Standard Energy Efficiency Data (SEED) Platform: A Tool to Track and Transact Energy Data

Cody Taylor, U.S. Department of Energy
Paul Mathew, Lawrence Berkeley National Laboratory
George Hernandez, Pacific Northwest National Laboratory
Andrea Mercado, Lawrence Berkeley National Laboratory

ABSTRACT

Many state and local governments in the U.S. and private building owners are using or are desiring to use energy benchmarking to track the performance of their buildings, and would also like to collect and store program data that is specific to the buildings. Several jurisdictions have also instituted requirements that privately owned buildings must benchmark their buildings and disclose the results. A common feature of these benchmarking and disclosure laws is the disclosure of some building energy benchmarking data - to the jurisdiction, transaction partners, or the public. Jurisdictions passing benchmarking and disclosure laws need a way to effectively manage the data disclosed as a result of these laws, and track compliance.

The Department of Energy is creating an open source software platform for managing the process of collecting, storing, using, and sharing this building energy efficiency energy related data. The software will allow jurisdictions to easily import data from Portfolio Manager, store it according to a standardized data taxonomy and database structure, conduct basic reporting from the data, and make the data available to others. This will reduce the burden of data management, and allow third parties to build data analysis applications that function across multiple jurisdictions via a common data standard.

This paper discusses the value of standardization of building energy performance data. It describes SEED as a common platform, and the potential use of SEED by jurisdictions and building owners to manage and exchange data with Portfolio Manager, DOE's Building Performance Database, or other systems.

Introduction

Many entities wish to track, analyze, and compare data about the energy performance of buildings. This is done to meet a number of needs, including:

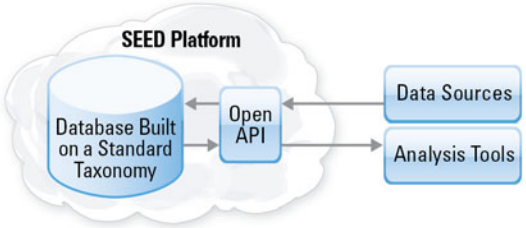
- tracking a single building's energy performance over time
- benchmarking energy performance against peers
- evaluating energy efficiency retrofit performance in a single building
- assessing opportunity for energy savings across a portfolio of buildings
- analyzing the typical impact of a particular energy savings measure for a given building type
- evaluating the persistence of savings from a given energy efficiency measure
- evaluating the cost effectiveness of an energy efficiency program
- forecasting load in a particular geographic region

Organizations undertaking these types of analyses have collected data in a variety of formats, using a variety of data definitions. Some differences in data collection are appropriate, while others may unnecessarily make it difficult or impossible to compare similar datasets from different programs, regions, or time periods. The energy efficiency stakeholder community can make more effective use of the data it already collects if it can agree to collect these data according to common data definitions, store them in a common data structure, and transact them in a common format.

Definition: The term "building energy performance data" is used in this paper to represent data that can be used to describe the energy performance of a building, a building system, or a collection of buildings. This can be subdivided into two primary categories, the usage data (e.g. therms of natural gas used in a given period, peak electric demand, etc.) and building characteristics data (e.g. building size, type, location, device descriptions, schedules, etc). Definitions for IT-specific terms can be found at the end of this paper.

The Department of Energy (DOE) is creating an open source software platform for managing the process of collecting, storing, using, and sharing building energy performance data. This software – the Standard Energy Efficiency Data (SEED) Platform - will allow a variety of users to store building energy performance data in a common structure, but within their own separate database implementations. SEED allows users to easily incorporate data on building benchmarking, energy audits, Demand Side Management (DSM) Program metrics or retrocommissioning, and store it in their own private databases according to a standardized data taxonomy and database structure. SEED also allows users to conduct basic reporting from the data and make selected data available to others via a standardized Application Programming Interface (API). This will reduce the burden of data management, and allow third parties to build data analysis applications that function across multiple jurisdictions via a common API. This paper describes the value of such software and its current execution by DOE.

Figure 1: SEED Overview



Background

There are a variety of entities that track, analyze, and compare data about the energy performance of buildings. These entities include building owners and managers, utilities, state and local governments, federal agencies, academic researchers, and vendors of products and services. Many of these entities collect building energy performance data today to meet their own needs. Often they have reasons to aggregate these data with similar data collected by others. For example, program managers may wish to compare the building stock in their territory to the building stock elsewhere in the country to understand whether an energy efficiency measure that has been effective elsewhere will be effective in their territory. However,

these comparisons can often be difficult because when entities collect building energy performance data, they often do so in ways that are dissimilar enough to hinder comparison between datasets.

"Energy Benchmarking" is a standardized way to measure building efficiency using a tool like ENERGY STAR Portfolio Manager to track a building's energy performance over time or compare it to other buildings. Many state and local governments in the U.S. and private building owners are using energy benchmarking to track the performance of their buildings. Seven jurisdictions have also instituted requirements that privately owned commercial buildings above a size threshold must benchmark their buildings and disclose the results (IMT 2012). Many others have voluntary programs to benchmark residential and commercial buildings. A common feature of these benchmarking and disclosure laws is the disclosure of some building energy benchmarking data - to the jurisdiction, transaction partners, or the public. Jurisdictions passing benchmarking and disclosure laws need a way to effectively manage the data disclosed as a result of these laws and track compliance.

DOE has created a data taxonomy for SEED that can be used to define building energy performance data. This taxonomy was created as a result of the process of designing the DOE Buildings Performance Database (BPD) (DOE 2011) and is the result of synthesizing the data structures of more than six existing databases, as well as review and consideration of over a dozen related efforts such as the Industry Foundation Classes (IFC 2012), OmniClass systems, ASTM BEPA checklist (ASTM 2011), ASHRAE Audit Procedures Checklist (ASHRAE 2011), etc. Where appropriate, data fields and enumerated types from these taxonomies were mapped into the BPD taxonomy. The BPD taxonomy also underwent a formal public review via a Request for Information (RFI) published in the Federal Register. This taxonomy represents an attempt to create a rationalized structure for building energy performance data that can accommodate a wide range of use cases requiring analysis of empirical building performance data. This data taxonomy is likely to evolve over time as it is used, and provides a reasonable starting point for a common national structure for storing building energy performance data.

Even with a common taxonomy and structure for storing building energy performance data, there will still be a need for adaptation when combining different data sets. Building energy performance data are collected for a variety of reasons by a variety of different entities, each of whom may have different data collection methods, validation mechanisms, and tolerance for error. Thus, the use of a common taxonomy does not mean that all data collected according to that taxonomy will necessarily be appropriate for all applications.

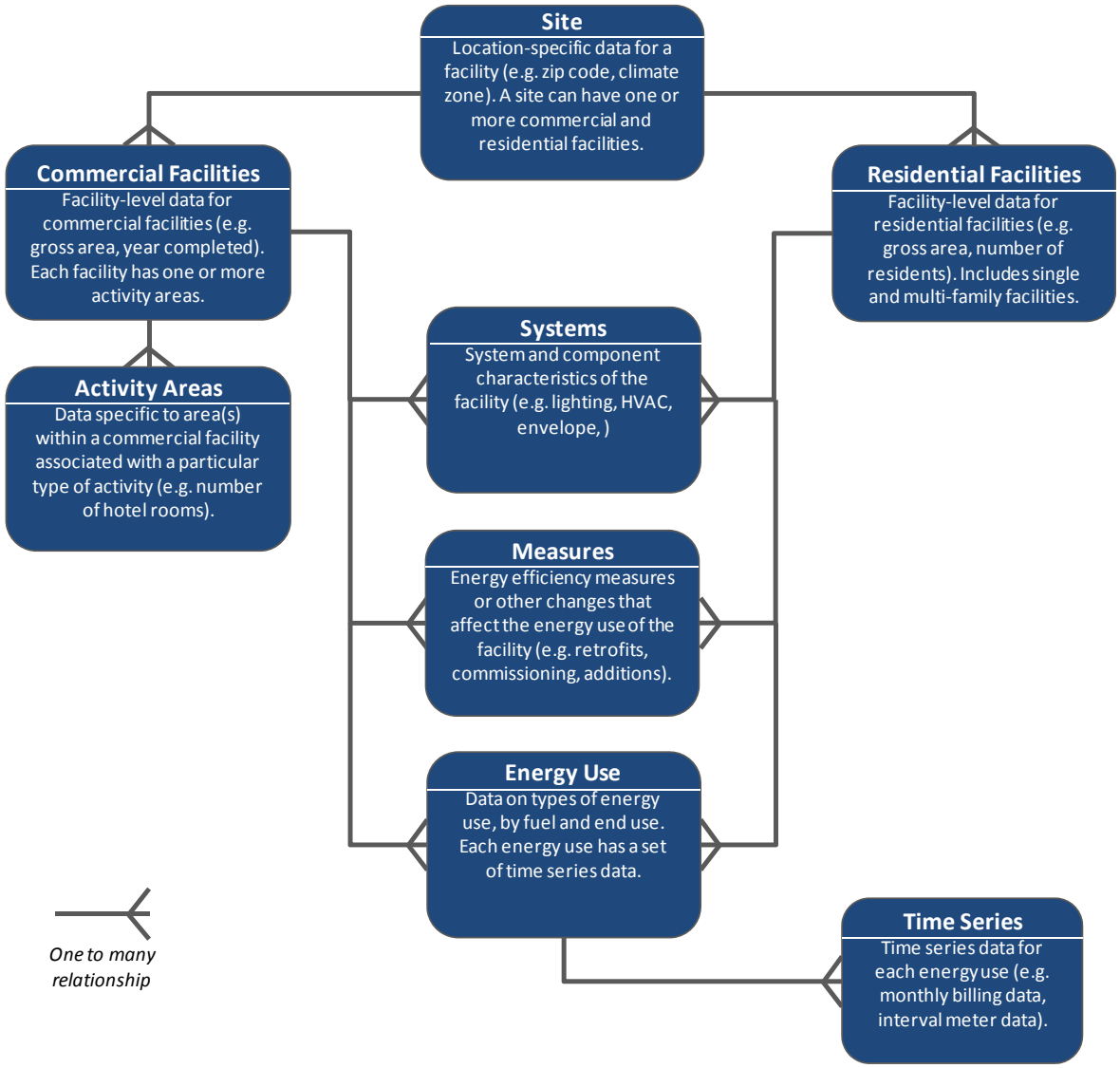
Approach

DOE began with the data taxonomy created for the DOE Buildings Performance Database. As explained above, this taxonomy was designed to support the integration of data from many different sources into one database. This taxonomy consists of a data model and data dictionary. The data model describes the logical relationships among entities and fields in the taxonomy, and the data dictionary offers definitions of each entity and field. These definitions are key to ensuring that various organizations using the taxonomy all have a common understanding of its terms, so that data can be comparable.

For the SEED effort, DOE expanded this taxonomy to include additional fields appropriate for typical building energy performance tracking activities such as energy audits, energy benchmarking, DSM Program metrics and retrocommissioning. The data model is

designed from a top-down perspective, currently describing the top levels of a hierarchical approach to building energy characteristics. We anticipate that the SEED taxonomy will evolve to include more detailed description of building characteristics over time. Examples of such characteristics include system efficiency, component efficiency, and component descriptors such as ballast factor or SEER. The taxonomy also contains structure to describe time-series energy usage data.

Figure 2: SEED Data Model



Note: This figure describes the data model at the highest level, noting that a site is made up of facilities (residential and/or commercial), which have activity areas, systems, energy efficiency measures, and energy usage data associated with them.

With the taxonomy developed, DOE then developed a database application – SEED – that is structured to capture data according to the taxonomy. The SEED Platform is an open source, MySQL-based database application built to allow anyone to store building energy performance data according to the same common taxonomy, facilitating the comparison, sharing,

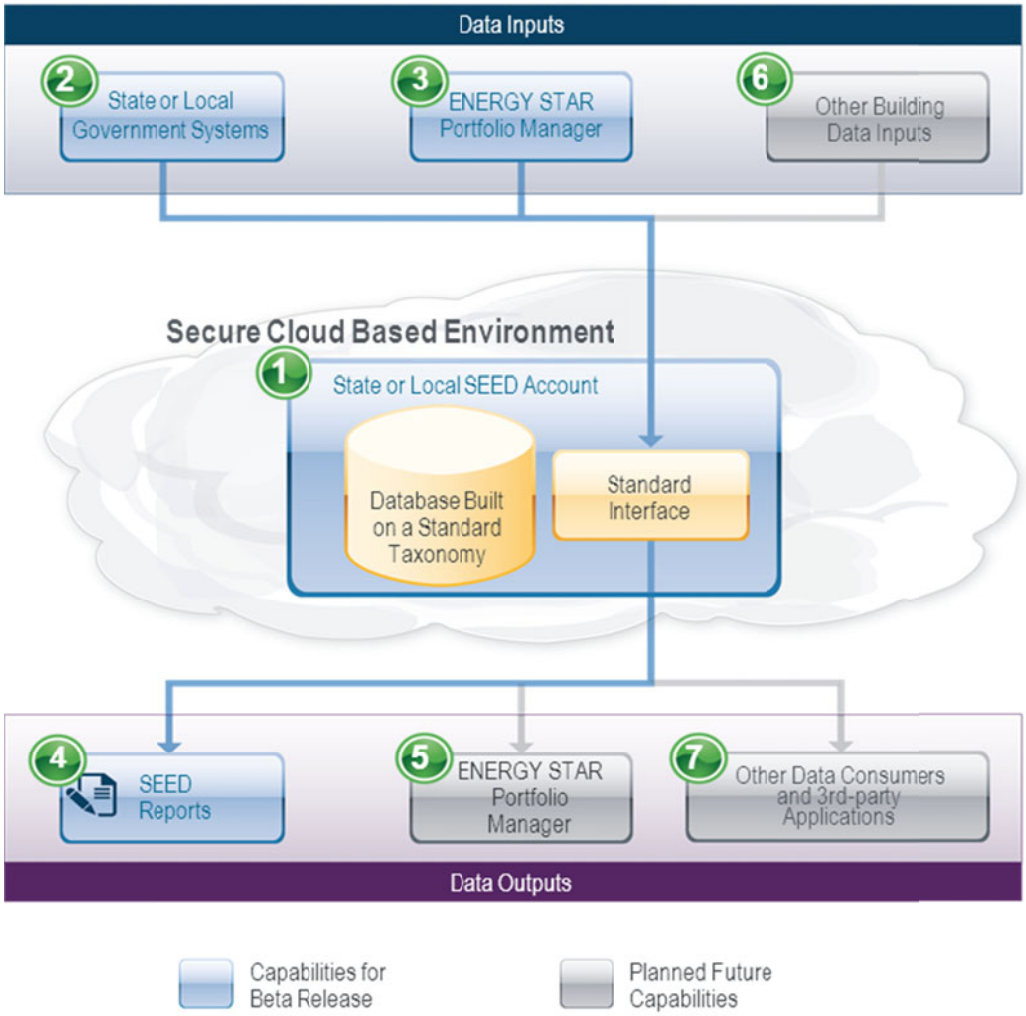
and aggregations of such data. SEED is designed to run in a cloud environment, minimizing the time it takes an organization to install and start using it. It can also be adapted to run on an organization's internal IT infrastructure. SEED is built with basic reporting features so that users can create customized reports from the data in their database.

In addition to the database structure itself, the SEED application contains built-in functionality to facilitate data import and export. SEED is structured to access ENERGY STAR Portfolio Manager – a commonly used application for benchmarking the energy performance of commercial buildings – to retrieve data for many buildings at once through Portfolio Manager's Automated Benchmarking Service (ABS) function or its custom reports function. SEED can also import data from an uploaded CSV file, and can be configured to collect data directly via custom web forms. Because the DOE's Buildings Performance Database is built around a subset of the same data taxonomy, SEED data can also be exported for easy inclusion in the national database.

The SEED platform is designed to significantly reduce the time and resources required by jurisdictions to implement a database for disclosure requirements. This is especially valuable for many smaller jurisdictions such as cities, whose limited IT personnel and budgets may preclude the development of custom databases from scratch. Creating and populating a new instance of a database linked to ENERGYSTAR Portfolio Manager (ESPM) involves the following steps:

1. Install a new instance of SEED in a commercially available server cloud
2. Import basic building data into SEED from existing sources such as tax records, etc., and assign each building an ID.
    a. Provide each building owners with their ID and ask them to enter this ID into their ESPM account
3. Download data from ESPM. (SEED will use the unique building ID to match buildings in SEED with the data from ESPM)
4. Use SEED reporting capabilities to track compliance

Figure 3: SEED Concept of Operations



Finally, SEED is being designed with an open API to facilitate data exchange. The API will allow the administrator of a SEED database to make select portions of the database accessible via the internet for other software to query. This can allow secure exchange of data within an organization. It can also allow organizations that wish to publish building energy performance data to do so in a standardized way that can be accessed programmatically by other software. This allows jurisdictions that wish to publish the energy performance data of buildings to do so in a structured format. Rather than individual jurisdictions posting data files on their website in their own formats, they can use SEED to publish full or partial data sets online via a common API. If a number of jurisdictions choose to publish building energy performance data via the same API, it becomes much easier for third party applications to aggregate this open data and use it in applications that deliver analysis or visualization to individuals who need it, at the time and in the way that is most compelling.

DOE is distributing SEED as a free and open source application to allow anyone to use it and modify as necessary. This does not necessarily dilute the standardization of the underlying data structure. While the data structure could be changed, most users are likely to leave the structure in place, facilitating data exchange with other parties. DOE expects third parties to

develop additional tools for data collection and analysis to extend the base application. For example, this might include an easy web interface for collecting a specific subset of data import to a particular organization and using language familiar to members of that organization, or the creation of enhanced analysis and reporting tools that allow users to draw conclusions from the data more easily.

Results and Discussion

DOE began developing SEED in late 2011 and launched a beta version of SEED in February 2012. Initial users included several state and local governments who are collecting building energy performance data for public and private buildings. These jurisdictions are using ENERGY STAR Portfolio Manager to collect benchmarking data for many buildings, and then importing the resulting data en masse to their own local SEED databases for analysis. Some are also building customized web forms linked directly to their own SEED database to collect data on building energy audits.

DOE expects to continue development of SEED through spring and summer 2012. The API will be developed during this time, and results of its use will be reported at a later date.

Because SEED users will be storing their data in a format that is closely aligned to the DOE Buildings Performance Database (BPD) – a national repository for empirical data and associated analysis tools – it will be easy for SEED users to include their data in the national database. The BPD allows users to access building energy performance data for a large number of buildings and perform various analyses of these buildings using associated tools. Any organization may submit its data to the BPD and perform analysis of its data in the context of the larger national dataset. This requires the owner of the data to work with DOE to map that organization's data into the BPD data taxonomy. However, for an organization that uses SEED to track building energy performance data, the data will already be in a format compatible with the BPD, and the process of using it with the BPD is thus much simpler because no new mapping is required.

DOE has developed SEED as open source software and expects to have a diminishing role in making continued upgrades after 2012. As the user base grows in 2013 and beyond, DOE intends to reduce its involvement. It is expected that the user community will collaborate to fund the development of desired enhancements to the SEED Platform. DOE expects that the taxonomy will also continue to evolve, and is investigating the best way to engage stakeholders in making any updates to the taxonomy so that it best serves the needs of all and facilitates energy efficiency. DOE expects to update the data structure of its own national Buildings Performance Database in parallel with the evolution of the SEED taxonomy.

Challenges

Even though the SEED taxonomy has a clear set of definitions for all data fields, the existence of these standard definitions doesn't necessarily mean they will be adhered to by all users in all cases. For organizations who wish to use SEED to track and share similar data, there will still be a need for business rules to ensure that they are indeed adhering to these definitions. Organizations who wish to share data will also need to clearly describe the methods used in data collection and reliability of the sources to ensure comparability between data collected by different parties or through different programs.

Conclusions

By moving toward the use of a common format for building energy performance data, organizations around the country will be better able to use their data in new ways. Balkanized data sets in different formats may be individually useful, but the transaction costs involved in mapping them to each other can inhibit comparisons. However, if disparate data sets use common definitions and are stored in a common format, it becomes much easier to join them for analysis.

The creation of the SEED Platform is a step toward widespread use of a common national format for storing building energy performance data. Organizations who wish to use and improve SEED are encouraged to collaborate with DOE to do so.

Acknowledgments

The authors would like to thank the SEED beta testers who have provided valuable feedback in the development of SEED. This work was supported by the Office of Energy Efficiency and Renewable Energy, Building Technologies Program, of the U.S. Department of Energy.

Glossary

Application programming interface (API) – a standard set of methods for different pieces of software to communicate with each-other.

Cloud environment – used here to refer to the fact that the software can be run on one of the services in which private vendors "rent" access to servers, rather than forcing a user to buy their own hardware and install software there. This can significantly reduce the up-front time, skill, and cost needed to begin using the software.

Open source software – computer software that is available in source code form, under a license that permits users to study, change, improve and at times also to distribute the software.

Taxonomy – a system of classification, with multiple classes arranged in a clear structure

References

ASHRAE 2011. Procedures for Commercial Building Energy Audits. ASHRAE. 2011.

ASTM 2011. Standard Practice for Building Energy performance Assessment for a Building Involved in a Real Estate Transaction. ASTM International. 2011.

DOE 2011. The DOE Building Performance Database. http://www.buildingsperformance.net/

IMT 2012. Rating and Disclosing the Energy Performance of Buildings: A Market-Based Solution to Unlock Commercial Energy Efficiency Opportunities. White paper. Institute for Market Transformation. http://www.imt.org/files/FileUpload/files/Benchmark/IMT_Rating_Policy_White_Paper.pdf. Accessed March 2012.

IFC 2012. Industry Foundation Classes. BuildingSmart. http://buildingsmart.com/standards/ifc. Accessed March 2012.

Omniclass 2012. OmniClass Construction Classification System. http://www.omniclass.org/