?
A17	<p>It is written in the TOR document. English and Arabic.</p>
Q18	Can we use an online LLM model and customize it, or is it required for the LLM to run locally?
A18	<p>It is required for the LLM to run locally. Therefore, only downloadable pre-trained models can be used for this PoC.</p>
Q19	Can a company with one year of experience participate in the RFI?
A19	<p>Yes. We have updated our TOR and eased the requirements for an IT company. A company with at least one (1) year experience can apply for this RFI.</p>
Q20	Does ASEZA already have the required infrastructure (including GPUs) to support model fine-tuning and deployment, or will additional hardware be procured as part of the PoC?
A20	<p>ASEZA already has IT infrastructure including server room and network equipment, but they don't have GPU for fine-tuning the model. The GPU will be purchased as a part of PoC. See the TOR document for the suggested GPU.</p> <p>Note that this is a more research-oriented PoC with a very limited budget and is not a full-scale development of LLM. Therefore, the GPU to be purchased must be bare minimum. See also the answer of Q9.</p>
Q21	Given that fine-tuning an LLM is resource-intensive, has ASEZA considered adopting a Retrieval-Augmented Generation (RAG) approach? This would allow the system to dynamically update and

	retrieve information from an external knowledge base, offering scalable, real-time updates without frequent retraining.
A21	It is up to applicant's proposal whether to use RAG or not as described in the TOR.
Q22	Data Size and Type: What is the estimated data size that ASEZA will provide for training the LLM?
A22	There is no information of estimated data size at this stage.
Q23	Data Size and Type: Are the datasets structured (e.g., tables, JSON) or unstructured (e.g., raw text, PDFs)?
A23	Some data are structured. Others are unstructured.
Q24	Data Cleaning and Preparation: Will ASEZA provide pre-cleaned datasets, or will the Joint Team be responsible for data cleaning and preprocessing?
A24	The Joint Team should be responsible for data cleaning and preprocessing.
Q25	Data Cleaning and Preparation: Are any labeled datasets available, or will the Joint Team need to label data as part of the process?
A25	No labeled dataset is available. The Joint Team should do the labeling (if required).
Q26	Data Cleaning and Preparation: Who will verify the correctness and quality of the datasets (ASEZA or the Joint Team)?
A26	Primarily the Joint Team should verify the correctness, but ASEZA will also review the data.
Q27	Language: Should the LLM prioritize specific dialects of Arabic in addition to Modern Standard Arabic?
A27	Only the Modern Standard Arabic should be prioritized in this PoC.
Q28	Language: Should the LLM be designed to support both Arabic and English languages?
A28	Yes.
Q29	Responsibility for Infrastructure:

	Will ASEZA or JICA provide the infrastructure for LLM training, or is the Joint Team expected to manage and fund it independently?
A29	See the answer of Q9 and Q20.
Q30	<p>Existing or Planned Infrastructure:</p> <p>Does ASEZA or JICA already have the infrastructure available for training the LLM? If yes:</p> <p>i. What are the hardware specifications (e.g., type of GPUs, amount of RAM, storage capacity)?</p> <p>ii. Where is the infrastructure located (onsite in Aqaba or remotely accessible)?</p> <p>iii. If new infrastructure will be procured, will it be set up onsite in Aqaba or provided through cloud services or VPN?</p>
A30	See the answer of Q9 and Q20. All hardware will be onsite at ASEZA.
Q31	<p>Research Opportunities Post-Completion</p> <p>a. After completing the project and analyzing the findings, can we submit a research paper based on the outcomes?</p> <p>b. If yes, will there be any restrictions on the choice of journals, conferences, or other publication platforms?</p>
A31	Yes, you can submit a research paper based on the outcome of PoC, but you must consult ASEZA for any disclosure of data or system information because it might affect the information security of ASEZA. There is no restriction on the choice of journals at this time.
Q32	<p>Team Presence and Logistics:</p> <p>a. Will the Joint Team be required to conduct onsite meetings or workshops in Aqaba?</p> <p>b. Are there specific travel or accommodation arrangements for onsite team members, or should the Joint Team budget for this independently?</p>
A32	Most meetings can be done online. Onsite visits would be required for deployment of the PoC system and knowledge transfer training for ASEZA staff members.
Q33	<p>Knowledge Transfer:</p> <p>Will the Joint Team be required to provide post-training technical support, and if so, for how long?</p>
A33	Refer to “2.3.6. Continuous Update and Ongoing Maintenance” in the TOR.
Q34	Budget and Procurement:

	<p>a. Will ASEZA or JICA provide funding for additional tools, cloud services, or hardware required for the PoC?</p> <p>b. Are there predefined budget constraints or guidelines for procurement?</p>
A34	JICA provides funding for expenses for developing the LLM but will not cover the cost of hardware (such as development PC, server, network equipment, etc.) or software tools to develop the LLM. The only available funding for hardware is a GPU. Server infrastructure will be prepared by ASEZA. There is a predefined budget cap in this PoC which cannot be disclosed according to JICA procurement rule.
Q35	<p>Testing Environment:</p> <p>Will ASEZA provide a dedicated testing environment for the LLM (e.g., staging servers)?</p>
A35	It would be negotiable with ASEZA.
Q36	<p>Testing Environment:</p> <p>Are there specific ASEZA systems (e.g., web portals, APIs) that the LLM must integrate with?</p>
A36	Integration to ASEZA's web portal is mandatory.
Q37	<p>Public Testing:</p> <p>If public testing is part of the PoC, who will manage and monitor user feedback? Will ASEZA or the Joint Team handle this?</p>
A37	The Joint Team as well as ASEZA should manage and monitor user feedback.
Q38	<p>Onsite Support:</p> <p>Will ASEZA staff or resources be available onsite in Aqaba to provide technical and operational support during the project</p>
A38	ASEZA IT team is doing regular technical support for their infrastructure, but the technical and operational support must be done by the Joint Team.
Q39	<p>Remote or Hybrid Work Support Options:</p> <p>Can the Joint Teamwork entirely remotely during the project, or will onsite visits be required?</p>
A39	See the answer of Q32.

Q40	Regarding software infrastructure and compatibility, is the Joint Team allowed to use any mathematical framework, e.g., JAX vs TensorFlow vs PyTorch, or would ASEZA staff prefer a specific one?
A40	Yes. The Joint Team can use any framework.
Q41	Data-curation: The ToR suggests that the PoC should train from existing documents without collecting data from humans. However, would we be allowed to ask university students register web-pages (for crawl starting-points) and give high-level summaries — in Arabic and in English? This could be useful for fine-tuning LLMs for summary-generation. If ASEZA desires, we can share this data among competitors submitting PoCs.
A41	Yes. It is up to the proposal of applicant.
Q42	Citing generated-sentences. While the ToR asks for quantifying misinformation (via offline metrics on evaluation sets), we additionally inquire if we should cite every generated text-piece, e.g., show a “cit” link that presents the evidence in quotations with a link to the source. Knowing this upfront may entirely change our direction.
A42	It is recommended to add “cit” link to the generated text.
Q43	Sample Query Patterns. While ASEZA is willing to provide documents as training data (containing information about tourism, economical, legal, operational, etc.), will ASEZA share a list of realistic user queries? We can think of some (e.g., regarding tourism, investments), but we trust that ASEZA better understands the intended audience. Only the queries suffice — even without example responses.
A43	It could be negotiable with ASEZA to provide such user queries in the testing phase.
Q44	Are we allowed to publish our work in Machine Learning Venues? If so, are we allowed to collaborate with ASEZA staff for steering, or are we expected to work independently on scientific advancement while ASEZA in the engagement prioritizes solution-delivery (reports and handover).
A44	It is allowed to publish an article or a paper regarding the methodologies used in the work of PoC, but it is not allowed to publish the created LLM. The applicant should discuss with ASEZA on how to collaborate for the publishing. See also the answer of Q31.
Q45	Model Specifications 1. Should the model be limited to text processing, or should a vision branch be integrated (e.g., for processing images or geospatial data)?
A45	This PoC is primarily focused on text processing. However, some of the data may be in PDF format or contain images. To maximize the utility of this data, it could be beneficial (but optional) to integrate some (OCR) techniques to extract and incorporate relevant text into the training process. Scanned document of hand-written text is not covered in the scope of this PoC.

Q46	Model Specifications 2. Does ASEZA envision additional functionalities or features for the LLM beyond typical question answering, such as advanced reasoning, sentiment analysis, or text summarization?
A46	This PoC focuses on typical question answering only.
Q47	Data Clarifications 3. Can you provide a detailed overview of ASEZA's existing data assets, including: a. Types of data (e.g., webpages, PDF documents, structured databases)? b. Approximate volume of the data (e.g., number of characters, records, or file sizes)? c. Formats and quality of the data (e.g., structured, semi-structured, or unstructured)?
A47	No information at this point.
Q48	Data Clarifications 4. Does the data include any non-textual elements, such as images, geospatial datasets, or multimedia files?
A48	Some existing data at ASEZA may include non-textual elements. Also see the answer of Q45.
Q49	Data Clarifications 5. Are there specific tools or standards ASEZA uses for data management and versioning?
A49	No.
Q50	Queries and Use Cases 6. Is there an existing set of common user queries? If so: a. What is the approximate size of this set (number of queries)? b. Are these queries categorized or labeled by type (e.g., tourism, investment)?
A50	No.
Q51	Queries and Use Cases 7. Should the LLM prioritize answering predefined queries, or should it aim for open ended conversational capabilities?
A51	This PoC prioritizes open ended conversational capabilities.
Q52	Infrastructure and Deployment

	8. Should the trained LLM be hosted on ASEZA’s local servers, or is a cloud-based hosting solution acceptable?
A52	The LLM should be hosted on ASEZA’s local server.
Q53	Infrastructure and Deployment If the model is hosted locally, does ASEZA have the necessary hardware (e.g., GPUs such as NVIDIA A100) to support training and inference? If additional resources are required, would the cost be covered by our team?
A53	See the answer of Q9 and Q20.
Q54	Infrastructure and Deployment Are there any restrictions on data storage or model access due to security, compliance, or operational policies? Specifically, is it permissible to transfer ASEZA's data to our own AI infrastructure (servers) for training the LLM model?
A54	It might be possible if the data security is fully ensured, and the applicant signs NDA.
Q55	Training and Scalability Should the model incorporate real-time data streams (e.g., RSS feeds or APIs) as part of its training pipeline?
A55	Yes. Refer to “2.3.3 Expected Inputs (data and source) to the PoC Solution” in the TOR.
Q56	Training and Scalability What are the expected scalability requirements post-PoC, in terms of: a. User volume and query frequency? b. Data updates and retraining intervals?
A56	Post-PoC requirements are out of scope of this PoC.
Q57	Testing and Validation Are there specific benchmarks or ASEZA testsets uses to evaluate the model's performance during testing?
A57	No. There is no existing test sets.
Q58	Testing and Validation How will user feedback be collected, and should it be integrated into the model refinement process in real-time or after predefined intervals?

A58	As described in Figure 3 of the TOR, user feedback should be collected through the Web interface of LLM. The feedback will not be integrated automatically (as a training data) to the LLM, and will be used by the Joint Team and ASEZA for the reference of further improvement.
Q59	Is there any requirements related to work location? Onsite or remote
A59	No requirements. The Joint Team could basically work remotely.
Q60	What are the POC project payments terms and milestones
A60	Basically, the payment will be at the completion of the PoC. There is no advance payment. The detailed payment conditions will be described in RFP tender documents after this RFI.
Q61	Who will be pay the costs hardware? GBU's, servers, storage etc
A61	Cost of required GPU must be included in the quotation of RFP. Servers and other infrastructure should be borne by ASEZA.
Q62	Scope Is the project scope limited to Aqaba only?
A62	Yes for the PoC.
Q63	Scope What technology will be used to collect public data?
A63	Refer to "2.3.3 Expected Inputs (data and source) to the PoC Solution" in the TOR.
Q64	Scope What will be the system's primary language and the language used for data gathering?
A64	English and Arabic.
Q65	Scope Will the 8-month project timeline include data gathering, or does it cover both data collection and system development?
A65	8-month is for both data collection and system development.
Q66	Scope What is the token scale of the system?
A66	The applicant should propose it. Note that this is a PoC and is not a full development project.
Q67	Scope What is the inference capacity?
A67	The applicant should propose it. Note that this is a PoC and is not a full development project.
Q68	Scope Is front-end development within the project scope?

A68	Yes. See Figure 3 of the TOR.
Q69	Scope What is the data ownership policy (private or public)?
A69	Refer to “Authority Matrix for Data” described in “2.3.3 Expected Inputs (data and source) to the PoC Solution” of the TOR.
Q70	Scope Does the scope include data cleansing and preparation activities?
A70	Yes.
Q71	Scope Will security-related activities, such as penetration testing (PT) and implementing security standards, be part of the scope?
A71	Yes.
Q72	Scope Is there a requirement to build a mobile application?
A72	No, but the Web interface must adapt to mobile devices.
Q73	Scope Are there any specific systems, platforms, or tools that the project needs to integrate with?
A73	ASEZA’s Web site. API-based Web page embedding should be supported by the system as described in the TOR.
Q74	Data Where will the data be stored (cloud database or on-premise)?
A74	On-premise at ASEZA’s local server.
Q75	Data If the AI system is integrated into the ASEZA website, how will data be categorized and managed (public data, private data, government data)?
A75	Refer to “Authority Matrix for Data” described in “2.3.3 Expected Inputs (data and source) to the PoC Solution” of the TOR.
Q76	Data Is the ASEZA website compatible with AI system integration?
A76	Only the LLM’s web interface will be integrated with ASEZA website. LLM should run on a separate server.
Q77	Data Does ASEZA have the capacity for knowledge transfer?
A77	The person in charge of the PoC should be the knowledgeable person.
Q78	Data Where will the platform be hosted?
A78	At ASEZA’s server room.

Q79	Data What specific sources will be used for regular data collection to keep the model updated?
A79	Specific sources of regular data collection should be discussed and agreed with ASEZA.
Q80	Data Can you provide more specific details about the "non-sensitive data" from the Tourism and Investment directorates that will be used for training the LLM?
A80	"Non-sensitive data" means data with no personal information or sensitive content.
Q81	Publishing Who will have access to the system (public access, restricted to whitelisted users, etc.)?
A81	Public access (for testing only). Note that the system will not be officially published within the scope of PoC and all access is granted with clear indication of disclaimer for prototype.
Q82	Publishing Which stakeholders will be involved in the project (technical teams and data gathering contributors)?
A82	ASEZA, AI Steering Committee of MoDEE, JICA expert team, JICA project secretariat. There may be other optional stakeholders.
Q83	Implementation Considerations Can the required company age be reduced to 1 year?
A83	Yes. We have reduced the company's experience requirement to 1 year.
Q84	Implementation Considerations Are you open to collaboration with FANAR?
A84	No. All stakeholders must be local in Jordan.
Q85	Implementation Considerations Will the client's existing hardware scale effectively, or will the system be cloud-based?
A85	Client's existing hardware is not scalable. Note that the system should be designed to be scalable in the future but it does not mean that the ASEZA's existing infrastructure is scalable. Within the scope of this PoC, the LLM will run on a single server. But the LLM must be designed to be scalable if ASEZA will expand the use of LLM after the PoC.
Q86	Implementation Considerations Can the infrastructure handle deploying and testing multiple models simultaneously?
A86	No, but it might be negotiable with ASEZA if they have spare capacity.
Q87	Implementation Considerations What criteria will determine when to use a knowledge base, retrain the model, or adopt a retrieval-augmented generation (RAG) approach?
A87	The applicant should propose these.
Q88	Implementation Considerations Do you have a preferred open-source pretrained LLM (e.g., Gemini, LLaMA, QWEN, BLOOM, JAIS, Aalam)?

A88	No. The applicant should propose one.
Q89	Implementation Considerations Do you prefer starting implementation with one pillar and gradually expanding, or using a RAG approach from the start?
A89	No preference. The applicant should propose the methodology.
Q90	Performance and Benchmarking What are the benchmarks and targets needed to measure the model's maturity and acceptance?
A90	Refer to "2.4 Key Performance Indicators (KPIs)" in the TOR.
Q91	Performance and Benchmarking What benchmarking techniques will ensure output integrity?
A91	Refer to "2.4 Key Performance Indicators (KPIs)" in the TOR.
Q92	Performance and Benchmarking What additional metrics, beyond the KPIs listed in the tender, should be adopted?
A92	Refer to "2.4.1 Additional KPIs for Addressing Specific Risks" in the TOR.
Q93	Maintenance and Updates What is the preferred schedule for routine data collection and updates to keep the LLM content current?
A93	Weekly or daily is preferred in this PoC.
Q94	Maintenance and Updates How will the dedicated team from ASEZA evaluate user satisfaction and response consistency?
A94	The dedicated team from ASEZA will test the LLM and give scores.
Q95	From the statement: "The selected model must also be compatible with ASEZA's current servers." Could you provide details about the capabilities and specifications of the servers currently available at ASEZA?
A95	No information can be shared at this RFI stage. Please also see Q9 and Q20.
Q96	Will the model's responses be tailored or restricted based on user roles (e.g. ASEZA employees versus other users)? If so, could you elaborate on how role-based restrictions will be implemented?
A96	Role-based restriction is not needed in this PoC.
Q97	From the statement: "Initial Setup: An initial research phase should compare the available models using practical testing with some real data from ASEZA to evaluate their accuracy and performance." How many open-source LLMs are expected to be included in the initial research and evaluation phase?
A97	A few (at least three).
Q98	From the statement: "Initial Setup: An initial research phase should compare the available models using practical testing with some real data from ASEZA to evaluate their accuracy and performance." What specific criteria will be used to select the best LLM for completing the project?

A98	The criteria will be determined later at RFP stage.
Q99	From the statement: "Initial Setup: An initial research phase should compare the available models using practical testing with some real data from ASEZA to evaluate their accuracy and performance." What specific criteria will be used to select the best LLM for completing the project?
A99	The criteria will be determined later at RFP stage.
Q100	From the statement: "Figure 2: Overview of the system to develop in this PoC." In the system overview diagram, 'etc.' is used to describe various data sources and types (e.g., public data, internal data). Could you clarify the intended scope and limits of 'etc.' for each category of data
A100	"etc." means any additional data that all parties (including the Joint Team) agree to add.
Q101	Please your kindly support to accept our request for 2-weeks extension for submission deadline.
A101	No. Note that this is RFI (Request for Information) and is not RFP. So you don't have to submit any proposal at this point. Just submit your information through the provided online form.
Q102	What are the primary success criteria for this PoC program?
A102	Refer to "2.4 Key Performance Indicators (KPIs)" in the TOR.
Q103	How will the scalability and extendability of the LLM prototype be evaluated?
A103	By technical adaptability to horizontal scalability (such as easiness of adding a computing node, etc.). Also refer to "2.3.7 Scalability Risk Plan and Mitigation Strategies" in the TOR.
Q104	What are the exact datasets ASEZA will provide for training the model?
A104	No detailed information at this stage of RFI.
Q105	Are there any specific KPIs that the solution must achieve?
A105	Refer to "2.4 Key Performance Indicators (KPIs)" in the TOR.
Q106	Will ASEZA provide any support in gathering publicly available data or should the Joint Team handle this independently?
A106	The Joint Team should handle it based on discussion with ASEZA.
Q107	Are there specific open-source models that ASEZA or JICA prefers to use? Can we request additional clarification workshops or meetings during the project if necessary?
A107	No preferences. Clarification meetings can be requested during the PoC program.
Q108	Data Availability and Access: 1. What specific data assets within ASEZA will be made available for training the LLM? (e.g., databases, documents, reports, website content)
A108	No detailed information at this stage of RFI.
Q109	Data Availability and Access: 2. How will access to this data be provided? (e.g., APIs, database dumps, direct access

A109	Mostly via database dumps and direct access.
Q110	Data Availability and Access: 3. Are there any data privacy or security concerns that need to be considered?
A110	See the answer of Q8.
Q111	The "readiness" of the data 4. Does ASEZA have a comprehensive inventory of its data assets, including metadata describing the data format, structure, and content? Can this inventory be shared?
A111	No, they don't have them.
Q112	The "readiness" of the data 5. Can ASEZA provide examples of the different data formats used (e.g., CSV, JSON, databases)?
A112	Might be possible at the stage of RFP.
Q113	The "readiness" of the data 6. Are there consistent schemas or data dictionaries available?
A113	No.
Q114	The "readiness" of the data 7. How easily accessible is the data? Are there existing APIs, databases, or data warehouses that can be used?
A114	Some databases exist but others are raw data (Excel document, PDF document, etc.).
Q115	The "readiness" of the data 8. What level of support and collaboration can the Joint Team expect from ASEZA in terms of data access, clarification, and validation during the data preparation phase?
A115	Basic support and collaboration only.
Q116	User Experience Enhancement 9. Who are the target users of the LLM? (e.g., tourists, investors, ASEZA staff)
A116	The primary user is ASEZA staff. The secondary user is general public who visited ASEZA web site (for testing purpose only).
Q117	User Experience Enhancement 10. Which specific tasks or processes are expected to be automated by the LLM?
A117	The applicant should propose the mechanism for the automation.
Q118	Tourism and Investment Focus 11. Can ASEZA provide specific examples of the types of queries the LLM should be able to handle within the tourism and investment domains?
A118	No. At least not at this time, maybe in later implementation phases.
Q119	Tourism and Investment Focus

	12. Are there any existing knowledge bases or FAQs that can be used as a starting point for training the LLM?
A119	No information at this stage.
Q120	Data & Model Update Schedule: 13. How will ASEZA prioritize data sources for updates based on criticality and time-sensitivity (e.g., daily updates for tourism data)?
A120	The applicant should determine the priority by discussion with ASEZA.
Q121	Scalability and Extendability 14. What are the criteria for determining the "success" of the PoC in terms of scalability and extendability?
A121	See answer of Q103.
Q122	Prototype vs. Production System 15. What are the plans for transitioning the prototype to a production system after the PoC is complete? What are the success metrics for the PoC that would lead to a full implementation?
A122	IMPORTANT: There is no immediate plan for transitioning to full implementation for now. This is <u>an independent PoC program</u> within the scope of JICA project, and it might or might not lead to future full-implementation if the PoC achieves satisfactory result AND the ASEZA makes decision to proceed with it.
Q123	Infrastructure Feasibility 16. Do the current servers in ASEZA support the addition of new GPUs, or are there any hardware limitations that need to be addressed?
A123	No information at this stage.
Q124	Infrastructure Feasibility 17. Can ASEZA provide more details about their current IT infrastructure? (e.g., servers, cloud resources, network bandwidth)
A124	Not now at RFI but probably at RFP.
Q125	Infrastructure Feasibility 18. Are there any specific infrastructure requirements or constraints that need to be considered?
A125	No information at this stage.
Q126	Hardware Requirements 19. Can ASEZA elaborate on the current server specifications (CPU, RAM) to assess compatibility with the proposed GPU (NVIDIA Tesla A100)?
A126	Not now at RFI but probably at RFP.
Q127	Hardware Requirements

	20. Will the final prototype be deployed on ASEZA's existing infrastructure, or will a <i>dedicated</i> new server (physical or virtual) be provided for the PoC?
A127	Possibly on a new server.
Q128	Hardware Requirements 21. If a new server is to be provided, who is responsible for procuring, configuring, and maintaining it?
A128	ASEZA
Q129	Hardware Requirements 22. Will ASEZA cover the cost of any additional hardware required for training and inference during the PoC program?
A129	JICA will cover the cost of GPU. See also Q34.
Q130	Consortium and Joint Team 23. Will JICA provide a template or guidelines for the consortium agreement between the private IT company and the academic AI researcher? This is crucial for defining roles, responsibilities, and intellectual property ownership within the team.
A130	No. It's up to the agreement among the Joint Team. Please refer to Table 1 of TOR for expected roles.
Q131	JICA Procurement Rules: 24. Can you provide access to the relevant JICA procurement rules?
A131	There are multiple volumes of the rule, and the most relevant one is here. https://www.jica.go.jp/about/announce/manual/guideline/consultant/entrust.html Note that the document is written in Japanese language.
Q132	Access Control and Authentication: 25. Who is responsible for managing user access to the LLM and implementing appropriate authentication mechanisms?
A132	Authentication mechanism is not required for this PoC program.
Q133	Cloud Deployment (Alternative): 26. Is cloud deployment (e.g., AWS, Azure, Google Cloud) an option for the PoC? If so, which cloud provider is preferred, and who is responsible for managing the cloud resources?
A133	Training of the LLM could be done on cloud based on the agreement with ASEZA, but the trained LLM must be deployed and hosted locally at ASEZA.