



California's Government Operations Agency

Large Language Models (LLMs)

**Intent to Present at the Generative Artificial Intelligence (GenAI) LLM
Developer Showcase**

Abt Global, LLC
6130 Executive Boulevard
Rockville, MD 20852
abtglobal.com



Submission Date: September 16, 2024

Principal Technical Point of Contact:

Matthew Gillingham
VP, AI and Digital Innovation
Matthew.Gillingham@abtglobal.com

Additional Point of Contact:

Natalie Matthews
Science and Research Director
Natalie.Matthews@abtglobal.com

The State of California is looking for partners with the capability to bring new and innovative GenAI technologies and services to bear on unique challenges confronting the State's work to address and support Californians experiencing homelessness. Abt Global is pleased to respond to this Request for Information (RFI), focusing on Challenges 5, 6, and 7. We approach these Challenges from the perspective of a digital services provider with an established AI practice, deep cross-domain expertise in homelessness, behavioral health, and housing development, as well as a shared mission to combat homelessness. Our team provides integral information technology support and guidance to California state agencies and federal agencies to leverage HMIS for data collection and reporting efforts, including:

Direct Use of California Homelessness Data: Abt currently provides technology support and technical assistance in response to the passage of AB977, which mandates the reporting of data from a broad range of programs into the local HMIS. We also routinely conduct research utilizing California's health and housing data systems and collect qualitative data from people experiencing homelessness. We routinely engage local governments and local homelessness and service partners for our work with the California Interagency Council on Homelessness (Cal ICH) and Department of Social Services.

IT and Data Support for HMIS: For more than 20 years, Abt Global has led data collection, analysis and reporting efforts for HUD's Annual Homeless Assessment Report (AHAR) to Congress and stood up HMIS data repositories for the VA and HHS/RHY to ensure the timely, comprehensive and accurate submission of critical data on persons experiencing homelessness and served by these federal agencies. Our team of experts have worked with every CoC in the nation on making improvements to their HMIS data collection efforts, through the AHAR data cleaning process, and has provided more intensive support to CoCs in Los Angeles and San Diego to improve the quality and strength of their HMIS implementations.

AI Past Performance: Our track record of developing and deploying more than forty AI technologies and services across federal civilian agencies – including development of sophisticated AI transcription tools for HUD that automate the process of pulling actionable insights from audio field interviews – means Abt Global has the AI technology accelerators and cloud-hosted platform, as well as the past performance and experience with GenAI that accelerates velocity of delivery at the high-levels of quality, rigor, and risk management. We bring a dedicated AI practice with a defined corporate Governance model and risk management framework that prioritizes federal and State of California Executive Orders around Responsible AI, an established MLOps toolchain, and adherence to a repeatable Test and Evaluation (T&E) framework that includes pre- and post-production deployment processes. Our development methodology and AI lifecycle utilize the latest Experience and Human Centered Design practices to drive the tailored design and development of AI tools grounded in the lived experience of the people we are serving, from case managers to people experiencing homelessness.

Based on our background and experience, we believe the development of a Domain-Specific LLM (DS LLM) using specialized training data focused on housing and homelessness is the critical next step in addressing Challenges 5 and 6 because this DS LLM serves as the foundation for the development of tools like chatbots and other AI services that summarize text and audio data, generate analytic insights and visualizations, and reduce workload burden and accelerate velocity of assessing programmatic impact using state reports.

Addressing Challenges 5, 6, and 7

With the rapid integration of data from the behavioral health system (CalMHSA) into the HDIS as a result of AB977 and launch of a statewide Electronic Health Record (EHR) for CalMHSA, opportunities abound for improved knowledge sharing and warm referrals, assisted by AI or LLM. Accessing treatment from a homeless program, or unsheltered homelessness remains a persistent challenge.

Many eligibility factors, or triage questions could be feasibly completed through an AI chatbot, such as understanding admissions criteria for hospitalization and other behavioral health programs; when to call mobile crisis; etc. There are hundreds of programs and resources at the local level for which it is impossible to expect case managers and street outreach workers to fully understand and know how to access. We would explore whether AI could generate bed availability information in advance of recommending referrals to reduce frustration on an already overburdened workforce.

We propose leveraging Abt's library of AI accelerators and our experience and framework for developing DS LLMs to develop tools that take natural language requests from users and translate them into structured queries, prompting the user for additional information if required to determine the appropriate context. It then subsequently returns to the user the requested data along with a source citation link, directing users back to the original dataset(s) and provides the annotated function performed by the tool to access the data presented in both natural language and SQL code formats. Creating secure, AI-powered tools and algorithms embedded into workflows requires a deep understanding of the complex and fluid regulatory landscape surrounding the use of AI technologies, the types of data used across housing and homelessness, equitable and trustworthy AI considerations, and the end users these tools and constituencies are designed to serve, interact with, and represent. Abt understands that the range of shelter and interim housing programs in each locality is continually evolving. AI offers great promise to promote instant knowledge sharing of resources and to reduce time and frustration spent by people experiencing homelessness themselves and service providers seeking resources.

Across the 58 counties and 482 municipalities in California, local planning departments are routinely understaffed and called upon to generate more detailed reports to the state, which is rightly focused on understanding statewide progress in increasing housing supply. Rather than rely on annual reports from the municipalities that come in well after the fact, AI could facilitate more frequent data reporting on building permits, synthesize information from master plans, zoning or building code changes, and could prompt state officials to explore trends, progress, and/or roadblocks to development in a timely enough manner for the state to provide assistance and options.

Approach

Abt's data science lifecycle methodology is based on a industry standard processes and provides a structured framework. This methodology encompasses a series of interconnected stages that guide data scientists through the process of problem formulation, data collection, analysis, modeling, evaluation, and deployment of insights. Abt's data science delivery framework can be tailored to meet the needs of the project but traditionally consists of the eight phases and incorporates a "human in the loop" as values cannot be built into AI models, and any data used to train AI is inherently biased in one or more ways. The only way to ensure alignment is through deliberate planning from the initial stages of development, as evidenced by Abt's AI lifecycle approach.

Phase 1: Business Understanding – Define the problem, project objectives, and requirements from a business perspective, including the formulation of any initial hypotheses and key performance metrics.

Phase 2: Data Gathering, Exploration and Preparation – Gain an understanding of the available data and assess quality. This includes gathering new data, exploring existing datasets, identifying potential data access or quality issues, cleaning data, identifying outliers and inconsistencies, constructing new variables, and integrating datasets.

Phase 3: Modeling – Build and select appropriate models addressing identified agency objectives by selecting modeling techniques, splitting data into training and testing sets, building models on the training data, and evaluating and validating models using performance metrics.

Phase 4: Evaluation – Evaluate results of the modeling process by reviewing the model's effectiveness, validating findings, and identifying next steps or adjustments that need to be made and evaluated prior to deployment.

Phase 5: Deployment – Deploy the model into production and continue to evaluate performance for model drift.

Phase 6: Maintenance, Monitoring, and Calibration – continuously monitor the model's performance and adjust as needed, reviewing to ensure it continues to meet business objectives and adapts to changes. Regular maintenance is performed to keep the model operating as expected and address new data or changes in the operating environment.

Phase 7 Retirement and Decommissioning – Determine when the model is no longer delivering value or is obsolete. Once that decision has been made, activities include planning for transitioning to a new system or updating existing solutions, and then safely retiring the model, ensuring that data and model components are handled in a compliant and secure manner.