Pacific Gas and Electric Company

Ali Moazed
Director, Electric Operations
Data Management & Analytics

77 Beale Street
San Francisco, CA 94105
Ali.Moazed@pge.com

June 8, 2022

VIA OFFICE OF ENERGY INFRASTRUCTURE SAFETY E-FILING SYSTEM

Stephen P. Lai
Data Manager, Data Analytics Division
Office of Energy Infrastructure Safety
California Natural Resources Agency
715 P Street, 20th Floor
Sacramento, CA 95814

Re: **PG&E Comments on Geographic Information Systems Data Reporting Standard version 2.2 Draft Guidelines**
2022 GIS DRs Docket (2022-GIS-DRs)

Dear Mr. Lai:

Pacific Gas and Electric Company (PG&E) appreciates the opportunity to again provide comments for the Office of Energy Infrastructure Safety's (OEIS or Energy Safety) Geographical Information System (GIS) Data Standard Version 2.2 as it is being presented as Guidelines for adoption. PG&E has been using Version 2.2 of the GIS Data Standard since the Q4 2021 submission on February 1, 2022, and, as such, offers the following feedback regarding the Version 2.2 release of the GIS Data Standard.

1. Data Completeness and Maintaining a Phased Approach with Clear Prioritization

PG&E acknowledges that there are remaining data gaps needing to be closed to meet compliance requirements of the GIS Data Standard. PG&E recommends that efforts to close the outstanding gaps be approached in a phased manner based on value of the data to Energy Safety's objectives and utility business operations. PG&E approaches its GIS Data Standard submissions with an understanding that providing best available data is preferable to withholding data that may not be complete when compared against the GIS Data Standard schema requirements. Each quarter, PG&E diligently pursues methods to incorporate net new data into our GIS Data Standard submissions, which often involves generating reasonable assumptions or applying approximations needed to transform the data available into Energy Safety's required reporting schemas. The majority of data not reported on in our Spatial QDR are due to data not being collected or unknown, not stored in a source system, requires clarification from Energy Safety, or are not stored in a manner that is conducive to the GIS Data Standard's schema. Closing reporting gaps would require several large-scale, multi-year projects[1] with significant

1 For example, PG&E does not collect much of the information being requested in the 3.6.1 Other Power Line Connection Location feature class regarding the other line information (e.g., OtherConductorMaterial) for private line owners. PG&E does not keep record of customer owned facilities and views private or customer line owners as separately accountable to compliance with electric line regulations. Collecting this information would require considerable support and coordination with private owners.

resourcing and may come at increased cost to customers. This would also require reprioritizing resources away from other wildfire mitigation related work. In addition, Version 2.2 represents the fourth version of the GIS Data Standard, each of which is subject to change in requirements. [2] The evolving nature of the GIS Data Standard creates complexities around prioritization of efforts to address reporting gaps, especially given that a future version change may modify or remove requirements.

PG&E urges a phased approach be applied to electrical corporation adoption of Version 2.2 requirements in recognition that certain data are subject to technical limitations and prioritization will be needed to ensure effective use of resources and funds. Energy Safety referenced this phased approach in Version 2.1 of the GIS Data Standard, which was subsequently removed in Version 2.2:

> Considering existing limitations with electrical corporation data capabilities and differing business processes that support the collection, treatment, and storage of GIS data, Energy Safety is employing a phased approach for full implementation of this standard.[3]

2. Compliance Use of Submission Data

Energy Safety states in Version 2.2 of the GIS Data Standard that submission data is used to monitor and evaluate utility safety, including wildfire risk reduction and to support compliance activities.[4] PG&E wants to assist Energy Safety in optimizing their compliance objectives and believes continued collaboration is key for success. As stated in the above section, data in the quarterly submissions has been transformed and packaged with underlying assumptions and approximations that work best for an individual field in a feature class. Although these approximations and assumptions are explained in the metadata and status reports, it is not possible to capture all limitations surrounding the use of this data (e.g., combining for analysis). Due to these complexities, PG&E urges Energy Safety to continue collaborating with the electrical corporations to align on understandings of data submitted, use limitations, and methods in which data is derived. This collaborative approach can contribute towards Energy Safety's pursuit of using data to assist with monitoring and evaluating utility safety, wildfire risk reduction, and compliance activities.

It takes time for data to undergo processes required for integration into a primary system of record. PG&E's GIS source system represents assets that have been received and mapped. Newly installed assets may not yet be mapped in GIS as completed jobs must undergo several processing steps and additional field verification before being mapped. Therefore, until a project is completed and mapped, detailed information remains in the design systems and paper job packages. For the GIS Data Standard submission, this results in differences between the 3.5

---

[2] Draft Version 1 of the GIS Data Standard (issued 8/5/2020) was in place in 2021 from January 1 to February 3. Version 2 (issued 2/4/2021) was effective from February 4 to September 6, 2021, Version 2.1 (issued 9/7/2021) was in place from September 7 to December 17, 2021. Version 2.2 became the latest standard when issued on December 17, 2021 and finalized January 2022.

[3] *See* GIS Data Standard V2.1, p. 1 (Sep. 7, 2021) (https://efiling.energysafety.ca.gov/Lists/DocketLog.aspx?docketnumber=2021-GIS-DRS).

[4] "Electrical corporations' standardized GIS data submissions will provide Energy Safety with important asset and risk data that will be used to monitor and evaluate utility safety, wildfire risk reduction, and compliance activities." *See* GIS Data Standard V2.2, p. 1 (Jan. 14, 2021) (https://energysafety.ca.gov/who-we-are/department-organization/electrical-infrastructure-directorate/data-analytics-division/).

Initiative feature datasets, which primarily leverage project managers' trackers for real-time job information, and the 3.1 Asset Point and 3.2 Asset Line feature datasets, which leverages data from PG&E's GIS source system. Due to the timing structure of the Data Standard report and the need to process completed work, conducting compliance activities such as field verifications will be challenged if relying on combining different data points across different feature classes.

PG&E has several initiatives underway that will contribute to more complete data and enhanced data quality seen in the GIS Data Standard submission data.[5] However, since each feature class may experience processing steps before data is available in primary source systems and have underlying use limitations that were created for the individual feature class and not intended for combined overlays of multiple feature classes, data users should be cautious in generating inferences from the data.

3. Challenges of Aligning GIS Data Standard Submission with the Quarterly Initiative Update (QIU) Tabular Reporting

PG&E is challenged in aligning GIS Data Standard with tabular reporting such as the Quarterly Initiative Update (QIU) and restates the need to implement a phased approach (as outlined above in section 1) to partner and prioritize alignment efforts with Energy Safety. Alignment challenges are primarily due to the following: (i) differentials in technical and schematic requirements, and (ii) differentials in timing of data readiness. PG&E reiterates its statement on the technical challenges involved with fully aligning the spatial Quarterly Data Report (QDR) with the QIU report, as outlined in PG&E's response to data request *OEIS to PGE- Quarterly Data Submission-20211026*:

> [D]ata included in the GIS Data Standard submission must meet specific technical criteria for inclusion — including, but not limited to, the ability to transform data from PG&E's internal data architecture into the FGDB required data architecture and display these data in a spatial format…. Tabular reports such as the QIU are not subject to such data schema/architectural technical requirements and thus allow more initiative types to be reported.[6]

Since the QIU updates provide written progress on how the WMP initiatives are progressing, the supporting evidence relied on does not require geometry. For example, data supporting QIU reporting can be in the form of vendor invoices, contractor progress reports, and field crew trackers. Though these data types can be integrated into tabular reporting, they do not meet requirements for transformation into the GIS Data Standard schema.

Additionally, timing differences exist between collection of initiative data and the population of said data into a geospatial format/database (GIS) due to the processes needed to document data, verify work performance, and update geospatial records. Tabular reporting on miles completed or otherwise can be readily collected through field updates and/or work tracking tools, leading to differentials in timing for which data can be used for the QIU versus GIS Data

[5] For example, PG&E is collecting transmission critical component data including age and installation date though our Asset Information Collection (AIC) effort. As this information is collected and stored in system of records, PG&E's installation date fields will gradually be more complete.

[6] *See* Energy Safety data request entitled "OEIS to PGE- Quarterly Data Submission-20211026" (Oct. 26, 2021).

Standard submissions. PG&E needs final datasets to be available two to three weeks ahead of the actual GIS Data Standard submission deadlines to consolidate and transform data into the prescribed File geodatabase (FGDB) format, further contributing to differentials in timelines when compared to tabular reporting.

4. Confidentiality Requirements in Section 2.3.9

PG&E continues to have reservations around the feasibility and practicability of providing confidential information for each record submitted. In the latest Q1 2022 submission, approximately 14.7 million records were submitted. Reviewing each of these records individually would not be feasible or practicable. Furthermore, the interconnected aspect of feature class data and geospatial representation of the data create complexities in identifying the confidentiality of individual records and introduces additional risk for error.

Therefore, PG&E will continue to take the approach by applying an attribute level confidentiality designation for each feature class in the FGDB. If any attribute in each requested feature class is deemed to contain what is or could be "partially"[7] confidential information, PG&E will continue to apply the "yes" confidentiality designation to each record to mitigate against the risk of mislabeling individual records.

5. Technical Workgroups for Collaboration and Consistent Implementation

PG&E appreciates that Energy Safety held a GIS data discussion with the utilities on May 17, 2022 and is committed to continuing and facilitating these quarterly working sessions. These working sessions provide an opportunity to shape modifications and drive consistent implementation of the GIS Data Standard.

PG&E continues to achieve improvements in data quantity and/or quality on a quarterly basis since the implementation of the GIS Data Standard in Q3 2020. Additional enhancement opportunities will largely require more involved operational and technological changes, and a significant investment of resources and time to collect, curate, and organize the Data Standard submissions on a recurring basis. Given the estimated level of effort required to meet the standard, regular collaboration with Energy Safety is needed to align on expectations, prioritization of data and information, technical feasibility issues, and shape modifications to the schema.

PG&E continues to suggest that future investor-owned utility and Energy Safety technical workshops be focused on each of the six feature datasets, with the exception of Feature Dataset 3.5 (Initiatives) for which workshops should be segmented according to work types performed including: Asset Inspections, Grid Hardening, and Vegetation Management. In addition, a workshop to review confidentiality designations would allow more consistent application across utility submissions while considering the complexities introduced through the interconnected aspect of feature class data and geospatial representation as described in section 4 above. PG&E requests these sessions continue to be held two to three weeks after the submission deadlines to allow time for preparation and help ensure productive discussions. Leading up to the

[7] PG&E considers transmission line data partially confidential. This confidentiality determination is dependent on whether the line is at or above 115kv in which case information is identified as confidential and protected. At that point it is considered physical facility, cyber-security sensitive, or critical energy infrastructure data and is protected from disclosure. *See* 18 C.F.R. § 388.113, see also Govt. Code § 6254(k), (ab); 6 U.S.C. § 131; 6 CFR § 29.2.

submission deadline, PG&E teams are in process of collecting, transforming, performing quality assurance, and preparing submission information including our status report, metadata, and cover letter.

6. Simultaneous Release of the PDF, GDB File, and Excel Status Report Template

For future iterations of the GIS Data Standard, PG&E reiterates our request that the PDF, Excel, and FGDB files be released simultaneously to assist with the analysis of proposed changes and the impacts to existing processes or data collection/curation techniques.[8] The impacts of Energy Safety's schema changes cannot be determined by viewing the PDF alone; full analysis requires release of supporting files. Simultaneously releasing the documents allows review of impacts to domain values and fields. Each version release has been subject to discrepancies between documents, challenging data development and automation processes for reporting.[9] Historically, PG&E's resources have spent hundreds of hours socializing, assessing, and implementing version changes across various teams and technology platforms to update the transformation logic needed only to discover that the PDF does not always follow the same schema as the GDB which results in rework.

7. Typographical Errors

In addition to the suggested revisions PG&E shared in the first version 2.2 reply comments,[10] PG&E has discovered additional typographical errors that would benefit from corrections including: (1) adding "Vegetation management and inspections" to the domain choices of the "WMPInitiativeCategory" field; and (2) that "ChangeOrder", "ChangeOrderDate", "ChangeOrderType", and "ChangeOrderTypeComment" are missing from the 3.5.5.2 Other Initiative Log table in the PDF, but present in the file geodatabase. The PDF instead places all the change order fields for the "Other Initiative" reporting in all the associated geometry feature classes which is also inconsistent with 3.5 Initiative reporting structure.

***

PG&E appreciates the opportunity to provide feedback on the GIS Data Standard and to continue to work with Energy Safety to promote wildfire safety.

Sincerely,

/s/ Ali Moazed

Ali Moazed
Director, Electric Operations Data Management & Analytics
77 Beale Street
San Francisco, CA 94105

[8] *See* PG&E Comments on GIS Data Standard V2.1 at p. 3 (Aug. 31, 2021) (https://efiling.energysafety.ca.gov/Lists/DocketLog.aspx?docketnumber=2021-GIS-DRS).

[9] An example of misalignment across guidance documents can be found in: PDF document, Feature Class 3.3.6.2 (PSPS Event Damage Point) contains fields for FuelBed and FuelBedDescription – which are not included in the FGDB. A second example of misalignment across guidance documents can be found in: Feature Class 3.3.6.2 (PSPS Event Damage Point): FGDB contains a field for "AssetID" (a net new field in V2.1) which is not present in the PDF document.

[10] *See* PG&E Comment on Draft GIS Data Reporting Standard Version 2.2 at p. 4 (Dec. 27, 2021). (https://efiling.energysafety.ca.gov/Lists/DocketLog.aspx?docketnumber=2021-GIS-DRS).