Data Warehouse RFP 2014
Answers to Requests for Clarification

1Q. What platforms (database, ETL and BI), does Westchester Co currently use for data warehousing/What is the version of Oracle, SQL Server etc. expected?
1A. At least Oracle 11g or SQL Server 2012 is expected, no ETL or Business Intelligence (BI)

2Q. Are there incumbents already providing services?
2A. No, there are no incumbents already providing services.

3Q. If so, will they also bid?
3A. N/A

4Q. If awarded, we propose to work on the project in a collaborative mode utilising our affiliate’s resources and expertise from outside USA. Is this acceptable?
4A. If you propose to subcontract any part(s) of this work to any entity/ies, you must disclose that in your proposal.

5Q. What % of work is expected onsite and which of the tasks?
5A. Most work can be performed offsite. There may be required periodic face-to-face meetings with relevant program people in the CoC Partnership.

6Q. Our References will be from private entities that can vouch for our capability to work in a complex DWH environment, is this acceptable?
6A. As described in Schedule “A-2”, you must provide three (3) references that can speak about you with regard to each of the seven (7) specified qualifications. Private entities that can vouch for your capability to work in a complex data warehouse environment are fine as references, as long as those entities you list can speak about you for the specified qualification(s) for which you list them.

7Q. Is there a budget allocated for the requirements as part of this RFP?
7A. As specified in the “Evaluation Criteria” portion of the RFP, while the costs associated with the services will be one of the criteria used to evaluate proposals, it is not the sole criterion. The County will select the proposal from a responsible proposer that is most advantageous to the County. In general, it is the policy of the County to procure services of maximum quality at the lowest possible cost under the circumstances. Accordingly, the County encourages you to propose your best possible pricing.

8Q. Is this a single award or multiple vendor award?
8A. The County anticipates making a single award. However, it reserves the right to enter into one or more agreements, for all or only portions of the services solicited by this RFP, with one or more of the proposers, or to not to enter into an agreement for any of the services solicited by this RFP.

9Q. Is the data to be organized available ready at the County?
9A. The initial data sources to be used are located at the County or maintained by an application service provider which can be exported.

10A. What specific documentation are required by the County as part of Development phase?
10A. Examples of documentation to be included in this phase would be descriptions of tables and data elements in the data warehouse that can be used for reports, data flow diagrams between data sources and the data warehouse, and report inventory with detailed report descriptions.
11Q. In the Response requirements on Page 22 of RFP, is it required that we provide specific Dept. of Social Service, CoC Partnership related experiences or can similar DWH functional experiences be provided?

11A. For page 22 (Schedule “A-1”), specific Department of Social Services-related and/or CoC Partnership-related experience is not required. You must simply provide a submission that is responsive to the questions that are posed and provide the information that is requested. To the extent that you want to reference prior experience in your response, you are welcome to reference similar data warehouse experience.

12Q. Are there defined report requirement document available. If so, can that be shared?

12A. Currently there are no formal defined report requirement documents.

13Q. What is the indicative number of reports/dimensions/measures to be incorporated into the BI?

13A. The successful proposer will have to determine this in the course of providing the services.

14Q. Are there any reporting tools already being used/planned to be used/preferred for BI?

14A. There are no reporting tools used/planned/preferred at this time. We will leave it to the successful proposer to make specific recommendations.

15Q. In addition to following, are there any other aspects of BI that are to be considered:

- Standard and adhoc reporting
- Exporting features
- Reports distribution
- Web based user interface
- Security and user administration

15A. These aspects must be considered during the Development Services term.

16Q. Please provide an indicative number for the following:

- Source database systems to be integrated (Oracle/SQL Server/any other)

16A. There are currently two source database systems to be integrated.

- Indicative number of source tables

16A. There are approximately 50 source tables.

- Indicative data size across all the different sources/ What kind of data volume is expected within this solution?

16A. There are between 15000-20000 client records and subsidiary data. Approximately 5-10% growth per year is anticipated, but may vary.

17Q. Please confirm on the data load frequency, daily/hourly/weekly/How often should we extract the data from source systems into the Data warehouse? Will it be on a daily, weekly, monthly or real time basis?

17A. The data should be extracted on a daily basis.

18Q. Please provide a high level overview of user base i.e. user roles, indicative number of users/How many people (Managers, Supervisors, and Key Members) require training?

18A. Initially, there will be an estimated 10-20 data warehouse users. These users will generally be comprised of key stakeholders and decision makers in the CoC Partnership. It is anticipated that the estimated number will require training.

19Q. Is server sizing, hardware, software setup part of the scope?

19A. Yes, server sizing, hardware and software setup are all part of the scope of work.

20Q. Is the proposed BI system planned to be hosted in-house?

20A. No, the proposed BI system is not currently planned to be hosted in-house.
21Q. What are the different formats of data that are being captured in the source systems (OLTP applications)?
21A. Currently, the source data is relational.

22Q. Please state the number of source systems?
22A. There are currently 2 known source systems.

23Q. What are the different types of technologies/applications used to capture the source systems?
23A. The source data is captured using web-based applications which are connected to relational backends.

24Q. In order to build the successful Data Warehouse and Business Intelligence (BI) reporting application, it is our assumption that the common standards are followed by the participating states, otherwise there will be an impact on the project delivery. Please correct if otherwise.
24A. There is a common data standard that is published by the United States Department of Housing and Urban Development (“HUD”). This common data standard can be changed by HUD periodically (every 2-3 years), but has historically not been changed significantly. The CoC Partnership is required to follow HUD’s common data standard (including any changes) for reporting and data entry purposes. This standard will naturally be extended with elements that are critical to the CoC Partnership based on reporting and analysis requirements.

25Q. What will be the expected volume of data (sizing of data) that will be extracted into the Data Warehouse on daily basis?
25A. The expected volume of data daily to be extracted would be small. This can be measured in the number of client records that would be created or updated. An estimated upper limit would be 100 client records.

26Q. What will be the expected growth of the Data Warehouse over the next three years – in terms of the volume of data? We would like to know as new states are added to this application, the users and the volume of data will change.
26A. Anticipated growth in volume per year is approximately 5-10%, but may vary.

27Q. Will SBAC be providing any test cases?
27A. It is not clear what SBAC stands for. Accordingly, your question cannot be answered.

28Q. What’s percentage of existing data quality is the County anticipating?
28A. The County anticipates a minimum of approximately 90-95% of the existing data to be of sufficient quality for use.

29Q. How “clean” are the data? Do they need to be transformed or cleansed before they are uploaded to the data warehouse?
29A. Because the data is being input through web-based systems, it is relatively clean. Because we are combining data across 2 databases, there will likely be some transformation needed to standardize the data.

30Q. Will the County have a technical team to support the vendor on gathering requirements and to discuss technical details?
30A. Yes.
31Q. Does the County have any predefined Key Performance Indicators planned or is the selected Vendor expected to provide this?
31A. This is not clear right now.

32Q. Is the County expecting any reports on Analysis Services? If yes, how many?
32A. This is not clear right now.

33Q. Does the County have any idea regarding the number of reports to be designed? If yes, how many?
33A. This is not clear right now.

34Q. Does the County have any specific Business Intelligence (BI) tools in mind?
34A. There are no reporting tools used/planned/preferred at this time. The County will leave it to the successful proposer to make specific recommendations.

35Q. What kind of reporting tools are the County currently using?
35A. The County uses Crystal Reports, but not as part of the CoC Partnership’s operations.

36Q. How many users (System Admin, DBA, power users) will connect to the servers (database servers, reporting servers, and domain servers)? Please provide an estimated total.
36A. Initially, there will be an estimated 10-20 users of the data warehouse.

37Q. Once the Data Warehouse meets the success criteria, what other problems might it solve for the County?
37A. As specified in the RFP, the successful proposer will have to provide the Post-Development Services. Exactly what those services might particularly involve is not clear right now.

38Q. What kind of applications are the County currently using to access reports, create reports, create measurements for the homeless?
38A. The County currently uses a Homeless Management Information System (HMIS) and another in-house shelter system to access and create reports. The HMIS generates some measurements and there is an online reporting system that HUD maintains that provides some measurements based on reported data.

39Q. How many legacy systems will be used in creating the Data Warehouse?
39A. There will be 2 legacy systems used in creating the Data Warehouse.

40Q. If the Data Warehouse is down or corrupted, what is the recovery time required by the County?
40A. Recovery time should be no worse than 48 hours.

41Q. Page 2 of the RFP it states “All requests for clarification MUST be submitted, as set forth below, no later than 1:00 pm on April 9, 2014” Did the County mean April 15 by 1pm?
41A. Initially, yes. This date was extended to May 12, 2014 at 1 PM. Please see posted RFP.

42Q. Has Westchester Co identified the DB platform for this data warehouse or are you looking for recommendations?
42A. We are looking for recommendations.

43Q. Has Westchester Co identified the ETL solution for this data warehouse or are you looking for recommendations?
43A. We are looking for recommendations.
44Q. Has Westchester Co identified the BI platform for this analytics solution or are you looking for recommendations?
44A. We are looking for recommendations.

45Q. Beyond the traditional DW/BI components (DB, ETL and BI layers), are the requested MDM and data profiling features required or optional?
45A. The requested master data management and data profiling features are required.

46Q. Please confirm that you desire assistance establishing a data governance program as part of this effort. If not, please clarify.
46A. The CoC Partnership has a data governance program in place for the Homeless Management Information System (HMIS). The CoC Partnership will need assistance from the successful proposer in enhancing this program to take into account the data warehouse.

47Q. Please provide details/specifs around the targeted source systems within scope, including any available software names and versions.
47A. The HMIS (Homeless Management Information System) is a HUD-mandated system hosted by an application service provider which is being used by the CoC Partnership to manage enrollment and service delivery to homeless families and singles in programs funded by HUD. The County Department of Social Services also maintains a shelter system used for managing placement activity of homeless families and singles.

48Q. For business requirements, please share any identified subject areas and estimated numbers of dimensions and facts to assist us in sizing the effort. Any other supporting requirements or details would be welcome as well.
48A. We have not developed these requirements yet. The vendor who is selected will need to work closely with the CoC Partnership to develop these requirements.

49Q. What is the current environment? How is data being gathered, organized and analyzed now? How many data source systems exist?
49A. There are two web-based systems used for data collection and some canned reporting.

50Q. What are the proposed plans for the data warehouse, in terms of hardware infrastructure? Is there a preferred list of o/s and database s/w?
50A. Hosted solution and SQL Server as a DB would be preferred.

51Q. Reporting - What is the current reporting environment? Is there a preferred list of BI tools?
51A. Canned reporting from each of the two current applications. One system has some better integration with Microsoft reporting tools, but this is not being leveraged at the current time. There is not a preferred list of BI tools.

52Q. ETL - What is the proposed SLA for loading the data from source to the data warehouse - is it real time or batch jobs?
52A. Batch is sufficient.

53Q. Is there any data governance process in place?
53A. The CoC Partnership has a data governance program in place for the Homeless Management Information System (HMIS).
54Q. What is the architecture framework followed currently?
54A. Application architecture is web-based, multi-tier. One system is hosted. Systems are developed with JEE and ASP.NET technologies. One system uses Oracle as a backend, the other uses SQL Server.

55Q. Do you have to sunset any existing systems when you have the DW in place? If so what is the timeline?
55A. The County does not anticipate having to sunset any existing systems when the data warehouse is in place.

56Q. RFP format - Is there any proposal template or guidelines to be used for RFP submission for the Data warehouse project?
56A. The minimum requirements (including a listing of the required certification/forms/schedules) and applicable guidelines for proposal content are located in the RFP in Part 1, Section D, entitled “Proposal Content.” As long as these minimum requirements are met, and the applicable guidelines are followed, a prospective proposer can format its proposal however it chooses.