

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17

STATE OF NEW HAMPSHIRE
BEFORE THE
PUBLIC UTILITIES COMMISSION

Docket No. DE 19-197

Electric and Natural Gas Utilities
Development of a Statewide, Multi-use Online Energy Data
Platform

TESTIMONY OF
ETHAN GOLDMAN

August 17, 2020

On behalf of Clean Energy NH

STATE OF NEW HAMPSHIRE

BEFORE THE PUBLIC UTILITIES COMMISSION

DIRECT TESTIMONY OF ETHAN GOLDMAN

DEVELOPMENT OF A STATEWIDE, MULTI-USE ONLINE ENERGY

DATA PLATFORM

August 17, 2020

Docket No. DE 19-197

Table of Contents

I.	Glossary of Terms	3
II.	Introduction & Witness Qualifications	4
III.	Summary of Policy Context & Legislative Objectives	6
IV.	An Agile Process for Defining the Functional Requirements of the Energy Data Platform	9
V.	Virtual Data Platform & Data Platform Hub	13
VI.	Governance: Data Platform Council	24
VII.	Conclusion	31
	Exhibit CENH-1: User Stories Narratives	33

1 **I. GLOSSARY OF TERMS**

2 **Application Programming Interface (API):** An application programming interface (API) is
3 a computing interface which defines interactions between multiple software intermediaries. It
4 defines the kinds of calls or requests that can be made, how to make them, the data formats that
5 should be used, the conventions to follow, etc. It can also provide extension mechanisms so that
6 users can extend existing functionality in various ways and to varying degrees. An API can be
7 entirely custom, specific to a component, or it can be designed based on an industry-standard to
8 ensure interoperability.

9 **Data Sources:** Any entity or system that provides data through the statewide, multi-use line
10 energy data platform (Energy Data Platform or Data Platform). Examples include electric and
11 gas utilities; third party energy service providers (e.g., distributed energy resource provider);
12 competitive energy suppliers; and Community Power Aggregators.

13 **Data Users:** Any entity or system that requests and receives data from the Data Platform.
14 Examples include: residential, commercial, municipal energy consumers; Community Power
15 Aggregations (CPAs); third party service providers; researchers; utilities; state government
16 agencies.

17 **Data Platform Hub:** A centralized web-based directory of approved and available data sets, the
18 location of all approved Data Sources, including documentation for their APIs and the shared
19 logical data model on which they are based. The Hub is a separate and distinct component from
20 all the individual Data Sources. The Hub and all Data Sources collectively comprise the Data
21 Platform.

22 **Data Platform Council:** The proposed governance body tasked with: (1) Approving standards
23 for publication on the Data Platform Hub, including shared logical data model, API standards,

1 and standards for authentication and authorization; (2) Ensuring that new Data Sources meet
2 established standards in order to be listed on the Data Platform Hub; and (3) Evaluating the
3 ongoing performance of Data Platform and its component Data Sources to ensure it is meeting its
4 goals (e.g., enabling priority user stories listed in Exhibit CENH-1.).

5

6 **II. INTRODUCTION & WITNESS QUALIFICATIONS**

7

8 **Q. State your name, the organization you work for, your position, and your business**
9 **address.**

10 **A.** My name is Ethan Goldman, I am the founder of Resilient Edge, LLC, a consulting firm
11 which is located at 5 Pavilion Ave in South Burlington, VT.

12

13 **Q. Describe your background and qualifications.**

14 **A.** I have been working as an information technology professional for more than 20 years,
15 and specifically in energy data analytics for more than 10 years. I have designed, developed,
16 delivered, and supported software systems that analyzed utility billing data as well as non-utility-
17 meter energy use data and grid system data from Independent System Operators (ISOs). Notably,
18 I was the technical subject matter expert for Efficiency Vermont's planning and deployment of a
19 state-wide multi-utility Advanced Metering Infrastructure (AMI) data warehouse. I hold multiple
20 patents for software algorithms that analyze data from energy meters and other sensors in order
21 to provide insights about energy use in buildings. I have published multiple peer-reviewed
22 articles on the topic of energy data analysis and have spoken at numerous national conferences
23 on this topic. I have participated in both regional and national committees on energy efficiency

1 measurement and verification techniques and practices, such as the Lawrence Berkeley National
2 Labs M&V 2.0 National Stakeholder Group. I earned an MS from Carnegie Mellon University,
3 where I researched non-intrusive load monitoring, a machine learning technique for
4 disaggregating multiple electric loads from whole-house meter data using machine learning
5 techniques.

6 **Q. Have you previously testified before the New Hampshire Public Utilities**
7 **Commission (NHPUC) or other regulatory bodies?**

8 **A.** I have not.

9

10 **Q. On whose behalf are you testifying?**

11 **A.** I am testifying on behalf of Clean Energy NH (CENH).

12

13 **Q. Explain CENH's interest in the creation of a statewide, multi-use online energy data**
14 **platform.**

15 **A.** CENH is a member-based 501(c)(3) nonprofit organization whose members include 26
16 New Hampshire local governments (cities, towns, and counties), over 130 businesses providing a
17 breadth and depth of energy related services, and hundreds of individual members. CENH acts
18 on behalf of both municipal and residential energy customers, and the third-party energy
19 companies offering in-state energy solutions to those energy consumers. A successful data
20 platform will cater to the needs of both the market actors and municipal and residential
21 customers of CENH's membership.¹

¹ Clean Energy NH's local government and business membership may be viewed at www.cleanenergynh.org/our-members.

1

2 **Q. Describe your involvement in DE 19-197 up until this point.**

3 **A.** I have participated actively in many PUC-led technical sessions and in informal
4 conversations with utility and non-utility stakeholders since early in this docket process.

5

1 **III. SUMMARY OF POLICY CONTEXT & LEGISLATIVE**

2 **OBJECTIVES**

3
4 **Q. What is the purpose of your testimony?**

5 **A.** The purpose of my testimony is to share expertise in energy data systems and platforms
6 with the Commission to support informed decision-making and the successful implementation of
7 Senate Bill 284 and RSA 378:51. My testimony covers both the functionality of the Data
8 Platform, and the importance of a thoughtful approach to data platform governance.

9
10 **Q. Briefly summarize the key points of your testimony.**

11 **A.** My testimony explains and endorses a virtual approach to establishment of the Energy
12 Data Platform. I describe the importance of shared standards for Data Sources (including electric
13 and gas utilities, as well as other Data Sources). Shared standards relate to: a logical data model;
14 Application Programming Interfaces (APIs); authentication (i.e., “are you who you say you
15 are?”); and authorization (“is this account allowed to access this data?”). I explain why a
16 successful Data Platform will have a single, unified point of access to energy data from multiple
17 Data Sources and for multiple Data Users, which I refer to as the Data Platform Hub. I describe
18 how a limited governance process providing oversight to both the establishment and enforcement
19 of shared standards, and the unified point of access to energy data, will allow for the agile
20 development and ongoing evolution of the Energy Data Platform.

1 **Q. What are the legislative objectives of Senate Bill 284 and RSA 378:51?**

2 **A.** The New Hampshire General Court adopted Senate Bill 284² to further accomplish the
3 purposes of electric utility restructuring under RSA 374-F,³ the most compelling of which is to
4 foster market competition in both retail and wholesale electricity markets, and to meet state
5 energy policy goals under RSA 378:37.⁴ Senate Bill 284 also acknowledges the importance of
6 energy data access, both as individual data and as aggregated and anonymized data, in fostering
7 innovative business applications and enabling Community Power Aggregations (CPAs) under
8 RSA 53-E.⁵ The State also aims to empower customers and to make the State's energy systems
9 more distributed, responsive, dynamic, and customer focused. The State legislature recognized
10 that in order to achieve these objectives, it is necessary to provide consumers and stakeholders
11 with safe, secure access to granular information about their energy usage.

12

13 **Q. How does the establishment of a statewide, multi-use online energy data platform**
14 **relate to electric grid modernization and NHPUC Docket No. IR 15-296?**

15 **A.** I have not been a participant in New Hampshire's grid modernization proceedings, so I
16 cannot speak to the particulars of that specific docket. The objectives of Electric Grid
17 Modernization in New Hampshire as established by the Order No. 25,877⁶ include: integrate

² Senate Bill 284 and NH RSA 378:51, Multi-use Online Energy Data Platform. Retrieved from:

<https://www.gencourt.state.nh.us/rsa/html/XXXIV/378/378-51.htm>

³ NH RSA 374-F, Electric Utility Restructuring. Retrieved from:

<http://www.gencourt.state.nh.us/rsa/html/XXXIV/374-F/374-F-mrg.htm>

⁴ NH RSA 378:37, Least Cost Energy Planning. Retrieved from:

<https://www.gencourt.state.nh.us/rsa/html/XXXIV/378/378-37.htm>

⁵ NH RSA 53-E, Aggregation of Electric Customers by Municipality and County. Retrieved from:

<https://www.gencourt.state.nh.us/rsa/html/III/53-E/53-E-mrg.htm>

⁶ NHPUC Docket No. IR 15-296 Investigation into Grid Modernization, Order on Scope and Process. April 1, 2016. Page 2. Retrieved from: https://www.puc.nh.gov/Regulatory/Docketbk/2015/15-296/ORDERS/15-296_2016-04-01_ORDER_25877.PDF

1 distributed energy resources; empower customers to use electricity more efficiently and to lower
2 their electric bills; reduce generation, transmission, and distribution system costs; and improve
3 reliability, resiliency, and operational efficiency of the grid. These objectives appear to be similar
4 to the objectives of Senate Bill 284, electric utility restructuring under RSA 374-F, and state
5 policy goals under RSA 378:37.

6 The Grid Modernization Working Group Report to the New Hampshire Public Utilities
7 Commission⁷ of March 2017 emphasizes the importance and the necessity of data access and
8 more dynamic rate design to empowering customers and modernizing the grid.

9

10 **Q. Please discuss RSA 53-E, relative to Aggregation of Electric Customers by**
11 **Municipalities and Counties, and its relationship to the Energy Data Platform.**

12 **A.** In order to provide electricity and other retail products and services to their residents and
13 businesses, Community Power Aggregations (CPAs) require ongoing access to individual and
14 aggregated customer energy data. A successful and modern Energy Data Platform will be of
15 great use to these CPAs.

⁷ New Hampshire Grid Modernization Working Group Report to the New Hampshire Public Utilities Commission. March 20, 2017. Retrieved from: https://www.puc.nh.gov/Regulatory/Docketbk/2015/15-296/LETTERS-MEMOS-TARIFFS/15-296_2017-03-20_NH_GRID_MOD_GRP_FINAL_RPT.PDF

1 **IV. AN AGILE PROCESS FOR DEFINING FUNCTIONAL**
2 **REQUIREMENTS OF THE ENERGY DATA PLATFORM**

3
4 **Q. At a high level, who do you anticipate would use the Energy Data Platform, and for**
5 **what purposes?**

6 **A.** The statewide Energy Data Platform should contain multiple types of energy data
7 (electric and gas usage, and potentially other types in the future) from multiple different sources.
8 I anticipate that the Data Platform will be accessed by a variety of different types of users,
9 including individual customers, authorized third-party service providers, community planners,
10 and researchers, to name just a few. While the Platform would support individual energy
11 customers or staff from a third-party service provider manually seeking out a particular data set
12 in order to perform a discrete analysis task, the true value of the Data Platform is enabling
13 automated software that makes use of energy data. The goal of the Platform is to move away
14 from infrequent individual data requests that require utility staff time as well as expert analyst
15 time, and toward an ecosystem of energy-aware market-based energy services that drive
16 economic growth in New Hampshire while lowering costs for ratepayers.

17
18 **Q. Should the Commission Order specify the exact functional requirements and**
19 **standards of the Data Platform? What is an agile software development process?**

20 **A.** No. It is not practical to specify at the outset all of the functional requirements and
21 standards that will enable the Data Platform to meet the obligations described in RSA 378:51,
22 nor would such an approach be advisable to attempt. Rather, modern software development best
23 practices point to an “agile” approach to projects such as this, where functionality is developed

1 incrementally over time so that feedback from users can be incorporated into the design. An agile
2 software development approach saves time and money that would otherwise be spent building
3 unnecessary features, and delivers a more useful product more quickly. Adjudicative regulatory
4 processes are slow, adversarial, costly in time, money and resources, and in many ways the
5 antithesis of agile software development. It is not appropriate to adjudicate every iteration of a
6 Data Platform that needs to grow and evolve over time.

7

8 **Q. Is it a valid alternative to direct and expect each utility to determine the detailed**
9 **functionality and standards of their portion of the Platform over time?**

10 **A.** No. I do not believe it is wise or reasonable to fully delegate the authority and
11 responsibility for detailing functionality and standards to monopoly investor-owned utilities
12 without empowered market and stakeholder engagement and oversight. The purpose of the
13 Energy Data Platform is to foster market competition through access to information. Investor-
14 owned monopolies are not good at making their markets competitive. This approach is not likely
15 to produce a successful Data Platform.

16 If each utility is left to detail its own functionality and standards, the outcome is likely to
17 be a disparate, uncoordinated, difficult to use series of separate data access points.

18

19 **Q. If you do not recommend detailing functional requirements in the Commission**
20 **Order, and you do not recommend each utility detail its own functional requirements**
21 **independently, what would you recommend is the best approach to detailing functional**
22 **requirements?**

1 **A.** Functional requirements should be detailed through establishment of a simple and limited
2 governance process, subject to final oversight by the Commission, which for the sake of this
3 testimony we will call the “Data Platform Council.” The Data Platform Council should be
4 responsible for three narrowly defined roles:

- 5
- 6 1. Approving standards for publication on the Data Platform Hub, including shared logical
7 data model, API standards, and standards for authentication and authorization;
 - 8 2. Ensuring that new Data Sources meet established standards in order to be included in the
9 Data Platform Hub;⁸
 - 10 3. Evaluating the ongoing performance of Data Platform to ensure it is meeting its goals
11 (e.g., enabling priority user stories listed in Exhibit CENH-1.).
- 12

13 These three roles are described in more detail in **Section VI. Governance.**

14

15 **Q.** **Why is governance an important part of the Energy Data Platform?**

16 **A.** A thoughtful governance mechanism is critical for several reasons. First, it is necessary to
17 provide oversight to the agile software development process, establishing of standards, and
18 detailing of functional requirements that will evolve over time. Second, it is the most efficient
19 way to monitor ongoing performance. Both ongoing development of incremental features and
20 maintenance of existing features require oversight to ensure that the results are delivered on time
21 and to the required level of quality. A third function of governance is to assess the completion of

⁸ The Data Platform Hub is a centralized web-based directory of approved and available data sets, the location of various Data Sources, including documentation for APIs and the shared logical data model on which they are based. It is described in further detail in **Section V. Virtual Data Platform & Data Platform Hub.**

1 new functionality and bug fixes to confirm that API services and data results meet all approved
2 standards. In this way, the role of the Data Platform Council is to serve as technical expert acting
3 in the interest of the marketplace, representing the interests of the “customers” (including both
4 ratepayers and third parties) in answering the questions: “is the Data Platform working? Does it
5 meet my needs? Is there a better technical solution to enabling data access?” A simple and
6 streamlined governance approach paired with the appropriate software and data expertise, and
7 subject to final Commission oversight, will not only be able to provide faster feedback than
8 would be possible through traditional regulatory review processes, it will also be able to
9 effectively assess the results of the data providers’ development process, balancing the level of
10 functionality and the level of effort.

11 The final role of the Data Platform Council is to carry a small stick: in the event that a
12 Data Source does not comply with the shared logical data model or API standards, the Data
13 Platform Council can exclude that Data Source from being listed on the Data Platform Hub.

14 Commission adjudication is the right process to set up governance, but not the right
15 process to oversee agile software development or develop technical requirements in an ongoing
16 fashion. Good governance will allow for Data Platform extensibility, the quality of being
17 designed to allow the addition of new capabilities or functionality over time.

1 **V. VIRTUAL DATA PLATFORM & DATA PLATFORM HUB**

2

3 **Q. Please describe, at a high level, what you would recommend as the appropriate**
4 **structure and management of the Data Platform.**

5 **A.** The defining feature of the Data Platform is that it provides a single, unified point of
6 access to energy data from multiple Data Sources and for multiple Data Users. This means that
7 there is one place to find a list of all available New Hampshire energy data, along with
8 instructions for creating an account and requesting or granting permission to access particular
9 data files, and documented procedures for retrieving standard-format data files. While multiple
10 Data Sources will all be providing data from a variety of software systems, users will receive
11 data that conforms to a single logical data model via consistent Application Programming
12 Interfaces (API), regardless of the Data Source. We refer to this centralized, standardizing
13 repository as the “Data Platform Hub.” Without this centralized, standardized hub for Data
14 Users, there is no Data Platform, but rather series of disparate data sources with disparate and
15 uncoordinated user interfaces.

16 In order to unify the various energy Data Sources into a cohesive Platform, there must be
17 a governance mechanism that sets standards for data formats and APIs, then validates that
18 participating Data Sources meet that standard in order to be part of the Platform.

19

20 **Q. Simply define the Data Platform Hub.**

21 **A.** A centralized web-based directory of approved and available data sets, the location of
22 various Data Sources, including documentation for APIs and the shared logical data model on
23 which they are based.

1 **Q. Explain what is meant by “virtual platform” and/or “hybrid model”.**

2 **A.** During the technical sessions, the terms “virtual platform” and “hybrid model” were used
3 synonymously. While neither term has a standard meaning in the industry, I will use the former
4 in my testimony as I believe that “virtual platform” more aptly describes the solution that was
5 discussed in those sessions: something that acts like a platform from the perspective of users, but
6 could be built in a distributed fashion so there is no singular “platform” installed on a server
7 somewhere.

8 Through the technical sessions held by PUC staff over the last several months, it seems
9 unlikely that there is appetite for a “physical” data platform that would ingest data from all
10 sources, transform it into a common data format, and then provide standardized data to
11 authorized users. Instead, there seemed to be consensus among utility and non-utility parties in
12 support of a “virtual” platform providing standard-format energy data from a single point of
13 access by routing those requests to the appropriate sources in real-time.

14 A good analogy is the way Amazon.com allows many merchants to sell their products
15 through their marketplace: it has a single point of entry where you can find products, you only
16 need to create one customer account with your payment info and shipping address, but each
17 product in your order might ultimately be fulfilled by a different vendor. If there was no front
18 page with categories to browse and a site-wide search feature, there would be no marketplace at
19 all. If each vendor used a different page layout to explain their product and required you to create
20 a separate account, the shopping experience would be much more cumbersome and hardly better
21 than visiting multiple sites to compare products and place your order.

22 This virtual platform approach would serve well the interests of Clean Energy NH,
23 provided it includes an accompanying governance process for ensuring Data Sources adhere to a

1 shared logical data model and other standards, as well as a Data Platform Hub that is
2 independently managed.

3

4 **Q. Explain further the role of centralized documentation for all data feeds that make**
5 **up the Data Platform. How does the Energy Data Platform work?**

6 **A.** First, note that the technical details of how all the individual Data Sources implement
7 their services is not only beyond the scope of this testimony, it is, by definition, immaterial to the
8 functioning of the virtual platform. The only centrally-managed portion of the virtual platform
9 that must be built outside of all Data Sources' IT systems is the documentation in the Data
10 Platform Hub. This would simply be a small website with instructions on where to find different
11 kinds of data and how to access it. This Data Platform Hub should include documentation about
12 how to access the different APIs that provide energy data and how to interpret the data files
13 returned by those APIs, as well as documentation for the logical data model that is used to
14 organize data across all the platform's Data Sources.

15 Some additional technology that could tie together the various Data Sources' APIs is a
16 thin layer that routes data queries to the appropriate Data Source. That would allow, for example,
17 an energy service provider to fetch customer data from any utility by sending an API request to
18 (hypothetically) `energy-data.nh.gov/electricity/get-data` and include the utility name and
19 customer account number in the request. The Platform would then redirect that request to the
20 appropriate utility's server, where it would be authenticated and (if the requestor was authorized)
21 the utility's server would then respond with the appropriate data. The exact details of how this
22 would work need to be determined in coordination with representatives of both the Data Source
23 and users of the system, but the overarching philosophy should be to minimize the role of the

1 Platform and allow the Data Source to maintain control over their systems, so long as they
2 conform to the established standards. If this same level of seamless access to data from different
3 Data Sources can be achieved with no additional technology beyond what is provided by the
4 Data Sources, that solution would be preferable.

5

6 **Q. What is an API?**

7 **A.** The acronym “API” stands for Application Programming Interface. It refers to an
8 interface between software applications that clearly defines the inputs and outputs, including
9 protocols, data formats, error messages, etc. The use of APIs allows different computer systems
10 to interact with one another in predictable ways. Rather than requiring explicit coordination
11 between developers on all components that need to exchange data or integrate functionality, the
12 API allows different software teams to create systems that work together simply by complying
13 with the standards defined by the API. This reusable, modular development approach has
14 facilitated an explosion of richly featured software that can be built much more quickly by
15 leveraging data and services provided by APIs.

16 One example is the proliferation of websites that include some sort of mapping
17 functionality to find locations near you and map the route to them, whether for a national chain
18 of retail stores or community meet-ups. While very few entities have the resources to create and
19 maintain systems that collect and organize these huge, complicated sets of geospatial data,
20 because they offer APIs to access the data through defined searches that return maps and routing
21 information, this data has facilitated or enhanced numerous websites and apps.

22 In a similar way, today’s access to energy data is analogous to requesting a local map
23 from a city’s chamber of commerce, then manually reading the legend to figure out how to

1 interpret the particular symbols they used in order to figure out where to get lunch nearby and
2 whether there is a bus that takes you there. If energy data is provided through APIs, it will be
3 possible for a wide variety of enterprises to build new functionality that incorporates that energy
4 data with their services, whether providing estimates for efficiency or solar PV projects,
5 estimating a building's carbon footprint, aggregating demand response, stimulating a friendly
6 competition to lower energy consumption, etc. Many of the apps that use mapping data today
7 would have been hard to conceive of 20 years ago, and would probably not have been built if
8 only Garmin and Delorme were providing services with that data.

9

10 **Q. Are APIs secure against data breaches or cyber-attacks?**

11 **A.** Yes. While APIs may openly publish the documentation for how to submit a data request
12 and how to interpret the result, it does not mean that they must also provide all the data to any
13 software application that submits a request. APIs can require authentication, as do the ones used
14 by online financial management software to access a user's bank records. In fact, not only can
15 APIs require authentication before returning particular records, they can even require that the
16 software application making the request has registered with the Data Source and accepted the
17 terms of service, including security and privacy requirements, before being allowed to even issue
18 requests. APIs can operate over secure connections, protecting both the data and the user
19 credentials from the request.

20 **Q. Are there other security considerations for the design of the Data Platform?**

21 **A.** Access to energy data through the platform requires two components: authentication
22 ("Are you the owner of this account?") and authorization ("Is this account allowed to access this
23 data?"). Because the Virtual Data Platform should feel like a single, cohesive point of access to

1 all of New Hampshire energy data, even though it relies on multiple Data Sources, the Platform
2 should use a Federated Identity Management system that allows users to create a single set of
3 authentication credentials (i.e. username and password) that can be linked to all of their utility
4 (and other Data Source) accounts. This would also apply to third parties, who would also have a
5 single set of credentials across all Data Sources. This would simplify tasks such as an energy
6 customer granting or revoking authorization for a third-party energy service provider to all of
7 their energy data (e.g. from both electricity and gas utilities, or for accounts from multiple
8 buildings). It would also make it easier for the Data Platform to enforce stronger security rules
9 regarding password strength, scheduled password changes, or even multi-factor authentication,
10 because all participants would be using a single Federated Identity Management system.

11

12 **Q. Explain what is meant by “logical data model.” Explain how the logical data model**
13 **relates to the “virtual platform” and is responsible for keeping the separate data APIs**
14 **coordinated.**

15 **A.** A logical data model is an abstract representation of the structure and meaning of data
16 that will be delivered through the Data Platform. It does not specify how the database or other
17 storage mechanism will be implemented; there are multiple ways to satisfy the requirements
18 specified by the logical data model, based on the existing infrastructure and other constraints of
19 each data provider who will be conforming to it.

20 The logical data model defines data entities such as customer, premises, meter, and
21 reading. It also defines the attributes of those entities, such as customer name, premises service
22 address, meter scaling factor, and reading period end date, which ensures that there is a common
23 definition of terms and means of encoding information. For example, is the rate code (which can

1 change over time) an attribute of a meter reading, meter service point, or premises? Is rate code
2 instead an attribute of a service contract entity linking a customer to a meter, and which has start
3 and end dates? These questions should have consistent answers that apply to all Data Sources.

4 This brings us to the last component of a logical data model: the relationships. These
5 describe how to link entities to one another, and whether they are one-to-one, one-to-many, or
6 many-to-many. This is a critical distinction when trying to determine which accounts belong to a
7 single customer, or when assembling the total energy use across all the meters on a single
8 building.

9 Again, since the logical data model is not a description of how any system will be
10 implemented, there are many ways that utilities and other data providers can input and store their
11 data, so long as it can be delivered in a format that matches the logical data model. In my past
12 experience, utility data systems organize billing data in many different ways. Some have an
13 explicit customer record, while others simply duplicate the billing address for each premises;
14 some associate meter readings with a “service point” that can, over time, have multiple meters
15 with distinct IDs installed in it. In fact, I have even encountered utility data where different
16 customer accounts were configured according to different logic, based on how each customer
17 wanted to receive their bills (e.g., all bills sent to one address as opposed to each bill sent to the
18 service address). In order for users of the Data Platform to analyze and interpret data about
19 multiple buildings or customers from different utilities, it is critical that all the data providers are
20 able to map their source data to the logical data model. Not only does this require that the logical
21 data model is clearly defined so that it can be unambiguously interpreted by both Data Sources
22 (such as utilities) and Data Users, it also requires that the logical data model be carefully defined
23 so that it can accurately represent data from all possible sources.

1

2 **Q. What kind of Data Sources will the logical data model apply to?**

3 **A.** The logical data model will need to encompass not only electric and gas utility meter and
4 billing data, it will also need to be able to expand to include other types of data, if not in the first
5 version of the Data Platform, then in a future version. This might include Distribute Energy
6 Resource (DER) installation and operation data, distribution system and grid operations data,
7 market data such as utility tariffs and/or data from ISO New England, and various other energy
8 data that can help customers, service providers, and communities understand and make decisions
9 about their energy use. The logical data model is not static; it will evolve with the Platform.

10

11 **Q. Why is a shared logical data model critical to the success of the Data Platform?**

12 **A.** Without a shared logical data model, Data Users will be forced to attempt their own
13 interpretation and translation of data from different sources. The absence of a shared logical data
14 model imposes additional costs on types of activities the statewide Data Platform is meant to
15 enable. The absence of a shared logical data model would also mean that analysis and
16 interpretation by different users cannot be compared without first unpacking all of the data-
17 mapping steps that each analysis used. Rather than being able to draw a straight line from an
18 academic analysis through a regulatory ruling to an energy service offering and finally to
19 individual project results, every calculation performed with energy data will be an island unto
20 itself.

21 It is essential that all Data Sources participating in the Energy Data Platform adhere to
22 the logical data model so that users of the Platform can be assured that data will be consistent
23 and clearly defined. This requirement must also be backed up with a testing regiment. If a Data

1 Source's APIs do not pass the required tests then they will be removed from the Platform,
2 meaning that they will no longer be listed in the Data Platform Hub. Note there could be a grace
3 period, during which they can correct the error. If the Platform maintained a redirection function
4 to route requests to the Data Sources, then it could cut off requests to non-conforming Data
5 Sources as well. For entities like gas and electric utilities that are required to feed data into the
6 Platform, the PUC could levy penalties if they are no longer supplying data through the NH
7 Energy Data Platform as required because they failed to conform to the logical data model in a
8 timely fashion.

9

10 **Q. Should a Commission order in this docket prescribe the logical data model? What**
11 **should the Commission order specify as it relates to the logical data model?**

12 **A.** A Commission order should not prescribe the logical data model, because the logical data
13 model will need to evolve through an agile software development process over time. An order
14 should, however, require that all Data Sources to the Energy Data Platform conform to a single
15 logical data model, and that the logical data model be defined and approved by a Data Platform
16 Council. The Commission could establish the charter for the Data Platform Council, defining its
17 composition and decision-making process, any requirements for public input on proposed
18 changes to standards or participation rules, etc.

19 The logical data model will need to change over time as new types of energy data are
20 added to the Platform, or when participants identify problems or opportunities to improve the
21 logical data model. The Commission should establish a streamlined and limited governance
22 process under which a Data Platform Council defines and approves the details of the logical data

1 model and other standards. This will ensure good governance, transparency, democratic
2 engagement, and agility in the process of creating the logical data model.

3

4 **Q. What is the appropriate way for System Data to interface with the data platform?**

5 **How can we ensure security with regard to system data?**

6 **A.** System data refers to electric distribution system data related to demand, capacity, circuit
7 loading, voltage, circuit mapping and grid topography, power quality, hosting capacity, etc. We
8 can think of two categories of system data: (1) Public (system data that poses no security risks
9 and is publicly available to everyone); and (2) Permission-protected (system data that poses
10 security risks and is available only through special permission). Since utility customer use data,
11 which will almost certainly be part of the Platform, is also considered private when individually
12 identified, the Platform will necessarily include security features that control what data is
13 available to which users, which could also protect non-public system data. In the future, system
14 data of both types (public and permission-protected) could be made available through the Data
15 Platform. This may be a topic more appropriate to be addressed through future proceedings such
16 as Electric Grid Modernization.

17

18 **Q. Could you please summarize your recommendation for the structure of the Data**

19 **Platform?**

20 **A.** We recommend a virtual platform, in which Data Sources remain distributed, but users
21 experience a unified Platform with a single login, a central hub with a directory of all Data
22 Sources, and a shared logical data model (the Data Platform Hub). This will allow data from all
23 sources to be accessed, analyzed, and interpreted consistently. All Data Sources that comply with

1 the logical data model, API standards, and other standards will be listed on a single centralized
2 Data Platform Hub. This can include not only energy use and billing data from electricity and
3 gas utilities but could also expand to include other types of energy data such as distribution
4 system data (e.g., hosting capacity), DER data, ISO New England data, utility tariff data, and
5 other types of data. While the technical implementation details of each Data Source are left up to
6 that Data Source (e.g., an electric or gas utility, or another Data Source), the functional
7 specifications are established and enforced through the governance of the Data Platform Council.
8

1 **VI. GOVERNANCE: DATA PLATFORM COUNCIL**

2
3 **Q. What is an appropriate governance framework to ensure successful implementation**
4 **and ongoing operational performance of the Data Platform?**

5 **A.** NHPUC Docket No. DE 19-197 is an appropriate forum for furthering the strategic
6 direction of the Data Platform in accordance with legislative objectives of Senate Bill 284. As
7 discussed previously, adjudicated regulation is not the appropriate framework to guide the agile
8 implementation and ongoing operational performance of the Data Platform. A more appropriate
9 forum would be the creation of a “Data Platform Council” that would be responsible for three
10 core functions:

- 11
12 1. Approving standards for publication on the Data Platform Hub, including shared logical
13 data model, API standards, and standards for authentication and authorization;
- 14 2. Ensuring that new Data Sources meet established standards in order to be included in the
15 Data Platform Hub;
- 16 3. Evaluating the ongoing performance of Data Platform to ensure it is meeting its goals
17 (e.g., enabling priority user stories listed in Exhibit CENH-1.).

18
19 The Data Platform Council would oversee implementation and ensure successful ongoing
20 operation of the Data Platform, in accordance with the direction set by PUC and subject to final
21 oversight by PUC.

1 **Q. Should the utilities be responsible for building, managing, and governing the Data**
2 **Platform?**

3 **A.** Not without meaningful oversight. While each of the utilities has an important role to
4 play in creating the Virtual Data Platform, no individual utility will be responsible for creating an
5 overarching organizational structure that ties all the diverse data structures into a cohesive
6 experience. If SB 284 had simply directed each utility to provide energy use data to customers
7 and their authorized agents through a standard interface, each utility could just set up Green
8 Button Connect and demonstrate compliance to the PUC. However, the law directed the PUC to
9 create a statewide, multi-use online energy data platform, which implies a system where multiple
10 energy data suppliers and consumers can integrate their functionality using shared standards. If
11 the Platform is going to develop and enforce these shared standards, there must be an entity that
12 can incorporate both the desires of Data Users and the technical limitations of Data Sources into
13 a set of functional requirements that represents a cost-benefit optimized design, which all
14 participants can then implement their software systems against in whatever way they see fit. In
15 this way, the utilities (and other Data Sources) can build and manage their portions of the Virtual
16 Platform, but the governance must be carried out by an independent body.

17

18 **Q: Should the PUC provide management and oversight for the Platform?**

19 **A:** The PUC does have an important role to play in the Data Platform: they have been
20 directed to determine if this Platform is a good investment by the ratepayers of NH, which will
21 require the PUC to articulate the structure and rules for operating the platform. However,
22 evaluation of the Platform should not be based on an annual pass/fail review, nor should it be
23 evaluated by adjudication. Design decisions are nuanced and we will get a better outcome if the

1 process is iterative on a shorter time-scale than PUC adjudication. Ongoing oversight and input
2 from technical experts representing market interests will avoid utility delays and can guard
3 against utility overspending. Non-utility actors (either Data Sources or consumers) that are
4 involved with the Platform are not subject to PUC regulation. However, if a Data Platform
5 Council is established and endowed with the authority and responsibilities to set data standards
6 and to enforce them by barring non-complying data services from participation, that Council can
7 engage with stakeholders and ensure compliance with and evolution of standards, even as it is
8 accountable to the PUC.

9

10 **Q. Who would be the members of the Data Platform Council?**

11 **A.** The Data Platform Council should have representation from diverse groups that represent
12 the market, including public and private sectors, as well as representatives with technical
13 familiarity with the subject matter. Candidates for Data Platform Council membership could
14 include:

- 15 ● One or more seats for Data Sources (including utilities)
- 16 ● One or more seats for state government (PUC, OCA, State Energy Manager)
- 17 ● One or more seats for local government
- 18 ● One or more seats for academia and other researchers
- 19 ● One or more seats for advocacy groups
- 20 ● One or more seats for third party energy service providers and DER representatives

21 Note that it is expected that all parties participating in the Data Platform Council have
22 adequate proficiency to participate in technical conversations about the functional requirements
23 of the Platform and the tradeoffs inherent in different options. It would be acceptable to allow

1 Council members to designate a technical expert to participate in proceedings on their behalf, or
2 to accompany the voting member at meetings to help parse the implications of different choices.

3

4 **Q. How would the members of the Data Platform Council be selected?**

5 **A.** Members of the Data Platform Council could be selected through an
6 application/nomination process to be vetted and approved by the PUC. It is recommended that
7 terms are staggered so that the Council maintains institutional memory through inter-term
8 transitions.

9

10 **Q. What is the role of utilities in the Data Platform Council? Should utilities have a**
11 **voting role?**

12 **A.** The utilities should play an active role in the process for defining and updating the shared
13 logical data model and other standards. Whether or not utilities should have voting seats on the
14 Data Platform Council is an interesting question. While the utilities are not the owners of the
15 Data Platform, they are a major Data Source. It would be valuable to have the utilities (and
16 representatives from other Data Sources) closely involved with setting these standards so that
17 they can help to avoid requirements that would be impossible or unduly expensive to meet, and
18 instead to look for ways to leverage existing data systems and functionality.

19 However, SB 284 seems to suggest that regulated gas and electric investor-owned
20 utilities would be required to provide some key types of data through the Data Platform and
21 would therefore be in the position of seeking cost recovery for those efforts. If the decisions of
22 the Data Platform Council have implications regarding the cost of development efforts needed to

1 meet those requirements, this could potentially represent a conflict-of-interest for the utilities if
2 they are also allowed to vote.

3

4 **Q: What does the Data Platform Hub and its governance cost?**

5 **A:** Given that participation in the Council will only require a few hours a month (after an
6 initial push to establish the draft standards) and that all participants will all be representing
7 organizations that will likely be willing to let them participate “on the clock” I would assume
8 that Council members will not require compensation. The majority of the Virtual Data Platform
9 will be built by the Data Sources, but the Data Platform Hub should be hosted independently of
10 any of the Data Sources. The Data Platform Hub is simply a repository of information about the
11 Data Platform, including documentation about the logical data model, how to create credentials
12 from the Federated Identity Management system, where to find APIs for each Data Source, etc.
13 The cost of hosting a small website and hiring someone to keep it updated whenever the Council
14 adds or updates a standard should only cost a few thousand dollars per year.

15 It is possible that the Federated Identity Management system will also need to be hosted
16 independently of all the Data Sources, so that function could also be an expense of the Platform
17 Hub. This would likely cost less than \$20,000 per year. If this and the previously-described web
18 hosting functionality could be accomplished by an existing NH State IT department, the cost
19 could be even lower.

20 The last budget category to consider is the occasional need to hire an expert consultant
21 who can provide independent advice to the Council regarding database structure, API
22 mechanisms, security models, etc. This expertise should primarily be the responsibility of the

1 Council members, so the need for such outside consulting contracts should be the exception, not
2 the rule.

3 One way to fund these costs for the Data Platform would be to charge each Data Source a
4 modest amount for the use of the platform, which the utilities could presumably rate-base if the
5 Commission deemed such an expense prudent.

6

7 **Q. How would you describe start-up costs vs. operating costs for the Data Platform,
8 and how might those costs be approved and recovered?**

9 **A.** Some investment will be needed for the initial start-up cost of the Data Platform. The
10 Commission could approve a specific limited budget for each utility that it deems is reasonable
11 to meet the legislative objectives of SB 284. The utilities could then be confident that these
12 investments are prudent, and will be recovered. It also ensures that spending is capped. If, upon
13 working with the Data Platform Council, parties determine additional funds are required, those
14 parties could seek Commission approval for modification of those budgets. Once the Data
15 Platform is established, operating costs could then be recovered under a performance-based
16 ratemaking approach.

17

18 **Q. Do you recommend traditional cost-recovery approaches for Data Platform costs?**

19 **A.** No. I do not believe it is the best approach for regulated utilities to recover costs they
20 incur to meet the requirements imposed by the Platform through traditional models. If the
21 Platform is treated as an “OpEx” pass-through cost, the motivation is to do as little work as
22 possible on it so their IT resources are available for other work. If it is treated as a “CapEx”
23 investment, it benefits utilities to invest more in the Platform, but spending more on software is

1 not consistently correlated with producing a more valuable result. Instead, I recommend the use
2 of performance-based ratemaking (PBR) to compensate utilities for implementing their portions
3 of the Data Platform.

4

5 **Q. What is the role of performance-based ratemaking in the operation of the data**
6 **platform?**

7 **A.** To create incentives for utilities to focus on ways to make the Platform truly successful, I
8 recommend creating a performance-based regulation framework for compensating utilities if the
9 Platform meets performance metrics, which could include: Platform traffic or amount of use;
10 timely responsiveness to requests for third-party authorization and data delivery; prevalence of
11 reported bugs (and ability to fix quickly); ability to keep the API services online with minimal
12 downtime; or other metrics. These metrics should all be quantifiable through automated means
13 (such as logging server activity, bug-tracking systems, etc.) so that up-to-date metrics are
14 available on a monthly basis. This ensures that any systemic issues can be quickly identified and
15 addressed, and that the utility has a clear incentive to do so.

16 In keeping with other models for utility compensation, the PBR could be structured so that
17 utilities were able to recover their implementation and maintenance costs if they meet the basic
18 requirements established by the Data Platform Council and remained a Data Source in good
19 standing. If they also meet other goals that measure the ratepayer value and impact of the
20 platform, such as number of datasets delivered, that would be factored into additional
21 compensation that would benefit utility shareholders. This creates both a disincentive (stick) for
22 allowing the platform to fail, as well as an incentive (carrot) for making it not just functional, but
23 effective and valuable to ratepayers.

1 **VII. CONCLUSION**

2 **Q. Please summarize the key recommendations of your testimony.**

3 **A.** I recommend the Commission direct for the implementation of a virtual Data Platform,
4 one that allows for Data Sources to remain distributed, and allows for Data Sources to implement
5 standard-compliant systems as they see fit.

6 I recommend the Commission require that the Data Platform and its Data Sources adhere
7 to a shared logical data model.

8 I recommend the Commission direct for the creation of a Data Platform Hub, defined as:
9 a centralized web-based directory of approved and available data sets, the location of various
10 Data Sources, documentation for APIs, and the shared logical data model on which they are
11 based. The Commission should also identify a funding mechanism to allow the Hub to be hosted
12 independently of any of the Data Sources.

13 I recommend the Commission establish a Data Platform Council to serve as the
14 governance body tasked with: (1) Approving standards for the Data Platform Hub, including
15 shared logical data model, API standards, and standards for authentication and authorization; (2)
16 Ensuring that new Data Sources meet established standards in order to be included in the Data
17 Platform Hub; and (3) Evaluating the ongoing performance of Data Platform to ensure it is
18 meeting its goals (e.g., enabling priority user stories listed in Exhibit CENH-1.).

19 I recommend that the Commission direct for the design and operation of the Data
20 Platform in such a way that allows for a diversity of Data Sources (beyond regulated electric and
21 gas utilities) to provide data through the Platform, so long as they are in compliance with the
22 logical data model and other standards.

1 I recommend creating a performance-based regulation framework for compensating
2 utilities if their participation in the Platform meets performance metrics, which could include:
3 Platform traffic or amount of use; timely responsiveness to requests for third-party authorization
4 and data delivery; prevalence of reported bugs (and ability to fix quickly); ability to keep the API
5 services online with minimal downtime; or other metrics. These metrics should all be
6 quantifiable through automated means (such as logging server activity, bug-tracking systems,
7 etc.) so that up-to-date metrics are available on a monthly basis.

8 I recommend the Commission include in its ruling a listing of user stories to capture the
9 intended outcomes and establish clear, testable goals for what the Data Platform will enable (i.e.,
10 if any of the required user stories are not made possible by the Data Platform, the Data Platform
11 is not accomplishing its intended purpose). (See Exhibit CENH-1.)

12

Exhibit List

1
2
3

Exhibit CENH-1

User Stories Narratives