G A O
Accountability * Integrity * Reliability

Comptroller General
of the United States

United States General Accounting Office
Washington, DC 20548

Decision

Matter of: The November Group, Inc.

File: B-292483

Date: September 30, 2003

Alison L. Doyle, Esq., and Jeffrey R. Boodman, Esq., McKenna Long & Aldridge, for the protester.
Scott C. Briles, Esq., Department of Health and Human Services, Centers for Disease Control and Prevention, for the agency.
Susan K. McAuliffe, Esq., and Christine S. Melody, Esq., Office of the General Counsel, GAO, participated in the preparation of the decision.

DIGEST

Protest alleging that agency improperly released solicitation materials containing protester's proprietary information is denied where protester fails to provide clear and convincing evidence rebutting agency's determination that materials at issue were based on publicly available information and general quality control concepts commonly used in epidemiological studies.

DECISION

The November Group, Inc. (TNG) protests the terms of request for proposals (RFP) No. 2003-N-00781, issued by the Department of Health and Human Services, Centers for Disease Control and Prevention, for the development of a database system and related data collection and reporting assistance services to track the clinical use of assisted reproductive technology (ART) procedures in the treatment of infertility. The protester is the database system subcontractor of the incumbent contractor of ART data collection services for the agency, the Society for Assisted Reproductive Technology (SART). TNG alleges that certain solicitation materials issued by the agency (namely, a cleaning criteria document released in response to an offeror's question about, among other things, edit specifications for checking the accuracy of data) disclosed proprietary quality control and programming features of its ART data collection software. The protester contends that, since the released agency edit specifications apply to the protester's software, they reveal proprietary software coding and logic rules by stating, for example, the inverse of TNG's allegedly proprietary logic rules. TNG argues that with the released information other firms will be able to replicate its software and easily access its proprietary database, which, according to TNG, deprives it of the competitive advantage it expected to

enjoy as the subcontractor providing the current ART data collection services for the incumbent contractor. The protester seeks a sole-source contract under the RFP or a requirement that other offerors purchase a license from TNG for use of the protester's ART database system.

We deny the protest.

The RFP, issued on May 19, 2003, contemplates the award of a cost-reimbursement type contract for a 7-year period to the firm submitting the proposal considered to offer the best value to the government. RFP at 34, 81. The contractor is to develop a standardized data collection system to track the use of ART procedures by clinics and medical practices in the United States and its territories. Among other things, the contractor is to provide: software to tabulate ART data and import the data into the agency's reporting system; software distribution and instruction; a paper data abstraction form; and a quality assurance program for the assessment of the quality and completeness of the data received, including detection of logic errors between data set elements, and out of range or otherwise questionable values for each data set element. RFP at 10. The challenged solicitation information released by the agency relates mostly to the quality assurance edit checks used by the agency to detect logic errors and questionable data under the current system.[1]

The RFP attachments define each data element to be reported by the clinics and provide information as to the relationships among specific data sub-elements. For instance, some data sub-elements are mutually exclusive, indicating that an affirmative response for one should preclude an affirmative response for another,

---

[1] The agency does not claim to have proprietary rights to the protester's software or its database system, which is entitled SART Clinic Outcome Reporting System (SART CORS). The agency reports that prior to this protest, it believed that SARTCORS was proprietary to its prime contractor, SART, which organization currently provides ART data collection services to the agency and subcontracts for the database system and related services with the protester. As stated above, we view TNG's protest as essentially limited to the cleaning criteria (or edit checks) released by the agency, since, although the protester initially challenged the agency's release of its data field names and data description, the protester conceded in its comments responding to the agency's report that the data descriptions are based on public information, namely, the solicitation's list of the required data elements. We similarly believe that the same public information clearly served as the basis for the generic truncated field names used by TNG (and released by the agency) as the variable name that merely identifies the publicly described data. Since the protester has not supported its general claim that the release of its shortened data element names alone will give others improper access to its software, and has otherwise provided insufficient basis for us to consider its claim of proprietary rights to the information, we do not consider the challenge of the truncated fields further in this decision.

while others are collective in that the sum of certain factors or subfactors may define a data element. For instance, under the data element of "gravidity" (meaning the number of the patient's prior pregnancies), each clinic was to separately report related sub-elements of such information, such as the number of the patient's full and pre-term births, and prior spontaneous abortions. Another example of the extensive data collection system information released in the solicitation materials, and unchallenged by the protester, is the sample data documentation table provided to illustrate how the data could be recorded by variable name (in a shortened form), data type (whether reported numerically or otherwise, such as by date or time of the reported procedure), variable description (defining the required data element as set out in the solicitation), data format and codes (identifying the specific software program code assigned by the programmer to the data element), and quality control issues (concerning the logic rules to apply to checking the accuracy of the data collected for that data element).[2] For quality control purposes, for instance, an ART patient's date of birth reported outside the range of 18 to 60 years of age would be questionable as illogical for falling outside of a determined general childbearing age range. The RFP also provided detailed data descriptions and substantive medical definitions through its express incorporation of publicly available documents, such as the agency's recent annual ART data report, and a detailed Federal Register notice, 65 Fed. Reg. 53310, 53312-6 (Sept. 1, 2000), containing the agency's comprehensive definitions and explanations of relevant ART terms and procedures, and providing substantial subject matter information, including information about the relationships among the cited ART data elements to be reported.

Many questions were received from prospective offerors about the terms of the current system. The agency explained, however, that it did not have proprietary rights in that system and thus could not release its software or proprietary terms. In response to a request for more information about edit specifications, the agency explained that logic and range checks were to be provided and that they were to be as comprehensive as the type of edit checks the agency uses to assess the quality of the current contractor's data. To illustrate the type of edit checks desired, the agency posted at the solicitation's Internet site a copy of its cleaning criteria document, including the data edit checks it runs to assess the accuracy of data reported by its current ART data collection contractor, SART. In that document, for instance, the agency's edit checks for the data element of "gravidity" (the number of the patient's prior pregnancies) include questioning data recorded as less than 0 or greater than 11 pregnancies for a patient; under the agency's cleaning criteria for the accuracy of gravidity data, the number of prior pregnancies reported is also checked for accuracy if it is recorded as a number less than the sum of separately reported

---

[2] Despite a noted similarity in format of this sample data documentation table and the format of the protester's own data field map, TNG has not challenged the agency's release of this table.

data accounting for the number of the patient's full and pre-term births, plus the number of spontaneous abortions.

After this information was posted, the protester promptly notified the agency that, in its view, the cleaning criteria document revealed proprietary information about TNG's software. As support for its assertion, TNG pointed to certain similarities between its and the agency's quality control edit terms. For example, in the protester's allegedly proprietary preliminary field map, TNG provides a similar edit check for "gravidity" that includes confirmation that the sum of a patient's reported full and pre-term births, plus spontaneous abortions is less than or equal to the number of reported prior pregnancies. Citing similarities such as this between the terms it uses and the information released by the agency, and contending that the agency's cleaning criteria merely inverted the terms of some of its edit checks, the protester sought relief from the agency for allegedly releasing the protester's proprietary information.

The agency disagreed with the protester's assertion of proprietary rights, noting that the released edit checks were disclosed only as examples of the agency's efforts (and the desired comprehensiveness of any offeror's proposed efforts) to check the accuracy of the ART data reported. The agency further explained that its cleaning criteria were the results of its own work over the course of several years checking the accuracy of data reported on the current (SART CORS) system. The agency reasoned that the release of its cleaning criteria is unobjectionable not only because they are typical of the type of edit criteria commonly used in similar epidemiological studies, but because they are based on publicly available information about ART and the agency's data collection requirements, as well as general principles of logic and statistics relevant to data collection studies. Due to the protester's continued insistence, and as a precautionary measure, the agency removed the challenged cleaning criteria document from its solicitation materials on the Internet 16 days after it had been posted; the public was instructed to destroy any copies of it. This protest followed.

We have recognized that a firm may protect its proprietary data from improper exposure in a solicitation where its material was marked proprietary or confidential, or was disclosed to the government in confidence, and where it involved significant time and expense in preparation and contained material or concepts that could not be independently obtained from publicly available information or common knowledge. The Source, B-266362, Feb. 7, 1996, 96-1 CPD ¶ 48 at 2. To prevail on such a claim, the protester must prove by clear and convincing evidence that its proprietary rights have been violated. Zodiac of North America, Inc., B-220012, Nov. 25, 1985, 85-2 CPD ¶ 595 at 3. TNG has not met this standard here. On the contrary, our review of the record confirms the reasonableness of the agency's position that TNG has not provided sufficient evidence to establish a proprietary right to the information released in the agency cleaning criteria document.

As stated above, TNG argues that a comparison of its preliminary data field map to the agency's cleaning criteria document reveals similarities in terms; according to TNG, since its field map was developed first, the agency should be found to have improperly derived its cleaning criteria from the protester's proprietary quality control terms. Our review of the two documents, however, provides no basis to sustain the protest on this ground. As an initial matter, as the agency points out, there are substantial differences in the two documents. For instance, the protester's document is not as comprehensive as the agency's document, and it does not include as many quality control specifications as those identified in the agency's document. Also, contrary to the protester's suggestion, the agency's document does not reveal the protester's software's extensive coding of data elements and sub-elements that constitute its underlying programming and reasonably would be required for others to replicate the protester's database system. In this regard, we cannot find that the agency's indication in a limited number of cleaning criteria of the number of sub-elements included in some of TNG's data fields constitutes the improper release of proprietary information, since the solicitation and above-referenced Federal Register notice specifically and publicly already identified the required data elements and sub-elements to be tracked. The mere number of items tracked by the protester, therefore, in our opinion, reveals insufficient information to warrant additional protections or other relief, since the proprietary coding of how the information is tracked and recorded has not been released.

Most important to the resolution of this protest, however, is the substantial amount of detailed public information that is readily available about ART procedures and approaches, as well as this agency's specific reporting requirements in the ART field. Extensive information remains publicly available about the specific data elements to be reported, and their related sub-elements; as stated above, offerors were also referred in the solicitation to a comprehensive Federal Register notice that further defined all of the data elements to be reported and how they relate to each other. Clearly, to satisfy the solicitation's requirements for comprehensive quality assurance checks for the data reported, any offeror would reasonably be expected to utilize this specific information and terminology in crafting its data edit checks. Thus, contrary to the protester's contention that any similarity in the terms of the agency's cleaning criteria to its own shows that the agency misappropriated the protester's data, we believe the record more reasonably supports the agency's position that the similarity in terms reflects the reasonable, yet independent, application of common data cleaning rules to the specific data elements and medical terminology at hand here. In this regard, using the above stated example of "gravidity," a data element under the RFP defined as the number of the patient's prior pregnancies, we cannot find reasonable the protester's position that its approach was unique in designing an edit check for this data element on the basis of the sum of stated underlying factors related to the outcomes of that number of pregnancies, such as full and pre-term births, and spontaneous abortions. On the contrary, the solicitation itself implied this edit check was appropriate through its incorporation of the Federal Register notice definition of "gravidity" which expressly

includes the factors (full and pre-term births, and spontaneous abortions) used by both the protester and the agency in developing their cleaning criteria terms.

In short, the protester has not met its burden of proving by clear and convincing evidence that its proprietary rights have been violated by CDC's disclosure in the solicitation materials.[3] Rather, at issue here is material or concepts that we believe can reasonably be independently derived from public information. This includes publicly available scientific and medical information expressly incorporated by reference in the solicitation, describing not only the ART subject matter, but also the specific data elements to be reported and the relationship among the sub-elements for purposes of checking the accuracy of such data. Further, the record shows that the final step in the process is application of commonly known principles of logic and statistics typically used to assess the accuracy of information gathered in similar types of epidemiological studies. Accordingly, we conclude that the protester has failed not only to demonstrate sufficient uniqueness in the limited material it claims is reflected in both its preliminary data field map's quality control terms and the agency's cleaning criteria document, but also has not sufficiently supported its assertions of the proprietary nature of the challenged information.[4]

As a final matter, we note that even assuming the protester had adequately supported its contention of proprietary rights in the released material, there would be no basis to recommend the relief sought by TNG. The record is clear, by the terms of the challenged document itself, that the agency released the cleansing criteria for informational purposes only as a guide for use by offerors in the preparation of their own proposed edit specifications and quality control terms. The information simply was not released as a requirement for offerors to propose the allegedly proprietary material (which, even if proprietary, would, at best, still only reflect some aspect of the protester's approach to meeting a term of the RFP). See Vinnell Corp., B-230919, June 30, 1988, 88-2 CPD ¶ 4 at 3. As such, even if the protester had established proprietary rights in the material, there would be no reasonable basis for us to conclude that the extraordinary remedies sought by the protester--a recommendation for a sole-source award or a requirement for each

---

[3] Our Office held a hearing on the protest to receive testimony from the parties regarding the allegedly proprietary nature of the material at issue. The hearing testimony supports our conclusion that the protester has failed to make the required showing that its proprietary rights have been violated.

[4] Given our conclusion that the protester has not established the proprietary nature of the material at issue, we need not reach the other prongs of the legal standard in this area. See The Source, supra. We note, however, that, based on the record, questions remain as to the adequacy of the protester's actions to protect the alleged proprietary material and the degree of collaboration between the agency and the protester in developing the challenged material.

offeror to purchase a license from the protester--would be appropriate. See Sentel Corp., B-244991, Dec. 6, 1991, 91-2 CPD ¶ 519 at 3, recon. den., B-244991.2, May 5, 1992, 92-1 CPD ¶ 419 at 3.

The protest is denied.

Anthony H. Gamboa
General Counsel