



STATE OF MINNESOTA

OFFICE OF THE ATTORNEY GENERAL

LORI SWANSON
ATTORNEY GENERAL

June 6, 2014

SUITE 1400
445 MINNESOTA STREET
ST. PAUL, MN 55101-2131
TELEPHONE: (651) 296-7575

CEUD Workgroup

Re: CEUD Comments on May 16, 2014 Meeting

Dear CEUD Members:

As requested by Chief Judge Pust, this letter comments on issues raised during the workgroup's May 16, 2014 meeting and follow-up telephone conference held on May 21, 2014, and on the recent release of the *Decision Adopting Rules to Provide Access to Energy Usage and Usage-Related Data While Protecting Privacy of Personal Data* by the Public Utilities Commission of the State of California ("California Decision").¹

In its May 5, 2014 letter to the workgroup, the OAG recommended that (1) individual customer data not be disclosed absent customer consent, and (2) the Commission obtain technical analysis to ensure that any release of aggregated or anonymized data is done in a way that appropriately protects individual customers. Commission staff has indicated, however, that obtaining a technical analysis on aggregation/anonymization techniques is unlikely at this time. Recognizing that it is unlikely to obtain technical analysis at this time, the workgroup has begun to discuss potential alternatives that could allow certain workgroup participants to obtain the energy-usage data they desire without compromising customer privacy.

As the OAG understands the current discussions, certain workgroup members have formulated a hypothetical four-part model to govern customer energy-usage disclosure. This model has not been specifically proposed by a workgroup participant, but is rather the OAG's interpretation of the workgroup's discussions from the May 16, 2014 meeting and May 21, 2014 telephone conference. The OAG understands that the four part model would operate as follows: First, individual-customer data would not be released under any circumstances absent customer consent. Second, utilities would gather their customers' energy-usage data and either make the data public according to specific aggregation/geographical standards established by the Commission, or would provide the individual-customer data for their entire service territory to a

¹ See *Order Instituting Rulemaking to Consider Smart Grid Technologies Pursuant to Federal Legislation and on the Commission's Own Motion to Actively Guide Policy in California's Development of a Smart Grid System*, Rulemaking 08-12-009.

centralized repository for standardized publication.² Third, owners of commercial buildings would be allowed to obtain from utilities aggregated energy-usage data of their properties for benchmarking purposes. Fourth, additional data requests (not contemplated by these three categories) would be referred to the PUC on an individual basis for further scrutiny.

The OAG agrees with the first and fourth components of this model. The OAG has concerns, however, with the second and third components of the model and cautions that the details of these disclosures are critical to ensure adequate protection of consumer data.

Public Access to Energy-Usage Data by Geographic Region

Standardizing the type of aggregated data that may be made public by pre-determined geographic boundaries *in lieu* of developing standards for utilities to respond to individual data requests could, if done properly, provide several benefits for protecting privacy. First, by standardizing the data that is produced by geographic boundaries, and *only* producing the specific, standardized data that is published, the Commission could curtail the ability of requestors to obtain data on individual customers by “layering” requests. In other words, standardizing the energy-usage data available to individuals by publishing certain data removes control from requestors over the specific data they receive, and limits the types of data they may obtain. Second, regulators would know exactly what energy-usage data is available to the public, and can make changes to the types of data available as needed. Third, as some utilities indicated during the May 16, 2014 meeting and May 21, 2014 telephone conference, the burden on utilities of responding to numerous individualized data requests may be reduced considerably if they can limit production to specific parameters.

The effectiveness of utilizing standard, geographic parameters for aggregating data to protect customer privacy is dependent on many important details regarding the published data. As indicated above, this proposal was developed as a potential alternative to hiring a technical expert who would report on appropriate levels of aggregation/anonymization for responding to individual data requests. Therefore, the data sets available under this model will likely not be verified as “safe” by statistical or other technical analysis. For example, while standardizing the data available to requestors may make it more difficult to identify individual customers through multiple queries, it does little to prevent attempts to identify individual customers by combining energy-usage data with other pre-existing information on an individual customer.³ Therefore, the Commission should proceed cautiously if it chooses to develop a plan to allow for publication of specific energy-usage information. Specifically, the Commission should limit publication of

² Currently, it is unclear who would manage this centralized repository. The group’s discussions appear to indicate that the repository would be managed by the Commission or Department of Commerce.

³ See California Decision at 22 (“EFF’s Technical Memo identifies two main security risks associated with releasing aggregated data: 1) privacy attacks using multiple queries on data and 2) privacy attacks using pre-existing information about an individual customer.”)

energy-usage data to information that is both “non-granular” and includes a large number of customers.

First, with regard to granularity, and as the OAG noted in its May 5, 2014 comments, no party has requested energy-usage data more granular than monthly. Accordingly, any publication of energy-usage data should be limited to data that is either monthly or based on a longer term.

Second, with respect to the number of customers needed to make energy-usage data public, the Commission should require the use of broad geographic regions that encompass large numbers of customers in order to meaningfully protect against the risk of customers being “re-identified” from the data sets. Absent the technical analysis recommended by the workgroup, the specific aggregation level needed to provide adequate protection of customers’ privacy cannot be known. Different aggregation levels have been established by regulators seeking to protect individual privacy when aggregated data is released.⁴ While some parties have requested aggregation levels as “low” as 4/80⁵ or by census block, such a standard—without supporting technical analysis—appears insufficient in light of other standards to protect customer privacy and the potential for re-identification of specific customers. Moreover, several utilities indicated during the May 21, 2014 call, that they do not maintain customer data by census block.

Finally, if the Commission elects to develop a standard for aggregating and releasing public utility data by geographic region, this standard should be applied to utilities directly, rather than requiring utilities to produce “raw” individual-customer data to a centralized repository, including the government, for aggregation and production. Any entity outside of the utility, including a government agency, should not have direct access to individual customer

⁴ For example, the California Decision discussed by the workgroup allowed for data aggregation at the zip code level, unless a specific zip code failed to meet minimal aggregation thresholds established for each customer class. For residential customers, aggregated data cannot be published unless at least 100 customers are included in the data set. California Decision at 26. In setting its aggregation levels, the California Public Utilities Commission acknowledged that “through review of the Working Group Report, technical and legal memos supplied by EFF, and the opening and reply comments by various parties, *it has become clear that there does not yet exist a set of ‘best practices’ that describe what data is sufficiently aggregated so as to prevent linkage to specific individuals.*” *Id.* at 22 (emphasis added). The Health Insurance Portability and Accountability Act (“HIPAA”) provides another aggregation standard regarding healthcare data for comparison. HIPAA requires that, absent an appropriate technical analysis, data must be removed of all unique identifiers and aggregated to the level of a zip code’s first three digits, with no fewer than 20,000 people, in order for the data to be considered not individual identifiable health information. *See* 45 CFR § 164.514.

⁵ A “4/80” standard would prohibit a release of energy-usage data unless it included at least four customers and no single customer could account for more than 80 percent of the group’s energy consumption.

energy-usage data without customers' knowledge and consent or other proper legal justification. Moreover, the OAG notes that, during the May 21, 2014 call, utilities appeared to indicate that producing data in a manner that complies with specified aggregation standards themselves may be less burdensome than providing raw individual customer data to a third-party. Accordingly, if the Commission elects to allow public release of non-individually identifiable energy-usage data aggregated by geographic boundaries, the OAG recommends that these boundaries be set broadly with a high number of individual customers, and that the aggregated, de-identified data be produced directly by the utility.

Building Benchmarking

Based on the workgroup's discussions, the OAG understands that releasing data to building owners for purposes of energy benchmarking is the only "use case" that would not be addressed either by obtaining individual customer consent or through the proposal to publish aggregated data at standardized parameters. In addition to general policy discussions, workgroup participants supporting the release of whole-building data have focused on both the city ordinance recently passed in Minneapolis and statutory language recently adopted in California.⁶ In considering the potential release of whole-building data and these laws, the following points should be considered.

First, both the Minneapolis ordinance and California statute limit the obligation to produce whole-building energy-usage data for benchmarking purposes to non-residential properties.⁷ Further, the workgroup has not identified a legal obligation for production of energy-usage data for residential properties. Accordingly, to the extent that the Commission allows production of aggregated whole-building energy usage data, it should explicitly restrict this production to non-residential properties.

Second, the group of large industrial companies participating in the workgroup has raised significant concerns regarding the possibility that important trade secret information could be compromised by the release of aggregated data. It is unclear whether any of the large industrial companies participating in this workgroup would be affected by the release of commercial whole-building data. Regardless of whether specific parties to the workgroup would be affected by a release of whole-building aggregated data, the type of concern expressed by these industrial companies should be fully considered when determining if and how whole-building data should be made available.

⁶ See MINNEAPOLIS, MN., CODE § 47.190; CAL. PRC CODE § 25402.10. While the statute recently passed in California obviously creates no legal obligation in Minnesota, the workgroup has discussed it.

⁷ MINNEAPOLIS, MN., CODE § 47.190 (a) (defining "covered building" as one "of an occupancy use other than residential or industrial"); CAL. PRC CODE § 25402.10 (a).

As indicated above, while the OAG agrees with the first and fourth components of the model currently being discussed by the workgroup, it would urge caution by the workgroup in developing the second and third components. The workgroup's consideration of the many difficult and important issues raised makes clear that careful evaluation of the details of these components is critical before such a platform is presented to the Commission.

Sincerely,



IAN DOBSON
Assistant Attorney General

(651) 757-1432 (Voice)

(651) 296-9663 (Fax)