DRAFT STATEMENT OF WORK

TITLE: CANNABIS-DERIVED PRODUCT DATA AND RELATED SERVICES

1. PURPOSE

The purpose of this requirement is to acquire access to data and services to better understand consumers and potential health risks from cannabis-derived products (CDPs) as new entries continue to flood the US market and usage surges. Interdisciplinary big data surveillance on multiple social media platforms and other potential online data sources (including dark web, e-commerce, and special interest websites), using machine learning (ML) and natural language processing (NLP) to identify and characterize online activity that other sources do not capture will be needed. Access to such data and services will help inform and strengthen surveillance and development of policy, guide prevention efforts, and target education/outreach initiatives.

2. BACKGROUND

Implementation of the Agriculture Improvement Act of 2018 (also known as the 2018 Farm Bill) resulted in removal of hemp from the definition of marijuana in the Controlled Substance Act (CSA). As a result, cannabis plants and derivatives containing no more than 0.3% delta-9-tetrahydrocannabinol (THC) on a dry weight basis would no longer fall under the purview of the CSA. Since the 2018 Farm Bill enactment, both the availability of and demand for CDPs have surged, prompting the FDA's Cannabis Product Committee (CPC) to identify critical needs to address the Agency's evidence gaps in the Cannabis-Derived Products Data Acceleration Plan (DAP) (available at this link: https://www.fda.gov/news-events/public-health-focus/cannabis-derived-products-data-acceleration-plan). The DAP calls for data about product problems (e.g., quality concerns), adverse events, and misinformation sources based on real world data (RWD) to help fill knowledge gaps and meet the new and emerging challenges these novel products and markets present with regard to advancing understanding of quality, safety, and surveillance, and informing development of policy, prevention, education, and outreach efforts.

The above-described data and services will help address FDA's data gaps by providing access to information about new and emerging risks and safety signals identified in consumer-reported adverse event experiences with CDPs from multiple social media platforms and other potential online data sources (including dark web, e-commerce, and special interest websites), using machine learning (ML) and natural language processing (NLP) to identify and characterize online activity that other sources do not collect. Such data and services will include analyzing the captured data, as well as providing information on the current scientific literature in general and through a health equity lens in particular to identify gaps, helping the Agency target resources appropriately in terms of potential actions to address the gaps, such as calls for research; postmarket surveillance refinement; sampling prioritization; and development of guidances, alerts, advisories, and other education and outreach. These data and services will also include summarized and FDA-cleared findings by due dates and in formats for dissemination to be determined in consultation with FDA, including but not limited to publications in peer-reviewed journals and meeting presentations/webinars, to promote engagement and participation among the state and local government, scientific, academic, consumer, industry, and public health

stakeholder community. In these ways, access to these data and services will help inform surveillance and development of policy, guide prevention efforts, and target education/outreach initiatives.

3. SCOPE AND OBJECTIVE

The scope and objective of work for this requirement are the same: to help address FDA-identified evidence gaps contributing to strengthening FDA's surveillance and safety signal detection, prevention, education/outreach, and public health protection activities. The goals and expectations with regard to the scope and objective will include, but will not be limited to, addressing the following aims:

- To collect information about new and emerging risks, and safety signals identified in online data sources (including dark web, e-commerce, and special interest websites), and in consumer-reported adverse event experiences with CDPs from multiple social media platforms, using data mining, machine learning (ML) and natural language processing (NLP) to identify and characterize online activity that other sources do not capture.

- To analyze the captured data on new and emerging risks, an safety signals, as well as to provide information on the current scientific literature in general and through a health equity lens in particular (e.g., describing what user groups are targeted by CDP marketing and identifying user groups that interact with CDP online content and communities), to identify gaps, helping the Agency target resources appropriately in terms of potential actions to address the gaps, such as calls for research; postmarket surveillance refinement; sampling prioritization; and development of guidances, alerts, advisories, and other education and outreach.

- To develop and deliver reports on findings by due dates and in formats to be determined in consultation with FDA, including summaries, manuscripts, and presentations/webinars synthesizing the considered data, with clearance obtained through FDA management for sharing with diverse stakeholders as determined appropriate by FDA to promote engagement and participation among the state and local government, scientific, academic, consumer, industry, and public health stakeholder community.

4. TASKS

The Contractor shall provide the following data and services:

4.1 **(Kick-off meeting):** The Contractor shall conduct a kick-off meeting with the Contracting Officer Representative (COR) and FDA staff within ten (10) business days following the effective date of award to provide a plan to include data access, training, and support, and discussion of aims to be addressed for the year, milestones for each task, roles and responsibilities, and timeframes/deadlines.

4.2 **Itemized Timelines/Monthly Progress Reports:** The Contractor shall develop and provide progress reports on a monthly basis to include itemized timelines and budgets by research aim (described above in the Scope and Objective section) to organize and track costs and progress, discuss project deliverables, and allow time for iterative discussion and refinement by FDA staff on project deliverables and objectives.

4.3 **Data collection services:** Using different data mining approaches, including available application programming interface (API) streams and through web data collection approaches, the contractor shall collect unstructured data from popular social media platforms Twitter, YouTube, Yelp, Instagram, TikTok, Reddit, and Telegram filtered for keywords associated with CDPs, adverse event conditions, and associated CDP slang terms. They shall use keywords derived from existing scientific literature in addition to those generated in our own prior studies examining CDPs such as Delta-8 THC. For platforms where keyword filtering is not feasible, or results are limited, they shall use keywords converted to hashtags or generate data based on selection of user groups or community channels of interest to CDP topics. Furthermore, they shall institute a multi-step process to geolocate social media messages that are within a geographic boundary of zip code, city, state, country, and region if location-specific publicly available data is collected. In addition to the text of user-generated posts and comments, publicly available metadata associated with such content shall be collected and analyzed. For select online data sources, they shall meet and confer with FDA CORs to determine if additional websites should be monitored. Additional websites or online marketplaces could include dark web sellers (for purposes of identifying potential illegal or unregulated CDPs that could lead to adverse events) or specific user community forums (e.g., drugbuyersguide.net).

4.4 **Data analysis and synthesis of the literature on CDP adverse events:** The Contractor shall conduct data analysis that will: (a) utilize NLP and ML to characterize the types of messages, content, products, users (including possible health equity factors related to online users or where CDP marketing is targeted), and user networks; and (b) content-code themes that are specifically associated with CDP-associated adverse events as well as the possible source or cause of adverse health impacts (e.g., CDP product implicated, behavioral characteristics, poly use, etc.). Following topic modeling and data exploration, specific posts/messages will be manually annotated to identify a representative sample of messages for specific characteristics of interest that will then be stratified for further analysis. Specifically, data exploration and classification will use a combination of NLP, topic models, and supervised ML approaches to detect “signal” (i.e., user self-reported discussions of CDP adverse events) from textual and image data. In addition to topic models, they shall also assess the utility of using supervised ML classifiers and deep learning to classify larger volumes of texts/posts using training data derived from the topic modeling phase. When inferring demographics such as age, education, and household income, they may aggregate data and overlay on U.S. Census Data joined at the zip code or census track level to better understand the distribution of user perception in specific communities if geolocation data is available. They shall also utilize trained expert human coders to validate signal data and classify content into types of CDPs. They shall incorporate results from the project’s data analysis into a synthesis of the current scientific literature on CDP adverse events, identifying gaps and novel data sources to better characterize adverse events using multiple data sources. They shall utilize existing Python packages and computing infrastructure including but not limited to dedicated customized deep learning server and cloud computing capabilities to complete these study aims.

4.5 Research dissemination deliverables: The Contractor shall lead the development of research abstracts, presentations, and webinar materials, potential scientific symposia, and submission of publications to peer-reviewed journals as part of the research dissemination aim of this project. Specifically, the Contractor shall develop at least 3 research abstracts for submission to academic conferences (e.g., American Public Health Association Annual Meeting, The College on Problems of Drug Dependence, etc.), also considering CDP-related industry conferences, as determined in consultation with FDA staff. The Contractor shall also explore the development of special panel session submissions on the topic of CDPs and adverse events, as well as topics that intersect between CDPs and health equity. Finally, the Contractor Shall lead the drafting, submission, review, and publication of at least 2 original research articles that shall be submitted to a scientific journal as determined appropriate and cleared by FDA staff. The Contractor shall work with FDA staff to appropriately select journals for publication of project results. If determined appropriate and cleared by FDA staff, the Contractor shall also develop rapidly deployable data visualizations to generate additional insights on project results for internal and external stakeholders.

5. DELIVERABLES

The Contractor shall provide the deliverables below for the twelve (12)-month base year period of performance (POP) and each of the subsequent POPs under Option Years 1 through 3 if exercised thereafter:

Deliverables Table
Cannabis-derived Product Data and Related Services

Task	Description	Due Date	Recipients and Format
4.1	Kick-off meeting covering access to data for FDA Cannabis Product Committee (CPC) staff and introductory orientation	Access for the entire period of performance, with the introductory orientation meeting within ten (10) business days after the effective award date	Multiple FDA CPC users via a virtual platform, with Word, PDF, PowerPoint, and/or Excel files provided by email
4.2	Itemized Timelines/Monthly Progress Reports	Due no later than ten (10) days after the end of each month continuing until completion of contract period(s) of performance. Upon completion of contract, contractor will provide a final progress report of all activities completed during the period of performance.	Multiple FDA CPC users via a virtual platform, with Word, PDF, PowerPoint, and/or Excel files provided by email

Deliverables Table (continued) Cannabis-derived Product Data and Related Services			
4.3	Data collection	Data collection will prioritize needs and will commence within 5 business days of start of contract, and will continue until completion of contract period(s) of performance.	De-identified publicly available data generated for this project is available upon request by COR in a format as reasonably requested by FDA (e.g., excel, JSON, etc.). Data will be shared with COR and other relevant CPC users via agreed upon data repository or delivery format.
4.4	Analysis of captured data and literature	Due as determined in consultation with FDA staff. Contractor shall make progress on analysis of captured data and synthesis of literature and present analysis as part of monthly progress reports outlined in section 4.2 above.	Preliminary data analysis results will be provided to FDA COR via a virtual platform, with Word, PDF, PowerPoint, and/or Excel files provided by email
4.5	Development, clearance, and delivery of findings in formats and by due dates determined in consultation with FDA, including but not limited to summaries/reports/manuscripts/presentations /webinars	Contractor and COR will mutually agree upon development, authorship, and timeline of delivery for specific products. It is anticipated that 1-5 summaries and/or research abstracts and at least 2 draft manuscripts shall be made available to COR for review by month 6 of the contract and the remaining deliverables shall be made available prior to month 12 of the contract.	Multiple FDA CPC users via a Word document and associated data files (e.g., Excel, figures in PDF, etc.) provided by email. All research abstract, symposia submissions, and journal articles will be formatted in accordance with the target audience, venue, or journal/publisher.

6. SECTION 508 COMPLIANCE

Section 508 of the Rehabilitation Act of 1973 (29. U.S.C. 794d), as amended by the Workforce Investment Act 1998, requires that all electronic and information technology (EIT) products and services developed, acquired, maintained, or used under this award must comply with the "Electronic and Information Technology Accessibility Provisions" set forth by the Architectural and Transportation Barriers Compliance Board (also referred to as the "Access Board") in 36 CFR part 1194. Information about Section 508 is available at http://www.access-board.gov/guidelines-andstandards/commications-and-it/about-the-section-508-standards. This requirement is shall comply with the following standards.

- E101.2 Equivalent Facilitation (Appendix A, Application and Scoping Requirements)
- E203 Access to Functionality (Appendix A, Application and Scoping Requirements)
- E204 Functional Performance (Appendix A, Application and Scoping Requirements)
- E205 Electronic Content (Appendix A, Application and Scoping Requirements)
- 302 Functional Performance (Appendix C, Functional Performance Criteria and Technical Requirements)
- Electronic content must be accessible to HHS acceptance criteria. Checklist for various formats are available at http://508.hhs.gov/, or from the Section 508 Coordinator listed a https://www.hhs.gov/web/section-508/additional-resources/section-508-contacts/index.html. Materials that are final items for delivery should be accompanied by the appropriate checklist, except up approval of the Contract Officer or Representative.
- E207 Software (Appendix A, Application and Scoping Requirements)
- E208 Support Documentation and Services (Appendix A, Application and Scoping Requirements)
- Chapter 5 Software (Appendix C, Functional Performance Criteria and Technical Requirements)
- Chapter 6 Support Documentation and Services (Appendix C, Functional Performance Criteria and Technical Requirements)

All EIT must meet the applicable accessibility standards at 36 CFR 1194, unless an agency exception to this requirement exists. 36 CFR 1194 implements Section 508 of the Rehabilitation Act of 1973, as amended, and is viewable at: https://www.section508.gov/.

7. PLACE OF PERFORMANCE

8. The place of performance for this purchase order shall be primarily at the Contractor's place of business as noted below:

PERIOD OF PERFORMANCE

The period of performance shall include a period of twelve (12) months from the effective award date for the base year and a subsequent period of twelve (12) months for each subsequent option year thereafter. All tasks shall be done in the respective periods of performance.

9. CONTRACTING OFFICER'S REPRESENTATIVE

TBD