

August 27, 2021

Via E-Mail

Stephen.Lai@energysafety.ca.gov

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 Response to OEIS Geographic Information Systems Data Standard, Version 2.1

Dear Mr. Lai:

Pacific Gas and Electric Company (PG&E) appreciates the revisions made by the Office of Energy Infrastructure Safety (OEIS) to its Geographic Information Systems Data Standard (GIS Data Standard) and the opportunity to provide comments. PG&E offers the following feedback regarding the Version 2.1 release of the GIS Data Standard and also proposes a working group be established between OEIS and the electrical corporations to allow for collaboration and consistent implementation on the GIS Data Standard.

1. Working Group for Collaboration and Consistent Implementation

PG&E requests a working group be established to strengthen collaboration and feedback mechanisms between OEIS and the electrical corporations. In particular, the facilitation of a working group prior to the release of future GIS Data Standard revisions would help to shape modifications and drive consistent implementation.

2. Redline Version with Updates Showing Changes to Data Requirements

PG&E requests that an updated PDF version of the GIS Data Standard be produced with redlines applied to feature classes and related tables to highlight changes to data requirements. PG&E's preliminary analysis indicates there were over 100 new or removed fields in the feature classes and related tables, which are not captured in the redline. Rather, the redline draft captures the changes in the fields for agency name from "WSD" to "Energy Safety" and the new "UtilityInitiativeTrackingID" fields. If an updated PDF with redlines is released, PG&E requests that additional time be provided to allow for comments from the other electrical corporations. Even minor changes to the data schema have potentially large implications, as existing code or scripts used to transform PG&E source system data into the required FGDB schema format require revisions. Highlighting these changes through redline would be appreciated for future GIS Data Standard revisions. Of the field changes, 55 are related to confidentiality, which was redlined in the overall guidance section but not via the data schema, and 12 of the field changes are related to Utility ID, which can be easily incorporated.

3. 3.5 Initiatives: Targets, Quarterly Progress, Cumulative Progress, Target Units

PG&E requests that consideration be given to removal of the following fields in Feature Dataset 3.5 Initiatives: (1) InitiativeTarget; (2) QuarterlyProgress; (3) CumulativeProgress; and (4) InitiativeTargetUnits. Though majority of these fields were included in the former V2 of the GIS Data Standards, this comment period presents the first formal forum in which to provide feedback on revisions to the GIS Data Standard requirements. Quarterly progress for specific initiatives is reported through PG&E's Quarterly Initiative Update (QIU). The data included in the FGDB of the GIS Data Standard submission must meet specific 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.

Initiative work is often stored across various internal source systems, each of which have unique data architectures that were established for operational purposes, which result in significant reporting complexities. In addition, errors can occur when incorporating latitudinal and longitudinal (lat/long) coordinates due to human error in data capture or technical challenges in capturing this data in remote conditions that lack network connectivity. Transforming the Initiative data or displaying it spatially through this submission can result in data loss when compared to tabular (e.g. excel) based reporting that are not subject to the same data schema/requirements. Some of these challenges were described earlier in PG&E's Q2 2021 Guidance-10 narrative¹: "PG&E's existing data and system architecture were developed over decades to address specific operational uses and lack integration capability and a cohesive data schema. This presents significant challenges to accessing and aligning data to meet OEIS' GIS Data Standard.... PG&E's submissions of the requested Status Report and Data Submission (collectively referred to as "GIS Data Standard Submission") are not fully complete as we do not have all the requested data or have all the data in the format requested." That this process would take time to accomplish and that all data would not be immediately available was anticipated by OEIS, who noted on page 5, Section 2.8 of V2 of its GIS Data Standard² that:

"Realistically, the WSD understands that electrical corporations are at different stages of their data journeys and employ differing business practices, which may impact certain electrical corporations' abilities to fully comply with the requirements in this document. The WSD looks forward to working collaboratively with electrical corporations and other stakeholders to determine appropriate and feasible submission schedules for regular reporting of GIS data."

4. Typographical Errors

PG&E identified several inadvertent typographical errors that should be corrected. Specifically: (1) data schema "3.1.1 Camera (Feature Class)" is missing from table of contents; (2) data schema "3.1.3 Customer Meter (Feature Class)" is listed as "2.1.3 Customer Meter (Feature Class)" in the table of contents and body of the document; and (3) data schema "3.3 PSPS Event (Feature Dataset)" is missing the connection points included with other entity relationship diagrams (ERD).

¹ Quarterly Report on 2020 Wildfire Mitigation Plan for Second Quarter 2021 – August 2, 2021. WMP Condition Guidance 10: Data Issues. Available at <https://www.pge.com/pge_global/common/pdfs/safety/emergency-preparedness/natural-disaster/wildfires/wildfire-mitigation-plan/PGE-WildfireMitigationPlans-QuarterlyReport-Q2-2021.pdf>

² Wildfire Safety Division. Geographic Information System (GIS) Data Reporting Standard for California Electric Corporations – V2. Released February 4, 2021.

5. Release of the PDF, GDB File, Excel Status Report Template, HTML File

PG&E requests that the PDF, Excel, FGDB, and HTML files be released simultaneously to assist with the analysis of the proposed changes to the GIS Data Standard and the impacts to existing processes or data collection/curation techniques. The full domain value rules are not visible through the PDF alone and could be analyzed if the supporting files were released.

Releasing the supporting files would also be helpful because, in the former GIS Data Standard update (V2), there were inconsistencies across the domain values of these documents. These inconsistencies led to difficulties in the development of automation processes and script developments.³

6. Additional Time to Assess the Changes

PG&E requests 10 business days be provided to the electrical corporations for review of future GIS Data Standard version updates. Revisions to the Data Standard require significant assessment from the companies' teams supporting this compliance requirement in order to identify impacts to existing data collection methods. This assessment requires input from data owners, GIS, information technology (IT) subject matter experts, as well as the regulatory and legal teams. The current 5 day review period only allows for a high-level analysis of the changes and associated impacts and is not sufficient for detailed assessment and follow up comments.

7. Revisions to the Confidentiality Requirements in Section 2.3.8

Section 2.3.8 of the GIS Data Standard states: "Each feature class and table in the template GDB, with the exception of the 'Red Flag Warning Day' feature, has a field named 'Confidential' which accepts values of 'Yes' or 'No.' Utilities should note all records they consider confidential in the attributes using this field." PG&E appreciates the ability to incorporate confidentiality designations directly into its FGDB data submission. However, PG&E has reservations around the feasibility and practicability of providing confidential information for each record submitted. Based on a preliminary assessment, PG&E submitted a total of approximately 13 million records in its Q1 2021 report and approximately 15 million records in its Q2 2021 FGDB submission.⁴ These datasets are not static and change each quarter. Reviewing each record individually is not feasible nor practicable given the scale of this quarterly submission.

In addition, 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. For example, the Transmission Outages ID (TOutageID) field is directly applied to five feature classes and can be indirectly linked back to an additional feature class – Risk Event Photo Log (Wires Down includes a TOutageID – that can be traced to the Risk Event Photo Log). TOutageID is confidential when linked to Transmission Lines over 115Kv.⁵ While Transmission Lines under 115Kv may not be considered confidential at the

³ Example: In the former GIS Data Standards (V2), the Ignition Size field domains were not aligned across the PDF and HTML files when compared to the FGDB template.

⁴ Preliminary Assessment: PG&E derived record counts directly from its FGDB submissions. A quality check of these counts has not been performed due to the condensed timelines allotted for the comment period.

⁵ Physical facility, cyber-security sensitive, or critical energy infrastructure data 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.)

individual asset level, there is additional information in all six referenced datasets that would need to be carefully assessed at the record level for other confidential data, such as locational data on substation or other equipment connectivity, customer related data, critical facility, SCADA connectivity, etc. that could impact the entire records' confidentiality designation.^{6 7}

PG&E has provided confidentiality designations in its Status Report at the field level for each feature class and plans to update its FGDB file with these designations across each record within a feature class. For data formerly labeled "partially confidential" in the Status Report, PG&E plans to apply the "yes" confidentiality designation to each record it is associated with, to mitigate against the risk of mislabeling individual records. In addition, PG&E endorses the comments contained in Southern California Edison's Q2 2021 QDR cover letter, which describe the importance of data confidentiality in relation to the submission of its GIS Data Standard:

"As SCE has discussed with Energy Safety, we continue to have reservations regarding the provision of confidential data. Release of the precise location, age, and other attributes of SCE's assets alongside the precise location of critical facilities may significantly increase safety risk to the public. For example, knowledge of underground line routes and electrical equipment serving a critical facility could facilitate an attack on that critical facility's power supply. Also, knowledge of the location of specific SCE assets in areas with historical high-fire weather could make them vulnerable to attack during the worst possible time. Further, the precise locations of SCE's high voltage transmission lines and substations alongside the above-mentioned confidential information, as well as the non-confidential information requested increases risk to the bulk power transmission system. The Commission has recognized the importance of safeguarding critical energy infrastructure information and although maps of varying detail of SCE's transmission system may be publicly available from other sources, SCE does not believe it is prudent to further propagate that information, in this level of detail, accompanying other information that, taken together, could prove to be useful to a bad actor. Notwithstanding these reasons, SCE has preliminarily designated confidentiality at the data field level even though it believes confidentiality should be applied at the feature class level for each provided dataset."⁸

8. Integration of a Field for Tree Species in Specific Feature Classes

PG&E requests clarification on whether OEIS will require scientific names (species/genus) for select Vegetation Management fields, and if so, a timeline for when this change will be applied. Item PG&E 21-22 from the 2021 WMP Action Statement, "[i]ncomplete

⁶ Physical facility, cyber-security sensitive, or critical energy infrastructure data 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.)

⁷ Customer-specific data, which may include demand, loads, names, addresses, and billing data (Protected under PUC § 8380; Civ. Code §§ 1798 et seq.; Govt. Code § 6254; Public Util. Code § 8380; Decisions (D.) 14-05-016, 04-08-055, 06-12-029)

⁸ Southern California Edison Q2 2021 Quarterly Data Report at p. 5..

<<https://www.sce.com/sites/default/files/AEM/Wildfire%20Mitigation%20Plan/2021/SCE%20Q2%202021%20QDR.pdf>>

identification of vegetation species and record keeping,” includes the following requirements for GIS Data Standard Reporting⁹:

PG&E must:

- 1) Use scientific names in its reporting (as opposed to common names). This change will be reflected in the upcoming updates to Energy Safety GIS Reporting Standard.
- ...
- 3) Identify the genus and species of a tree that has caused an outage or ignition in the Quarterly Data Reports (QDRs) (in these cases, an unknown “sp.” designation is not acceptable).¹⁰

The updated GIS Data Standard (V2.1) feature classes for outages include species identification fields (“3.4.5: Transmission Vegetation Caused Unplanned Outage;” “3.4.7: Distribution Vegetation Caused Unplanned Outage”). However, feature class “3.4.3: Ignition” lacks a field to identify tree species. Other vegetation related feature classes request the data in common name format (“3.5.1.3: Vegetation Inspection Point,” “3.6.5: Major Woody Stem Exempt Tree Point”). PG&E requests clarification on when a field to identify tree species will be integrated into the “3.4.3: Ignition” feature class, and whether this field will be required for other vegetation related feature classes (“3.5.1.3: Vegetation Inspection Point,” “3.6.5: Major Woody Stem Exempt Tree Point”).

PG&E appreciates the opportunity to provide feedback on the GIS standard and to continue to work with OEIS to promote wildfire safety.

Sincerely,

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

⁹ Office of Energy Infrastructure Safety’s Draft Evaluation of 2021 Wildfire Mitigation Plan Update: Draft Action Statement on 2021 Wildfire Mitigation Plan Update – Pacific Gas and Electric. Utility-# PG&E 21-22.

¹⁰ OEIS Draft Evaluation of 2021 Wildfire Mitigation Plan Update at pp. 80-81.