

Government of West Bengal
Finance Department
Nabanna, Howrah

Memo No. 3554-F(e-Gov)

Date: 25.09.2025

Memorandum

Sub: Addendum to the Expression of Interest (EOI No. 2830-F(e-Gov) dated 30.07.2025) [Ref. No. EOI/2025/001/1239311] for Selection of Firms for Development of AI-Based Solutions

The following details may be noted by interested bidders in connection with the Expression of Interest (EOI No. 2830-F(e-Gov) dated 30.07.2025) [Ref. No. EOI/2025/001/1239311] for Selection of Firms for Development of AI-Based Solutions.

1. Clarifications of Queries:

The detailed clarifications to the queries raised have been annexed with this Memorandum for ready reference.

2. Single Point of Contact (SPoC):

The following senior officers have been designated as **Single Point of Contact (SPOC)** for the respective Use Cases as mentioned against their name to facilitate communication, coordination, and query resolution of the participating firms:

Use Case No.	Title	Name	Email ID
Use Case 1	Enterprise AI Model for the Scientific Framework of Valuation Assessment	Shri Anup Kumar Mandal , Joint Commissioner of Revenue, Directorate of Registration & Stamp Revenue, Finance Department	acsrberac.igr-wb@nic.in
Use Case 2	Digitization of old Registration Records using AI and use of AI for chain deed preparation for bringing registration records up-to-date for all properties	Shri Abhijit Bera , Joint Commissioner of Revenue, Directorate of Registration & Stamp Revenue, Finance Department	acsrperac.igr-wb@nic.in
Use Case 3	Establishing a Unified Social Registry for enhanced delivery of public welfare services	Shri Atanu Majumdar , IAS, OSD & ex-officio Secretary, Finance Department	majumder.ctax@nic.in
Use Case 4	Finance Department Enterprise AI Model for file analysis and decision making.	Shri Panicker Harishanker , IAS, Special Secretary, Finance Department	panicker.h@ias.nic.in

Note 1: The SPoCs may be contacted for their specific Use Cases after an interested bidder has qualified for PoC development, i.e. by participating firms. "Participating Firms" are those interested bidders whose Bids have been evaluated by the Technical Bid Evaluation Committee and the Finance Department has confirmed that they have qualified for PoC development of specific Use Case(s).

Note 2: Emails from participating firms addressed to the respective SPoCs may also be copied to aieoifd2025@gmail.com, pawan.kadyan@ias.nic.in, and randhir.ias@gov.in for necessary information of the Finance Department.

3. Timelines:

- **Bid Submission deadline:** Firms may apply for one or more use cases as per their expertise. Intending bidders may submit their bid on any working day **between 11:00 AM to 5:00 PM up to 28th November, 2025 (Friday) in hard copy in a Sealed Envelope** clearly marked "EOI for AI Use Case [Sl. No. and Name of the Use Case(s)]" and submit it at **Room No. 1210, 12th Floor, Finance Department, Nabanna** as per the details mentioned under Section 10 of the EOI.
- **Bid Evaluation:** The Technical Bid Evaluation Committee designated by the Finance Department shall evaluate the Qualification Criteria of a Bid within 7 working days after the receipt of a Bid in hard copy in Sealed Envelope addressed to the Finance Department as per details mentioned in the EOI.
- **Commencement of PoC development:** After completion of evaluation of a Bid by the Technical Bid Evaluation Committee, the bidder will be informed the result of such Bid Evaluation through email from aieoifd2025@gmail.com. A bidder, if qualified, will be informed to participate in the PoC. The email id "aieoifd2025@gmail.com" may be whitelisted/ marked as trusted by the respective bidders so that they can receive the email update. The qualified bidders i.e. 'participating firms' are expected to commence the PoC development without delay. All participating firms would be required to submit their PoC's result within the Deadline of submission of PoC results.
- **Deadline for submission of PoC results:** All participating firms shall be required to share their PoC's results with the Finance Department within **30th December, 2025 (Tuesday)** irrespective of the date on which they may had submitted their bids. The manner of submission of the PoC's results will be informed to the qualified bidders separately.

4. Mandatory Supporting Documents for ascertaining Qualification Criteria of the EOI:

The participating firms are may note that **submission of appropriate supporting documents is mandatory** for establishing eligibility under the Qualification Criteria as mentioned in the EOI.

The following documents must be enclosed with the bid submission by a participating firm:

Criteria		Document Required
Legal Status	Must be a registered entity (company, partnership, or consortium) with at least 2 years of operation.	Trade License / Certificate of Incorporation / Partnership Deed.

			In case of consortium, all companies forming the consortium will need to provide their individual Certificates of Incorporation / Trade Licenses. etc
	Valid GST registration		Copy of GST Registration Certificate
	Valid PAN		Copy of PAN Card
Technical Expertise :	Minimum of 3 completed projects in AI/ML, preferably in the public sector or government domain	o For Use Case 1: Experience in geospatial analytics, econometrics, or valuation models. o For Use Case 2: Expertise in OCR, NLP, computer vision, or graph neural networks. For Use Case 3: Experience in federated databases, data virtualization, or DBT systems. 1. Expertise in AI/ML-Based Beneficiary Analytics and Dynamic Eligibility Engines 2. Experience with Geo-Spatial Intelligence and GIS-AI Integration 3. Expertise in AI/NLP for Data Interoperability and Semantic Mapping 4. Proven Capability in Data Harmonization and De-Duplication across Heterogeneous Systems 5. Hands-on Experience in Grievance Analytics and Citizen Feedback System 6. Demonstrated Knowledge of Data Privacy, Security Protocols, and Regulatory Compliance o For Use Case 4: Expertise in knowledge management systems or enterprise AI and enterprise GPT solutions	Relevant Work Orders / Completion Certificates
Financial Capability:	Average annual turnover of INR 5 crore or more for IT & ITeS related activities over the last 2		Firms should provide CA-certified turnover certificates and ITR

	Financial Years.		documents of the ITRs filed for respective last 2 Financial Years
Team Composition:	Dedicated team with expertise in AI/ML, data science, and domain-specific knowledge for the applied use case(s).		Self-certified CVs of key team members + supporting evidence (employment letters, project experience letters)
	At least one Project Manager with 10+ years of experience in AI/ML projects		
Compliance:	No blacklisting by any government or public sector entity in the last 2 years.		Self-certified declaration on company letterhead
	Cybersecurity requirements should align with ISO 27001/CERT-IN standards, particularly for federated systems		Valid ISO 27001 / CERT-IN certificate copy

All interested bidders are hereby informed that any further queries, clarifications or requests for additional information, and communications relating to the Expression of Interest (EOI) may be communicated to Finance Department through email to aieoifd2025@gmail.com with a copy to pawan.kadyan@ias.nic.in and randhir.ias@gov.in.

This Addendum shall be treated as an integral part of the aforementioned EOI. All interested bidders, including those who have sought pre-bid clarifications and others, are advised to take note of this Addendum and make their submissions accordingly.


 Secretary
 Finance Department
 Government of West Bengal

Annexure to Finance Department Memo No. 3554-F(eGov) dated 25.09.2025

Sl. No.	Heading	Category/ Reference (if any)	Queries	Replies
1		Average annual turnover of INR 5 crore or more (IT & ITeS related activities) for the last 2 financial years (CA-certified).	As per the Government of India's Startup-friendly procurement norms (Manual for Procurement of Goods, 2022 & 2024), DPIIT-recognised and State-supported Startups are eligible for certain relaxations, including: - Exemption from prior turnover and experience requirements, - Exemption from submission of Earnest Money Deposit (EMD), and - Flexibility in technical evaluation to promote innovation. Accordingly, we respectfully request an exemption from the following eligibility criteria: - Average Annual Turnover of INR 5 crore or more (for IT & ITeS-related activities) for the last two financial years,	There is no EMD or Financial Bid sought in the PoC. The intending bidder may please examine the EOI in more detail. The details related to relaxations for Startups will be informed later, if any.
2		Bid Submission deadline: Firms may apply for one or more use cases as per their expertise. Intending bidders may submit their bids up to 90days from the date of publication of this EOI.	Should we consider the deadline as Oct 30th?	The EOI was published on 30th July 2025. The last date for submission of Bids has been extended to 28th November 2025, has been mentioned under the sub-heading "3. Timelines" in this Addendum along with other details and deadlines. Late submissions will not be considered. The bidders will be informed any updates, if any, through the Finance Department website.
3		No blacklisting by any government or public sector entity in the last 2 years.	As per current industry standards followed in most RFPs and tenders floated by government and public sector entities, compliance with non-blacklisting or non-debarment criteria is typically established through a self-declaration submitted on the bidder's letterhead, duly signed by an authorized signatory. Will such a self-declaration be accepted as sufficient proof for meeting this qualification criterion?	Yes. For the purpose of this EOI, compliance with the non-blacklisting or non-debarment requirement, as specified under Section 6(c) – Compliance, may be established through a self-declaration on the bidder's official letterhead, duly signed and stamped by an authorized signatory of the bidding entity. The Finance Department reserves the right to seek additional supporting information or carry out independent verification if deemed necessary.
4			As per standard government procurement practices, two primary evaluation methodologies are commonly followed: - Quality and Cost-Based Selection (QCBS) - Least Cost Selection (LCS) In this context, we seek clarification on the following: - Is this EOI being evaluated under QCBS or LCS methodology? - If QCBS is being adopted, kindly specify the marking scheme, including the weightage assigned to Pre-Qualification (PQ), Technical Qualification (TQ), and Financial Evaluation. If LCS is being followed, please confirm that only technically qualified bidders will be considered and that the lowest financial bid will be the sole basis for final selection.	This stage is an Expression of Interest (EOI) and does not seek a financial bid. The current evaluation is limited to qualification and technical capability assessment for shortlisting firms to undertake the Proof of Concept (PoC) development, as outlined in Section 5 (Selection Process) and Section 9 (PoC Evaluation). The question of whether evaluation is under QCBS or LCS does not arise in this EOI's PoC evaluation.
5			Is there a specific template to be used for preparing our bid document?	No specific bid template has been prescribed under this EOI.
6			If no template exists, please confirm the complete list of information and documents to be included in the submission.	The details of mandatory documents are mentioned in this Addendum under the sub-heading "4. Mandatory Supporting Documents for ascertaining Qualification Criteria of the EOI". Bidders must ensure that their submission covers all requirements listed in the EOI, particularly under: Section 5 (Selection Process) – Bid submission requirements, including company profile, relevant experience, approach and methodology, team composition, CVs of key personnel, and financial statements. Section 6 (Qualification Criteria) – Eligibility, technical expertise, financial capability, team composition, and compliance requirements. Bids must be submitted in hard copy, in sealed envelopes, clearly marked "EOI for AI Use Case [Sl. No. and Name of the Use Case(s)]", and addressed to the Secretary, Finance Department, as per the details provided in Section 10 (Contact Information).
7			At this stage, is there an estimate for the following for each use case: Volume of data to be processed Number of documents involved Expected number of end-users Number and type of system integrations required for final implementation	For PoC development, the volume of data, documents, etc. would depend on the robustness of the Model deployed and results achieved. Data and documents are available in sufficiently large nos. How much is required for a participant for developing a PoC is a subjective question and cannot be objectively answered at this stage. Further specific details will be discussed with participating firms during PoC development. The no. of end-users during PoC will be very limited only, say not more than 10 per use case. Any integrations with existing systems will be through standard APIs only, and will not be during PoC development stage.
8			Is there a rollout plan for the full solutions post-PoC?	In principle, yes. The firms selected through this EOI will develop PoCs for the use cases (as defined in Section 2 of the EOI). If the PoCs developed meet the desired criteria for full-fledged rollout then the Finance Department may onboard firms who qualify in such PoC development to rollout these full-fledged AI solutions for the Finance Department, Government of West Bengal.
9			Any indication of overall timelines and departments to be covered would be useful for planning.	As per the EOI (Section 6 – Timelines), the indicative timelines are as follows: Pre-bid Meeting: 11th August 2025 at 3:00 PM IST, Finance Department, Nabanna, Howrah. Bid Submission Deadline and PoC Submission Deadline are mentioned in this Addendum under the sub-heading "3. Timelines". Commencement of PoC Development: Immediately upon notification of qualification in the bid evaluation. Regarding departments to be covered: Each Use Case specifies the concerned Government entities. Separate senior officers designated as Single Point of Contact (SPOC) for each Use Case are mentioned in this Addendum separately. They would be coordinating with required Departments internally within the State Government for your convenience. For instance, Use Case 3 (Unified Social Registry) may involve coordination with different scheme-administering departments such as Women & Child Development, Agriculture, Food & Supplies, Backward Classes Welfare, School Education, etc., as outlined in the EOI. For other use cases, the primary coverage and possible interactions with other departments as required will also be done for any data or validation requirements by the SPOC.
10			Can a firm apply jointly with another entity (e.g., as a technology + domain partnership)?	Yes, Consortium Bids are permitted. As per the eligibility criteria, the bidder may be a consortium with at least 2 years of operations of its consortium members. In case of a consortium bid, relevant documents and experience of each partner should be furnished. The following details may be noted for Consortium Bids: - Consortium as a whole shall have to meet the eligibility criteria and turnover requirements specified in the EOI. - The Lead Member of the Consortium should have to have at least 50% of the minimum turnover as specified in the EOI. - A Consortium Bid may use the credentials of its Associates for the purpose of qualification under this EOI, provided however that the Consortium Bid shall ensure that such Associate continues to remain an Associate throughout the Contract Period.
11			Since the PoC development cost is to be borne entirely by the bidder, will there be any reimbursement or incentives provided to firms whose PoCs are shortlisted or accepted?	As per Section 7(2), firms shall bear all costs for PoC development. The EOI does not provide for reimbursement or incentives at the PoC stage. Participation in the PoC will be acknowledged officially for all selected firms. However, after PoC development how many firms qualify is not possible to be known at this stage and is also beyond the scope of this EOI. However, the PoC development is akin to a technical bid before being considered for full-fledged rollout. A full-fledged rollout would invariably seek submission of financial bids from firms that are successful in PoC development.

12		Are there any clauses regarding formation of Consortium / SPV ?	1. Consortium Formation a. Maximum three (3) members permitted. b. Members may form a Special Purpose Vehicle (SPV) for project execution post-award. 2. Lead Member a. One member designated as Lead Member. b. Lead Member shall hold minimum 51% shareholding / participation in the consortium or SPV. c. Acts as Single Point of Contact with the Finance Department. 3. Other Members a. Each remaining member to hold minimum 10% shareholding in the consortium/SPV. b. Shareholding / participation percentages to be clearly stated in the Consortium Agreement. 4. SPV Structure a. If an SPV is formed, consortium shareholding shall mirror the agreed participation percentages unless otherwise approved by the Finance Department. b. Lead Member to retain controlling interest (≥51%) in SPV throughout the project duration. 5. Joint & Several Liability a. All members jointly and severally liable for performance, delivery of PoC, and project obligations. 6. No Change in Composition a. No change in members or shareholding after bid submission without prior written approval. 7. Consortium Agreement a. Must be executed on non-judicial stamp paper before bid submission, specifying Roles & responsibilities, Shareholding pattern, Cost/profit sharing, Dispute resolution. 8. Authorisation a. Power of Attorney from all members authorising Lead Member to act on their behalf.
13		If a firm qualifies for more than one use case, will those be evaluated and awarded independently or bundled under a single contract?	Each use case will be evaluated and awarded independently irrespective of whether a firm has applied for one or more use cases.
14	Data Protection and Legal Concerns:	Can the department share specific compliance standards or legal frameworks (e.g., Personal Data Protection Bill) expected to be followed during PoC?	Bidders must comply with applicable national and state data protection laws. Compliance with ISO 27001/CERT-IN standards is mandatory, as stated in Section 6(e). NDA, confidentiality and masking of sensitive data related requirements will have to be adhered to as per established standards.
15	Pre-Bid Meeting & Communication:	Can queries be submitted post pre-bid if new clarifications are required?	Queries received till now from all interested bidders have been compiled and their responses published in this Addendum while not disclosing the names of concerned interested bidders. The EOI does not mandate replies to any subsequent queries.
16	Post-PoC Phase:	What is the anticipated timeline and budgetary framework for scaling up the PoC into a full-fledged deployment, if selected?	Timelines and budgets for full-scale deployment will be determined in the subsequent RFP Stage based on the scope and complexity of the developed PoCs that qualify as selected solution(s) for full-fledged deployment.
17		Will there be a separate RFP or commercial negotiation round post-PoC for implementation, or is this EOI process final?	Yes. The EOI is for shortlisting firms and PoC development and is akin to technical qualification for post-POC stages. Post-PoC, depending on PoC results, the Finance Department may issue a separate limited RFP inviting financial bids from those firms whose developed PoCs qualify in the evaluation by the PoC Evaluation Committee.
18		Will there be any further clarification on the PoC scope for individual use cases? The current EOI describes the scope for a full deployment; PoC-level expectations are not clearly differentiated.	The scope of work and deliverables of the PoC for individual use cases is clearly outlined in Section 3 and Section 4 of the EOI. For the sake of complete clarity during PoC development, the problem statements covering the full-fledged scope of deployment is informed to ensure better PoC development. However, the PoC does not include integration with existing systems which could be required for full deployment. Shortlisted bidders will be provided detailed inputs about the outputs to be demonstrated within PoC timelines.
19		For each use case, what data and documents will the government provide to support PoC execution?	All necessary data and documents will be provided as deemed required and per availability to ensure successful PoC execution/development. NDA, confidentiality and masking of sensitive data related requirements will have to be adhered to by the PoC participants as per established Govt standards.
20		Will a well-defined set of success criteria be provided for each use case?	The PoC Evaluation Committee composed of experts in related fields will evaluate the PoC after its development based on well-defined criteria.
21		The EOI's current success metrics appear high-level — will more granular, measurable criteria be issued before PoC commencement?	Any such updates at any later stages will be informed to all firms who qualify for developing the PoCs.
22		Is there a preferred technology stack for the solutions?	The EOI does not mandate a fixed technology stack. Solutions may be developed using any suitable technology stack that meets the functional, performance, and security requirements as per industry standards.
23		Will the preferred stack vary across use cases?	There is no mandated variation; however, technology choices should be optimal for the specific functional requirements of each use case and align with interoperability and integration needs as per industry standards.
24		Are there any restrictions on using third-party AI models or cloud-based AI services? If so, please specify.	As such there are no overarching restrictions specified in the EOI. However, the following may be noted: As per Section 7(3) of the EOI, all intellectual property rights of developed solutions will vest with the Government of West Bengal. NDA, confidentiality and masking of sensitive data related requirements will have to be adhered to by the PoC participants as per established Govt standards. Moreover, the solution should be capable of being hosted and deployed at the West Bengal State Data Centre.
25	General queries	In what form the PoC results needs to be showcased ? Kindly explain	The PoC results are expected to be demonstrated in a clear and practical manner that showcases the effectiveness, feasibility and applicability of the proposed solution. Bidders are encouraged to prepare a live demo, accompanied by a presentation and brief documentation explaining the approach, architecture, and expected outcomes. Any further updates related to showcasing PoC results will be informed to firms who qualify for developing the PoCs.
26		What will be selection success criteria for the use cases ?	Selection will be done in the following stages: Eligibility & Technical Qualification – Firms must meet qualification criteria as detailed in Section 6 of the EOI. PoC development – Shortlisted firms will develop Proofs of Concept. PoC Evaluation - The PoC Evaluation Committee composed of experts in related fields will evaluate the PoC after its development based on well-defined criteria. Only firms that qualify the assessment of the PoC Evaluation Committee will be considered for further post-POC stage. Any further updates at any later stages will be informed to all firms who qualify for developing the PoCs.
27		Under Technical Expertise the completed Projects has to meet all asked criterias or selective criterias will also do ?	As per Section 6(b), Page 6 of the EOI, bidders must demonstrate a minimum of three completed projects in AI/ML, with relevance to any one or more of the listed technical domains (e.g., Computer Vision, NLP, LLMs, etc.). It is not mandatory for each project to cover all listed technologies. Demonstrating capability in areas among those mentioned will be considered valid.
28		SPOC will be designated for every use case will be post PoC evaluation qualification ? Please clarify	The details of each designated Single Point of Contact (SPOC) have been provided under the sub-heading "2. Single Point of Contact (SPOC)" of this Addendum.
29		As part of Qualification criteria, can a firm provide CS certificate for Turnover.	No
30	3. Scope of Work Page No. 2, section 3	Please provide us details on the preferred mode of deployment for use case solution(s) (cloud or on-premise)	The solution should be capable of being hosted and deployed at the West Bengal State Data Centre.
31	3. Scope of Work Page No. 2, section 3	Incuse solutions are to be deployed on cloud, please provide details on procurement, pricing and maintainance of infrastructure	NA
32	3. Scope of Work Page No. 2, section 3	Does the scope of work include maintainance and support of the developed solotion(s). If yes, please provide details on duration of maintainance and support and implementing changes during maintainance and support period	At the Expression of Interest (EOI) stage, the current scope is limited to the development of Proofs of Concept (PoCs) for the identified use cases. The questions of maintenance and support of the developed solution(s) will be examined only Post-POC whenever the limited Request for Proposal (RFP) among successful firms who qualify the PoC is floated for a later stage.
33	3. Scope of Work Page No. 2, section 3	Please provide details of different data source for all use cases (databases, flat files, websites, pdf, excel etc.)	All such data sources of each use case are there. Such updates will be informed to all participating firms at later stages.

34	General Queries	3. Scope of Work Page No. 2, section 3	Please provide details on timeline for implementing the solution once qualified the PoC evaluation for all use cases	The EOI provides this in Section 6: Timelines: “Commencement of PoC development: The qualifying bidders will be informed the result of such Bid Evaluation thereafter through email to participate in the PoC. Upon receipt of such email from Finance Department, they are expected to commence the PoC development without delay.” As per “3. Timelines” of the Addendum order, all entities shall be required to share their PoC’s results with the Finance Department within 30th December, 2025, irrespective of the date on which they may have submitted their bids.” PoC Evaluation Committee will examine the PoCs once developed and submitted. Mentioning further timelines are beyond the scope of this EOI.
35		3. Scope of Work Page No. 2, section 3	Please provide details on approximate number of users for the solution developed for all use cases	During PoC stage, for any particular Use Case the no. of users will be limited and would not exceed say 10 users. Once developed, the users may range in thousands. However, that is Post-POC, and such usage is not covered in the scope of this EOI. Details will be discussed with participating firms during PoC development.
36		6. Qualification Criteria: b) Technical Expertise Page No. 6, Section 6.b Technical Expertise: Minimum of 3 completed projects in AI/ML, preferably in the public sector or government domain.	Requesting to include on-going projects	No exemption or relaxation will be provided in respect of the experience requirement specified in the EOI. Only completed projects that meet the eligibility criteria will be considered for evaluation. Ongoing projects may be mentioned in the bid for reference but will not be counted towards fulfilling the minimum experience requirement.
37		A. General Queries	Can a bidder apply for multiple use cases under a single proposal? If so, should a consolidated or separate approach be followed?	Yes. As per Section 1 and Section 6 of the EOI, bidders may apply for one or more use cases. For ease of evaluation, a separate technical approach and methodology should be provided for each use case applied for, even if submitted under a single cover.
38			Will sandbox/test environments or sample datasets be provided for PoC development?	No dedicated sandbox / test environment will be provided. However, data and documents deemed necessary for PoC development will be shared with qualified bidders. Participants will be required to adhere to Non-Disclosure Agreement (NDA) provisions, confidentiality obligations, and masking of sensitive data, in line with established Government standards and applicable data privacy and security requirements.
39			Will existing systems or APIs (like MVDB, registration, MIS) be accessible for integration during PoC?	Direct access to live production systems would not be provided during PoC. However, any integration would be through standard APIs where deemed necessary after PoC development. Details would be discussed during PoC development as & when required.
40			Will a Non-Disclosure Agreement (NDA) be executed prior to the release of any sensitive datasets?	Yes
41			What level of infrastructure (e.g., servers, hosting environments) will the Finance Department provide for PoC testing?	The data will be hosted on West Bengal State Data Centre. However, the PoC is to be self-financed by the participating bidder.
42			Will clarifications raised during the pre-bid be shared publicly with all prospective bidders?	Yes. All pre-bid clarifications will be compiled and shared with all interested bidders for transparency.
43			Will extension of submission deadline be considered if critical clarifications are delayed or data access is pending?	Any extension, other than those expressly mentioned in this Addendum order will be at the discretion of the Finance Department and will be communicated via addendum, if granted.
44		C. PoC Evaluation and Metrics	Will the evaluation matrix (precision, recall, F1-score, MAE, MSE, Perplexity) have defined minimum thresholds for qualification?	PoC evaluation criteria will be the same across bidders. Any further updates will be informed to all participating firms at later stages.
45			Will the same dataset be used across bidders for fair benchmarking in PoC evaluation?	PoC evaluation criteria will be the same across bidders.
46			Is there any preference or bonus for efficiency metrics like latency, memory, or compute performance?	Efficiency metrics will be examined during PoC evaluation.
47		D. Eligibility & Qualification Criteria	How will robustness and ethical bias be quantitatively assessed across models?	Similar to extant industry practices.
48			Will international experience (outside India) in public sector AI projects be considered valid for eligibility?	Yes, subject to validation of the documents submitted.
49			Are domain-specific CVs evaluated during the EOI stage or only at the RFP stage?	Yes, CVs of key personnel are required at EOI stage.
50			Can legal and financial documents be submitted digitally first and physically later due to postal delays?	No. As per EOI (Section 10) bids must be submitted in sealed physical form only.
51		E. Contractual & Legal Terms	Will intellectual property developed using pre-existing proprietary modules also vest with GoWB?	As per Section 7, Page 8 of the EOI, all Intellectual Property Rights (IPR) arising out of the work carried out under this engagement—including PoC development—shall be owned by the Finance Department, Government of West Bengal. However, the pre-existing proprietary platforms, tools, or models used by the vendor for building the solution remain the property of the vendor, unless specifically customized or licensed for this engagement. Any further clarity on IPR, if required, would be addressed in due course.
52			Can firms retain ownership of their base platforms while transferring usage rights of custom solutions?	subject to granting GoWB perpetual usage rights for deployed solutions.
53			What is the maximum liability cap under the terms of the PoC agreement?	It is not explicitly defined in the EOI as it does not seek financial bids. However, breach of trust, confidentiality or any other terms by a bidder may lead to immediate blacklisting and other penalties by the Government. However, such details will be informed to participating firms only at a later stage.
54			Will there be a formal agreement signed for PoC development to safeguard investments?	Yes
55			What is the expected response time from the Department for approval/clarification during PoC phase?	The SPOC for each Use Case would be responsible to clarify any matters during PoC development phase. He may be reached out anytime. As the query is subjective, a more specific response is not possible at this stage.
56	F. Submission Process		What is the exact format for marking envelopes (e.g., “EOI for AI Use Case No. ___”)?	Mark as: “EOI for AI Use Case [Sl. No. and Name of Use Case(s)]” as per Section 8.
57			Is there a limit on the size/volume of annexures or technical documents to be submitted?	No specific page limit; bidders should ensure submissions are clear and well-organized.
58		Overall Goal and Objectives	What are the primary business objectives the Finance Department hopes to achieve by implementing these AI/ML solutions? Beyond the stated “enhancing governance, improving service delivery, and promoting transparency”, what specific pain points or strategic goals are paramount for each use case?	The EOI has mentioned the details of use case-specific objectives and strategic goals in Sections 2 and at Appendix A.
59		Overall Goal and Objectives	Are there any pre-defined Key Performance Indicators (KPIs) or success metrics for measuring the impact and ROI (Return on Investment) of these AI initiatives, both at the Proof of Concept (PoC) stage and for full-fledged solutions?	Yes.
60		Current Infrastructure & AI Readiness	Can you provide a high-level overview of the Finance Department's current IT infrastructure, including server environments (physical/virtual), network architecture, and data center facilities?	Required IT infrastructure and facilities are available. The query seems to be beyond the scope of the current EOI. However, necessary details would be discussed during PoC development as & when required.
61		Current Infrastructure & AI Readiness	Are there any existing AI/ML platforms, Big Data tools, or specific software solutions currently in use within the Government of West Bengal that the proposed solutions should integrate with or build upon?	Yes. Integration is not a part of the scope of the PoC. However, any integration would be through standard APIs where deemed necessary after PoC development. Details would be discussed during PoC development as & when required.
62		Current Infrastructure & AI Readiness	What is the current state of data governance policies and practices within the Finance Department and other relevant departments?	Necessary policies and practices as per Government Rules are in place. However, the query does not seem to be in the scope of PoC development of this EOI.
63		Current Infrastructure & AI Readiness	What is the current level of AI/ML expertise and readiness within the Finance Department and other collaborating departments in terms of in-house technical staff, data scientists, or previous AI project experience?	The query seems to be beyond the scope of the current EOI. However, necessary details would be discussed during PoC development as & when required.
64		Computing & Storage Resources	What are the current data storage capacities and types (e.g., Storage Area Network (SAN), Network Attached Storage (NAS), object storage) available for potentially hosting the large datasets associated with these AI models?	Not relevant for PoC. However, sufficient storage space capacity of necessary type is available in the West Bengal State Data Centre to host and deploy the full fledged solutions at scale.
65		Networking & Accessibility	What is the current network security posture and architecture, specifically concerning data transfer between different departmental silos and any potential external interfaces?	As per SDC standards and IT guidelines of the Government.

66	Cloud Infrastructure	Does the Government of West Bengal have a preferred cloud strategy (e.g., specific public cloud provider like AWS, Azure, GCP; private cloud; hybrid cloud model)?	The solution should be capable of being hosted and deployed at the West Bengal State Data Centre. Further details will be shared during PoC development.
67	Cloud Infrastructure	Are there any specific data residency requirements that mandate data to be stored exclusively within India or within government-controlled data centers?	Yes
68	Security & Compliance	Beyond the stated ISO 27001/CERT-IN standards, are there any other specific cybersecurity frameworks, government security policies, or audit requirements that the solutions must adhere to?	Solutions must also adhere to applicable national cyber security guidelines, sectoral regulations and State-specific security policies in force at the time of implementation.
69	Security & Compliance	What is the expected incident response plan or security breach notification protocol that firms must integrate with or adhere to?	As per IT guidelines of the Government.
70	User Base & Scalability	What are the long-term scalability expectations for each solution in terms of anticipated data growth, user growth, and integration with future government systems?	The query seems to be beyond the scope of the current EOI. However, necessary details would be discussed during PoC development as & when required.
71	Technology Preferences & Strategy	Are there any preferences for open-source versus proprietary software components in the proposed AI solutions?	This is open to innovation. There are no recommended solutions. However, efficiency metrics including compute performance and cost of operation may form a part of the PoC evaluation criteria of the PoC Evaluation Committee. Also, it may be kept in mind that as per Section 7, Page 8 of the EOI, all Intellectual Property Rights (IPR) arising out of the work carried out under this engagement—including PoC development—shall be owned by the Finance Department, Government of West Bengal. Any further clarity on IPR, if required, would be addressed in due course.
72	Technology Preferences & Strategy	What is the long-term maintenance and support strategy envisioned for the full-fledged AI solutions (e.g., in-house team management, vendor support contract, hybrid model)?	The query is beyond the scope of the EOI. The strategy would typically include a hybrid approach. However, it will be known whenever post-POC RFP is floated.
73	Project Management & Support (Technical Aspects)	What level of technical documentation (beyond user manuals and training materials) is expected for the AI models, including model architecture, training data specifics, and performance benchmarks?	Yes, all these details are expected.
74	A. General Queries	Will the 120-day PoC clock start after documents are shared, or from EOI publication . PoC timeline will be dependent on the information which are required as a part of due diligence .	The extended dates are mentioned under the sub-heading "3. Timelines" in this Addendum.
75	General - Capability-Oriented Framing	1. Broader Recognition of AI/ML Delivery Experience Given the emerging nature of AI/ML deployments in the government sector, would the Department consider relevant experience from other regulated industries (such as BFSI, healthcare, utilities), where projects have operated under comparable compliance, security, and scale constraints, as qualifying towards the "completed projects" requirement?	Yes. While preference will be given to projects implemented in the public sector or government domain, the Department recognizes that AI/ML expertise in other regulated industries—such as Banking, Financial Services and Insurance (BFSI), healthcare, utilities, or similarly compliance-intensive sectors—can be highly relevant, provided such projects have operated under comparable compliance, data privacy, security and scale constraints.
76		2. Eligibility Criteria – Minimum of 3 Completed AI/ML Projects The EoI specifies a minimum of 3 completed projects in AI/ML, preferably in the public sector or government domain. Would the Department consider relaxing this requirement, including allowing the use of consortium partner credentials to collectively meet the "3 completed projects" threshold, where the bidding team can demonstrate strong domain expertise, relevant technical capability, and readiness to prove competence through the mandated PoC stage?	Consortium Bids are permitted. As per the eligibility criteria, the bidder may be a consortium with at least 2 years of operations of its consortium members. In case of a consortium bid, relevant documents and experience of each partner should be furnished. The following details may be noted for Consortium Bids: • Consortium as a whole shall have to meet the eligibility criteria and turnover requirements specified in the EOI. • The Lead Member of the Consortium should have to have at least 50% of the minimum turnover as specified in the EOI. • A Consortium Bid may use the credentials of its Associates for the purpose of qualification under this EOI, provided however that the Consortium Bid shall ensure that such Associate continues to remain an Associate throughout the Contract Period.
77		3. Inclusion of PoCs and Large-Scale Prototypes Many innovative AI/ML solutions, particularly in public-sector-like environments, begin as large-scale pilots or PoCs before transitioning to production. Would the Department consider successfully delivered PoCs/prototypes - meeting all functional and technical objectives for reputed organizations - as eligible experience, even if the ongoing live operations were managed by the client's in-house teams?	Yes. However, only completed projects meeting the stated criteria will be considered.
78		4. Demonstrating Capability via PoC Since the selection process includes a PoC stage, would the Department consider permitting technically strong but government-new bidders to qualify for PoC participation, so that evaluation is based on actual solution performance against the Department's use cases?	Technical selection will be based on qualification criteria as mentioned in the EOI.
79		5. Team Experience Criteria For the requirement of a Project Manager with 10+ years of AI/ML project experience, would the Department accept combined credentials where a senior transformation leader has extensive program delivery experience and is supported by AI/ML technical leads with proven implementation track records?	No. Such relaxations, if any, will be informed later by the Finance Department.
80		6. Consortium Partner Credentials If the lead bidder possesses strong government-domain delivery experience but limited in-house AI/ML capability, can the consortium partner's AI/ML project credentials be leveraged to meet the technical eligibility criteria?	Consortium Bids are permitted. As per the eligibility criteria, the bidder may be a consortium with at least 2 years of operations of its consortium members. In case of a consortium bid, relevant documents and experience of each partner should be furnished. The following details may be noted for Consortium Bids: • Consortium as a whole shall have to meet the eligibility criteria and turnover requirements specified in the EOI. • The Lead Member of the Consortium should have to have at least 50% of the minimum turnover as specified in the EOI. • A Consortium Bid may use the credentials of its Associates for the purpose of qualification under this EOI, provided however that the Consortium Bid shall ensure that such Associate continues to remain an Associate throughout the Contract Period.
81		7. Eligibility Criteria – Use Case 1: Experience in Geospatial Analytics, Econometrics, or Valuation Models The EoI specifies prior experience in geospatial analytics, econometrics, or valuation models as part of the eligibility criteria for Use Case 1. Would the Department consider relaxing this to also include bidders who have demonstrated capability in developing AI/ML solutions for large-scale data integration and predictive modelling, and who can bring in the required geospatial/econometric expertise through a consortium partner or domain specialist during the PoC and implementation phases?	Consortium Bids are permitted. As per the eligibility criteria, the bidder may be a consortium with at least 2 years of operations of its consortium members. In case of a consortium bid, relevant documents and experience of each partner should be furnished. The following details may be noted for Consortium Bids: • Consortium as a whole shall have to meet the eligibility criteria and turnover requirements specified in the EOI. • The Lead Member of the Consortium should have to have at least 50% of the minimum turnover as specified in the EOI. • A Consortium Bid may use the credentials of its Associates for the purpose of qualification under this EOI, provided however that the Consortium Bid shall ensure that such Associate continues to remain an Associate throughout the Contract Period.
82		8. Eligibility Criteria – Use Case 3: Federated Databases, Geo-Spatial Intelligence, and Grievance Analytics The EoI specifies prior experience in (i) Geo-Spatial Intelligence and GIS-AI integration, and (ii) hands-on experience in grievance analytics and citizen feedback systems. Given that these are diverse and highly specialized skill areas, would the Department consider allowing bidders to meet this requirement through a combination of in-house expertise and consortium partnerships, or by demonstrating equivalent capabilities through related large-scale AI/ML implementations, with the opportunity to validate these capabilities during the PoC stage?	Consortium Bids are permitted. As per the eligibility criteria, the bidder may be a consortium with at least 2 years of operations of its consortium members. In case of a consortium bid, relevant documents and experience of each partner should be furnished. The following details may be noted for Consortium Bids: • Consortium as a whole shall have to meet the eligibility criteria and turnover requirements specified in the EOI. • The Lead Member of the Consortium should have to have at least 50% of the minimum turnover as specified in the EOI. • A Consortium Bid may use the credentials of its Associates for the purpose of qualification under this EOI, provided however that the Consortium Bid shall ensure that such Associate continues to remain an Associate throughout the Contract Period.

83		5. Selection Process a) Bid Submission Financial Statements of the last 3 Financial Years 6. Qualification Criteria: c) Financial Capability: - Average annual turnover of INR 5 crore or more for IT & ITeS related activities over the last 2 financial years. - Positive net worth in the last 2 financial years.	Under 5. Selection Process, a) Bid Submission, the financial statements are sought for last 3 Financial Years. However, under 6. Qualification Criteria, c) Financial Capability, financial statements are sought for last 2 Financial years	Please refer to the sub-heading "4. Mandatory Supporting Documents for ascertaining Qualification Criteria of the EOI" of this Addendum.
84			For the AI-led Valuation Assessment System (Usecase 1), what datasets related to recent property transactions are available internally, and how would you prefer we validate outputs given the known gap between guideline values and actual market rates?	Sufficient data on transactions of properties through registered deeds are available. However, identification of gap between guideline value and value in actual market is also a deliverable of the use case. Therefore, the bidder shall have to identify suitable mechanism for identification of the gaps as well as validation of the gaps so identified.
85			Are there any data sensitivity or classification issues (Restricted zones etc)	All firms are advised to build their solutions with privacy-by-design principles and to ensure full compliance with ISO 27001/CERT-IN standards wherever applicable. However, PoC development would not be affected due to any restricted zones.
86			Will the bidder have access to any existing GIS infrastructure?	Integration during PoC development is not envisaged. However, GIS data will be provided as per requirement where available.
87			Will past transaction data and geospatial datasets be shared during the PoC phase?	Yes. However, the participating firm is also expected to complement this with publicly available geospatial data and tools, where necessary, for model enhancement, validation, and ground-truthing. Further, as the question is open ended, a definitive answer to the query is not possible till selection of firm for participation in the PoC.
88			What accuracy benchmarks or confidence thresholds are expected for zone delineation and valuation prediction?	Specific thresholds are not prescribed in the EOI. However, higher accuracy and confidence will positively influence evaluation scores. Any further updates will be informed to all participating firms at later stages.
89		Enterprise AI Model for Scientific Valuation Assessment	Will crawling public real estate listings for training/validation be permitted under compliance guidelines?	Permitted if compliant with applicable laws and terms of source sites.
90			Will validation require field surveys conducted independently by bidders or jointly with the government?	Field surveys data used will be of those surveys that are conducted by Government / jointly only.
91			Is integration with the live Market Valuation Database required at PoC or only during final implementation?	Integration is not required during PoC stage.
92			Are there any recommended models (e.g., hedonic pricing) or is the approach open to innovation?	Approach is open to innovation. There are no recommended models. Any approaches may be used if felt appropriate.
93		Data Environment (Specific to Use Cases)	Use Case 1: What are the typical file formats and the data freshness/update frequency for the "transaction records, GIS layers, and Market Valuation Database (MVDB)"?	The databases are 'Live' and new data keeps getting added daily. Formats are compatible with industry standards.
94		Data Environment (Specific to Use Cases)	Use Case 1: What is the estimated volume and historical depth of these datasets (e.g., number of property transactions per year, size of GIS layers)?	The volume and historical depth is high. Necessary details would be discussed during PoC development as & when required depending on a PoC's requirements.
95		User Base & Scalability	Use Case 1 (Property Valuation): What is the estimated number of concurrent users for the valuation model within the Directorate of Registration & Stamp Revenue, and what is the expected transaction volume (e.g., number of valuations/revisions) per day/month?	For PoC development, there won't be any such issues. Post-PoC, the concurrent users may be around 300 and the expected transaction volume per day may be upto 20000.
96		Use Case Specific Technical Clarifications	Use Case 1 (Property Valuation): Can you provide sample datasets or detailed schema for the "transaction records, GIS layers, and Market Valuation Database (MVDB)" for a specific representative region to aid in solution design?	Yes, details will be provided during PoC development.
97			What exact datasets will be shared for the PoC? · Transaction records (sale date, sale price, property type, area, floor, age, etc.) · Survey data (amenities, building condition) · GIS layers (ward boundaries, zones, coordinates) · Existing MVDB data.	Transaction records and MVDB data samples will be shared. Data and documents deemed necessary for PoC development will be provided. NDA, confidentiality and masking of sensitive data related requirements will have to be adhered to by the PoC participants as per established Govt standards and applicable data privacy and security requirements.
98			Will datasets include historical data or only recent records? How far back do records go?	Historical data is available for as long as genuinely required. It will be shared depending on the PoC Model's requirement. Data and documents deemed necessary for PoC development will be provided.
99			Are datasets digitized and clean, or will significant preprocessing be needed?	Datasets are digitized. Significant preprocessing may not be needed. However, some data cleaning and preprocessing may still be crucial for meaningful analysis.
100			Will unique property identifiers be consistent across datasets?	Unique property identifiers may not always be consistent across all datasets.
101			Will sensitive data (owner names, IDs) be anonymized before sharing?	Bidders must sign an NDA and comply with privacy-by-design principles, ISO 27001/CERT-IN standards and applicable data protection laws to ensure that any sensitive information remains secure throughout the PoC development.
102			How and when will we get dataset access — before PoC start or after award?	Data / datasets will be shared during PoC development.
103			Will data be shared via API, bulk CSV/Excel, or secure database access?	Data will be shared in formats suited to the nature of each dataset — this may include bulk CSV/Excel files, scanned documents, or secure database extracts. For any integration needs, standard APIs may be provided, but API-based access will not be part of the PoC development stage and would be considered only at the post-PoC implementation stage.
104		Use Case 1: Enterprise AI Model for Scientific Valuation Assessment	Can we integrate external public datasets (e.g., govt open data, Google Places, remote sensing imagery) in the PoC?	Yes, the details may be discussed as part of proposed process flow and PoC Model proposed by a firm during PoC development. Bidders may integrate external public datasets into the PoC, provided that: The data is publicly available and its use complies with applicable laws and the terms of the source sites. Any integration enhances model performance, validation and ground-truthing without violating data privacy, security, or intellectual property rights.
105			Should the valuation model predict transaction price or update official MVDB rates?	The PoC for the valuation model is expected to focus on scientific valuation assessment and prediction of property values based on the provided datasets and other relevant inputs. Mechanism or broad guideline of the procedure to be followed for updation should be a deliverable of the PoC so as to understand the feasibility of the model. Updating official Market Valuation Database (MVDB) rates is not within the PoC scope.
106	Use Case 1		Are we expected to deliver zonal valuation maps as part of the PoC?	Not maps, but valuation data based on identification of zones is expected. The details will be discussed during PoC development.
107			What level of spatial granularity is expected — ward, street, building etc.?	Property level granurity is expected.
108			Should the model handle forecasting (future valuation trends) or only current values?	Forecasting may intrinsically be a part of the exercise of evaluation of current values. Such details may depend on the PoC model proposed by respective firms and can be discussed during PoC development.
109			Are there weights for different factors (accuracy vs explainability vs speed)?	Yes
110			Should the PoC integrate with an existing government application (e.g., property registration system)?	Integration is not envisaged in PoC stage. Post-POC, integration during full-fledged deployment is required. So necessary APIs may be needed to be showcased during PoC development that can push data when required.
111			In what format should results be delivered — dashboard, API, reports, GIS shapefiles?	All these formats may be required. The PoC results are expected to be demonstrated in a clear and practical manner. Firms are encouraged to prepare a live demo, accompanied by a presentation and brief documentation.
112			Is there a preferred tech stack or hosting environment (on-premises, government cloud, external cloud allowed)?	There is no mandated technology stack in the EOI. Firms may choose any stack that meets the functional, performance, and security requirements. They are expected to discuss the details with the SPOC for feedback during PoC development. However, the solution must be capable of being hosted and deployed at the West Bengal State Data Centre.

113	Enterprise AI Model for the Scientific Framework of Valuation Assessment	Q1. To develop a supervised model, labelled data is crucial. Can the Department confirm that the provided transaction records will contain the final property valuation or sale price, which can serve as the "label"? Is any manual data labelling or annotation activity anticipated to be part of the scope?	Labelled data will be provided. However, the query is not clear. Details will be discussed during PoC development.
114		Q2. Section 1 of Appendix A mentions providing existing databases like "transaction records, GIS layers, and the Market Valuation Database (MVDB)." What will be the formats in which this data will be provided (e.g., CSV, database dump, API)?	Data will be shared in formats suited to the nature of each dataset — this may include bulk CSV/Excel files, scanned documents, or secure database extracts. For any integration needs, standard APIs may be provided, but API-based access will not be part of the PoC development stage and would be considered only at the post-PoC implementation stage.
115		Any known data quality issues, such as missing values or inconsistencies, that we should be aware of for the data cleaning phase?	The Finance Department, Government of West Bengal, maintains multiple datasets relevant to the proposed AI use cases, sourced from various departmental systems and historical records. Given the heterogeneous nature of these datasets, some quality issues such as incomplete fields, outdated entries, inconsistencies in formats (e.g., date formats, spellings) and duplication may exist, particularly in legacy datasets.
116		Q3. The scope mentions a mechanism for "periodic revision." Could the Department provide an indication of the expected frequency of these updates (e.g., quarterly, semi-annually, on-demand) to help us design a suitable system architecture?	There is no predefined or fixed frequency. However, feature for uploading updates by pre-designated users may be kept. Access to central repository through APIs would only also be provided in post-POC stage.
117		Q4. Page 9 lists "User Interface Design" as a PoC evaluation criterion. Could you please clarify the expected scope and key functionalities of the user interface for the final, full-fledged solution beyond the Proof of Concept? For instance, will it require dashboards, reporting features, or interactive map-based visualizations?	The PoC results are expected to be demonstrated in a clear and practical manner. Firms are encouraged to prepare a live demo, accompanied by a presentation and brief documentation.
118		Q5. The EOI states that the firm should use AI to delineate zones based on "statistically significant variables." does the Department have any pre-identified or high-priority parameters that must be included in the model?	Yes, some details are there. However, they may not the full set of required details and the bidders will be required to identify the full set of variables as such statistical modelling is a requirement and deliverable of the PoC.
119		Q6. For ground-truthing, the EOI mentions crawling "real estate website listings." Does the Department have a list of preferred or indicative websites that are considered reliable sources for property data in West Bengal? Are there any specific websites we should avoid?	There is no such preference. Publicly accessible website data is expected to be web crawled. Separate permissions are not envisaged in the scope of the EOI. For any such data pulling from a website, the participating firm shall have to check that it is permissible as per the policy of the concerned website as well as under the Law.
120		Availability & Sources of Transaction Data	Clarify if recent and historical property transaction records (including sale deeds) are available in digital form and at what granularity (property, zone, district). Necessary details upto Property level of granularity in digital form are available.
121		GIS Data Specifications	Confirm what GIS layers (administrative boundaries, land use, transport networks, satellite imagery) are available, including file formats, coordinate systems, and update frequency. All these GIS layers are available. File formats are as per industry standards. Update frequency varies depending on the type of GIS layer, however, it is as per industry standards.
122		Market Valuation Database (MVDB) Access	Specify the fields, historical range, and update schedule of MVDB data, and whether historical factor weights are available. Necessary details of fields, update schedule and historical data are available. The details will be discussed during PoC development stage when required.
123	External Data Sources	Clarify if integration with web-scraped property listings and market portals is allowed, and whether any preferred APIs or data vendors are specified.	Yes. Integration with publicly available web-scraped property listings and market portals is permitted for validation and benchmarking, subject to compliance with applicable laws, data privacy, and terms of use. The Department has not specified any preferred APIs or data vendors; bidders are free to propose suitable sources in their technical approach.
124		Ground Truth & Validation	Field survey data will not be provided by the Department. Bidders are required to conduct field surveys for validation purposes. The sampling methodology must be proposed by the bidder in their technical approach. The Department may provide broad guidance on parameters, but the design and execution of the survey will be the bidder's responsibility.
125		Factor Identification Methodology	Request existing documentation or previous studies on factors influencing property valuation in the state.(Rules and factor needed) The details of earlier studies and Rules will be discussed during PoC development when required.
126		Frequency of Model Updates	Clarify the expected frequency for revising valuation zones, base values, and factor weights (quarterly, annually). The "Expression of Interest" document states that the objective is to "develop an AI/ML-based mechanism for periodic revision of zones, base values, and factor weights". It is expected that the bidder will propose a suitable methodology and frequency for these revisions as part of their solution.
127		Integration Requirements	Determine whether the AI model must integrate directly into existing valuation software or will run as a standalone module in the PoC. For the PoC, the AI model is expected to run as a standalone module, working with datasets provided or collected by the bidder. Direct integration with existing valuation software is not required at the PoC stage. However, the proposed solution should be designed to allow seamless integration with the Department's valuation systems at the implementation stage.
128			For the AI-led Deed Digitisation and Chain of Title system, what proportion of physical deeds have already been scanned, and at what typical resolution or quality levels? Also, since many deeds may include local code words or terms specific to West Bengal, what kind of support can we expect from local registration offices during the PoC to help validate and interpret these documents correctly A large no. of deeds, in crores, have already been scanned. The resolution is 200 dpi at least. Bidders are expected to have resources for basic output validation of the Bengali language. However, for any technical terms or codes or any particular matters requiring specific guidance, it would be provided during PoC development.
129		Technical Clarifications:	Should the conversational interface and OCR modules support both Bengali and English during PoC, or is proof of concept in English sufficient?
130		Digitization of Handwritten Records	Will there be some Bengali specific data sets that will be shared by the client?
131		Document restoration	Are there any resolution requirements for restoring degraded documents?
132			Will the Finance Department provide a set of manually annotated legacy Bengali deeds for OCR training?
133	Digitization of Old Registration Records & Chain Deed Preparation	What character recognition accuracy is expected of the OCR model for PoC qualification?	Specific thresholds are not prescribed in the EOI. However, higher accuracy and confidence will positively influence evaluation scores. Any further updates will be informed to all participating firms at later stages.
134		Are there any pre-defined metadata fields (like the 60 types mentioned) available as a schema for reference?	Yes, the list of different types of metadata will be provided during PoC development.
135		Can hybrid models (AI + human validation) be proposed during metadata validation?	Techniques like RLHF, etc. may be used.
136		Is the use of Graph Neural Networks (GNN) mandatory for deed linkage, or are other approaches acceptable?	Approach is open to innovation. There are no recommended models. Any approaches may be used if felt appropriate.
137		Should the deed generation model provide legal validation or just structural templates based on metadata?	It should create legally sound structural templates based on metadata. Providing legal validation is not required even though such a feature could enhance the evaluation score.
138		Data Environment (Specific to Use Cases)	Use Case 2: What is the estimated total number of legacy handwritten Bengali registration records requiring digitization, and what is their approximate physical condition?
139		Data Environment (Specific to Use Cases)	Use Case 2: Are there any sample sets of these legacy documents available for initial OCR model training and testing?
140	User Base & Scalability	Use Case 2 (Digitization): What is the expected daily/weekly throughput for digitized documents and metadata extraction? How many users will typically interact with the metadata validation and chain deed reconstruction tools?	The participants in the PoC do not have to scan the legacy physical records as the scanned samples of such records in sufficient numbers would be made available to the participants. The legacy records are physically preserved in Record Rooms. Their condition is as can be expected after the passage of many years and the resultant natural wear and tear. Their total number is not required for PoC development though it is in crores. During PoC development, these are not immediate concerns. However, Post-PoC, the daily throughput for digitized documents and metadata extraction should deliver outputs for at least be 50k documents, with each document having around 7-10 pages. Multiple instances of the solution may be run simultaneously to scale up the throughput to such levels and higher. The users who may interact daily may be around 1000 and the expected interactions per day may be upto 100000.
141		What volume of deed documents will be provided for the PoC (number of pages, number of deeds)?	There are sufficient no. of documents available. The no.s that will be provided will depend on the PoC Model's requirements.
142		What document types will be included — handwritten, typed, printed, or mixed?	It'll be a mix of scanned copies of handwritten and typed documents. A large majority of the documents will be handwritten. Some would be typed.
143		Will samples cover different time periods(e.g., old-degraded ink or modern printed forms)?	Yes, both types of samples will be there.
144		In addition to Bengali, will there be any probability of other languages and scripts appear in the documents?	Apart from Bengali, there will be some English as well. Any other language's notable text is not expected to be there in the metadata to be captured.
145		Will scanned images be high resolution (minimum DPI) and will there be colour scans or only black-and-white?	Minimum 200 dpi resolution scanned documents will be provided. Images will be primarily black & white.
146		Will metadata (like registration year, location, property number) be provided alongside the scanned images?	A 14 digit unique code may be provided alongside scanned images.
147		Is the Government providing any existing labelled datasets (image + ground truth text) for Bengali cursive OCR?	No

148	Use Case 2	Use Case 2: Digitization of Old Registration Records & Chain Deed Preparation	For translation/transliteration — should the output be English only or also retain the original Bengali text?	The output is required in English primarily. However, it is advisable that the PoC output retains both — the translated English content for broader accessibility, and the original Bengali text for reference and higher administrative scrutiny and accuracy. Doing so will also support evaluation aspects like accuracy, inclusiveness, and usability.	
149			Are there accuracy targets for OCR and translation (e.g., word-level accuracy $\geq X\%$)?	Specific targets are not prescribed in the EOI. However, higher accuracy and confidence will positively influence evaluation scores.	
150			Should OCR handle handwritten signatures and seals, or can those be excluded?	Any further updates will be informed to all participating firms at later stages.	
151			Should we restore degraded/damaged documents for archival, or it will be used for other purposes?	Can be excluded.	
152			Are we expected to deliver image enhancement/restoration tools as part of the PoC?	Restored documents are required for other purposes.	
153			What data fields should be used to identify chain of ownership (e.g., property ID, owner name, address, deed number)?	It is inherently required to achieve the required deliverables though no dedicated image enhancement/restoration tool delivery is explicitly mandated in the EOI. However, bidders are expected to process degraded documents sufficiently to meet the PoC's accuracy and legibility requirements (e.g., ≥ 200 dpi for OCR)	
154			Will there be a unique property identifier across all deeds, or will we rely on fuzzy matching?	Primarily Buyer & Seller Names and Property Addresses. Innovations for chain identification are encouraged and expected.	
155			Is the chain linking limited to current ownership history, or should it reconstruct full historical ownership from the earliest available record?	Unique Property identifiers will not be there in all cases. Other techniques and innovations would be required.	
156			Should digitized deeds and metadata be stored in a searchable database with full-text search?	Reconstruction of historical ownership from the earliest available record is desired in the Post-POC stage. The PoC needs to be developed accordingly.	
157			Should we provide GIS integration to map properties on a map interface?	Yes. Further details will be discussed during PoC development.	
158			What is the preferred output format — PDF with searchable text, structured JSON/XML, or database entry?	Not required in this Use Case.	
159			Should the PoC integrate with any existing land records / property registration system?	All these formats would be required to be incorporated post-POC. The output format should be selected to best demonstrate the proposed solution's functionality, interoperability, and ease of integration in a potential full-scale deployment. Acceptable formats include, but are not limited to, PDF with searchable text, structured JSON/XML, and database-ready formats. The PoC should be designed with flexibility to support multiple formats if required in later stages.	
160			What evaluation metrics will be applied — OCR accuracy, field-level extraction F1 score, chain linking correctness?	Integration is not envisaged in PoC stage. Post-POC, integration during full-fledged deployment is required. So necessary APIs may be needed to be showcased during PoC development that can push data when required.	
161			Is there an expected turnaround time per document for processing?	These are other evaluation metrics will be examined. PoC evaluation criteria will be the same across bidders. Any further updates will be informed to all participating firms at later stages.	
162			Digitization of Old Registration Records and AI-Driven Chain Deed Preparation	Q1. For developing the Bengali-specific OCR and NLP models, does the Department possess or can it provide any linguistic resources, such as dictionaries of historical legal terms, gazetteers, or any pre-labeled (ground truth) data that could accelerate model training?	Specific targets are not prescribed in the EOI. However, higher accuracy and lower error thresholds will positively influence evaluation scores. Checking for accuracy of character recognition will have to be done by participating firms during PoC development. Some checks would be done by the Department to check veracity of the claimed accuracy during PoC development.
163				Q2. Page 12 mentions the challenge of digitizing cursive Bengali. What is the Department's target accuracy for the OCR? How will this accuracy be measured and validated?	There is no universal unique plot identifier.
164				Q3. What are the key data points or business rules that should be used to link two deeds as being for the same property? For example, is there a unique plot identifier, or must the model rely on fuzzy matching of owner names, property descriptions, and neighbour details?	Acceptable formats include, but are not limited to, PDF with searchable text, structured JSON/XML, and database-ready formats. The PoC should be designed with flexibility to support multiple formats if required in later stages.
165	Q4. What is the desired format for the output of the chain deed reconstruction? Is it a visual graph showing ownership flow, a chronological list of linked deed numbers, or another format?	Existing pretrained OCR frameworks may be leveraged, subject to compliance with licensing terms. However, bidders are expected to carry out domain-specific fine-tuning and model adaptation to handle cursive Bengali scripts, degraded handwritten records, and mixed-script documents. The evaluation will focus on accuracy and usability of the output rather than development of a new OCR engine.			
166	Digitization Approach	Confirm if the Bengali-specific OCR model must be built from scratch or can leverage existing pretrained OCR frameworks with domain-specific fine-tuning for cursive Bengali scripts.	Standardized metadata fields to be extracted. Yes.		
167	Metadata Requirements	Clarify the exact metadata fields to be extracted (e.g., owner names, dates, property IDs, survey numbers) and whether transliteration/translation should be standardized across outputs.	For degraded documents, the bidder's solution should demonstrate measurable quality improvements. Minimum expected standards include resolution of 200 dpi equivalent		
168	Document Restoration Standards:	Define the expected quality improvement metrics for degraded documents (e.g., resolution, legibility score) and the acceptable restoration methods (e.g., AI-based super-resolution, de-noising).	Historical linkage is not limited to certain decades or deed types; the solution should be capable of handling all digitized deeds. While prioritization may be given to high-volume categories and more recent decades, the system should extend across the entire archive. In cases of missing or ambiguous linkages, the solution should not force assignments but instead flag such instances with confidence scores, leaving them for manual review or curation by the Department.		
169	Chain Deed Reconstruction Scope:	Specify whether historical linkage is limited to certain decades/record types or covers all digitized deeds; confirm how missing or ambiguous linkages should be handled.	The objective is to assist officials through automation while ensuring legal accuracy through human oversight.		
170	Deed Preparation Automation	Clarify the degree of automation — whether the AI should prepare fully ready-to-register deeds or generate drafts for manual verification.	The solution must be capable of handling a wide range of handwriting styles, including period-specific Bengali and English scripts, variations in ink quality, and non-standard character forms. The Department will provide sample deed scans for PoC, but annotated handwriting datasets will not be supplied in full. Bidders are expected to undertake domain-specific fine-tuning and, where necessary, create annotated datasets for training. Limited annotated samples may be shared by the Department for evaluation purposes only.		
171	Handwriting Variability	Clarify the range of handwriting styles to be covered — e.g., period-specific scripts, variations in ink quality, and non-standard character forms — and whether the dataset will include annotated handwriting samples for model training.	The chain deeds establishing the chain of change of property ownership will need to be prepared during the PoC development. Such a chain will lead to the latest ownership status in registration records and make the property 'mutation ready' to enable mutation seamlessly. The process of automated mutation in such a scenario has already been developed and would not be required to be developed as a part of the PoC. Further details of the process and acceptable errors, if any, will be discussed during PoC development. The ideal error rate is zero.		
172	Mutation Automation Boundaries	Specify which parts of the mutation process will be automated versus human-verified. Define acceptable error rates before a chain deed is deemed valid for mutation.	The details will be discussed during PoC development stage. Integration is not a part of the scope of the PoC. However, any integration would be through standard APIs where deemed necessary after PoC development. Details would be discussed during PoC development as & when required. For any integration related aspects, standard APIs may be provided even though API-based access to Live production environment will not be permitted during PoC development stage and would be considered only at the post-PoC implementation stage.		
173	Integration with Existing Systems	Confirm the technical standards for integrating OCR outputs, metadata, and reconstructed deeds into current registration department systems. Define API/data exchange format requirements.	Bidders must comply with all applicable national and state data protection laws. Compliance with ISO 27001/CERT-IN standards is mandatory. Bidders are required to adhere to Non-Disclosure Agreement (NDA), confidentiality, and data masking requirements as per established government standards.		
174	Data Privacy & Security	Outline the legal and compliance framework for handling sensitive ownership records, including encryption standards and archival retention policies.	The bidder is expected to propose a detailed methodology for validating the accuracy of the model's output in their technical proposal. This should include the proposed metrics and approach for ensuring the quality of the digitized deeds and the correctness of the generated metadata and historical linkages.		
175	Validation Process	Specify the methodology for validating OCR accuracy, metadata correctness, and chain deed integrity (e.g., sampling size, human-in-loop verification, automated cross-checks).	The AI models must be designed to support continuous retraining as new data and corrections become available. The solution will be hosted and maintained centrally, with district and sub-registrar offices serving as data entry and validation points. Centralized deployment will ensure uniformity of models and outputs across the State.		
176	Scalability & Maintenance	Clarify whether the AI models must be designed for continuous retraining with new data and if the system will be maintained centrally or at the district/sub-registrar office level.	Are there structured datasets currently available for training eligibility and segmentation models?		
177			Yes. However, as the question is open ended, a definitive answer to the query is not possible.		
178		Will historical beneficiary and inclusion/exclusion data be provided?	Yes. However, as the question is open ended, a definitive answer to the query is not possible.		

179		Will access to relevant scheme documents, MIS databases, and legacy records be provided for PoC?	Yes. However, as the question is open ended, a definitive answer to the query is not possible.
180		Is a governance structure defined for inter-departmental collaboration and conflict resolution regarding eligibility criteria?	Senior officers as Single Points of Contact (SPOCs) have been designated for each use case. However, as the question is not clear, a definitive answer to the query is not possible.
181		Which departments and schemes are prioritized for integration during the PoC phase?	For PoC purposes under Use Case 3 (Unified Social Registry), a select set of departments and schemes have been selected. Details will be communicated to shortlisted PoC participants.
182		Is there an existing Master Data Management (MDM) system or data dictionary framework in place?	Government maintains federated databases. One of the objective's of the PoC is to streamline its database management.
183		What standards or formats are mandated for semantic interoperability (e.g., JSON-LD, RDF, XSD)?	The PoC should be designed to accommodate widely accepted interoperability standards such as JSON, XML/XSD, or RDF, with flexibility to adopt specific standards to be finalized at the implementation stage.
184	Use Cases 3 & 4	Are geospatial datasets (tribal maps, census blocks, satellite imagery) available and accessible for area identification?	Yes. However, the participating firm is also expected to complement this with publicly available geospatial data and tools, where necessary, for model enhancement, validation, and ground-truthing. Further, as the question is open ended, a definitive answer to the query is not possible till selection of firm for participation in the PoC.
185		Will IP rights of PoC remain with the developer or be transferred to the government?	As per Section 7(3) of the EOI, all intellectual property rights of developed solutions will vest with the Government of West Bengal.
186		Who will provide the infrastructure?	The PoC is to be self-financed by the participating bidder.
187		Will the data be hosted exclusively on state data centers, or can secure cloud infrastructure (e.g., AWS, Azure, NIC cloud) be used?	The data will be hosted on West Bengal State Data Centre. However, the PoC is to be self-financed by the participating bidder.
188		Is there any data residency requirement (i.e., must all data reside within state infrastructure)?	Yes.
189		How many queries/conversation transactions is the department looking for in the first phase?	There are no phases. The query is not clearly understood. However, the PoC during development has to be tested for accuracy and coverage provided by the solution among other parameters. During development of the PoC, the Department would not be running any queries/conversations except for testing results of the PoC under development.
190		For initial phases of development and testing, can 3rd party enterprise solutions be used (not applicable for fraud detection)?	The PoC will be tested after development on its resource usage requirements among other parameters. However, the models and solutions deployed by a bidder are upto them.
191		For the Unified Social Registry, what type of real-time integrations with department MIS systems are expected during PoC?	No real-time integrations are expected during PoC stage.
192		Federated database management	The Finance Department will exercise strict access control and governance protocols while sharing any sensitive or classified data. NDA, confidentiality and masking of sensitive data related requirements will have to be adhered to by the PoC participants as per established Govt standards.
193		What level of controls will be exercised by the client in sharing sensitive data?	
194	Unified Social Registry for Enhanced Delivery of Welfare Services	Will synthetic or anonymized datasets be shared to simulate beneficiary identification, segmentation, and DBT use cases?	Yes
195		Will the architecture for federated databases need to be demonstrated in a live multi-departmental PoC setup?	Yes
196		What standards or taxonomies are to be followed for interoperability across schemes and departments?	Interoperability Framework for e-Governance (IFEG) – India may be followed.
197		What is the expected real-time performance benchmark for the eligibility engine?	Specific thresholds are not prescribed in the EOI. However, higher accuracy and confidence with low latency will positively influence evaluation scores.
198		Will the Department provide grievance redressal data for training the sentiment analysis model?	Any further updates will be informed to all participating firms at later stages.
199		Is compliance with the DPDP Act 2023 mandatory or will evolving standards suffice?	There is no grievance redressal portal of the Social Registry Portal as of now. However, an effort to provide data from a separate grievance redressal portal is being made.
200	Data Environment (Specific to Use Cases)	Use Case 3: What are the typical database technologies used by individual departmental schemes?	Yes, along with ISO 27001/CERT-IN alignment.
201	Data Environment (Specific to Use Cases)	Use Case 3: What is the estimated total volume of beneficiary data across all schemes, and what is the typical update frequency for this data?	The main database of the Social Registry Portal is in PostgreSQL. Different departments are using different databases, e.g., MSSQL, etc.
202	Data Environment (Specific to Use Cases)	Use Case 3: Are there any existing unique identifiers for citizens or households across different schemes that could aid in data de-duplication and harmonization?	Database is evolving. At present, it is around 22Crores. Each scheme has its own frequency, ranging from monthly to annual.
203	System Integration & Interoperability	For Use Case 3's "Federated Database Management" and "Interoperability Layer", what are the specific data sharing protocols or mechanisms currently in use between departments? What is the expected real-time synchronization requirement for data across departments?	There is no unique identifier across the databases of each department. However, in Social Registry Portal, Family ID is being used to identify the family. There are some beneficiaries who are not mapped to any Family ID due to lack of unique identifier.
204	Computing & Storage Resources	What are the data retention policies or archival requirements for the raw, processed, and generated data, particularly for sensitive information in Use Case 3?	Standard APIs are used for sharing data across different platforms of Departments. Further details will be discussed during PoC development stage with participating firms.
205	Networking & Accessibility	What are the network bandwidth and latency requirements for internal access within the Finance Department and between different departments, especially for real-time data access for Use Cases 3 and 4?	The data would be provided after anonymization, etc., for POC. Data should not be retained after POC.
206	Networking & Accessibility	Will the AI solutions be accessible only within the government's internal network, or is there a requirement for external (internet) access for certain user groups or services (e.g., citizen feedback for Use Case 3, deed drafting services for Use Case 2)?	Most of the systems of various departments are in SDC, so bandwidth issue is not there.
207	Security & Compliance	For handling sensitive data (especially in Use Case 3), what are the specific requirements for data anonymization, pseudonymization, or encryption (at rest and in transit)?	The AI solutions will be accessible over internet as well but any such access will be login based.
208	User Base & Scalability	Use Case 3 (Unified Social Registry): What is the estimated total number of beneficiaries across all 100+ schemes? What is the expected volume of real-time queries for eligibility checks or beneficiary segmentation? What is the anticipated volume of "structured and unstructured grievance inputs (calls, texts, social media)" per day for the NLP-based feedback loop?	Data would be provided in anonymized manner.
209	User Case Specific Technical Clarifications	Use Case 3 (Unified Social Registry): What is the precise definition of "near real-time" for beneficiary identification and update? For "anomaly detection in scheme utilization," what are the most critical types of "irregularities in benefit usage patterns" (e.g., specific fraud scenarios, duplicate claims, over-entitlement) that the system should prioritize detecting?	The dataset is evolving and the current count is 22 Crores. The queries for eligibility checks are not in live environment at present, so no estimate is available as of now. Grievance input volumes are not available as the same is not in the live environment.
210		Could you walk us through how your current architecture handles duplicate or inconsistent records across different schemes and departments?	"Near real time" refers to the time when the Social Registry gets updated. The "irregularities" would be intimated when the POC is being developed.
211		What challenges are you currently facing with data harmonization across your federated databases?	This would be detailed at the time of development of POC.
212		Are there any existing systems or tools you're using today for reconciling mismatched data across various programs?	Lack of unique identifier is the biggest challenge.
213		What are the database systems you are currently using (Used Case 3)	The database of the Ration Card is used as the base database for this purpose.
214		What level of accuracy and confidence do you expect from the de-duplication process?	PostgreSQL 16 is being used.
215		How do you currently expect the system to resolve conflicts when the same individual is represented differently across departments?	Specific thresholds are not prescribed in the EOI. However, higher accuracy and confidence will positively influence evaluation scores.
216		Are you looking for real-time(any streaming) de-duplication, or is a periodic batch process sufficient for your operations?	Any further updates will be informed to all participating firms at later stages.
217		How often do you expect harmonized data to be refreshed or revalidated across the federated network?	Base database of Ration Card is used.
218		If new schemes or departments are added in the future, how do you envision the harmonization process adapting?	Both real-time and periodic.
219		Are there specific formats or data structures that need to be supported immediately?	As per predetermined schedule.
220		1. Which departments and schemes will be part of the PoC (e.g., health, education, welfare, pensions)? 2. How many datasets will be provided for PoC integration, and what is their approximate size?	Yes
			Query is not specific.
			The specific departments and schemes for the PoC will be decided based on the use case and will be informed to shortlisted bidders at a later stage.
			The number and size of datasets will depend on the requirements of the PoC and availability. Necessary data will be provided as deemed required to ensure successful PoC execution.

221	Use Case 3: Unified Social Registry for Enhanced Delivery of Welfare Services	3. Will datasets include historical records or only current beneficiary lists?	Datasets may include historical records as well as current beneficiary lists, depending on the use case requirements and availability.
222		4. Will data be anonymized before sharing, or contain personally identifiable information (PII)?	Anonymized.
223		5. Will the same unique identifier (e.g., Aadhaar, ration card number, mobile) be present across datasets to enable matching?	Ration Card is used as base.
224		6. Will metadata/schema definitions be shared in advance for each dataset?	Will be discussed at the time of development for PoC.
225		1. How will data be shared — APIs, database connections, flat files?	All necessary data and documents for PoC execution will be shared in formats deemed appropriate by the Finance Department, depending on availability and security considerations.
226		2. Will departments expose live/federated APIs for the PoC, or will we work on static snapshots?	Direct access to live production systems would not be provided during PoC.
227		3. Is there a data dictionary available for each participating scheme's database?	No
228		4. What is the frequency of updates expected during PoC (real-time, daily batch, ad-hoc)?	During the PoC stage, there is no predefined frequency of live updates from departmental systems. Data will typically be provided as static datasets or controlled snapshots sufficient for PoC development. If periodic refreshes are necessary for testing, these may be shared on an ad-hoc basis at the discretion of the Finance Department. Real-time or daily live updates would only be relevant for post-PoC implementation.
229		1. For beneficiary segmentation — what are the primary grouping criteria (income, geography, demographics)?	Grouping is as per scheme criteria.
230		2. For GIS-based targeting — will geolocation data be provided in beneficiary records?	Not available with all the scheme data.
231		3. For interoperability — should the PoC demonstrate cross-department updates (e.g., change in one DB reflected in others)?	Would be discussed during development for PoC.
232		4. Should the system support duplicate detection and de-duplication across datasets in the PoC?	Will be discussed at the time of development for PoC.
233		1. Should anomaly detection be applied to eligibility fraud, beneficiary duplication, or transaction misuse?	Yes
234		2. For grievance analysis — will we receive past grievance text data, and in which languages?	The data would be provided as far as practicable.
235		3. Should sentiment analysis be included in grievance handling?	Would be discussed during development for PoC.
236	Unified Social Registry for Enhanced Delivery of Public Welfare Services	1. What are the success metrics — data match rate, eligibility accuracy, anomaly detection precision/recall, grievance classification accuracy?	Will be discussed at the time of development for PoC.
237		2. Will there be a hidden test dataset for evaluation?	No
238		3. What is the expected user interface — dashboard, API, reports?	The PoC results are expected to be demonstrated in a clear and practical manner that showcases the effectiveness, feasibility and applicability of the proposed solution. Bidders are encouraged to prepare a live demo, accompanied by a presentation and brief documentation explaining the approach, architecture, and expected outcomes. Any further updates at any later stages will be informed to all firms who qualify for developing the PoCs
239		Q1. The "Geo-Spatial Intelligence" module aims to identify underserved geographies. How does the Department currently define an "underserved" area?	As per scheme criteria
240		Q2. For the "AI/ML-based beneficiary identification and updation module," could you provide examples of the "dynamic socioeconomic indicators" and "scheme-specific criteria" that will be used?	Would be discussed during development for PoC.
241		Q3. For the "Anomaly Detection in Scheme Utilization" module, does the Department have a set of known fraud patterns or business rules (e.g., one person receiving benefits under mutually exclusive schemes) that can be used to train or guide the initial models?	Would be discussed during development for PoC.
242		Q5. How will the success of the PoC for this use case be measured? For example, will it be based on the percentage reduction in duplicate beneficiaries identified in a sample dataset, or the accuracy of the eligibility engine against a pre-verified beneficiary list?	The higher the confidence threshold and lower the error rate, the better it may be for favourable evaluation. There are no minimum acceptable benchmarks defined in the EOI. Further updates, if any, would be informed to all participating firms during PoC development
243		Q6. The EOI specifies high-level categories such as 'risk,' 'vulnerability,' and 'lifecycle stage.' While we understand that our role is to use machine learning to identify the optimal data parameters for this segmentation, could the Department provide any existing business rules, definitions, or frameworks it currently uses to define these categories? For example, is there a formal definition of what constitutes a 'high-risk' household?	Would be discussed during development for PoC.
244		Beneficiary Data Sources	Identify the departments/schemes providing beneficiary data and the formats in which they will be shared (CSV, API, database dump).
245		Data Update Frequency	Confirm how often beneficiary data is updated in each department, and whether real-time sync is possible.
246	Beneficiary Data Sources	Privacy & Data Sharing Protocols	Real time
247		Interoperability Requirements	All firms are advised to build their solutions with privacy-by-design principles and to ensure full compliance with ISO 27001/CERT-IN standards wherever applicable. However, PoC development would not be affected due to any restricted zones.
248		Segmentation Parameters	The detailed structure of existing MIS databases and data dictionaries will not be shared publicly in the EOI stage. For PoC, bidders will be provided with sample extracts. Full database schemas, field mappings and data dictionaries will be shared with selected bidders at the implementation stage, under NDA and subject to security clearance, to ensure accurate integration.
249		GIS Data Availability	The AI model is expected to categorize beneficiaries by risk, vulnerability, and life-cycle stage
250		Anomaly Detection Scope	Available data will be provided. However, it is expected that the bidder will propose a methodology for either using existing GIS data (if available) or creating geospatial coverage maps as part of the solution to achieve the stated objectives. The necessary data will be provided during the Proof of Concept (PoC) stage for the successful development of the solution.
251		Feedback & Grievance Sources	Would be discussed during development for PoC.
252		Integration Approach for PoC	Citizen feedback is collected via multiple platforms: helplines/call centres, online portals, mobile app etc. The necessary data and documents will be provided as required for the successful execution and development of the PoC.
253			No real-time integrations are expected during PoC stage
254			
255			
256	Preserving Institutional knowledge	For the Generative AI and Document Intelligence system for Finance (Usecase 4), is the primary objective accurate document-level retrieval with citation, or should the PoC also demonstrate deeper analysis like policy comparison or redundancy detection	The PoC is expected to demonstrate both accurate retrieval with citation and the feasibility of deeper analytical capabilities.
		Are there any recommendations from client on any techniques to be used for preserving Institutional knowledge?	As outlined in the EOI document.
		Should the analytics be performed in real time, near real time or in batch mode?	The type of analytics to be performed would depend on the user. However, real time and near real time analytics features would be required.
		Should the final solution be on-prem, cloud or hybrid setup?	The solution should be capable of being hosted and deployed at the West Bengal State Data Centre.

257	4. Deliverables: Use Case 4 Page No. 5, section 4 Automated proposal examination tool using Enterprise GPT with Conversational with context Chat-like interface for official users with technologies like NLP (Natural Language Processing) Layer which allows users to ask queries directly in English, and it should be able to handle reasonably Complex Queries with semantic understanding (not just keywords-based search) using latest technologies like LLM-RAG while ensuring it does not give confident wrong answers (i.e. control AI hallucinations).	LLM will be used to build the solutions. Please confirm if there is any preference for open source models or proprietary	This is open to innovation. There are no recommended solutions. However, efficiency metrics including compute performance and cost of operation may form a part of the PoC evaluation criteria of the PoC Evaluation Committee.
258	Enterprise AI Model for File Analysis and Decision Making	Will sample files or historical case records be shared for training and testing the proposal analysis module?	Yes
259		Should the Enterprise GPT model handle both English and Bengali inputs, especially for voice interfaces?	Yes
260		Are there any restrictions on the use of external LLM APIs (e.g., OpenAI, Cohere) or is on-premise hosting preferred?	As such there are no overarching restrictions specified in the EOI. However, the following may be noted: As per Section 7(3) of the EOI, all intellectual property rights of developed solutions will vest with the Government of West Bengal. NDA, confidentiality and masking of sensitive data related requirements will have to be adhered to by the PoC participants as per established Govt standards. Moreover, the solution should be capable of being hosted and deployed at the West Bengal State Data Centre.
261		What are the evaluation criteria for hallucination control in LLM outputs?	Hallucination control will be checked during PoC development also before evaluation. The Outputs may quote the references from the relevant context and/or underlying documents referred by the Model for generating the output. Other detailed criteria could be outlined by the PoC Evaluation Committee. Any further updates will be informed to all participating firms at later stages.
262		How frequently are circulars/OMs updated, and will a central repository be provided to integrate dynamic learning?	There is no predefined or fixed frequency when circulars/OMs get updated. However, feature for uploading updates by pre-designated users may be kept. Access to central repository through APIs would only also be provided in post-POC stage.
263		What are the expectations around citation and para-highlighting from the knowledge repository?	Yes. However, as the question is open ended, a definitive answer to the query is not possible.
264	Data Environment (Specific to Use Cases)	Use Case 4: Can you provide a more detailed inventory or categorization of the "Information/Data Sources" from the organization?	Yes, details will be provided during PoC development.
265	System Integration & Interoperability	Are there existing APIs or web services available for programmatic access and integration with departmental databases or other enterprise systems for any of the use cases, particularly for Use Case 4's "integration with existing database and APIs"?	Integration will be done at post-POC stage. For PoC development sample data based testing would suffice along with API development (but no integration) which can pull data from a trusted source/existing database whenever integration happens.
266	User Base & Scalability	Use Case 4 (File Analysis): What is the estimated number of unique officials who will utilize the Enterprise GPT chat-like interface within the Finance Department? What is the expected daily query volume for the AI model?	Around 500 unique users are there. Not everyone would utilise it daily perhaps. Expected daily query volume may be upto 10000.
267	Use Case Specific Technical Clarifications	Use Case 4 (File Analysis): Can you provide more concrete examples of "reasonably Complex Queries" that the Enterprise GPT chat-like interface should be able to handle with semantic understanding? What is the acceptable "confidence interval" or "error rate" for answers provided by the AI model, and how will "control AI hallucinations" be specifically measured or validated during the PoC evaluation?	Yes, concrete examples will be provided during PoC development. Some parts of the query form part of the PoC evaluation. Therefore, the higher the confidence threshold and lower the error rate, the better it may be for favourable evaluation. There are no minimum acceptable benchmarks defined in the EOI. Further updates, if any, would be informed to all participating firms during PoC development. Hallucination control will be checked during PoC development also before evaluation. The Outputs may quote the references from the relevant context and/or underlying documents referred by the Model for generating the output. Any other detailed criteria that would be outlined by the PoC Evaluation Committee. Any further updates will be informed to all participating firms at later stages.
268	Use Case 4	What types of documents will be included in the PoC (circulars, orders, manuals, case files, meeting notes, scanned PDFs)?	Circulars, Orders, Statutes, digitally readable and scanned PDFs are expected to be included. NDA, confidentiality, and masking of sensitive data will apply as per established Government standards. The exact list and formats will be shared with shortlisted bidders at the PoC development stage.
269		Will we receive both structured and unstructured content, or only unstructured text?	A mix of structured and unstructured content will be provided. However, a larger percentage will be unstructured than structured.
270		Will documents be provided in Bengali, English, or both?	Mostly English. However, some documents in Bengali will also be there.
271		What is the expected data volume for the PoC (number of documents/pages)?	The volume will be in GBs.
272		Will documents contain sensitive or classified information, and if so, will the redact versions be provided?	Sample data provided in PoC stage will not contain such information. However, NDA, confidentiality and masking of sensitive data related requirements will have to be adhered to by the firm as per established Government standards.
273		How will documents be shared — bulk download, API, secure file transfer?	Documents will be shared in formats as deemed necessary and as per availability to ensure successful PoC execution. NDA, confidentiality and masking of sensitive data related requirements will have to be adhered to as per established Government standards.
274		Will we get continuous updates during PoC, or only a static snapshot?	Direct access to live production systems would not be provided during PoC.
275		Should the system automatically index newly issued circulars/OMs in near-real-time?	Yes
276		Should the PoC demonstrate Retrieval-Augmented Generation (RAG) with citations for every answer?	The EOI does not mandate the use of Retrieval-Augmented Generation (RAG) for every answer. However, bidders are encouraged to incorporate mechanisms for citation and referencing from relevant context or underlying documents, as this will be evaluated positively.
277		For citations — should they link to specific paragraphs or entire documents?	Citations should be as specific as possible to the relevant source material paragraph/sentence/ etc. to ensure accuracy and traceability.
278		Is highlighting the exact answer text in the source document mandatory?	Enabling highlighting of the specific answer text or relevant excerpt within the source document would significantly enhance usability, transparency, and evaluation scores, especially for use cases involving knowledge retrieval.
279		Should the PoC support multi-document summarization?	Yes
280	Use Case 4: Enterprise AI Model for File Analysis and Decision Making	Is Bengali voice input required for PoC or can it be part of full implementation?	Bengali voice input is required for successful PoC development. The Enterprise GPT model should handle both English and Bengali inputs (including voice) where applicable. Implementing Bengali voice input at PoC stage could enhance evaluation scores. In the worst case, it may be proposed for the full implementation phase if not found feasible at all during PoC, provided the core multilingual capability is demonstrated in the PoC.
281		Should the system support multilingual queries (ask in Bengali, get English content, and vice versa)?	Yes. As indicated in earlier clarifications (e.g., Use Case 4), the system should be able to handle both English and Bengali inputs, including for voice interfaces where applicable. This would naturally extend to supporting cross-language queries — for example, asking in Bengali and retrieving English content and vice versa — provided the underlying data and translation modules meet the required accuracy, privacy and security standards.

282		What are the acceptance criteria for hallucination control — e.g., maximum X% incorrect answers without citation?	Hallucination control will be evaluated during PoC development and before final assessment based on similar metrics. Specific thresholds are not prescribed in the EOI. However, higher accuracy and confidence will positively influence evaluation scores. Any further updates will be informed to all participating firms at later stages.
283		Should the system refuse to answer if confidence is below a threshold?	Yes
284		Are there specific evaluation metrics (perplexity, grounding accuracy, citation recall) for this PoC?	PoC evaluation criteria will be the same across bidders. Any further updates will be informed to all participating firms at later stages.
285		Will we need to integrate feedback capture for incorrect answers during PoC?	Yes, ideally when feedback is given explicitly, it should help train the Model.
286		Should the PoC integrate into an existing government portal or be a standalone interface?	Integration I envisaged during PoC development. Post-PoC integration may be required if the solution is chosen for full-fledged rollout. However, any integration would be through standard APIs where deemed necessary after PoC development. Details would be discussed during PoC development as & when required.
287		Is there a preferred vector database or search engine (e.g., Elastic, Pinecone, Azure Cognitive Search)?	There is no preferred vector database or search engine. Bidder may use any suitable technology stack that meets the functional, performance and security requirements as per industry standards.
288		Are there data residency requirements— must all data and indexes stay within India?	Yes
289		Are there approved LLMs or providers, or can we choose freely as long as they meet security guidelines?	Bidder can choose freely as long as the solution meets security guidelines, adheres to NDA and confidentiality requirements and can be hosted and deployed at the West Bengal State Data Centre.
290		Should all LLM calls be self-hosted to avoid sending data to public APIs?	Yes, ideally. All LLM calls must comply with the Government's data security and confidentiality requirements. Solutions should be capable of being hosted and deployed at the West Bengal State Data Centre, so using public APIs is not preferred unless they fully meet these requirements and keep all data within approved infrastructure.
291		What is the size of the data to be processed (ballpark estimate of the number of pages across the set of documents) ?	For PoC development, the volume of data or number of pages to be processed will depend on the robustness of the model deployed and the results achieved. Data and documents are available in sufficiently large numbers to enable meaningful PoC execution. However, providing a fixed ballpark figure at this stage is not possible, as requirements will vary depending on the approach and methodology adopted by each participating firm. Post-PoC, the size of data to be processed is in GBs.
292		What are the different formats in which the current data is stored ? Is it just PDF (text), PDF (Image) etc. ? Or does it have other formats of documents also ?	The format is typically PDF (text), PDF (image), etc. There could be MS Office file formats also in a few cases. NDA, confidentiality and masking of sensitive data will apply as per established Government standards. The exact list and formats will be shared with shortlisted bidders at the PoC development stage.
293		Will OCR (Image to Text conversion) - Optical Character Recognition) for converting scanned/text be required ?	It may be required in some cases.
294		Q1. The scope includes automating the analysis of proposals. What are the key criteria the system should check for? For example, should it check for completeness against a checklist, adherence to financial codes, consistency with past decisions, or potential duplication with other proposals?	Yes, the system is expected to incorporate checks across multiple dimensions to assist Finance Department officials in proposal examination
295	Use Case 4: Finance Department Enterprise AI Model for File Analysis and Decision Making	Q2. Page 5 specifies the system should handle "reasonably Complex Queries with semantic understanding." Could the Department provide 2-3 examples of a simple query versus a complex query that an official might ask?	Simple Query (direct information retrieval): "Provide the latest order on TA/DA rules." Complex Query (semantic understanding and multi-step reasoning): "Has the Finance Department issued any order in the last two years regarding relaxation of austerity measure on computer procurement
296		Q3. The EOI mentions a "Learning system and Admin View so that the system improves as more data passes through it." Could you clarify the expectation for this "learning"? Is it based on explicit user feedback (e.g., a thumbs up/down on an answer), or should the model learn implicitly from user query patterns? What key metrics should the Admin View display?	The "learning system" is expected to encompass both explicit and implicit feedback mechanisms, within the constraints of Government data privacy and security policies
297	Use of Cloud-based LLM Models	Whether cloud-hosted Large Language Models (e.g., OpenAI, Azure OpenAI) can be used for the enterprise GPT-like chat interface, or if the deployment must be fully on-premises due to data sensitivity and policy restrictions.	As such there are no overarching restrictions specified in the EOI. However, the following may be noted: As per Section 7(3) of the EOI, all intellectual property rights of developed solutions will vest with the Government of West Bengal. NDA, confidentiality and masking of sensitive data related requirements will have to be adhered to by the PoC participants as per established Govt standards. Moreover, the solution should be capable of being hosted and deployed at the West Bengal State Data Centre.
298	Document Types & Formats	Clarify the types of documents to be processed (PDF, DOCX, XLSX, scanned images, emails) and whether they include embedded images, tables, or charts.	Solutions must be capable of handling all type of data format, both structured and unstructured content will be there. It must support text, tabular and image-based information.
299	Document Quality & Structure	Confirm if the majority of documents are well-formatted and machine-readable, or if significant preprocessing (OCR, layout detection, cleaning) will be required.	the majority of documents for Use Case 4 are likely well-formatted and machine-readable. Some files may contain scanned annexures, signatures or older scanned references
300	Evaluation Flow & Criteria for Proposals	Request any SOPs, manuals, or official guidelines that define proposal evaluation workflows, and details of the criteria or rules used to identify inconsistencies or gaps in proposals.	At present, the Finance Department does not have a single consolidated SOP/manual governing proposal evaluation workflows. Proposal evaluation is carried out by relevant Groups, Branches, Cells, and Directorates, each following departmental practices, circulars, OMs, and policy guidelines applicable to their domain.
301	Proposal Categorization Scope for PoC	Confirm whether proposals can be grouped into categories with similar workflows and if the PoC scope can be limited to a few priority categories before scaling up in full deployment.	Yes. Proposals handled by the Finance Department may be grouped into broad categories . Many of these categories share similar workflows and evaluation criteria. For the PoC stage, bidders may be asked to focus on a limited set of priority categories that are representative of typical workflows.
302	Knowledge Management System Requirements	Identify the types of institutional knowledge to be preserved (policy documents, case histories, manuals) and whether a digitized repository already exists or needs to be created from physical records.	For Use Case 4, the institutional knowledge to be preserved primarily includes: Policy documents, circulars and office memoranda (OMs). Proposal examination records, case histories and file notings. Departmental guidelines and manuals relating to financial management and administrative procedures. Clarifications and orders issued by the Finance Department over time. A significant portion of this knowledge already exists in digital form. However, these are currently siloed and fragmented rather than consolidated.
303	Integration with Existing Systems	List the departmental databases, MIS platforms, or APIs the PoC must connect to, and whether integration will be API-based or file-based during the PoC phase.	During the Proof of Concept (PoC) phase, no real-time integration with existing departmental databases, MIS platforms, or APIs is required