IndiaAI Financial Reporting Compliance Challenge

1. Introduction

IndiaAI, an Independent Business Division (IBD) under the Digital India Corporation (DIC) of the Ministry of Electronics and IT (MeitY), is the implementation agency of the IndiaAI Mission, which aims to democratise AI's benefits across all strata of society, bolster India's global leadership in AI, foster technological self-reliance, and ensure ethical and responsible use of AI.

As part of this Mission, the IndiaAI Application Development Initiative (IADI) aims to promote the development, deployment, and adoption of AI applications in critical sectors that have the potential to catalyse large-scale socio-economic transformation.

Towards this, IndiaAI is collaborating with the National Financial Reporting Authority (NFRA), the independent regulatory body responsible for overseeing the auditing and accounting standards of public interest entities (PIEs) in India. This partnership seeks to harness advanced AI technologies to improve overall financial reporting quality in the country. By combining NFRA's regulatory experience with IndiaAI's technical expertise, the IndiaAI Financial Reporting Compliance Challenge launched under IADI aims to augment and accelerate compliance monitoring and facilitate effective regulatory decision-making. This is a strategic effort to build a scalable regulatory technology infrastructure for the country in the interest of public trust and investor protection, while meticulously safeguarding confidentiality and ensuring adherence to statutory responsibilities.

The Challenge offers a critical platform to develop and scale practical, innovative solutions that streamline compliance processes and foster greater trust and transparency throughout the financial ecosystem.

2. Challenge Process and Participation Details

I. Stage 1 (Application):

a. Participation Requirement:

- Participants may participate as teams as per the eligibility criteria detailed in Section 7.

b. Participation Process

- All participants are required to utilise the IndiaAI portal to access the application form (draft of which is annexed hereto as Annexure II)

- A team leader will have to individually apply for the Challenge on the IndiaAI portal by clicking the submit link.
- While submitting the form, the Team Leader should list all the Team Members under the Management Team and complete all organisation details.
- Team Leaders are required to share the source code required to build, train, and test the submitted AI solution in a private repository. The Team Lead shall paste the link for their GitHub source code in the application form.
- The Team Leader will have to answer all additional questions, including uploading documents, and click 'Submit'.
- Interested applicants can apply within the specified time period from the Launch.
- Applicants are required to utilise publically available datasets to demonstrate their solution performance at this stage.
- Any edits to the Source Code post the final date of submission will lead to immediate disqualification of the application.

II. Stage 2 (Virtual Challenge)

- Shortlisted teams will be mandatorily required to sign a Non-Disclosure Agreement (NDA) to participate in both Stage 2 and Stage 3 of the evaluation process, and obtain access to the datasets.
- For Stage 2, up to 10 teams may be shortlisted on the basis of their applications and will receive INR 5 lakhs each to refine their solutions using a sample dataset curated by NFRA, including financial statements such as balance sheet, statement of profit and loss, cash flow statement, audit reports, company annual reports, and reference scenarios, which will be provided through AIKosh to evaluate solution performance.

III. Stage 3 (On-premises Round)

- For Stage 3, up to 3 shortlisted teams may be shortlisted for a 5-day on-premises round in New Delhi to refine and add functionalities to address the provided Problem Statement on a sample dataset made available via AIKosh or other chosen Cloud Environment. Details of the location and logistics will be communicated via email to the shortlisted teams. Shortlisted applicants will be required to travel to the venue on their own expense.
- *IndiaAI reserves the right to modify the number of qualifying solutions to ensure competition and operational efficiency.*

3. Problem Statement

I. Problem Statement: (Stage 1) - AI Engine for Compliance Gaps

Objective: AI-powered engine capable of extracting text, tables, hyperlinks, embedded data, and financial data from multi-format document sources in scanned and digital versions, segmenting them into logical, header-based sections and validating each section's completeness, integrity and compliance against a pre-defined framework. The solution should structure document metadata and support search, analytics, visualisation and indexing of source unstructured data to ensure efficient retrieval and referenceability.

While advanced infrastructure, such as vector databases for embeddings and encrypted storage systems to safeguard original documents, is not mandatory in this phase, solutions should be designed to accommodate these requirements in later stages, ensuring scalability, security, and compliance with data-handling standards.

Expected Output

i. Compliance Validation Report Generator: A generic Engine configurable for testing compliance across large volumes of structured or unstructured datasets with a corresponding set of regulations. The Engine should first extract and structure text, numerical and financial data from scanned and digital multi-format documents. It should segment documents into logical, header-based sections with appropriate metadata tagging, enabling structured access and navigation.

The Engine should be able to map the rules and regulations against the documents received providing an explainable output indicating the dataset's compliance and non-compliance with each provision of the rules and regulations. Examples of such rules/regulations/frameworks include Indian Accounting Standards and associated reference and educational materials, SEBI disclosure requirements, RBI Disclosure Norms for Financial Institutions, SEBI ESG Framework, among others[1]. Further, examples of datasets that reflect compliance with these rules/regulations include financial statements, statutory forms, XBRL filings, stock exchange filings,

[1] Examples of regulatory frameworks for compliance validation include Accounting Standards (here), Indian Accounting Standards - IndAs Vol 1 (here) and Vol 2 (here), Conceptual Framework - IndAS (here), Quick Referencer - IndAS (here), Educational Materials - IndAS (here), Exhaustive list of Auditing Standards (here), General Instruction for preparation of financial statements - (Refer to Schedule III here), SEBI Disclosure requirements (here), RBI Disclosure Norms for Financial Institutions (here), SEBI ESG Framework (here), and SEBI ESG Disclosures - BRSR Core (here).

and so on. The Engine should structure the regulations and datasets or guide the user in structuring them using built-in tools.

ii. Project Report: A structured document summarising the approach, methodologies, datasets used, validation results, and challenges encountered during the development of the AI Validator Tool.

II. Problem Statement (Stage 2 and Stage 3):

i. Automated Analytics Engine: To create and regularly update a structured database of financial statement and audit report-related datasets received by NFRA from the publicly identified sources. Based on this data, build tools to analyse financial performance, risk indicators, financial history, audit history, stock exchange-related activities, governance structures, significant developments, related parties, management actions and predictive analytics based on financial and non-financial information.

ii. Preliminary Examination Tool: Identify, structure and analyse real-time information regarding news, enforcement actions, legal cases, taxation-related non-compliance and whistle-blower allegations.

iii. NFRA Specific Insight Bot: Chatbot interface that offers an intuitive question–answer experience over NFRA's corpus by leveraging chunked documents and structured metadata, enabling users to query content, retrieve relevant documents and sections and explore audit findings efficiently. The bot will be enhanced to support NFRA's compliance analysis using compliance-matrix reports, manual inputs, past regulatory precedents and decisions, explanatory statements, financial-sector compliance rules, and real-time signals such as news, enforcement actions, legal cases, taxation-related non-compliance and whistle-blower allegations. The solution will be designed to meet NFRA's confidentiality and security requirements in line with applicable Indian and international data-protection laws and guidelines.

4. Application Requirements

I. Solution Code and Documentation (GitHub):

- Team Leaders are required to share the source code required to build, train, and test the submitted AI solution, report, readme file and ppt (optional) in a private repository. The Team Lead shall paste the link for their GitHub source code in the application form.
- Explanation of the key methodology and steps taken in solution development.
- Steps to grant access to your GitHub repository:

- Go to the main page of your GitHub repository.
- Click on the 'Settings' tab in the menu bar.
- In the left sidebar, select 'Collaborators'.
- Under the 'Manage Access' section, click on 'Add people'.
- In the text field, search for 'indiaaihackathon25' and add it as a collaborator.

II. Project Proposal (Application Form):

- Description of the solution, approach to addressing the problem statement, and core AI technologies used. Provide a summary of document completeness or integrity challenges. Present essential metrics (OCR accuracy, segmentation performance, etc.). Cost efficiency of the proposed OCR tool across different document types.
- Uniqueness and novelty of the solution along with replicability and scalability across public and private sectors.
- AI solution details including information about proprietary architecture and solution design, third party integration, data utilised for training and validation (including data provenance, coverage and size), and outcomes. Please outline the solution monitoring and enhancement strategy, including areas that require further refinement and the process to be adopted to integrate improvements.

5. Evaluation Process

The evaluation process for the Challenge will be overseen by a distinguished panel of jury members, comprising subject matter experts in machine learning, data science, financial reporting, and auditing. The jury would rigorously assess each submission based on predefined criteria as outlined in Annexure I to ensure a fair and comprehensive evaluation. The evaluation will ensure equitable weightage is given to both the Technical and General parameters.

Stage 1 (Application Stage)

- **Initial Screening**: Submissions would undergo an initial screening to ensure compliance with submission guidelines and solution functionality.
- **Technical Evaluation**: The jury would conduct a detailed technical evaluation of the solutions.

Stage 2 (Virtual Challenge)

- Up to top 10 shortlisted teams may be shortlisted for the Virtual Challenge Round, wherein they will receive INR 5 Lakhs each, and will be required to refine their solution on a sample dataset, provided via AIKosh

- Solution performance will be evaluated based the solution's ability to process:
    - Edge cases containing specific instances where compliance verification with selected framework(s) is particularly complex or prone to high false-positive rates.
    - Standard high-frequency workflows that the solution must automate with high reliability.

Note: *Shortlisted teams will be required to sign a Non-Disclosure Agreement (NDA) to obtain access to datasets for participation in Stage 2 and Stage 3 of the evaluation process.*

Stage 3 (On-Premises Round)

- Up to 3 Shortlisted teams may be invited to a 5-day on-premises development round in New Delhi to refine their solution on a sample dataset provided via AIKosh or other chosen Cloud Environment.
- Following the assessment of the results submitted by the shortlisted teams, 1 team may secure a chance to get a two-year work contract of up to INR 1 Crore to develop and deploy their solution for use by NFRA or the Government of India and its associated entities.
- The jury's decision would be final and binding.

IndiaAI and the jury reserve the right to modify the number of qualifying solutions at any stage to ensure competition and operational efficiency.

6. Opportunity for Applicants

- **Opportunity to Build for the Nation**: Contribute to developing innovative solutions that address critical challenges faced by the country, making a direct impact on society.
- **National Recognition**: Gain visibility and recognition from Government officials, industry leaders, and peers for your contributions and innovative ideas.
- **Networking Opportunities**: Connect with like-minded innovators, potential collaborators, and key stakeholders in the tech and innovation ecosystem.
- **Exposure to Real-world Challenges:** Work on pressing issues faced by the nation, providing practical experience and a deep understanding of real-world problems.
- **Support for Implementation**: The winning solution will get potential support in scaling and implementing the solution at a national level, bringing your ideas to life.

7. Eligibility

- **Indian Company**: The team may be an Indian company registered under the Companies Act, 2013. An Indian company must have 51% or more shareholding by Indian citizens or persons of Indian origin.
- **Start-up**: The team may be a start-up as defined in the latest notification by the Department for Promotion of Industry and Internal Trade (DPIIT), accessible at Startup India.

The participating entities shall provide proof of a proprietary solution submitting verifiable documentation (e.g., technical specifications, necessary certifications, IP and patents). Additionally, entities must demonstrate development and deployment experience of relevant solutions with previous engagements in the private and public sector.

8. Timeline

#	Activity	Timeline
1	Launch Date	TBC
2	Last Date for Online Submission	20-02-26
3	Announcement of Results of First Round	TBC
4	Round 2 (Virtual Challenge)	TBC
5	Round 3 (On-premises Round	TBC
6	Announcement of Winning Entity and pilot	TBC

9. Intellectual Property Rights

All Intellectual Property Rights (IPR) in the solution developed and submitted under the Challenge shall remain with the solution owner. NFRA and IndiaAI shall have non-exclusive, royalty-free, irrevocable and perpetual license to use, reproduce, modify, and deploy the awarded AI solution, including any IPR arising out of its use. All rights, title, and interest, including all IPR, in and to any data, datasets, inputs, or information provided by NFRA or IndiaAI for use in the development, training, testing, or deployment of the solution shall remain the exclusive property of NFRA and IndiaAI. Any derivative datasets, metadata, or outputs generated from such data and solution shall also belong exclusively to NFRA. The solution owner shall be deemed to have given a No Objection Certificate (NOC) for the above and shall also remain bound by the terms of the Non-Disclosure Agreement (NDA) with respect to such work.

10. Terms and Conditions

a. All participants must meet the outlined eligibility criteria (Section 7) and belong to legally registered entities as of launch date of the Challenge.

b. The award from this initiative can only be used by the participating teams for the purpose of AI solution development.

c. The winning entity will retain the rights to the solution/product developed subject to the intellectual property rights outlined in this document (Section 9).

d. The participants shall ensure that code is free from viruses, malware, botware or any other malicious software. The participants will not use this Challenge to do anything unlawful, misleading, malicious, or discriminatory.

e. The solution must not violate/breach/copy any copyrighted or patented concepts in the AI market.

f. The solution must not violate any data protection and governance regulations and policies.

g. The solution must be in adherence with related cybersecurity standards and guidelines of the Government of India.

h. The solution must adhere to ethical principles and guidelines for the development, deployment and use of AI technologies, including fairness, transparency, accountability, and non-discrimination.

i. The developed solution/product will be deployed in the chosen Cloud Environment and used for Union/State/UT Government entities.

j. Any new enhancements, features or innovation should be released on the chosen Cloud Environment. At all times, the updated source code shall be shared with the partnering institution for its free use.

k. The winning entity may receive a work contract of a fixed amount to support the solution development and deployment for at least two years from the issuance of the contract. The support includes manpower for end-to-end development, deployment, maintenance, and bug fixing across the entire application.

l. The winning entity shall submit progress-cum-achievement reports at quarterly intervals on the progress made on all aspects of the project, including expenditure incurred on various approved items during the two-year contract period. The scope of work, payment terms, milestones, and other contract conditions will be as agreed between IndiaAI, partnering institution i.e. NFRA and the winning entity. As both NFRA and IndiaAI are mandated to comply with the General Financial Rules (GFR) of the Government of India, the winning entity shall acknowledge and agree to comply with the requirements arising out of it. The final decision to enter

into the work contract for NFRA is at NFRA's discretion, based on the cost, quality, utility, and capabilities of the winning team. Once a contract is offered, the winning entity shall accept it based on the conditions in this document and as available in the GFR. Failure to do so may result in penalties, including blacklisting.

m. The winning entity is not allowed to entrust the implementation of this project for which the award is received to another institution, and to divert the award received from IndiaAI as assistance to the latter institution.

n. The winning entity should not enter into collaboration with a foreign party (individual/academic institution/industry) in execution of this project without prior approval of IndiaAI and NFRA.

o. The winning entity is free to market the product to any entity outside the Union/State/UT Government Organisations of India.

p. In case of any dispute on any other matter related to the project during the course of its implementation, the decision of the CEO, IndiaAI and NFRA shall be final and binding on the winning entity.

q. IndiaAI reserves the right to modify the terms and conditions of this challenge for operational feasibility and compliance to rules and regulations of the Government of India.

r. By participating in this Challenge, the participating entities understand and undertake the above commitments and agree to the terms and conditions.

2. Plagiarism and Ethics

a. Participants are expected to uphold the highest standards of ethics and integrity throughout the Challenge.

b. All work submitted must be original and developed by the participant or their team.

c. Plagiarism, or the use of someone else’s work without proper attribution, is strictly prohibited and would result in immediate disqualification.

d. Participants must ensure that their solutions are proprietary and not copied from existing projects or code repositories.

e. Moreover, the use of any external resources or pre-trained models should be clearly cited, and proper permissions should be obtained where necessary. Adherence to these ethical guidelines ensures a fair and competitive environment for all participants.

f. By registering for this Challenge, participants are giving an undertaking to adhere to all plagiarism and ethical guidelines set forth by the IndiaAI.

Annexure I: Evaluation Parameters

I. General

	Parameter	Description
1	**Approach Towards Problem Solving**	Product Idea, Methods adopted, Simplicity of Final Solution, Uniqueness of Idea, Novelty of Approach.
2	**Solution Technical Feasibility**	Product features, Scalability, Interoperability, Enhancement & Expansion, Underlying Technology Components & Stack and Futuristic Orientation, and System Integration plan.
3	**Product Cost**	Potential Cost to Build, Deploy and Maintain Product for two years from issuance of the contract, and the cost of enhancements, etc, for two years.
4	**Team Ability & Culture**	Prior Experience of the entity in developing and deploying similar solutions in private and public domains, Team Leader's Effectiveness (i.e. Understanding of subject matter, Ability to guide, Ability to present idea), Ability to scale up and market the product, Growth Potential of Organisation.
5	**Adherence to Responsible AI Principles**	Safety and Reliability, Equality, Inclusivity and Non-discrimination, Privacy and Security, Transparency, Accountability, Protection and Reinforcement of positive human values.
6	**Adherence to Data Policies and Cyber Security Guidelines**	Adherence to applicable Government of India policies, guidelines, regulations on Data Governance and Cyber Security. Regulatory compliance.

II. Technical

	Parameter	Description
1.	**Data preparation**	Participant has • Explored the data and removed unnecessary columns. • Checked if there is any skewness in the data and tried to mitigate it.

		• Performed a stratified train-test split successfully to create train & test datasets. • Ensured data coverage, size, and quality. • Integrated data management and governance policy
2.	**Solution Building**	• Participants have performed the required cross-validation and have built different solutions on raw data. • After evaluation on the raw dataset, Solution hyperparameters are tuned using correct principles and the approach is explained clearly. • A reasonable number and variety of different AI/ML technologies are attempted, and the best one is chosen based on key performance metrics.
3.	**Solution Evaluation**	• Solution evaluation is conducted using an appropriate metric. • Integrated Solution monitoring and enhancement strategy.
4.	**Code readability and conciseness**	Efficient, concise code is written. The code is well documented, and the analysis is explained in the report format with findings from the dataset.
5.	**Technical Robustness**	Applicants may refer to the below parameters to assess and share solution performance as part of the application process: - Data Extraction: a) Character Error Rate (CER) on standard public benchmarks for scanned documents, such as the ICDAR 2019 ArT dataset or the SROIE dataset, or equivalent. b) Key Information Extraction - Strict, entity-level F1-score on the FUNSD (Form Understanding in Noisy Scanned Documents) benchmark, or equivalent. c) Extraction Accuracy - % of correctly extracted text, tables, hyperlinks, and numerical data vs. ground truth d) Latency - Average processing time per page/document

		e) Configurability - Ability to adapt compliance tests f) Scalability - Performance with increased volume or new data source g) Financial data recognition rate - Accuracy in detecting and tagging financial values, tables, and XBRL elements - Document Segmentation: a) Mean Intersection over Union (mIoU) for segmenting document regions (e.g., tables, text, figures) on the PubLayNet or DocLayNet benchmarks or equivalent. - Generation of Structured Output like Flagging Outcomes & Summarisation: a) ROUGE-1, ROUGE-2, and ROUGE-L scores on a standard summarisation benchmark like CNN/DailyMail/Xsum or equivalent. Also, BERT Score. b) Macro-F1 and Matthews Correlation Coefficient (MCC). c) Provide confusion matrix, a table providing a detailed breakdown of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). d) Specify if the solution can provide explanation for a 'Non-Compliant' flag, especially, citing rules or text passages that led to its decision. Additional Criteria: Any other metrics as agreed upon by jury members.

Annexure 2: Application Form - IndiaAI Financial Reporting Compliance Challenge

Section 1: Organisation Information

1. Applicant Type (Startup/Company)*:
2. Organisation Information*:
    - Name of Organisation:
    - Registration Number:
    - Date of Incorporation:
    - Number of Employees:
    - Core Function of Organisation:
    - Address:
    - Website:
3. PoC Information*:
    - Full Name:
    - Designation:
    - Core Expertise Areas:
    - Email Address:
    - Phone Number:
    - LinkedIn Profile:
4. Organisation Member Information: List details of members to be involved in this challenges
    - Name
    - Role
    - Email
    - LinkedIn Profile
5. Prior Experience in relevant project implementation and research work*: Describe relevant AI solutions designed and developed, technologies used, and outcomes, relevant publications or patents, etc. (max 200 words).

6. Experience Collaborating with Government Entities*:
Specify partners, nature of engagement, and key results achieved (max 200 words).

7. Demonstrated expertise in Financial Data Analytics, if any*: (max 100 words)

Section 2: Project Proposal

1. Approach to Addressing the Problem Statement*: Explain how your solution utilises AI for:
    a) Data Extraction
    b) Document Segmentation
    c) Generation of Structured Output & Summarisation

    (Max 300 words)

2. Description of AI Solution*: Provide a comprehensive overview of your AI solution, including:
    - Functionality
    - Features
    - Core AI technologies used
    - Training and validation data used, highlighting data provenance, coverage, size and quality
    - Process and strategies adopted for data preparation, AI/ML technologies selection, training, hyper-parameter tuning, refinement, solution monitoring and enhancement
    - Solution replicability across multiple sectors for relevant use cases *(Max 300 words)*

3. Upload Solution Architecture Diagram*

4. Upload technical performance metrics measured (for each task such as Data Extraction, Document Segmentation and Generation of Structured Output & Summarisation based on standard public benchmarks), the methodologies used for measurement, and the outcomes in the format provided in Annexure III, Part A of the Submission Guidelines.* *Please refer to the 'Technical Robustness' section of the Evaluation Parameters in the Schema Document.*

5. Proprietary Solution*:
    - Is the AI solution developed in-house (not based on third-party pre-trained models)?
        - ☐ Yes ☐ No
    - If Yes, provide:
        - Details of a proprietary technology base *(max 100 words)*
        - If the solution is developed on open-source models, share details and customisation approach *(max 100 words)*
        - Details of proprietary data utilised for training and validation, along with explicit confirmation and evidence of adherence to all relevant Indian laws and compliance standards. *(max 100 words)*

- If No, provide:
  - Names of the third-party models or components used, share specific licensing agreements that govern their use *(max 100 words)*
  - refinement approach *(max 100 words)*
  - Data sourcing, coverage and validation approach *(max 100 words)*

6. Data Governance and Security*: Describe how data collection, confidentiality, encryption, storage, access control, retention and removal will be implemented. Include measures taken to ensure compliance with relevant regulations and standards for data privacy and security measures *(max 100 words)*

7. Scalability and Integration Readiness*
   Describe deployment mode, integration compatibility, offline operability, and future expansion capability. *(max 100 words)*

8. Compliance with Responsible AI Principles*
   Describe how the solution adheres to principles of fairness and transparency and adopts measures for solution interpretability, auditability, inclusivity and fairness. *(max 100 words)*

9. GitHub Link:

Section 3: Supporting Documents (Upload)

1. Certificate of Incorporation/Legal Registration*
2. Certificate of Recognition (for Startups)
3. Technical Documentation/Proof of IPs/Patents, if any
4. Ethics/Regulatory Clearance (if secured)
5. Solution demo video (2-3 minutes)*
6. Any additional documentation to strengthen the proposal such as solution accuracy proof etc.
7. Estimation of Initial manpower cost for a two-year deployment period after stage 2. Estimation of other Potential costs to build, deploy and maintain the solution, along with the System Integration plan *in the format provided in Annexure III, Part B of the Submission Guidelines*.*

Section 4: Declaration

Declaration by Team Leader:

- I/We declare that all the information provided in this application is true and complete to the best of our knowledge. I hereby also declare that I am authorised by my company/startup to participate in this Challenge, sign all the documents and agreements related to the Challenge and to commit resources.
- I/We confirm that we will abide by the conditions mentioned in the Challenge document in full and without any deviation.
- I/We shall observe confidentiality of all the information passed on to us in the course of the Challenge and shall not use the information for any other purpose than the current Challenge.
- I/We confirm that we have not been blacklisted/banned in the last three years by any State/Central Government organisations/Firms / Institutions/ Central PSU / PSE.

☐ I Accept

Annexure III: Template for Solution Technical Performance Matrix and Cost for Development, Deployment and Maintenance and Support

A. Template for Solution Technical Performance Matrix

1. Data Extraction

Metric	Definition	Value/Benchmark	Methodology Used
Metric 1: Character Error Rate (CER)	The percentage of incorrectly recognised characters (substitutions, insertions, deletions) out of the total number of characters in the ground truth text. A lower CER is better.	ICDAR 2019 ArT dataset or SROIE dataset, or equivalent	
Metric 2: Key Information Extraction (Strict, entity-level F1-score)	A metric that combines Precision and Recall to measure the accuracy of extracting specific entities (e.g., "Total Amount," "Vendor Name"). The F1-score is the harmonic mean of both, providing a single score for Solution performance.	FUNSD (Form Understanding in Noisy Scanned Documents) benchmark, or equivalent	
Metric 3: OCR Operational Cost	Cost per page character recognition		
Any other Metric			

2. Documentation Segmentation Metrics

Metric	Definition	Value/Benchmark	Methodology Used

Metric 1: Mean Intersection over Union (mIoU)	For segmenting document regions (e.g., tables, text, figures)	PubLayNet or DocLayNet benchmarks or equivalent	
Any other Metric			

3. Generation of Structured Output & Summarisation

Metric	Definition	Value/Benchmark	Methodology Used
Metric 1: ROUGE-1, ROUGE-2, ROUGE-L	Measures the overlap of unigrams (individual words), bigrams (pairs of consecutive words), and Longest Common Subsequence (LCS) between the generated summary and a set of reference summaries.	CNN/DailyMail/Xsum or equivalent	
Metric 2: BERT Score	An automatic evaluation metric that computes the similarity between tokens in a candidate summary and a reference summary using BERT embeddings, capturing semantic similarity beyond simple word overlap.	CNN/DailyMail/Xsum or equivalent	

Any other Metric			

B. **Template for Cost for Development, Deployment and Maintenance and Support**

Initial cost and Manpower cost for further development, deployment, maintenance and support for two years from issuance of the contract as per IndiaAI Financial Reporting Compliance Challenge.

(All in INR, excluding GST, and including any overheads, employee benefits, other taxes and levies)

1. Cost of development/refinement of the solution

Sl. No.	Cost Component	Rate	Total

2. Cost of deployment and further enhancements, operations, maintenance and support after deployment (including human resource, security compliance, licensing and overheads but excluding hardware, cloud, connectivity, and GST)

Sl. No.	Cost Component	Monthly Rate	Total estimated rate for two years