DEPARTMENT OF HEALTH & HUMAN SERVICES
Centers for Medicare & Medicaid Services
7111 Security Boulevard, Mail Stop B3-30-03
Baltimore, Maryland 21244-1850

Office of Acquisition and Grants Management

REQUEST FOR INFORMATION

Agency: Centers for Medicare & Medicaid Services

Office: Office of the Administrator (OA) in collaboration with the Office of Acquisition and Grants Management (OAGM)

Title: National Provider Directory (NPD) Work Areas A and E: Ground Truth Datasets and Data Aggregation, Quality & Assurance Services

The Centers for Medicare & Medicaid Services (CMS) is issuing this Request for Information (RFI) as a means of conducting market research to determine the availability and technical ability of vendors regarding the establishment of a trusted, accurate, and comprehensive source of healthcare provider information that serves as the foundation for improved care coordination, reduced administrative burden, and enhanced transparency across the U.S. healthcare system. The result of this market research will contribute to determining the agency's method of procuring such services in the future.

Purpose

A draft Statement of Objectives (SOO) including National Provider directory Ground Truth Datasets and Data Aggregation, Quality & Assurance Services are being provided as part of this RFI. CMS asks interested vendors review the SOO and address the questions based on their experience and expertise. The feedback from this RFI may be used to help refine the SOO and future Request for Quotation (RFQ).

Request for Information

THIS IS A REQUEST FOR INFORMATION (RFI) ONLY. This RFI is issued solely for CMS' information and planning purposes. This RFI does not constitute a Request for Proposal (RFP), RFQ, or a promise to issue an RFP or RFQ in the future. This RFI does not commit the Government to contract for any supply or service. CMS is not soliciting quotes/proposals. Respondents are advised that the U.S. Government will not pay for any information or administrative costs incurred in response to this RFI; all costs associated with responding to this RFI will be solely at the respondent's expense. Not responding to this RFI does not preclude participation in any future procurement. It is the responsibility of the interested vendors to monitor this site for additional information pertaining to this requirement. Information provided in the RFI is subject to change and is not binding on the Government. HHS, CMS, CM or OAGM have not made a commitment to procure any of the items or services discussed, and release of this RFI is not a commitment or an authorization to incur cost for which reimbursement would be allowed or sought. All submissions become Government property and will not be returned.

Responses to this RFI shall be submitted by July 10, 2026, via email to the Contracting Officer, Alex Jarema at alex.jarema@cms.hhs.gov and the Contract Specialist, Walker Hare at

walker.hare@cms.hhs.gov in a format that is readable in Microsoft Word or Adobe PDF formats. CMS will not accept hard-copy responses or other formats. CMS will consider only those responses that contain the information described below in the format requested. Respondents are requested to provide a response to this RFI that includes three sections as follows below.

Section I. General Vendor Information

- Organization Name
- Unique Entity Identification (UEI)
- Business Size and Small Business subcategory if applicable
- Mailing address
- Name, title, phone number and e-mail address of designated point of contact
- Provide a description of any licenses, specific expertise, any General Services Administration (GSA) credentials (if applicable), and information that is relevant to this RFI.
- Active GSA MAS, GWAC, or other government contract vehicles (list vehicle number and expiration)

Section II. The name and contact information of the organization whose views are represented in the submission, if different from the information provided in Section I.

Section III. Respondents are encouraged to provide complete but concise responses to the questions listed below. Please note that a response to every question is not required. For each question or remark, please include reference to the specific section(s) of the SOO to which the question/feedback applies following the format below (i.e., the format below assumes a table with columns associated with each element). Contractors should use a file format that will facilitate CMS merging questions from multiple respondents into a single file.

Question Number	SOO Section or Task and Page Number	Question (for Respondent use)	Response (for CMS use)

CMS will not consider additional information submitted beyond these sections.

Questions

1. CMS is seeking to establish direct licensing relationships with authoritative provider data sources across multiple categories. For each data category below, indicate:
   (a) whether you ARE the authoritative source for this data, whether you HAVE an active agreement with the authoritative source, or both;
   (b) whether federal government use is permitted;
   (c) whether CMS can freely redistribute derived data in a public directory without downstream licensing obligations; and

(d) whether any fields in this data category contain sensitive or restricted information that CMS could use internally but could not publish publicly.

Data Category	**Are You the Source / Do You Have an Agreement**	**Federal Use Permitted?**	**License-Free Redistribution Permitted?**	**Any Fields That Cannot Be Publicly Published?**
Clinician licensure (state licensing boards or FSMB)	Source / Agreement / Both / Neither	Y / N / Partial	Y / N / Requires negotiation	Y / N – if yes, describe briefly
Board certification (ABMS, ABPS, or equivalent)	Source / Agreement / Both / Neither	Y / N / Partial	Y / N / Requires negotiation	Y / N – if yes, describe briefly
Credentialing platforms (CAQH or equivalent)	Source / Agreement / Both / Neither	Y / N / Partial	Y / N / Requires negotiation	Y / N – if yes, describe briefly
Pharmacy data (NCPDP or equivalent)	Source / Agreement / Both / Neither	Y / N / Partial	Y / N / Requires negotiation	Y / N – if yes, describe briefly
Hospital/facility data (AHA or equivalent)	Source / Agreement / Both / Neither	Y / N / Partial	Y / N / Requires negotiation	Y / N – if yes, describe briefly
Organizational accreditation (TJC, NCQA, or equivalent)	Source / Agreement / Both / Neither	Y / N / Partial	Y / N / Requires negotiation	Y / N – if yes, describe briefly
Payer/health plan data	Source / Agreement /	Y / N /	Y / N / Requires	Y / N – if yes,

(NAIC or equivalent)	Both / Neither	Partial	negotiation	describe briefly
Provider Network Participation data	Source / Agreement / Both / Neither	Y / N / Partial	Y / N / Requires negotiation	Y / N – if yes, describe briefly

For state licensing boards specifically: list the number of states for which you currently are the source or hold active data access agreements.

***Note on sensitive fields: CMS understands that some data fields may be used for internal validation purposes only and may not be appropriate for public publication. Please flag any such fields in the final column — this will help CMS design appropriate data handling and publication policies.*

2. A core requirement for this contract is that all data licensing agreements must permit CMS to freely redistribute reconciled NPD data without downstream licensing obligations on end users. CMS recognizes that achieving this in practice may require negotiating field-level publication terms — some data fields within an otherwise licensable dataset may be appropriate for internal validation use only and may not be appropriate for public publication due to sensitivity, commercial restrictions, or legal constraints.

   Describe a specific engagement in which you negotiated a data licensing agreement that explicitly permitted a federal government client to freely redistribute or publicly publish derived data. Your response should address:

   (a) What type of data source was involved? What specific licensing restriction did you encounter? How did you resolve it and what was the outcome?

   (b) Were there any fields within the dataset that could not be publicly published even after successful negotiation — for example, due to sensitivity, commercial restrictions, or state legal constraints? If so, how did you handle the distinction between fields that could be published and fields that could only be used internally?

   (c) If you have encountered situations where full license-free redistribution was not achievable but partial redistribution was negotiated — covering most but not all fields — describe how you structured that arrangement and how you ensured that restricted fields were not inadvertently included in public output.

   (d) If you have not negotiated this type of agreement before, describe how you would approach it for the data categories listed in Question 1, including how you would identify and handle fields that may be sensitive or legally restricted from public publication.

3. Work Area A requires Initial Operating Capability within 30 days of award — defined as normalized data files delivered to CMS-designated storage for at least five (5) ground truth

sources, with all priority sources in active negotiation.

(a) State licensing boards are highly fragmented across 50+ jurisdictions. Describe your experience managing data relationships with state licensing boards at scale. How many active state board relationships do you currently hold? What is the most significant obstacle you have encountered in establishing a state board data agreement and how did you resolve it?

(b) Based on your existing agreements, which specific sources could you have operational within 30 days of award? What would you do before award to prepare, and what is the primary risk to the 30-day timeline?

(b) Are the IOC timelines realistic — 30 days for Work Area A and 45 days for Work Area E? If not, what would you recommend?

4. Work Area E is an outcome-based framework. CMS anticipates evaluating vendors on the measurable quality of their data corrections — not on correction volume. WA E operates at both ends of the NPD data lifecycle: validating data quality as it enters the pipeline from ground truth sources AND validating the accuracy of data published in the public-facing directory, with findings fed back to improve ground truth sourcing decisions.

   (a) Describe a specific, production-scale engagement in which you delivered documented improvements to provider directory data accuracy. Include: the starting baseline, the types of errors corrected, the quantified outcome, the scale (number of records), and the methodology used. PoC or pilot results alone are not sufficient — describe a production deployment.

   (b) Have you assessed data quality at both the input end (data entering a pipeline) and the output end (data published to end users) for the same client? If yes, describe that engagement. If no, describe how you would approach it for NPD.

   (c) What correction acceptance rates have you achieved in prior engagements? Provide specific data. How do you ensure your methodology does not generate high volumes of low-quality corrections? Can you fully document your methodology — including any AI/ML model specifications — in a format CMS can review, audit, and own? Are there proprietary elements that could not be transferred to a successor contractor?

5. Based on your review of the draft SOO:

   (a) Are there significant gaps, ambiguities, or over-specifications in the SOO that CMS should address before solicitation release?

   (b) Is the license-free output mandate achievable across all data categories described? Are there specific categories where this would be particularly difficult or costly — including categories where only partial field-level redistribution may be achievable?

(c) Does the boundary between Work Area A and Work Area E make sense from an industry perspective? Are there aspects of the scope that are difficult to separate cleanly between the two work areas?

(d) Are there any aspects of the scope that would significantly limit competition?

6. What is the estimated market cost range for services of this nature, and what factors most significantly drive cost variability for an effort of this scope? Your response should include an estimated annual cost range for WA A (data sourcing and licensing) and WA E (aggregation and quality assurance) separately.