HAWAII STATE DEPARTMENT OF EDUCATION

PROCUREMENT AND CONTRACTS BRANCH

September 18, 2023

ADDENDUM A

TO

REQUEST FOR PROPOSALS

RFP D24-023

SEALED PROPOSALS

TO PROVIDE ARTIFICIAL INTELLIGENCE STAKEHOLDER DEVELOPMENT

FOR CLASSROOM-BASED ASSESSMENTS

FOR THE HAWAII STATE DEPARTMENT OF EDUCATION

PART I: PRE-PROPOSAL CONFERENCE

The following is a brief summary of the Pre-Proposal Virtual Conference that occurred on August 31, 2023:

The Pre-Proposal Virtual Conference was held on August, 31 2023 at 9:00 a.m. HST. The conference provided an overview of the RFP's purpose to gain an understanding of the breadth and scope of work under this RFP. Questions, concerns, and/or clarifications were responded to informally and everyone was thanked for their participation and interest.

PART II: WRITTEN QUESTIONS

Requirements pertaining to written questions were stated in the original RFP. Written questions received by the Department are listed below. The following responses are hereby provided and incorporated into the RFP:

1. Reference Page 9, Section 2.1 Phase I: Proof of Concept:
   a) What kind of questions will the Virtual Student need to answer? Level of technical skills are required?

   Response: During Phase I, the Virtual Students will be required to interact with a diverse set of questions similar to those encountered in educational assessments, including but not limited to:

   1. Multiple-Choice Questions: Virtual Students should be able to provide responses to multiple-choice questions where they select one answer from a set of four options.
   2. Selected Response Questions: These questions may involve selecting multiple correct answers, matching items, or other selected-response formats.
   3. Short Answer Questions: Virtual Students will need to generate concise written responses to short-answer questions, which may require demonstrating comprehension and knowledge.

4. Essay Questions: Some assessments may include essay questions, for which Virtual Students should be capable of providing well-structured, coherent, and thoughtful written responses.

Level of technical skills required:

The technical skills required for this phase are substantial and encompass expertise in artificial intelligence (AI), natural language processing (NLP), and machine learning (ML) technologies. Additionally, a deep understanding of educational assessment principles and the ability to develop algorithms and models for generating realistic student responses are essential. The technical team should have experience in data analysis and the ability to process and interpret educational data to ensure that Virtual Students provide accurate and contextually relevant responses.

The STATE's aim is to create Virtual Students that closely mimic real students in Hawaii's educational landscape, both in terms of subject matter knowledge and the ability to respond to assessment items in a way that emulates real student responses.

b) Are we modeling the inability to answer questions also?

Response: As part of the comprehensive development of Virtual Students, the STATE does intend to model not only their ability to answer questions but also their inability to do so accurately in certain situations. This is a critical aspect of creating authentic and realistic Virtual Students.

To elaborate further, Virtual Students should exhibit a range of behaviors and responses that are representative of actual students. This includes scenarios where they may not be able to answer questions correctly due to various factors such as lack of knowledge, misunderstanding of the question, or difficulty in comprehending certain topics.

Modeling these situations is important for several reasons:

1. Realism: It enhances the authenticity of the Virtual Students by reflecting the diverse range of responses that real students exhibit during assessments.
2. Assessment Development: It aids in the development of high-quality assessment items by allowing for the testing of different item difficulty levels and providing insights into areas where students commonly struggle.
3. Research and Analysis: It supports research on student performance and helps identify areas where additional support or instruction may be needed.

The STATE's goal is to create Virtual Students that not only excel in answering questions appropriately but also demonstrate the nuances and challenges that educators encounter when assessing students in real-world educational settings.

c) What behaviors of the students need to be emulated?

Response: The STATE welcomes creative input in defining the behaviors of Virtual Students. While STATE has outlined some general expectations, the STATE encourages offerors to propose a range of student behaviors that you believe are essential for an effective educational assessment tool. These may include, but are not limited to, behaviors related to engagement, attentiveness, critical thinking, problem-solving, and responses to various question types. Please feel free to suggest behaviors that align with your vision for the Virtual Student concept. The STATE values your expertise in this area and look forward to your innovative proposals.

d) Do we have student responses corresponding to each of those behaviors?

Response: While the STATE has some student responses that can serve as references, the STATE is open to innovative ideas and solutions for generating responses that correspond to the desired Virtual Student behaviors. Proposals should aim to achieve the alignment between

behaviors and responses as you see fit, without specific requirements set forth by STATE. The STATE values your expertise and look forward to your creative approaches in this regard.

2. In Section 2, page 10, the STATE refers to a grade-level, classroom-based assessment which implies an assessment that has a primary purpose of guiding instruction. In Section 3.2.1, page 16, the STATE requests "various item formats, including multiple-choice, open-ended, and interactive items." In Section 3.2.3, page 17, the STATE requests a standard setting process. In Section 3.3.1, page 19, the STATE further clarifies that this is an adaptive assessment.

    a) Will the STATE provide an example of the assessment?

        Response: The STATE recognizes that the term 'grade-level, classroom-based assessment' in Section 2 may imply an assessment primarily designed to guide instruction. However, the STATE would like to clarify that the reference is not specific to any one type of assessment. The intent is to utilize AI to create virtual students, teachers, and stakeholders that can be used in the test development process. These virtual entities will facilitate the creation of various types of assessments, including but not limited to those that guide instruction. The specific assessments developed from the items field-tested using Virtual Students can vary widely and are not limited to a single format.

    b) What are the stakes associated with this assessment?

        Response: The assessments can encompass a broad spectrum, ranging from assessments with relatively low stakes, often referred to as 'formative type' assessments, which are primarily designed to guide instruction and provide feedback to educators, to high-stakes assessments that hold significance in student grading or other critical educational decisions.

        This intentional diversity in assessment types and associated stakes aligns with the STATE's objective of creating a versatile and adaptable system capable of accommodating a wide array of assessment requirements within the education landscape. The specific stakes associated with each assessment will be contingent upon its intended purpose and the context in which it will be employed. The STATE invites offerors to consider this diversity when formulating their proposals and to suggest solutions that can flexibly accommodate various stake levels for different types of assessments.

    c) Has the STATE adopted achievement levels for this assessment?

        Response: It is important to clarify that this project focuses on developing a platform to facilitate the creation of a variety of assessments, rather than a single standardized assessment. Consequently, there is no predefined set of achievement levels that applies universally to all assessments generated through this platform.

        The determination of achievement levels for each assessment will be contingent upon its specific objectives and the context in which it will be employed. While certain assessments may indeed incorporate traditional achievement levels that align with their intended objectives, others, particularly formative assessments, may not adhere to conventional achievement levels in the same manner as summative assessments.

        The STATE invites offerors to acknowledge the adaptable nature of the STATE's platform, which can cater to diverse assessment requirements. These may or may not involve the establishment of achievement levels, depending on the particular goals and design of the assessment.

    d) Does the STATE anticipate that all items for this test will be administered in a common sequence with common opportunities to learn preceding their administration?

        Response: The STATE does not envision a standardized sequence or uniform prerequisites for

all items, nor does the STATE dictate common opportunities for learning before their administration.

The inherent flexibility of this project enables the creation of a diverse array of assessments, each characterized by its individualized design, sequence, and prerequisites aligned with its intended objectives. While certain assessments may adhere to a shared sequence, others may not, as this will be driven by the unique requirements of each assessment.

e) Does the STATE conceive of this assessment as a series of assessment blocks within a content area that educators administer by selection?

Response: The concept of this project is adaptable to various approaches, including the potential utilization of a series of assessments within a content area, which educators can select for administration. Nevertheless, it's important to emphasize that the precise structure and administration approach may diverge based on the unique requirements of a particular assessment.

f) Is the STATE open to stage-adaptive assessments as well as traditional CAT assessments?

Response: See responses to questions 2a-e.

3. Reference Page 12, Section 3.1.2. Virtual Student Prototype:
   a) Will Hawaii State provide access to educational experts and psychometricians to coordinate the evaluation of the proof of concept?

   Response: While the STATE will provide access to our educational experts to assist in coordinating the evaluation of the proof of concept, it's important to note that the STATE does not employ psychometricians. However, the STATE is committed to facilitating a collaborative and supportive environment to ensure the successful evaluation of the Virtual Student Prototype.

   b) Do you have sufficient data available for the following features to be built? Problem-solving strategies and test-taking behaviors accurately reflecting the diverse demographic representations and educational needs present in Hawaii's public schools.

   Response: While the STATE has access to a wide range of educational data, it is important to emphasize that the success of this project relies on innovative approaches to create Virtual Students that accurately reflect diverse demographic representations and educational needs. The STATE encourages offerors to propose solutions that can leverage available data sources and AI technologies to build problem-solving strategies and test-taking behaviors that align with Hawaii's unique educational landscape.

   c) Virtual students will provide realistic responses to test items - Will the real student interact with the bot initially and help train the bot?

   Response: While the exact interaction process between real students and Virtual Students may vary depending on the proposed solution, the intention is to create Virtual Students capable of providing realistic responses to test items. The STATE is open to innovative approaches that may involve initial interactions between real students and Virtual Students to help train the AI model. However, this will ultimately be a decision made by the selected contractor in their proposal. The STATE's primary goal is to ensure that the Virtual Students can emulate real student behaviors effectively.

   d) What "cognitive processes" does the Virtual student need to exhibit? Does the model show how it reached that solution, in other words, exhibiting its thinking process?

Response: The Virtual Students should exhibit a range of cognitive processes that reflect real student behaviors when addressing assessment items. These processes may include but are not limited to problem-solving, critical thinking, decision-making, and reasoning.

Regarding whether the model should show how it reached a solution, the STATE is open to innovative solutions that may include exhibiting the thinking process. However, the extent to which this is implemented will depend on the proposed approach by the offeror. The ability to demonstrate the thinking process can enhance the authenticity of the Virtual Student, but the STATE also understands that different AI models may have varying capabilities in this regard. Therefore, the STATE encourages offerors to propose solutions that align with their AI model's capabilities and effectively emulate real student behaviors.

4. Reference Page 13, Section 3.1.4 Validation Study:
    Validation study to evaluate the authenticity, accuracy, and alignment of the Virtual Students with real-world counterparts in Hawaii's educational context - Does the contractor need to build a response-capturing system from real students, or are those responses already available from surveys conducted in the past?

    Response: The contractor will not be required to build a response-capturing system from real students. Responses from real students are not required to be collected for this project. Instead, the validation of the Virtual Students will focus on assessing their authenticity, accuracy, and alignment with real-world counterparts in Hawaii's educational context based on the Virtual Students' interactions and responses within the AI-driven simulation environment.

    The validation study will aim to ensure that the Virtual Students' behaviors closely emulate the characteristics and perspectives of actual students in Hawaii's public schools without directly involving real students. Therefore, there is no need to gather responses from real students through surveys or other means.

    The evaluation will be based on the performance and behavior of the Virtual Students within the AI environment, assessing their ability to effectively contribute to the item and test development processes while closely resembling real students in Hawaii's educational landscape.

5. Reference Pages 14-15, Section 3.2.1 Virtual Student Development; Page 19, Section 3.4.2. Virtual Student Prototype Development:
    a) Will the STATE provide access to state-maintained student databases? If so, what types of data are available? How many students? How many years?

        Response: The STATE will provide access to certain educational data that is relevant to the development of Virtual Students and Virtual Stakeholders for this project. However, it's important to clarify that this data will not include access to state-maintained student databases containing sensitive or personally identifiable information (PII) of real students.

        The specific types of data that will be made available for this project, the quantity of data, and the number of years will be subject to the resultant, executed contract between the awarded Contractor(s) and the STATE. The data made available will be limited to what is necessary and relevant for the development, testing, and validation of the Virtual Students and Stakeholders within the scope of this project. The data provided will comply with all applicable data privacy and security regulations.

        Please note that the data provided will be anonymized and de-identified to ensure the privacy and confidentiality of students and stakeholders, and it will be used solely for the purpose of creating realistic Virtual Students and Virtual Stakeholders. The specifics of data access and usage will be

outlined in the contractual agreements and data sharing protocols established between the Contractor(s) and the STATE.

The exact details regarding the types, quantity, and duration of data access will be determined in the resultant, executed contract between the awarded Contractor(s) and the STATE.

b) Does the STATE anticipate sampling from the current student population to provide representative demographic input data for the Student Bots (Virtual Students)?

Response: The STATE acknowledges the importance of ensuring that the Virtual Students accurately reflect the diversity of Hawaii's current student population, including demographic characteristics. While the STATE possesses a substantial historical dataset spanning approximately 10 years, the STATE understands that demographic trends and student populations may evolve over time.

To enhance the authenticity and representativeness of the Virtual Students, the STATE anticipates collaborating with the awarded Contractor(s) to consider various approaches, which may include sampling from the current student population in addition to leveraging historical data. The goal is to capture the most up-to-date demographic input data and ensure that the Virtual Students align with the current educational landscape in Hawaii.

The specifics of this data sampling and collection process will be discussed and refined during the project's planning and development phases in consultation with the Contractor(s). The STATE is committed to working closely to establish a robust methodology that accounts for demographic diversity, emerging trends, and the evolving needs of Hawaii's public schools.

Ultimately, the STATE's aim is to create Virtual Students that are not only grounded in historical data but also informed by current demographic representations, thus enhancing their authenticity and relevance in the educational context.

6. Reference Pages 14-15, Section 3.2.1 Virtual Student Development; Page 22, Section 3.5.1.1. Enhancements to Virtual Students:
Who does the STATE envision administering items to the Student Bots (Virtual Students)?

Response: The STATE envisions that the contractor, responsible for developing and enhancing the Virtual Students (Student Bots), will also administer newly developed items to these Virtual Students. This approach is integral to the iterative and collaborative nature of the project, allowing the contractor to not only create but also evaluate and refine the Virtual Students' capabilities in real-world testing scenarios.

The administration of items to the Virtual Student will primarily fall under the purview of the contractor and their team of experts. This may involve designing and deploying assessment items, observing and recording the Virtual Students' responses, and analyzing the results to assess the authenticity, accuracy, and alignment of the Virtual Students with real-world counterparts in Hawaii's educational context.

The collaborative and research-oriented nature of this endeavor implies that the contractor will actively engage with the Virtual Students to enhance their functionality, refine their responses, and ensure their alignment with the project's objectives. It also opens up opportunities for collaboration with educational experts, psychometricians, and other stakeholders who can contribute to the validation and enhancement process.

Overall, the contractor will play a central role in administering items to the Virtual Students as part of the ongoing development and evaluation process. This approach will help create a robust and adaptable platform for various assessment needs within the education landscape.

7. Reference Pages 15-16, Section 3.2.3 Virtual Community Member Development; Pages 23-24, Section 3.5.3 Development of Virtual Community Members (Virtual Community Member Bots): Does the STATE have data from past standard settings? If so, what is available in that data? Will the STATE provide access to the data to the selected vendor?

   Response: The STATE possesses data from past standard settings, which can offer valuable insights and information for the development of Virtual Community Members. While the STATE does have access to this historical data, it's essential to clarify that the specific details of this data, including its content and format, may vary depending on the assessments and standard settings conducted in the past.

   To ensure a transparent and collaborative approach, the STATE is open to providing the selected Contractor(s) with access to relevant historical data from past standard settings. The nature and extent of the data shared will be discussed and agreed upon in collaboration with the selected Contractor(s) during the project's implementation phase.

   We encourage potential offerors to consider this opportunity for data access and collaboration when proposing their methodologies and approaches for the development of Virtual Community Members. This collaborative effort aims to create a more effective and contextually relevant project outcome.

8. Reference Pages 17-18, Section 3.3.1. Virtual Field Testing:
   a) What would be the scope of the virtual field test (FT)? Please be specific about subjects (math, ELA, science, etc.), grades, content/ standards measured. Are some content areas prioritized over others?

      Response: The scope of the development of a virtual field testing platform is designed to be comprehensive, covering various subjects and grade levels within the education domain. The goal is to create a versatile and adaptable platform capable of supporting a wide range of assessments in subjects including but not limited to mathematics, English language arts (ELA), science, and other content areas.

      The specific subjects, grades, and content/standards to be measured may vary based on the assessment needs and priorities identified during the project's implementation. While there is no fixed prioritization of content areas over others, the intent is to create flexibility within the platform so that assessments can be tailored to the unique requirements and priorities of Hawaii's educational landscape.

      The exact scope of the virtual field test, including the subjects, grades, and content areas to be included, will be determined collaboratively between the selected vendor and the STATE during the project's implementation phase. This collaborative approach aims to ensure that the virtual field test aligns with the evolving needs and objectives of Hawaii's education system, allowing for the creation of assessments that are both contextually relevant and effective in supporting educational goals.

      The STATE encourages potential offerors to consider the adaptability and flexibility of the platform when proposing their methodologies and approaches for the virtual field testing phase. This flexibility is intended to accommodate a wide range of assessment subjects and grade levels, enabling the development of assessments that can address various educational needs within Hawaii.

   b) Does the STATE anticipate that these items will be placed on an existing or a unique difficulty scale?

      Response: The STATE anticipates that the items developed during the virtual field test phase may be placed on an existing or a unique difficulty scale, depending on the specific assessment's

requirements and context of use. The decision regarding the placement of items on a difficulty scale will be made in collaboration between the selected vendor and the STATE, taking into consideration the alignment with the intended purpose and goals of the assessment.

While there may be existing difficulty scales available, the flexibility of this project allows for the creation of unique difficulty scales if it is deemed necessary to better align with the assessment's objectives and the characteristics of the items being developed.

We encourage potential offerors to consider the adaptability of the platform and the potential need for both existing and unique difficulty scales when proposing their methodologies for the virtual field testing phase. This flexibility is intended to ensure that the assessments created can effectively measure student proficiency and provide valuable insights for educational decision-making.

c) Does the STATE anticipate that students can be mandated to participate in research studies to support the validity of the virtual FT and that sufficient numbers of students with disabilities and who are emerging English Learners will participate in the studies?

Response: The STATE acknowledges the importance of conducting research studies to support the validity of the virtual field test (FT). While we understand the value of student participation in research studies, we also recognize the need to ensure that research activities adhere to ethical and legal guidelines, including considerations related to students with disabilities and emerging English Learners.

The participation of students in research studies, including those with disabilities and emerging English Learners, should be voluntary and comply with all applicable laws and regulations governing human subjects research. It is our intent to work collaboratively with the selected vendor to design and implement research studies that adhere to these principles and respect the rights and privacy of all students involved.

The specific scope and approach of research studies, as well as the recruitment of participants, will be determined in consultation between the selected vendor and the STATE, ensuring that research activities are conducted ethically, responsibly, and in alignment with legal requirements.

We encourage potential offerors to consider these ethical and legal considerations when developing their proposals for the virtual field testing phase, and to propose methodologies that ensure the valid and reliable assessment of student proficiency while respecting the rights and needs of all student populations.

9. Reference Page 19, Section 3.4.2. Virtual Student Prototype Development:
What existing demographic and behavioral data will the STATE provide to support the development of prototype Student Bots (Virtual Students)?

Response: See response to question 5a.

10. Reference Page 20, Section 3.4.4.4. Test-taking behaviors:
a) What type of assessment data is available to capture the cognitive processes, knowledge and skills, problem-solving strategies, and test-taking behaviors of Hawaii's students (see also sections 3.1.2 and 3.1.4.4)? How many students, per grade and content area, will be reported in the data?

Response: See response to question 3b.

b) If data is not available, should the bidder propose a data collection design and procedure?

Response: If the specific data needed for capturing the cognitive processes, knowledge and skills, problem-solving strategies, and test-taking behaviors of Hawaii's students is not readily available, the STATE encourages offerors to propose a data collection design and procedure as part of their solution. This will allow for a comprehensive approach to gathering the necessary data to support the development and validation of the Virtual Students.

11. Reference Page 22, Section 3.5.1.4. Scalability and Performance:
    a) What is the estimated number of Virtual Students for the extensive load testing?

    Response: The estimated number of Virtual Students required for extensive load testing will vary based on the scalability and performance capabilities of the proposed solution. The precise number needed will be determined during the evaluation of proposals. We encourage offerors to propose a scalable solution that can effectively handle a significant number of Virtual Students to ensure robust performance under varying loads and conditions.

    b) What is the estimated number of field-test items to be administered?

    Response: The estimated number of field-test items to be administered is not fixed, as this project is primarily focused on developing a versatile platform that enables the field testing of assessment items using Virtual Students. The platform's design allows for the administration of an extensive range of field-test items, making it flexible and adaptable to accommodate various content areas and assessment needs. The specific number of items to be administered will depend on the test development and educational objectives pursued through the platform. The selected vendor will collaborate with the STATE to determine the scope and scale of field testing based on the project's evolving requirements.

    c) Is the goal of these requirements to develop a testing platform which can accommodate a large-scale assessment?

    Response: The primary goal of these requirements is to develop a versatile and adaptable platform specifically designed for field testing assessment items using virtual students. This platform will collect student responses, which will subsequently be used to calibrate the items. The calibrated items will then be used to create assessments that can provide human test takers with a proficiency level. While scalability and performance are important considerations, the primary focus is on creating a comprehensive tool to support the field testing process and item calibration, ultimately enhancing the quality and reliability of assessments for human test takers. This platform aims to accommodate a wide range of assessment scenarios, from small-scale field tests to larger-scale item calibration initiatives, all contributing to the development of high-quality assessments.

12. Reference Page 28-29, Section 3.6.4. Virtual Scoring of Real Student Work:
    a) What does the STATE see as the difference between scoring by virtual teachers and automated scoring?

    Response: Virtual teachers represent simulated educators who assess and score student work as real teachers would. They are designed to emulate the human grading process, considering not only the correctness of answers but also the quality, depth, and nuances of student responses. These virtual teachers aim to replicate the assessment grading experience realistically, including the capacity to provide student-specific qualitative feedback. They evaluate how closely a human student's response aligns with predefined criteria and standards, offering insights into how well the response meets or does not meet these criteria. This personalized feedback is a key feature of virtual teachers and can be valuable for students in understanding their strengths and areas for improvement.

Automated scoring, on the other hand, typically involves the use of algorithms and artificial intelligence to assess and score student work. While it can efficiently handle large volumes of assessments and objective questions (like multiple-choice), it may have limitations in evaluating subjective or open-ended responses in the same nuanced way as human graders or virtual teachers. Automated scoring often relies on predefined criteria and patterns in the data to assign scores and might not provide student-specific qualitative feedback to the same extent as virtual teachers.

The choice between these two approaches depends on the specific assessment needs, including the type of questions, the desired level of detail in scoring, and the balance between efficiency and the quality of feedback. The STATE is interested in exploring both options to determine which best aligns with the project's goals and objectives for different assessment scenarios.

b) Will the STATE provide real human student work samples?

Response: The STATE anticipates that real human student work samples will be valuable for several aspects of this project. These samples can help in the development and calibration of assessment items, the training and refinement of virtual teachers, and the overall validation of the system's performance.

While the specifics of providing such samples will depend on the project's needs and logistics, it is within the project's scope to collaborate with the STATE to access and utilize real student work samples as necessary for the various phases and components of the project.

The availability and use of real human student work samples will be determined in coordination with the selected contractor(s) to ensure the project's success and the achievement of its objectives.

13. Reference Page 29, Section 3.6.4.4. Reporting and Documentation:
    a) What types of student-level feedback and reporting does the STATE expect from this assessment?

    Response: Within this component, the primary goal is not to administer a traditional assessment or generate scores for individual students. Instead, the focus is on utilizing Virtual Teachers to assess real students' work samples and provide detailed, individualized feedback. This feedback aims to support and guide students in their learning journeys.

    The Reporting and Documentation task encompasses several key aspects:

    1. Comprehensive Reports: We will generate comprehensive reports and documentation detailing the implementation, performance, and outcomes of the Virtual Scoring of Real Student Work component. These reports will provide insights into the assessment process.
    2. Scoring Algorithms: Descriptions of the scoring algorithms used in the assessment process will be included, ensuring transparency in how assessments are evaluated.
    3. Statistical Analyses: Reports will contain information about the statistical analyses conducted as part of the assessment process, ensuring robustness and validity.
    4. Interpretation Guidelines: Clear interpretation guidelines for the assessment scores will be provided to facilitate understanding of the feedback provided to students.
    5. Recommendations for Refinements: We will also offer recommendations for any necessary refinements or improvements to enhance the scoring process, ensuring the continuous improvement of the assessment and feedback mechanisms.

    While these comprehensive reports and documentation will be generated, the core objective remains to provide meaningful, personalized feedback to students through the Virtual Scoring of Real Student Work component. The emphasis is on enhancing the learning experience and assisting students in improving their performance based on constructive feedback.

b) What types of classroom-level feedback and reporting does the STATE expect from this assessment?

Response: This project is primarily focused on the virtual scoring of real student work using Virtual Teachers. This means that the main objective is to evaluate real students' work samples and provide individualized feedback.

As detailed in our response to 13a, the emphasis is on generating comprehensive reports and documentation for the Virtual Scoring of Real Student Work component. This includes descriptions of scoring algorithms, statistical analyses, and interpretation guidelines for assessment scores.

Traditional classroom-level reporting, which typically involves aggregating data on student performance for instructional or administrative purposes, is not within the scope of this initiative. Instead, the project seeks to offer personalized feedback to real students based on their submitted work samples.

c) What types of district-level feedback and reporting does the STATE expect from this assessment?

Response: See responses to 13a/b.

d) What types of state-level feedback and reporting does the STATE expect from this assessment?

Response: See responses to 13a/b.

e) What types of student- and family-level feedback and reporting does the STATE expect from this assessment?

Response: See responses to 13a/b.

14. Reference Appendix E, General Conditions, Page 12, Section 26, Ownership Rights and Copyrights:
   a) This work will likely require the use of proprietary vendor technologies or, perhaps, inventing new methodologies that take advantage of proprietary technologies. Will the STATE consider the proprietary vendor technologies, or new methodologies, to be owned by the STATE as a "work made for hire"?

   Response: The following provision shall be incorporated into Appendix D, Contract Minimum and Special Conditions (refer to Part III, Changes to the RFP below):

   Ownership of Documents

   General Conditions paragraph 26 entitled "Ownership Rights and Copyright" is deleted entirely and replaced with the following:

   All documents and reports, and student data/records developed or generated under this contract shall be the sole property of the STATE. CONTRACTOR retains ownership of any proprietary or copyrighted materials, data, software, technologies, test items, test forms that are previously developed, owned and/or copyrighted by the CONTRACTOR and used or adapted for use under this contract.

   b) May Offerors propose a licensing agreement either to retain ownership of some deliverables (e.g., technologies and methods used to develop Virtual Students, Teachers, and Stakeholders) or to use these deliverables after completion of the contract?

Response: Reference Page 32, Section 4.8, Executive Summary, "Terms and Conditions" and "Deviations," and Section 4.11, Proposed Solution; Technical Proposal of the solicitation. In addition, the following provision shall be incorporated into Appendix D, Contract Minimum and Special Conditions (refer to Part III, Changes to the RFP below):

Ownership of Documents

General Conditions paragraph 26 entitled "Ownership Rights and Copyright" is deleted entirely and replaced with the following:

All documents and reports, and student data/records developed or generated under this contract shall be the sole property of the STATE. CONTRACTOR retains ownership of any proprietary or copyrighted materials, data, software, technologies, test items, test forms that are previously developed, owned and/or copyrighted by the CONTRACTOR and used or adapted for use under this contract.

15. May Offerors recommend plans for maintaining some of the project deliverables beyond the contract term (e.g., maintenance of the Virtual Students, Teachers, and Stakeholders)?

    Response: Offerors are welcome to recommend plans for maintaining some of the project deliverables beyond the contract term, such as the Virtual Students, Teachers, and Stakeholders. The STATE is open to considering proposals that outline sustainable strategies for the continued use and maintenance of project assets after the contract term.

16. Are there any pages limits?

    Response: There are no page limits. However, please note the file size when uploading your firm's electronic submitted offer in HIePRO (reference section 4.5, Proposal Submission and Format of the RFP).

17. Is there a recording of the pre-proposal conference or slide deck?

    Response: There is no Pre-Proposal Conference recording or slide deck available. The information shared during the pre-proposal conference should be considered an overview of the process only; it does not replace the Request for Proposal document and any official written response.

18. If the sample data is available, is it in digital format? And what kind of data is it?

    Response: The data available for this project is primarily in digital format. It encompasses various types of data related to student assessments, including assessment scores (including individual item responses) and student demographic information. The specific details of the data format and content can be discussed and refined in collaboration between the CONTRACTOR and the STATE to ensure its compatibility with the project's requirements.

19. Is there an Artificial Intelligence currently in use? If yes, please can you brief its usage in State.

    Response: The STATE currently employs Artificial Intelligence (AI) in its educational assessments. Specifically, AI technology is used for the machine scoring of short answer and essay items within the Smarter Balanced interim and summative assessments.

    The primary purpose of AI in this context is to efficiently and consistently evaluate and score student responses to these types of assessment items. By automating the scoring process, the STATE can provide prompt feedback to educators and students, helping to inform instruction and support student learning.

20. Is any data that needs to be converted/ captured from the paper document (PDF, Image)? If yes, please let us know the volume of data that needs to be captured.

    Response: Yes, the STATE envisions the capability for the Virtual Teacher to score student work samples that may be in various digital formats, including MS Office (Word, Excel, Powerpoint) documents, Google Docs, Sheets, etc., or PDF files.

    The volume of data that needs to be captured in these digital formats can vary depending on the specific assessments and tasks assigned. Since this project aims to create a flexible platform, the volume of data for conversion and capture will depend on the assessments administered and the quantity of student work samples generated in these digital formats. It is expected that the platform should be adaptable to accommodate varying volumes of digital documents.

21. Is the use of third-party LLM APIs permitted, where the data (vectorized) will be to the third-party servers?

    Response: Yes, the use of third-party Language Model (LLM) APIs is permitted for this project, and data may be sent to third-party servers when necessary for the fulfillment of project requirements. However, it's essential to emphasize that all data sharing must adhere to the project's data security and privacy requirements and comply with relevant regulations and agreements.

22. What are the types of assessments? Can we get sample of the assessments?

    Response: As mentioned earlier, the project aims to create a versatile platform capable of supporting various types of assessments, including but not limited to those that guide instruction. The specific assessments developed from the items field-tested using Virtual Students can vary widely and are not limited to a single format. Therefore, we do not have a single set of sample assessments to provide. The diversity of assessment types is intentional, as the goal is to create a flexible and adaptable platform that can accommodate a wide range of assessment needs within the education landscape.

23. Where will the datasets for these "virtual person" be pulled from?

    Response: The datasets for creating 'virtual persons,' such as Virtual Students, Virtual Teachers, and Virtual Community Members, will be generated using a combination of real-world educational data, machine learning techniques, and simulation. The specifics of data sources and the generation process may vary depending on the virtual persona in question. We do not pull data from existing individuals or entities; instead, we use data to simulate behavior and responses that align with the goals of the project. The emphasis is on creating realistic and representative virtual personas that can contribute effectively to the assessment development process while respecting privacy and data sharing requirements.

24. Where do all the cultural, demographic, and other details come from?

    Response: The cultural, demographic, and other details used in the creation of virtual personas, such as Virtual Students and Virtual Community Members, are sourced from a combination of publicly available data, educational databases, and simulations. These details are not necessarily extracted from real individuals or entities but are modeled based on aggregate data and patterns to ensure realism and relevance to the context of Hawaii's educational landscape. It's important to note that privacy and data security regulations are strictly adhered to throughout this process to safeguard sensitive information.

25. Proficiency was brought up a lot during the pre-proposal conference. Proficiency is all based on what the baseline will be derived from. Who has set these standards? Or is that part of these efforts?

Response: Typically, proficiency standards are defined by educational authorities, state departments of education, and relevant stakeholders. In the context of this project, the Virtual Community Members, including Virtual Stakeholders, will play a role in the standard setting process. They will participate in the process to help establish cut scores for specific assessments, which will enable the determination of individual student proficiency. This project aims to facilitate the calibration of items and the setting of standards for an assessment through the involvement of virtual entities.

26. What does success look like for Phase 1 since this has not really been done before?

Response: Success for Phase 1 of this project will be defined by several key milestones and outcomes which include, but are not limited to, the following:

1. Functional Virtual Students: The successful creation and deployment of functional Virtual Students who can interact with assessment items in a manner that closely emulates real students, demonstrating a range of behaviors and responses.
2. Proof of Concept: The successful execution of a proof of concept phase where the Virtual Students are tested with actual assessment items. This phase should provide valuable insights into the feasibility, effectiveness, and alignment of Virtual Students with the project's objectives.
3. Alignment with Project Objectives: The extent to which the Virtual Students align with the project's objectives, as outlined in the RFP. This includes their ability to create realistic student behaviors and contribute effectively to the assessment process.
4. Technical Feasibility: Demonstrating that the technical aspects of the prototype, such as compatibility with specifications and data processing capabilities, are robust and meet project requirements.
5. Performance Quality: The quality of Virtual Students' performance in simulating student responses and interactions, showcasing accuracy, reliability, and consistency across various assessment scenarios.
6. Validation Study: Initiation of a validation study that assesses the authenticity, accuracy, and alignment of Virtual Students with real-world student behaviors. While this study may not be completed in Phase 1, its successful commencement is a crucial step.

Success in Phase 1 will provide a strong foundation for subsequent phases of the project, building confidence in the capabilities of Virtual Students and their potential to enhance educational assessment in Hawaii.

27. What kind of historical data do we have on student? Psychometric Analysis Data, Assessment Scores, Student Profiles, Background etc.? What is the format in which this data will be available? How far back in number of years does the data go?

Response: The STATE maintains a diverse repository of historical student data, encompassing assessment scores, including individual item responses, and student demographic information. The format and accessibility of this data can exhibit variability, and it spans an extensive historical timeframe of nearly a decade.

The data typically exists in digital formats that align with prevalent data analysis tools and platforms. Precise format specifications will be collaboratively established with the STATE to ensure seamless integration. The STATE is dedicated to fostering a collaborative partnership with the chosen vendor to facilitate access to pertinent data for the effective development and testing of Virtual Students and Virtual Community Members.

28. Can we rely on the Subject Matter Experts at the Department of Education to learn more about the how assessments are done currently, what is tracked and how is it measured currently for success?

Response: The STATE welcomes collaboration with SMEs within the STATE to provide insights into current assessment practices, tracking mechanisms, and success measurement criteria. Their expertise will be invaluable in aligning this project with existing educational standards and practices, ensuring a seamless integration of Virtual Students and Virtual Community Members into the assessment process.

29. Can you please explain what a "Hawaii-like" response is?

    Response: A 'Hawaii-like' response refers to a response from the Virtual Students or Virtual Community Members that closely emulates the characteristics, behaviors, and perspectives of individuals typically found within Hawaii's diverse educational landscape. This term is used to emphasize the need for authenticity and realism in how these virtual entities interact, respond, and behave, ensuring that they mirror the cultural and demographic diversity of Hawaii's student population.

PART III: CHANGES TO THE RFP

The following changes are hereby made (additions in **bold/blue/underlined** text and deletions in **bold/strikethrough** text):

1. Page 10, section 2.1, Phase I: Proof of Concept:

   The STATE is seeking proposals to furnish and deliver a Proof of Concept. Once the Proof of Concept is complete, the STATE will further evaluate the Proof of Concept in accordance with the evaluation criteria set forth in section 5 to determine if the STATE will exercise its option to move forward with Phases II and III. If the STATE moves forward with Phases II and III, ~~one (1)~~ Contractor**(s)** will be selected to perform Phases II and III.

2. Page 10, section 2.2, Phase II: Full Research and Development:

   Upon successful completion of Phase I and based on the positive outcomes of the proof of concept, the Contractor(s) awarded in Phase I shall be further evaluated for selection to move forth with future Phase II, focusing on full-scale Research and Development efforts. ~~One (1)~~ Contractor**(s) with the highest number of points scored (no more than two (2) Contractors)** will be selected for Phases II and III. The aim is to develop robust and reliable Virtual Students and Virtual Stakeholders, ensuring efficient data collection, psychometric analysis, and scoring of real student work samples.

3. Page 11, section 2.3, Phase III: Deployment:

   With the successful completion of Phase II and validation of the AI Assessment Community, the Contractor**(s)** shall proceed to Phase III which will focus on the application of the developed technology for use in field testing items, scoring student work samples, and the setting of cut scores for common assessments.

4. Page 14, section 3.2, Phase II: Full Research and Development:

   During Phase II of the project, the selected Contractor**(s)** will embark on a comprehensive research and development (R&D) journey to bring the concept of an AI Assessment Community to fruition. Building upon the successful proof of concept developed in Phase I, the primary focus will be on the meticulous creation and refinement of Virtual Students, Virtual Teachers, and Virtual Community Members (Virtual Stakeholders). This phase marks a crucial stage in the development process,

where cutting-edge AI technologies and advanced algorithms will be harnessed to design and engineer these virtual entities with the highest levels of accuracy, reliability, and authenticity. The primary objective is to develop robust and reliable virtual entities that emulate real students, teachers, and stakeholders, ensuring seamless participation in the test development process and the accurate scoring of student work samples. Throughout this phase, rigorous testing, validation, and iterative refinement will be conducted to ensure the seamless integration of Virtual Students, Virtual Teachers, and Virtual Community Members into the assessment process, setting the stage for a groundbreaking approach to educational assessment that embraces innovation, efficiency, and fairness.

5. Page 34, section 4.9, Offeror Qualifications, "Information/Data Security and Confidentiality":

Information/Data Security and Confidentiality. The Offeror shall include specific information regarding its internal control environment, security history, legal compliance and confidentiality compliance (for example: SAS70/SSAE16/SOC2 reports) to ensure processes and policies related to data sharing are performed to STATE's satisfaction.

The Offeror shall confirm that it complies with State and Federal privacy laws and will follow the Department's requirements relating to privacy, including, but not limited to, the Family Educational Rights and Privacy Act (FERPA) and chapters 8-6 and 8-34, HAR. **Offeror shall also confirm that the solution will have continuous protection and will mitigate against AI vulnerabilities and attacks such as prompt injection.**

6. Page 41, section 5.7, Evaluation Process Upon Completion of Phase I, Proof of Concept:

Scoring for the prototype evaluation will be conducted using a numerical rubric. The evaluation will be based on the criteria outlined herein to each prototype and assign points based on their performance. The final selection of ~~one (1)~~ Contractor**(s)** for Phase II and III will be made based on the evaluation results of the Contractor**(s)** with the highest number of points **(no more than two (2) Contractors)**, and the alignment of the prototypes with Hawaii's educational needs and requirements.

7. Appendix D, Contract Minimum and Special Conditions – the following provision shall be incorporated:

Ownership of Documents

General Conditions paragraph 26 entitled "Ownership Rights and Copyright" is deleted entirely and replaced with the following:

All documents and reports and student data/records developed or generated under this contract shall be the sole property of the STATE. CONTRACTOR retains ownership of any proprietary or copyrighted materials, data, software, technologies, test items, test forms that are previously developed, owned and/or copyrighted by the CONTRACTOR and used or adapted for use under this contract.