



**ADDENDUM NUMBER 1: QUESTIONS AND ANSWERS**

**FOR**

**REQUEST FOR PROPOSALS FOR DATA ANALYTICS  
PLATFORM SERVICES**

**March 26, 2021**

**Question #1:** SVCE had issued a similar requirement earlier and the RFP was closed. Why is it being re-published?

**Answer #1:** SVCE released a DAISY 2.0 RFI in December 2020. This was intended to help SVCE gather more information to develop this RFP.

**Question #2:** Is there a preference for staying on GCP universe or migrating?

**Answer #2:** SVCE has found Google Cloud Platform (GCP) to be sufficient for its needs and use cases, but is open to migrating to another platform if there is sufficient reason to do so.

**Question #3:** Who is the current vendor for DAISY 1.0 and will they be bidding in this RFP?

**Answer #3:** SVCE is working with Camus Energy to develop and maintain DAISY 1.0. The current DAISY 1.0 vendor is permitted to bid on this RFP.

**Question #4:** Do you have a listing of the enabling tools that you have considered?

**Answer #4:** SVCE currently uses ArcGIS Pro for geocoding and spatial analysis. SVCE is aware of other geocoding tools such as Google Maps API and SmartyStreets, but has yet to fully evaluate these tools. SVCE seeks recommendations from the bidder on other enabling tools to consider.

**Question #5:** Is it possible to get a demo of DAISY 1.0 functionality so we can understand the basic expectations?

**Answer #5:** DAISY 1.0 is hosted on GCP and uses GCP services such as Google Cloud Storage, BigQuery, Data Studio, Compute Engine, and Cloud SQL. Commonly used functionalities include ad hoc queries to analyze and export data (CSV, Excel, Data Studio), visualize data and build dashboards and reports in Data Studio. DAISY 1.0 also includes a pipeline to receive and process meter data on a daily basis.

**Question #6:** Are there any quality checks for automated data ingestion in the current system?

**Answer #6:** Quality checks are in place in DAISY 1.0 and should be carried over to DAISY 2.0.

**Question #7:** What type of connection capabilities does DAISY 1.0 currently have?

**Answer #7:** DAISY 1.0 has a connection established with SVCE's meter data manager to receive daily meter and load data.

**Question #8:** About how much data are you processing regularly and what is the estimated range of the data warehouse size?

**Answer #8:** Currently, our BigQuery data warehouse contains around a dozen core tables and is around 5 TB in size, growing at 100s of GBs per year. The majority of this new data is 15- or 60-minute interval data from SVCE's 270,000 meters.

**Question #9:** What should be included in the ongoing support costs?

**Answer #9:** Please itemize the different contributions to the cost (e.g. licensing costs, cloud costs, etc.). Please estimate cloud costs based on the proposed services to be provided.

**Question #10:** How big is your user base?

**Answer #10:** Please refer to page 13 of the RFP for more information on SVCE's user tiers. User tier 4 can have up to 5 users, and tiers 1 through 3 can have up to 10 users. These numbers can grow over the years, especially tier 1 as it may include member agencies and SVCE customers in the future.

**Question #11:** Concurrent user requests that would be invoked at each user tier level?

**Answer #11:** In DAISY 2.0, we can expect 3-5 users performing analytical queries on the data warehouse concurrently. In addition, dashboards and integrated applications will pull regularly from the data warehouse.

**Question #12:** Business cases for the data?

**Answer #12:** SVCE has outlined a list of key applications to be integrated into DAISY 2.0 into the future. Example use cases include load forecasting, customer segmentation, virtual power plant program management, and targeted outreach and marketing.

**Question #13:** How many concurrent analytics programs do you expect to run concurrently?

**Answer #13:** SVCE expects to build up to 10 key dashboards that pull data on a regular basis. In addition, SVCE plans to integrate key applications for use cases such as load forecasting, customer segmentation, and virtual power plant program management, which will also pull data on a regular basis.

**Question #14:** Are you expecting to expect to connect DAISY 2.0 to tools like Jupyter Notebooks?

**Answer #14:** Yes. SVCE would like to integrate its Python and R workflows with DAISY 2.0 to the extent possible.

**Question #15:** Are there timeline expectations for implementation?

**Answer #15:** Please refer to Section 17 of the RFP for more information on timelines.

**Question #16:** Is there any US-specific regulations and certifications that will restrict a European company from delivering a new DAISY 2.0 platform?

**Answer #16:** No.

**Question #17:** Will consolidated pre-bid Q&A updates be posted?

**Answer #17:** Yes, this Q&A addendum includes all questions covered in the pre-proposal teleconference and any RFP-related questions sent to [innovation@svcleanenergy.org](mailto:innovation@svcleanenergy.org).

**Question #18:** What are your favorite pieces of DAISY 1.0, and any frustrations?

**Answer #18:** In DAISY 1.0, we appreciate the ability to quickly and efficiently process large amounts of data, being able to export query results directly into a visualization platform, being able to create and share dashboards and reports. SVCE is limited in our use cases because we have yet to integrate the relevant applications and tools. We expect DAISY 2.0 to provide a foundation for integrating these applications and tools.

**Question #19:** Are there budget expectations for DAISY 2.0, both in terms of development cost and ongoing support cost?

**Answer #19:** Given the open-ended nature of this RFP scope, SVCE would like to understand the cost of the bidder's solution and proposal. SVCE received a wide range of costs from the DAISY 2.0 RFI process.

**Question #20:** Are offshore staff resources able to provide services?

**Answer #20:** Yes.

**Question #21:** Do you have best practices around DAISY 1.0 that will be carried through to DAISY 2.0?

**Answer #21:** SVCE has ingested a number of datasets into core tables in DAISY 1.0. SVCE would like to preserve these core tables, data pipelines, and transformation rules in DAISY 2.0.

**Question #22:** When processing raw data, how are exceptions and data quality issues handled?

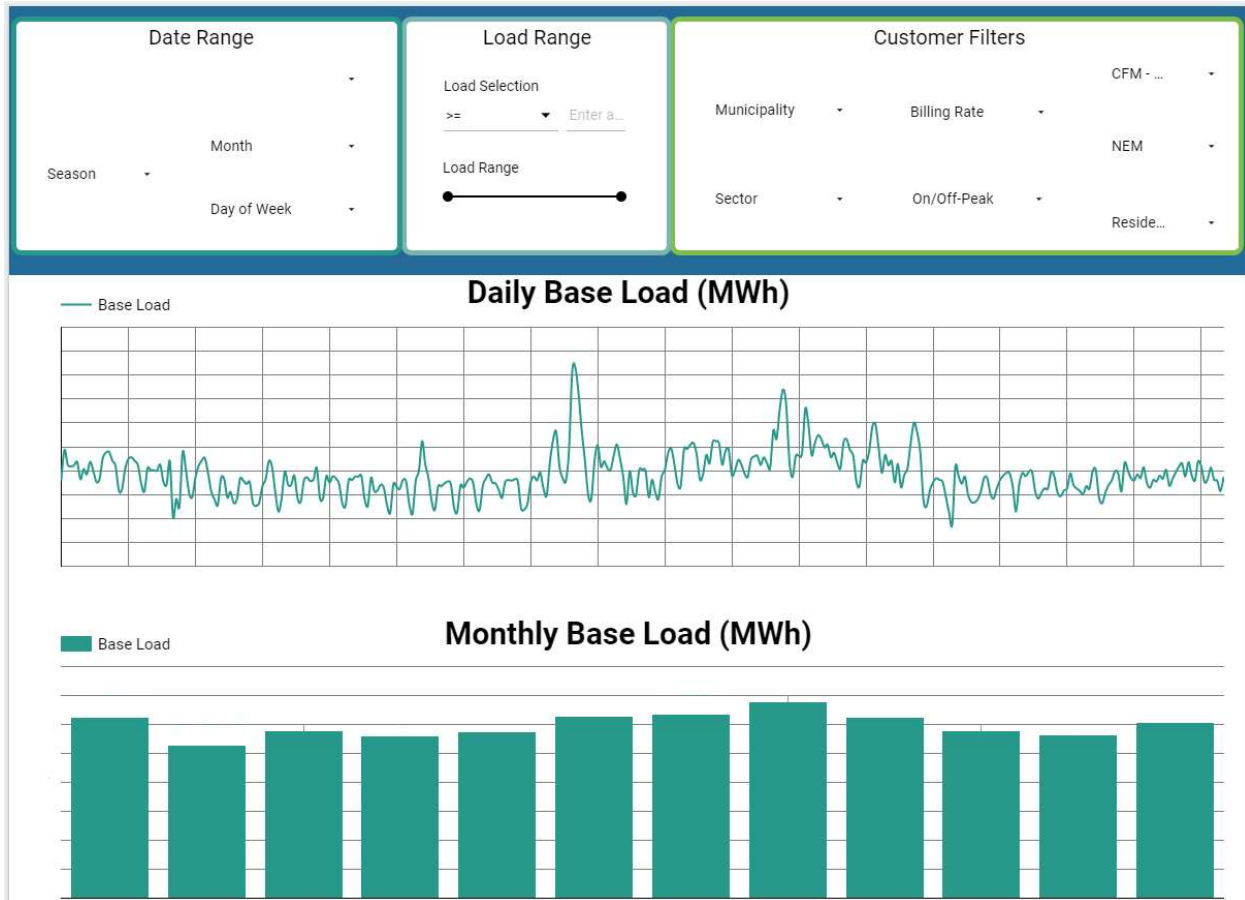
**Answer #22:** When new data quality issues and exceptions appear, SVCE collaborates with the vendor to resolve them.

**Question #23:** Why is there a preference for open-source? What would need to be true to go without an open-source option?

**Answer #23:** Open-source and permission-based data access are strategic priorities for SVCE and relate to SVCE's vision for how the grid and markets are going to evolve in the future. SVCE has a preference for open source in order to have a broader impact by developing solutions that are adopted and propagated in other areas. However, SVCE is open to non-open source solutions for DAISY 2.0 if there are sufficient benefits compared to alternatives.

**Question #24:** If possible, please provide screenshots or examples of DAISY 1.0's dashboard tools and ETL functionality as built.

**Answer #24:** Below is a prototype Data Studio dashboard for illustrative purposes. SVCE welcomes vendor suggestions and enhancements to any of our existing dashboards.



**Question #25:** Please describe in as much detail as possible the current formatting of all datasets currently integrated into DAISY 1.0 as described in Section 16A of the RFP. e.g. - for Customer attribute data, is the data stored in a vendor system, a spreadsheet, or a database, and in what format? Which datasets provide API access?

**Answer #25:** Raw files are in spreadsheet or CSV format and are stored in Google Cloud Storage prior to being ingested into a BigQuery data warehouse. For DAISY 2.0, we expect to leverage APIs for datasets such as CAISO data, weather data, and air quality data.

**Question #26:** What is the current technical infrastructure storing meter data in Calpine? Does an API currently exist to obtain meter data from Calpine?

**Answer #26:** Meter data is transferred to the DAISY 1.0 vendor via secure protocol, then uploaded to Google Cloud Storage, then processed and ingested into a BigQuery data warehouse.

**Question #27:** What is the current technical infrastructure storing any needed CAISO data? Does an API currently exist to obtain needed data from CAISO?

**Answer #27:** DAISY 1.0 currently does not have a connector to pull CAISO data. CAISO APIs do exist but would require further investigation to determine whether they provide access to the desired data.

**Question #28:** What is the current technical infrastructure storing any needed PG&E data? Does an API currently exist to obtain needed data from PG&E?

**Answer #28:** SVCE receives flat files from PG&E via secure protocol and transfers them to the DAISY 1.0 vendor via SFTP. The vendor then processes the PG&E data and ingests into BigQuery.

**Question #29:** Which of the mentioned "Enabling tools" will come under Task 2.a and Task 2.b? If existing tools found to be undeployable, shall the development of the tool commence during the time of Task 2.a?

**Answer #29:** SVCE is open to either integrating existing tools or developing new tools for any of the stated workflows. If no existing tools are planned to be deployed, then the subsequent task (i.e. tool development) can commence during that time.

**Question #30:** What metrics will be used to gauge success of the project, and at what threshold(s)?

**Answer #30:** SVCE seeks to maintain the same level of service as we are receiving with DAISY 1.0, incorporate additional datasets as listed in Section 16 of the RFP, streamline workflows via enabling tools, and provide functionalities for user tiers 1 through 4.

**Question #31:** Is there anything we can do to help speed up the data requirement gathering process?

**Answer #31:** SVCE is open to suggestions for how to streamline data requirement gathering.

**Question #32:** How involved (Knowledge Transfer and other technical questions to be answered) will SVCE be with the project as the project is evolved? Will SVCE be doing all the work or working along with other contractors or the client's staff?

**Answer #32:** SVCE will be involved in all stages of the project as it evolves. The current vendor for DAISY 1.0 will also be involved during the transition process, to the extent required.

**Question #33:** Is there any existing documentation or other introductory information for this project?

**Answer #33:** Please refer to the DAISY 2.0 RFI and this RFP for background information on DAISY 1.0.

**Question #34:** Who are the primary users of the product, and what is their technical level? Are they familiar with this technology already?

**Answer #34:** Please refer to page 13 of the RFP for more information on SVCE's user tiers.

**Question #35:** What's the environment (O/S, Web, cloud) in which this product will be used?

**Answer #35:** DAISY 1.0 is currently hosted in GCP. SVCE is open to staying on GCP or migrating to another cloud data platform if there is sufficient reason to do so.

**Question #36:** How will you monitor progress and performance on the account?

**Answer #36:** Sorry, but we do not understand the question, and therefore are unable to provide a response.

**Question #37:** How frequently can we connect to your subject experts and project resources?

**Answer #37:** SVCE and the vendor will hold regular check-in meetings and set up a channel for asynchronous communication.

**Question #38:** Are there any special circumstances or "hot buttons" of which we should be aware?

**Answer #38:** No.

**Question #39:** Who owns the ideas that we submit or present?



**Answer #39:** Please refer to Section 9 of the RFP.

**Question #40:** Please provide details on the input data information - size, data dictionary, layouts, connection type, where is it hosted, etc.

**Answer #40:** Please see Question #8 for details about SVCE's data warehouse.

**Question #41:** How much data will be anticipated annually in MBs, GBs or TBs?

**Answer #41:** Please see Question #8.

**Question #42:** Where the existing data is stored/managed – on-premise or cloud?

**Answer #42:** Existing data is stored on the cloud, either in GCP or on box.

**Question #43:** Can you please provide the list of technologies used in the current system?

**Answer #43:** DAISY 1.0 is hosted on GCP and uses GCP services such as Google Cloud Storage, BigQuery, Data Studio, Compute Engine, and Cloud SQL. SVCE also uses box for file storage.

**Question #44:** Can you please provide the number of users and their roles that will be interacting with the system?

**Answer #44:** Please see Question #10.

**Question #45:** Please specify an incumbent vendor?

**Answer #45:** Please see Question #3.

**Question #46:** What is the expected project start date?

**Answer #46:** Please see Sections 4 and 17 of the RFP for more information on timeline.

**Question #47:** To accomplish the project do you expect the vendor to deploy on-site resources or can the work be performed remotely?

**Answer #47:** Work can be performed remotely.

**Question #48:** Is the project is funded? Is there any budget cap we should consider?

**Answer #48:** Please see Question #19.

**Question #49:** Does the existing data platform (Daisy 1.0) has any custom workflows, AI/ML statistical / predictive models already built, which needs to be ported into Daisy 2.0 ?

**Answer #49:** DAISY 1.0 does not have additional models or workflows that need to be ported beyond what has been outlined in the RFP.

**Question #50:** Does the Ongoing services include solution design and delivery services (build, test, cutover) of identified use case implementation on Daily 2.0 platform? Are data science services included in scope? Will SVCE program manage the use case implementation project?

**Answer #50:** Bidders are free to propose solution design and delivery services for identified priority applications. However, the focus of this RFP is on the core services identified in the scope.

**Question #51:** Could you please share some of the key issues and challenges with the existing Daisy 1.0 platform?

**Answer #51:** Please see Question #18.

**Question #52:** Can you please provide details of ETL/BI/Data Science tools/software currently being utilized in Daisy 1.0?

**Answer #52:** ETL pipelines are written in Python. SVCE leverages Data Studio for dashboard and report creation. Outside of DAISY 1.0, SVCE leverages Python and R for data science.

**Question #53:** How many different environments does SVCE anticipate deploying Daisy 2.0 ? E.g Dev, Training, Test, Pre-prod, Prod etc.

**Answer #53:** SVCE will rely on the bidder's expertise and recommendation for environment deployment.

**Question #54:** Is the proposed platform expected to have capability to ingest data from devices in the field?

**Answer #54:** This is not a requirement of this RFP.

**Question #55:** Could you please share any data redundancy, disaster recovery and multi-region support requirements?

**Answer #55:** SVCE is not requiring a specific solution, but we are willing to work with the vendor to determine the standards for DAISY 2.0.

**Question #56:** What are your current systems for MDM, ETRM, CRM and GIS?

**Answer #56:** SVCE leverages ArcGIS Pro for ad hoc spatial analyses. MDM, ETRM, and CRM are not within the scope of this RFP.

**Question #57:** Please confirm your estimated Data Growth Rates per year

**Answer #57:** Please see Question #8.

**Question #58:** Please share your GCP installation details and architecture with the interfacing applications.

**Answer #58:** Please refer to the DAISY 2.0 RFI, Figure 1.

**Question #59:** Can you please share current Daisy 1.0 support details including SLA

**Answer #59:** Bidders are requested to propose in their response terms and support details including SLA.

**Question #60:** Do you have or are you targeting to implement a DER Management System? If yes, please indicate vendor.

**Answer #60:** SVCE does not currently have a DER Management System.

**Question #61:** Section 5, Item 4: The RFP sets the page limit of this section as 4 pages, but also indicates that there is a 20-slide maximum for the accompanying

PowerPoint. However, there are no other references to the PowerPoint in the RFP. Can SVCE please provide instructions and details for what should be included in the PowerPoint?

**Answer #61:** The page limit for Section 5, Item 4 is 9 pages. For the slides, please include information that is relevant to the bullets listed under Item 4.

**Question #62:** Other than what was obtained via the RFI, has SVCE met with any vendors to develop the requirements in the RFP? If so, which vendors?

**Answer #62:** The RFP was developed based on responses to the RFI.

**Question #63:** Has SVCE seen their ideal solution implemented at another agency or agencies? If so, which agency or agencies?

**Answer #63:** No.

**Question #64:** Let's assume, under Task 1 page 14 of RFP, that a Vendor is choosing to keep the Daisy 1.0 Infrastructure that sits on GCP. It is not clear what parts of this current infrastructure are inadequate or SVCE is seeking to replace? Can SVCE elaborate on what parts of the Daisy v1.0 infrastructure or processes they would like to keep or have improved?

**Answer #64:** SVCE is seeking to preserve all current DAISY 1.0 functionalities (data storage, ETL, data warehousing, data visualization, dashboard/report creation, user access permissions, data provenance). DAISY 2.0 enhancements include: the integration of enabling tools, improved ETL and data provenance capabilities as new datasets are integrated, and improved query libraries and query version control. SVCE would welcome other enhancements in addition to those listed here.

**Question #65:** Does SVCE prefer that the vendor keep the infrastructure on GCP? What are the advantages to SVCE and Vendor?

**Answer #65:** Please see Question #2.

**Question #66:** Under Current State Daisy v1.0 - SVCE mentions a number of tools and processes that exist today (GCP, BigQuery, ARCGIS, etc.). Can SVCE describe the additional toolchain (Data Analytics Tools? ETL Tools? Dashboard Tools? Other applicable tools or processes) that are in use today? Are there tools and/or known gaps in the toolchain SVCE would like to add?

**Answer #66:** SVCE uses ArcGIS Pro for geocoding and spatial analyses. We use Python, R, and Excel for data analysis, and Data Studio for dashboard/report creation. Currently, our ArcGIS, Python, R, and Excel workflows are not integrated with DAISY 1.0. We would like to integrate these workflows into DAISY 2.0 to the extent possible, and also welcome vendor recommendations for alternative tools.

**Question #67:** SVCE lists a number of data sources and pipelines: Weather Data, CAISO Data, Third Party Data, etc. Does SVCE have this data today and use it in DAISY v1.0? If so, what are the data sources? And please explain how data gets from these sources onto Daisy 1.0.

**Answer #67:** Please refer to Section 16, Item A of the RFP for a list of datasets that have been ingested into DAISY 1.0. The datasets listed in Section 16, Item B have not been ingested into DAISY 1.0. SVCE expects that DAISY 2.0 will have the capability to ingest these datasets, and is open to vendor recommendations for sources of weather data and air quality data.

**Question #68:** On page 13, SVCE describes 4 Tiers of Users. Can SVCE provide a user count for each Tier of User?

**Answer #68:** Please see Question #10.

**Question #69:** On page 14, SVCE describes needs around forecasting and analyses. Each of these seem to be specific use-cases with specific data sources. Can SVCE elaborate on the nature of each of these use-cases? (What are the data sources, who will be consuming this data, how will they be consuming this data?)

**Answer #69:** Use cases that are listed under "Future Applications" are not within the scope of this RFP.

**Question #70:** On Page 14 Task 1, SVCE mentions a 4-6-week timeline to take over the current DAISY 1.0 Infrastructure; Storage, Data Warehouse, and ETL. Can SVCE elaborate on this scope of work - what exactly needs to be done and how do we know it can be done in 4-6 weeks?

**Answer #70:** Detailed scope of work for vendor transition will be developed during contract negotiations.

**Question #71:** For Task 2a: Existing Tools Deployment, without knowing the number of tools to be deployed, what are the assumptions SVCE makes when the RFP calls for a duration of 2 - 4 weeks?

**Answer #71:** The tool deployment timeline is adjustable, and dependent on the proposed and negotiated scope with the vendor.

**Question #72:** Schema Information: Please provide the complete DDL that supports your data model(s) in your existing database.

**Answer #72:** Detailed information will be provided during contract negotiations. For the purposes of submitting a proposal, please indicate how your proposal (e.g. pricing, terms) depends on the identified variables.

**Question #73:** Schema Information: Please provide the number of rows (in Production) for each table referenced in the DDL.

**Answer #73:** Detailed information will be provided during contract negotiations. For the purposes of submitting a proposal, please indicate how your proposal (e.g. pricing, terms) depends on the identified variables.

**Question #74:** Schema Information: Sample data for the tables, roughly 1,000 rows

**Answer #74:** Detailed information will be provided during contract negotiations. For the purposes of submitting a proposal, please indicate how your proposal (e.g. pricing, terms) depends on the identified variables.

**Question #75:** Schema Information: Users which need to be created, their contact information and preferred username

**Answer #75:** Please refer to page 13 of the RFP for more information on SVCE's user tiers.

**Question #76:** How does SVCE use the Geo-functions today?

**Answer #76:** SVCE geocodes addresses to determine their lat/lon coordinates and assessor's parcel number (APN). The lat/lon coordinates allows us to join the data with other geographical information and the APN allows us to merge disparate address-level datasets.

**Question #77:** Can you provide any sample data/queries for us to take a look at?

**Answer #77:** Not at this time. Please indicate how your proposal (e.g. pricing, scope, terms, etc.) may change based on data/queries.

**Question #78:** Is Geospatial toolkit a dealbreaker for this project?

**Answer #78:** No.

**Question #79:** How many database objects involved in the POC?

1. # of Databases?
2. # of Schemas?
3. # of Tables?
4. # of Indexes?
5. # of Stored Procedures?
6. Any customer SQL Functions?

**Answer #79:** Please refer to Question #8 for the current size of the data warehouse and expected growth.

**Question #80:** How much data per table?

1. Data volume?
2. Row counts?
3. # of tables with high column counts (more than 200)?

**Answer #80:** Please refer to Question #8 for the current size of the data warehouse and expected growth.

**Question #81:** What use cases do you intend to test?

1. Clearly define each use case to test.
2. Clearly define why test this use case, what is the challenge/pain.
3. Clearly define what is current runtimes or issues.
4. What is success for the use case? Faster? More Concurrency?
5. When successful, what happens next?

**Answer #81:** Please refer to the RFP for details on the current scope and future applications.

**Question #82:** What is the incumbent configuration and are there any competitors?

1. # of CPU cores?
2. How much memory?
3. Rack space of foot print?

**Answer #82:** We are not able to respond to this question at this time.

**Question #83:** How many queries are run daily in the incumbent environment? Peak hour?

**Answer #83:** DAISY 1.0 querying is not reflective of the expected query load of DAISY 2.0.

**Question #84:** How many average concurrent users are active in the incumbent environment? Peak hour?

**Answer #84:** Please see Question #11.

**Question #85:** How is data loaded into the incumbent environment daily? Peak hour?

**Answer #85:** We currently process around 10 GB of data per day. This volume is expected to grow as new datasets (e.g. weather, CAISO, resource telemetry) are integrated into DAISY 2.0.

**Question #86:** How is data loaded and at what volume? Batch? Trickle? Microbatch?

**Answer #86:** SVCE is seeking batch (daily/weekly/monthly basis) processing and near real-time data processing.

**Question #87:** Will any ETL tools be tested?

**Answer #87:** We do not understand the question and are therefore unable to provide a response.

**Question #88:** Will any additional applications be tested (MSTR, Tableau, R, SAS)?



**Answer #88:** We do not understand the question and are therefore unable to provide a response.

**Question #89:** Query Information: Please provide any sample SQL queries, current run times, and returned row counts for known queries to be benchmarked.

**Answer #89:** Please indicate in your proposal how your response (e.g. pricing, scope, terms) may change based on queries, run time, etc.

**Question #90:** Query Information: Are there any stored procedures to be tested? How many stored procedures are there in the current environment? Sample code?

**Answer #90:** Please indicate in your proposal how your response (e.g. pricing, scope, terms) may change based on the response.

**Question #91:** Performance Testing: How many queries should be in to test set?

**Answer #91:** We do not understand the question and therefore unable to provide a response.

**Question #92:** Performance Testing: How many users should be simulated? (This should align with current volumes).

**Answer #92:** Please see Question #10.

**Question #93:** Performance Testing: How long will the test run?

**Answer #93:** We do not understand the question and are therefore unable to provide a response.

**Question #94:** Performance Testing: What does success look like?

**Answer #94:** We do not understand the question and are therefore unable to provide a response.

**Question #95:** Performance Testing: Will all competitors run the same tests?

**Answer #95:** We do not understand the question and are therefore unable to provide a response.

**Question #96:** Performance Testing: How will you account for cached results in the testing?

**Answer #96:** We do not understand the question and are therefore unable to provide a response.

**Question #97:** Section 17: What level of engagement in the project could we expect from SVCE? How many SVCE business subject-matter experts & technical resources could be allocated (and at approximately what % of FTE) to the project?

**Answer #97:** SVCE will be involved in all stages of the project as it evolves. The current vendor for DAISY 1.0 will also be involved during the transition process, to the extent required.

**Question #98:** Section 17: If any technical resources are to be allocated to the project what would their technical skillset be?

**Answer #98:** The DAISY 2.0 vendor will work closely with SVCE tier 4 users, who are familiar with GCP and its related services and are proficient in SQL, Python, R, and GIS.

**Question #99:** Section 17: What type of SLA is expected on the ongoing services team?

**Answer #99:** Bidders are requested to propose in their response the terms and support details including SLA.

**Question #100:** Section 17: Are there any onsite / onshore / nearshore / offshore requirements for the ongoing services team?

**Answer #100:** No.

**Question #101:** Section 17: For the three-years of ongoing administrative support and user coaching, would standard business hours (9am-5pm M-F excluding SVCE holidays) be acceptable?

**Answer #101:** Yes, standard business hours (9am-5pm Pacific M-F excluding SVCE holidays) are acceptable for ongoing coaching services.

**Question #102:** Section 17: What is the budget for the project?

**Answer #102:** Please see Question #19.

**Question #103:** Section 16A: What level of data cleansing is happening prior to being ingested into DAISY 1.0? Any specific tools used to do this beyond python scripts?

**Answer #103:** Python scripts are the primary tool used for data cleansing. Other pre-processing workflows prior to DAISY ingestion include geocoding addresses and spatial joins with map layers.

**Question #104:** Section 16A: How is data being retrieved from DAISY 1.0 to be used in ArcGIS Pro?

**Answer #104:** Data is exported into CSV format and then uploaded manually to ArcGIS Pro.

**Question #105:** Section 16A: How is DAISY 1.0 ingestion jobs currently being orchestrated and scheduled?

**Answer #105:** Data from SVCE's meter data manager is received and processed on a daily basis. Other datasets are updated on an ad hoc basis (ranging from weekly to every few months), depending on SVCE's needs.

**Question #106:** Section 16A: Are there aspects of the DAISY 1.0 technology that are prompting you to consider a different technology stack for DAISY 2.0?

**Answer #106:** Please see Question #18.

**Question #107:** Section 16B: What technology/language is used for transformations on the existing ETL pipelines? SQL? Python? Other?

**Answer #107:** Transformations are primarily done in Python.

**Question #108:** Section 16B: Can you share current data volumes for DAISY 1.0 existing data sources and Data Warehouse (in GCS and BigQuery)? Do you have estimates on the data size for the new data sources to be added as part of 2.0?

**Answer #108:** Please see Question #8.

**Question #109:** Section 16B: What is your definition of real-time data? (5 mins, 15mins, 1 hour, etc.)

**Answer #109:** Please indicate in your response how your proposal (e.g. pricing, scope, terms) may change based on the response to this question.

**Question #110:** Section 16B: What level of automation are you expecting for the ETL data pipelines?

**Answer #110:** The ETL pipelines should be automated enough to efficiently process the data at the desired rate (near real-time or batch, depending on the dataset).

**Question #111:** Section 17B: What tools (software, packages) are being used outside of DAISY 1.0 to provide the functionalities of the enabling tools? (e.g. Geocoding, Spatial analysis, Address standardization, Weather normalization)

**Answer #111:** SVCE uses ArcGIS Pro for geocoding and spatial analysis. Address standardization and weather normalization are performed using Python scripts or in Excel.

**Question #112:** Section 17 Task 1: What level of support (If any) would we be getting from the current vendor who developed DAISY 1.0?

**Answer #112:** The current vendor for DAISY 1.0 will be involved during the transition process, to the extent required.

**Question #113:** Are you comfortable working with multiple agencies for this work, or are you exclusively looking for one team to fulfill all of the scope?

**Answer #113:** SVCE is open to working with multiple entities for this work.

**Question #114:** For each of the new datasets, do you have a preferred data source identified? If yes, have you tested it for errors or should the scope include (a) identifying and/or building new APIs; and (b) cleaning the results before using within DAISY 2.0?

**Answer #114:** SVCE is open to vendor recommendations for sources of weather data and air quality data. SVCE has identified the data sources for the other new datasets listed in Section 16 Item B of the RFP, but has not yet tested, cleaned, or

ingested the data. Therefore the scope should include (a) identifying and/or building new APIs and (b) cleaning the results before using within DAISY 2.0.

**Question #115:** Could you share more detail on the DAISY 1.0 implementation? Is there a diagram that shows the relationships and data flows of the current datasets as listed in RFP Section 16.A. (PDF page 11)? If possible, please also share the languages and tools being used.

**Answer #115:** Please refer to the DAISY 2.0 RFI, Figure 1. Current state DAISY 1.0 follows a similar structure to the DAISY 2.0 schematic. The main difference is that DAISY 1.0 currently has no integrated applications and no Tier 2 users.

**Question #116:** Could you confirm that DAISY (both 1.0 and 2.0) is currently on the cloud and will remain there (as opposed to being hosted on-premise)?

**Answer #116:** That is correct.

**Question #117:** Regarding security, it is our understanding that the most sensitive data is customer attribute information. This is part of DAISY 1.0. Could you please share how that data is currently stored, processed, and secured? And does SVCE expect the winning bidder to mostly use current practices, or develop their own?

**Answer #117:** Data is transferred from SVCE to the DAISY 1.0 vendor via secure protocol. The data is then stored, processed, and secured in GCP. SVCE expects the winning bidder to either use current practices or develop their own, depending on the proposed solution.

**Question #118:** Does SVCE plan to use a third-party auditor to ensure cybersecurity policies such as SOC 2, or should the winning bidder plan to provide their own as part of the RFP response?

**Answer #118:** SVCE will use a third-party auditor that will work with the vendor to ensure that proper security measures are in place.

**Question #119:** During the pre-proposal teleconference, the size of the database was shared as 5 TB, with "hundreds of GB added per year." Understanding these are estimates, could you confirm this is the best guess as of today?

**Answer #119:** Please see Question #8.

**Question #120:** During the pre-proposal teleconference, it was shared that SVCE is happy with the cleaning rules for the data. Is this ruleset firm and final, or is it expected to change over the course of the project as more APIs are added? If the latter, could you share where the ruleset currently exists, and is there a process for the winning bidder to contribute to the code of the ruleset (ie, a pull request using git)?

**Answer #120:** Cleaning rules will need to be established for new datasets. DAISY 1.0 cleaning rules on existing datasets are expected to be carried over to DAISY 2.0, but these cleaning rules are subject to change upon request by SVCE.

**Question #121:** Is there anything about the current big data hosting solution (BigQuery + Google Cloud storage buckets) that SVCE isn't happy with?

**Answer #121:** Please see Question #18.

**Question #122:** Requirement: Provide overall organizational structure

Question: Could you please provide more context on what you are evaluating? We want to make sure we provide the right level of detail.

**Answer #122:** Please describe the organizational structure of your business, and how it relates to the services you are proposing to provide.

**Question #123:** Requirement: Current audited Financial Statements if possible credit rating reports from S&P Global Ratings and/or Fitch and/or Moody's.

Question: Is providing audited financials and credit ratings a hard requirement for selection? Can a letter from our board of directors and Venture Capital investors be used as evidence of financial strength?

**Answer #123:** If the bidder is not able to provide audited Financial Statements and credit rating reports, please provide other information that can serve as evidence of financial strength.

**Question #124:** Requirement: Provide an overview of your qualifications and previous experience on at least three similar or related projects. Include descriptions, costs, timeline and reference contact information.

Question: If we are unable to provide specific details about costs or related projects due to customer confidentiality, are more generalized examples still acceptable along with a specific customer contacts?

**Answer #124:** Yes.

**Question #125:** Does SVCE expects to build any Data Science Model as part of delivery in DAISY 2.0? How does SVCE expect to meet its Predictive analytics?

**Answer #125:** A data science model is not within the scope of this RFP.

**Question #126:** How many tables are required to be migrated in DAISY 2.0? What is the volume of data that needs to be migrated?

**Answer #126:** Please see Question #8.

**Question #127:** What kind of tools are currently used for Geocoding, Spatial Analytics, Address Standardization and Weather normalization?

**Answer #127:** Please see Question #111.

**Question #128:** Does DAISY2.0 would involve any data to be ingested from API or any Website for the above data sources?

**Answer #128:** Yes, DAISY 2.0 will involve data to be ingested from API or website.

**Question #129:** Does DAISY 2.0 need to ingest realtime time and streaming data form any data sources?

**Answer #129:** SVCE's needs will inevitably evolve over the course of the coming months and years. Please indicate in your response how your proposal (e.g. pricing, scope, terms) may change based on the response to this question.

**Question #130:** What are the third party services referred to here in RFP for settlement? What is the frequency of data, volume and format of data?

**Answer #130:** The details of the settlement data frequency, volume, and format have yet to be determined.

**Question #131:** What business values will be derived from the data ingested from Geocoding, Spatial data, Address standardization, weather normalization?

**Answer #131:** These workflows are important for strategic analyses such as load forecasting and customer segmentation. Please refer to the DAISY 2.0 RFI for additional detail on priority use cases to SVCE.

**Question #132:** What functionalities will be derived from query library and data lineage and processing?

**Answer #132:** A query library and data lineage are important for tracking data flows, data governance and risk management. A query library is also intended to improve collaboration between DAISY users.

**Question #133:** What kind of technology stack will be required, to build and customize the new tool?

**Answer #133:** The technology stack requirements are to be determined by the bidder.

**Question #134:** As per Page No 12 in RFP under section Current State Daisy 1.0. Analysis are done on data using Python,R and Excel. What exactly are this analysis and do we need to create this analysis in Daisy 2.0 as well?

**Answer #134:** The specific analyses that SVCE is seeking to integrate into DAISY 2.0 are covered in Section 16, Item B under "Enabling Tools." Apart from these analyses, SVCE is generally seeking to streamline the integration between DAISY 2.0 and Python/R/Excel.

**Question #135:** For the Below Mentioned How many Reports, Dashboards and Visualization/KPI's are expected:

- Data visualization, including mapping capability
- Dashboards: create, save, share, update, and versioning
- Reports: create, save, share, update, and versioning

**Answer #135:** Please indicate in your response how your proposal (e.g. pricing, scope, terms) may change based on the response to this question.

**Question #136:** As per our assumption reading Task 2a Page No 14. All the enabling tools functionality which were earlier executed outside DAISY 1.0 need to be a part DAISY 2.0 Cloud Deployment assuming functionality is workflow automation.

Are these the Tools that needed to be considered(Gecoding,Spatial Analysis,Address Standardization,Weather Normalization)?



**Answer #136:** Please refer to Section 16, Item B of the RFP for information on the types of enabling tools that SVCE is seeking to integrate into DAISY 2.0.

**Question #137:** Kindly List down the tools that need to be defined and executed as a part of Task 2b Page No 15 ? What is the tech stack of the tools? How many tools need to be customized?

**Answer #137:** The tools that SVCE seeks to integrate into DAISY 2.0 are covered in Section 16, Item B under "Enabling Tools." The technology stack is to be determined by the bidder.

**Question #138:** List down the Innovative component that need to be considered as a part of DAISY 2.0 ?

Also list down the data process which are executed outside Daisy 1.0?

**Answer #138:** We do not understand the question and are therefore unable to provide a response.

**Question #139:** Kindly list down the no. of ETL pipeline that need to be developed as a part of DAISY 2.0 along with its frequency of Execution?

(Daily,Hourly,Weekly,Monthly)

**Answer #139:** SVCE intends to carry over existing ETL pipelines from DAISY 1.0 to DAISY 2.0. In addition to maintaining these existing pipelines, the vendor is expected to develop new ETL pipelines for the datasets listed in Section 16, Item B of the RFP. Please see Question #86 for information on frequency of execution.

**Question #140:** There is a fairly strict set of page guidelines. Attaching full resumes of key staff will be challenging to fit in 4 pages when in combination with the other information requested in Section 3 "Organization Description and Qualifications". May we add full resumes in an Appendix not counting against the page count guidelines?

**Answer #140:** Yes.