

Public Sector Accounting Assessment

PULSE

IT SUPPORT SYSTEM FOR THE PUBLIC SECTOR
ACCOUNTING AND REPORTING PROGRAM (PULSAR)

INTELLIGENT AUGMENTATION SUPPORT

Pilots completed & Statistics

Institution: MFU
Ministry of Finance of Ukraine

Country: Ukraine

Institution: MoF
Ministry of Finance of Armenia

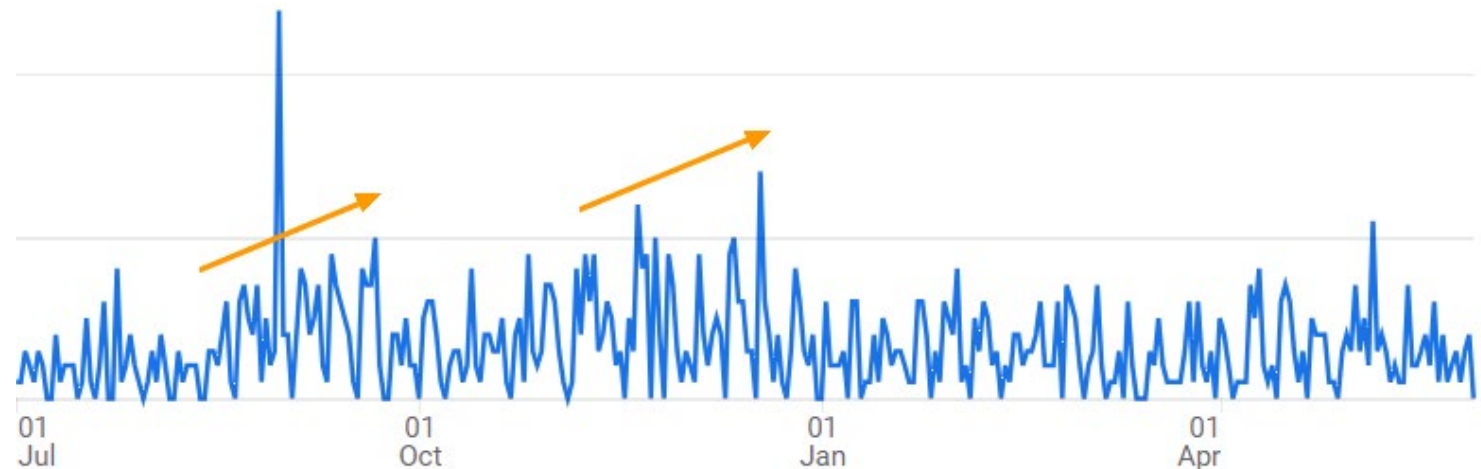
Country: Armenia

Users ?
226

New users ?
209

Average engagement time ?
34m 55s

Total revenue
\$0.0 >



































<u>COUNTRY</u>	<u>USERS</u>
United States	70
Armenia	17
Ukraine	14
Netherlands	9
China	8

New users per country

New methodology PULSAR 2024

Pulse Methodology Final March 2024

Pillars	Indicators	Dimensions
  I. PSA Framework	  PI-1. Conceptual Framework	  1.1. Conceptual Framework   1.2. Qualitative Characteristics   1.3. Reporting Entity   1.4. Elements of financial statements
	  PI-2. Accounting policies, estimations, errors and events after the reporting date	  2.1. Accounting policies   2.2. Estimations   2.3. Errors   2.4. Reporting of events after reporting date
	  PI-3. Foreign currency transactions and inflation effects	  3.1. Initial recognition of exchange differences resulting from foreign currency transactions   3.2. End of year reporting of foreign currencies transactions   3.3. Compliance with hyperinflation requirements

Future work (draft)

- Conversational interface for dimension review and revision.
- Support area with access to FAQ and option to send questions.
- Reassign tasks and operations from users that are no longer part of an assessment.

- AI/LLM (Large Language Models) supported enhancements:
 - Generate a draft of the concept note providing minimal information.
 - Validate the justifications and explanations for the dimension scoring as support for evaluators.
 - Provide initial recommendations for an indicator based on all its dimensions and their scores.

Motivation

PULSE assessments are inherently **complex**, time-intensive, and require specialized expertise in public sector accounting and financial reporting.

Significant effort and precision, presenting challenges in achieving timely and accurate evaluations.





LLMs

Integrating LLMs offers a solution to streamline the PULSE assessments by **automating** data collection and analysis, enhance accuracy and consistency, and reduce the reliance on specialized expertise.



LLMs

Large Language Models (LLMs) are advanced AI systems capable of **understanding, processing, and generating human language**, which can automate and enhance various tasks such as data analysis, report generation, and quality assurance in financial assessments.

- OpenAI ChatGPT
- Anthropic Claude
- Google Gemini
- Meta Llama



AI/LLMs do not Replace. They Complement

Efficiency in Data Collection: LLMs can quickly scan and extract relevant financial information, standards, and practices, significantly reducing the time and effort required by human team members.

Consistency and Accuracy: LLMs ensure data is uniformly evaluated according to the stringent standards set by the PULSE framework.

Automated Reporting: LLMs can ensure that the report is coherent, integrated, and follows the PULSE handbook structure without merely collating separate individual contributions.

Real-Time Updates and Revisions: LLMs can efficiently handle revisions, ensuring that the most current information is always used.



AI/LLMs do not Replace. They Complement

Enhanced QA Processes: LLMs can be programmed to cross-check data against IPSAS standards and other relevant benchmarks. They can identify discrepancies, flag potential issues, and suggest corrections.

Thorough Peer Review Integration: LLMs can track changes, generate matrices of comments and responses, and ensure feedback is systematically addressed.

Insight Generation: LLMs can provide advanced insights by analyzing trends and patterns within the data. This can help identifying areas of improvement that may not be immediately apparent through manual analysis.

Scenario Planning and Forecasting: LLMs can assist in creating various scenarios based on current data, helping governments to plan for future funding requirements and manage fiscal risks more effectively.



AI/LLMs do not Replace. They Complement

Training and Capacity Building: LLMs can serve as training tools for the assessment team, providing on-the-spot explanations and context for complex standards and methodologies.

Reduced Timeframes: By automating many of the time-consuming tasks, LLMs can significantly shorten the assessment process, adhering to the six-month timeframe recommended by PULSE.

Leveraging Global Knowledge: LLMs can access and analyze a vast array of case studies and best practices from around the world, providing valuable insights and benchmarks that can be applied to the local context.

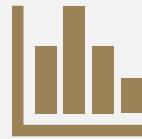
Practical Examples



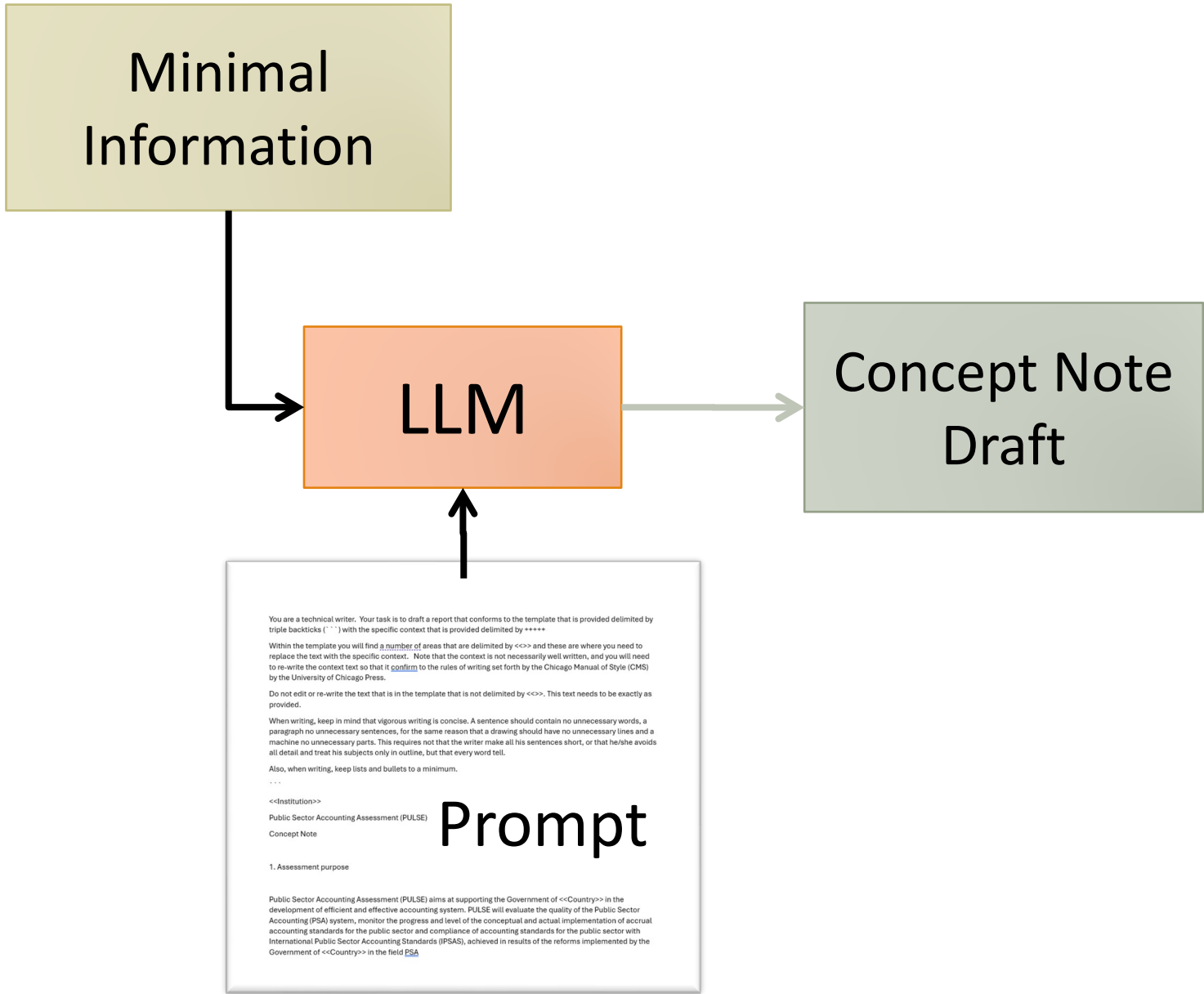
Generating a Draft Concept Note



Validation of justification for a
specific dimension



Preliminary Analysis,
Recommendations and Draft
Report



Generating a Draft Concept Note - Prompt

Ministry of Finance of Georgia Public Sector Accounting Assessment (PULSE) Concept Note

Assessment purpose

Public Sector Accounting Assessment (PULSE) aims at supporting the Government of Georgia in the development of an efficient and effective accounting system. PULSE will evaluate the quality of the Public Sector Accounting (PSA) system, monitor the progress and level of the conceptual and actual implementation of accrual accounting standards for the public sector and compliance of accounting standards for the public sector with International Public Sector Accounting Standards (IPSAS), achieved in results of the reforms implemented by the Government of Georgia in the field of PSA.

Background

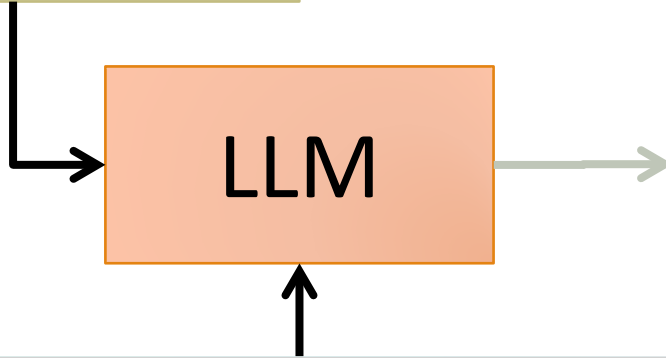
Georgia is a country in the Caucasus region with a population of approximately 3.7 million people (as of 2024), with around 40% residing in rural areas. Georgia's economy has demonstrated considerable resilience. The adoption of modern technologies and the integration of advanced practices in business processes have contributed to the stability and dynamic growth of the Georgian economy.

Throughout 2021 and 2022, the economy contended with the impacts of the COVID-19 pandemic, not yet fully recovering from previous regional conflicts. The Georgian government's anti-crisis and stimulus measures enabled successful operations amid quarantine restrictions and other external pressures, maintaining macroeconomic stability and positive growth trajectories. Georgia concluded 2022 with a record GDP of \$19 billion, \$1.2 billion in direct foreign investment, and a 15% increase in exports of goods and services. Real growth in the average nominal salary of full-time employees is 7%.

The ongoing regional instability, including tensions with neighboring countries, heightened the need for efficient and timely information processing to ensure effective state functioning. The challenges posed by geopolitical tensions have tested Georgia's state structures. The Georgian government has maintained macroeconomic stability through active engagement with international support. This engagement has helped mitigate large-scale production halts and the outflow of labor caused by internal displacements.

Generating a Draft Concept Note - Result

**Justification /
Explanation for a
dimension score**



**Assessment of
justification
against
methodology**

You are an expert assessor for PULSE. Scoring of the dimensions is the heart of the PULSE process. The scoring and assessment methodology has been aligned to the PEFA methodology, including a four-point calibration scale between A and D for each dimension.

Each dimension is scored separately on a four-point ordinal scale: A, B, C, or D, according to precise criteria established for each dimension. To justify a particular score for a dimension, every aspect specified in the scoring requirements must be fulfilled. If the requirements are only partly met or the criteria are not satisfied, a lower score is given that coincides with achievement of the requirements for the lower performance rating.

For those indicators referring to IPSAS, to achieve an "A" score, full compliance with the respective requirements of the standards is necessary. The "B" score is awarded for a high degree of compliance, but still certain gaps to the IPSAS requirements exist. A "C" score reflects the basic level of performance for each indicator and dimension, consistent with good international practices, which is aligned with what is established in the respective IPSAS for those indicators (see Table 6 for IPSAS references). A "D" score means that the feature being measured is nonexistent or presented at less than the basic level of performance.

The "D" score indicates performance that falls below the basic level. A "D" score is warranted when sufficient information is not available to establish the actual level of performance. The "D" score, due to insufficient information, is differentiated from "D" scores for lower-level performance by the use of an asterisk following the letter: "D*." The aggregation of an indicator with dimensions containing "D" scores is no different from aggregation with other "D" scores. Aggregate indicator scores do not include an asterisk, and thus cases of insufficiency of information are only denoted at the dimension level.

The assess performance level and evidence for scoring of each dimension, the assessment team members should focus on the analysis of performance against PULSE criteria, provided in the MG and the web-based tool itself. The Justification/Explanation section of each dimension should provide factual evidence, including quantitative data, if available, which has been used to substantiate the assessment. The information provided should be as specific as possible, for example, in terms of quantities, proportions, and time spans. Any issues in relation to the timeliness or reliability of data and evidence should be noted. If insufficient information has been obtained either for a whole indicator or one of its dimensions, the text should explicitly mention it. A general suggestion is that the wording provided in Justification/Explanation section could be directly used for drafting PULSE report.

Prompt

Validation of
justification for a
specific
dimension -
Prompt

The justification provided for the conceptual and actual scores of "A" for Dimension 7.1 appears to be insufficient. While the justification mentions that financial instruments or their components are presented as financial assets, financial liabilities, or equity instruments in accordance with the substance of the contractual agreement, it lacks specific details and evidence to substantiate full compliance with the IPSAS 28 standards.

Reasons for Insufficiency:

1. **Lack of Detailed Evidence:** The justification does not provide detailed evidence or examples of how the financial instruments are being presented in practice. It should include

Reasons for Insufficiency:

1. **Lack of Detailed Evidence:** The justification does not provide specific instances or documentation showing that the practice meets the definitions and requirements of IPSAS 28.
2. **Quantitative Data:** There is no quantitative data or specific examples of instruments that demonstrate the adherence to the standards. The justification should include the number of instruments reviewed, types of instruments, and

Ass

Conceptual Score: A

- **Disagreement:** Without detailed evidence, it's challenging to confirm that the comprehensive regulatory framework is fully aligned with IPSAS 28. The score might

- **Conceptual Score: A**

- **Disagreement:** Without detailed evidence, it's challenging to confirm that the comprehensive regulatory framework is fully aligned with IPSAS 28. The score might be considered too high without supporting evidence. A lower score, such as "B", might be more appropriate unless detailed documentation can be provided.

Validation of justification for a specific dimension - Result

Justifications for scores in all dimensions in indicator

LLM

Analysis & Recommendations for indicator

You are an expert assessor for PULSE. Scoring of the dimensions is the heart of the PULSE process. The scoring and assessment methodology has been aligned to the PEFA methodology, including a four-point calibration scale between A and D for each dimension.

Each dimension is scored separately on a four-point ordinal scale: A, B, C, or D, according to precise criteria established for each dimension. To justify a particular score for a dimension, every aspect specified in the scoring requirements must be fulfilled. If the requirements are only partly met or the criteria are not satisfied, a lower score is given that coincides with achievement of the requirements for the lower performance rating.

For those indicators referring to IPSAS, to achieve an "A" score, full compliance with the respective requirements of the standards is necessary. The "B" score is awarded for a high degree of compliance, but still certain gaps to the IPSAS requirements exist. A "C" score reflects the basic level of performance for each indicator and dimension, consistent with good international practices, which is aligned with what is established in the respective IPSAS for those indicators (see Table 6 for IPSAS references). A "D" score means that the feature being measured is nonexistent or presented at less than the basic level of performance.

The "D" score indicates performance that falls below the basic level. A "D" score is warranted when sufficient information is not available to establish the actual level of performance. The "D" score, due to insufficient information, is differentiated from "D" scores for lower-level performance by the use of an asterisk following the letter: "D*." The aggregation of an indicator with dimensions containing "D*" scores is no different from aggregation with other "D" scores. Aggregate indicator scores do not include an asterisk, and thus cases of insufficiency of information are only denoted at the dimension level.

The assess performance level and evidence for scoring of each dimension, the assessment team members should focus on the analysis of performance against PULSE criteria, provided in the MG and the web-based tool itself. The Justification/Explanation section of each dimension should provide factual evidence, including quantitative data, if available, which has been used to substantiate the assessment. The information provided should be as specific as possible, for example, in terms of quantities, proportions, and time spans. Any issues in relation to the timeliness or reliability of data and evidence should be noted. If insufficient information has been obtained either for a whole indicator or one of its dimensions, the text should explicitly mention it. A general suggestion is that the wording provided in Justification/Explanation section could be directly used for drafting PULSE report.

Prompt

Preliminary Analysis, Recommendations and Draft Report - Prompt

Dimension 19.3: Recognition of Tax Revenue

- **Score:** D (both conceptual and actual)
- **Findings:** The entity does not recognize tax revenue based on the occurrence of the taxable event or the triggering event. Additionally, the basic and additional requirements of recognition are not met.
- **Implications:** The failure to recognize tax revenue accurately according to IPSAS requirements indicates a significant gap in the entity's accounting practices. This can result in misstated financial statements and a lack of compliance with international standards.

• **Action:** Implement a robust system to recognize tax revenue based on the occurrence of taxable or triggering events and meet all the basic and additional requirements of recognition.

- **Benefit:** Addressing this major gap will enhance the reliability of financial statements, ensuring compliance with IPSAS standards and providing a clearer financial position of the entity.

2. Improve Recognition of Tax Revenue (Dimension 19.3)

- **Action:** Implement a robust system to recognize tax revenue based on the occurrence of taxable or triggering events and meet all the basic and additional requirements of recognition.
- **Benefit:** Addressing this major gap will enhance the reliability of financial statements, ensuring compliance with IPSAS standards and providing a clearer financial position of the entity.

- **Benefit:** Regular audits will help identify and correct any deviations from standards, ensuring ongoing compliance and improving the entity's scores across all dimensions.

Preliminary Analysis, Recommendations and Draft Report - Result

”

Thank you
