Cleaning, Curation and Communication of Home Metadata

Emma Casavant, and Katherine Bassett, *Graduate Researcher, Northeastern University*
Zoe Chappell, REU student, Middlesex Community College
Dr. Michael Kane, *Department of Civil and Environmental Engineering, Northeastern University*

Northeastern University
College of Engineering

Northeastern University
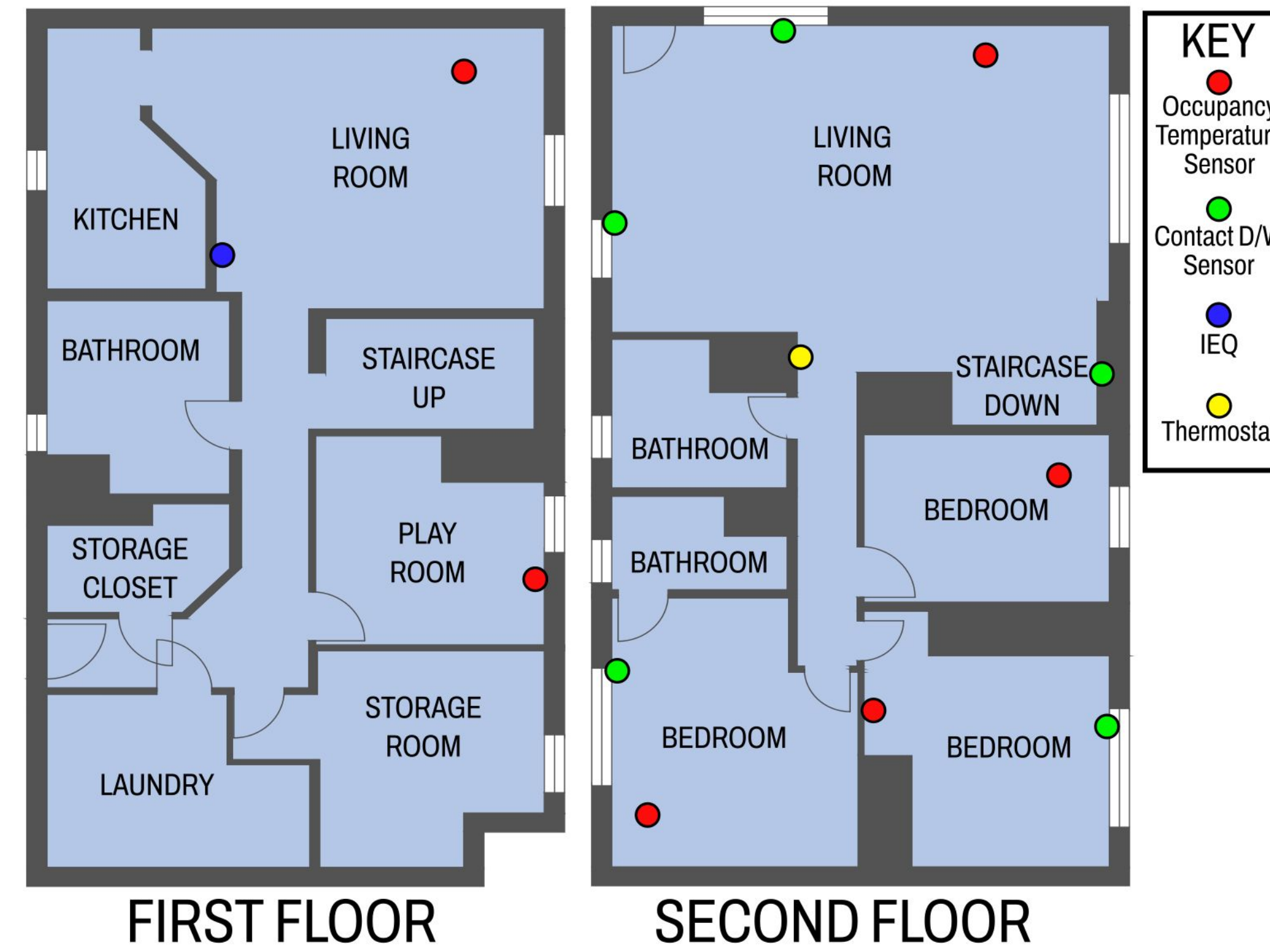**Center for STEM Education**

Abstract

The purpose of this project is to clean, curate and communicate the metadata collected on houses for the ABLE Lab 'Whole Energy Homes' study; a study dedicated to making an open data set on the interaction home occupants have with their thermostats and how this affects thermal comfort. The processes used in this project includes; editing CAD designs of home layouts into a format that more clearly represents the layouts of each home, including the locations of windows and doors, determining if parts of data sets are redundant, and organizing the collected home metadata into easily readable tables. This project's focus was on the ten homes located in Massachusetts. The results of this process were as follows, ten CAD layouts of homes, and a set of four tables. These tables identify the number of, and types of sensors in each home, various measurements of physical qualities of the homes, the number and type of occupant, and information regarding secondary methods of effecting home temperatures, such as additional cooling systems, or the existence and location of a dehumidifier.

Background

The figure to the left is a symbolic representation of much of the studies metadata, the physical home layouts, as well as the physical locations of the Sensors used to gather data.

- The purpose of the '**Whole Energy Homes**' study is to create an open dataset to aid fellow future researchers. This is a useful steppingstone in reducing climate change and switching to renewable energy sources.
- Since there is so much raw data, it must be organized and cleaned so that it is in a usable more accessible state, allowing for researchers to quickly recognize what types of data were gathered, and how they were gathered.
- In '**Cleaning, Curation and Communication of Home Metadata**', 'Metadata' refers to data that helps contextualize a wider data set. The curation of metadata is an important step in the data organization process, as it helps provide a framework upon which more specialized data can be categorized and organized.

Methodology

AutoCAD

A series of floor layouts with completed walls that do not overlap, that line up with each other at 90 degree angles, a 12" thick outer wall, with windows represented by open rectangles with a line through them.

After loading the file in AutoCAD and consulting a provided floorplan, with additional annotations showing the location of other unlabeled rooms such as closets, guidelines were applied to ensure each wall was on a 90-degree axis unless obviously not applicable. Walls were straightened, additional rooms shown in the annotation were added. Any walls that intersected with each other were shifted so they no longer did so, and any gaps left in the walls were closed. A 12" outer wall was added, as well as the rectangular windows and any other doors such as garage doors or sliding doors.
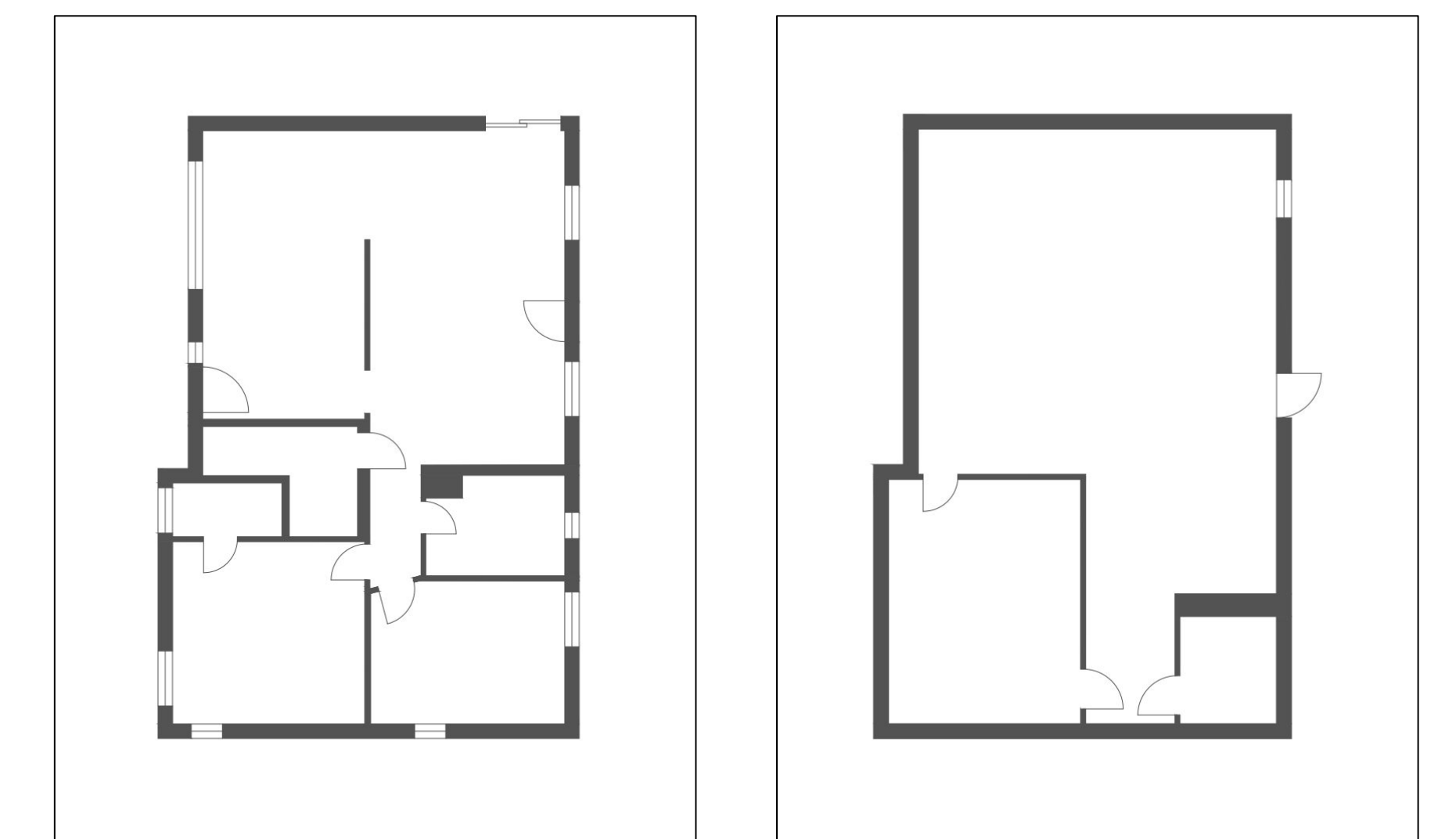
MA05 CAD layout.
Upper left: Basement
Upper right: Entrance Floor
Left: Garage

Methodology .Cont

Inkscape

A layout that fits within a phone screen, with the advent of scrolling, that shows all floors of a home, and indicates the areas of the home that an occupant is currently located in, while labeling all major rooms in each home

Inkscape image is created to fit on a phone screen, 780 x 1200 px. The height was able to be adjusted in cases where a home would not fit at a reasonable size, (*see house 10 layout*). Each floor of a home was imported into the program and positioned and sized to be easily visible and navigable. As each floor was its own layer they were re-named accordingly, so that both the '**Label**' and '**ID**' were the same, and all were placed on a parent layer named, '**FLOORPLAN**'. From there a '**BASE**' layer was created.

Rooms without sensors (*bathrooms, entry halls ect.*) were labeled and colored in grey. Rooms with sensors were outlined in blue. Button layers were created for these rooms, named '**ROOM_Button_On/Off**' that could be used to indicate if a room was 'selected' or not, and each room was labeled. All layer names reflected what they contained. This was done so that the Inkscape images could be uploaded into an app designed for gathering further data in the study using surveys.

Inkscape Drawing of MA10, where the 'Kitchen' is selected, and all other selectable rooms are in the 'off' position

Data Curation and Table Formation

Data Tables

A set of four tables. These tables identify the number of, and types of sensors in each home, various measurements of physical qualities of the homes, the number and type of occupant, and information regarding secondary methods of effecting home temperatures, such as additional cooling systems, or the existence and location of a dehumidifier.

Data was organized into 4 tables per state, based on information that was deemed as important for research. The common thread between the tables, is that they are all variables that an individual wishing to use the dataset would most likely want or need to know. As each house's *primary* method of heating was a **central furnace** and cooling was **central air conditioning**, this information was not specified on the tables.

MA Home Info							
House	Control Zones (Thermostats)	Floors	Age of Home (years)	Floor Area (Ft^2)	ELA (In^2)	Conditioned Area (Ft^3)	Elevation (Ft)
MA01	1	2	72	2,054	213.1	15,452.50	108
MA02	1	2	36	2,050	96	15,374.30	541
MA03	3	3	21	2,051	102.7	17,384.30	213
MA04	1	2	5	1,612	42.6	12,583.20	623
MA05	1	2	35	2,285	82.8	16,673.60	495
MA06	2	3	20	1,746	76.7	12,549.90	72
MA07	2	2	122	2,042	139.2	14,930.30	
MA08	1	3	26	2,882	104.1	22,683.10	
MA09	1	4	120	2,592	391.7	18,834.30	
MA10	3	3	40	6,164	267.9	46,517.90	30

Number of Sensors in MA Homes					
House	Occupancy	Humidity	Temperature	Contact Sensors	Smart Plugs
MA01	5	2	6	5	2
MA02	7	2	8	5	2
MA03	9	2	10	7	4
MA04	4	2	5	5	1
MA05	8	2	9	4	5
MA06	7	1	7	6	1
MA07	6	1	7	4	4
MA08	11	1	12	3	2
MA09	11	0	12	6	4
MA10	14	0	15	13	5

Occupant Info			
House	Full Time Occupants	Part Time Occupants	Pets
MA01	3	1	NA
MA02	5	NA	NA
MA03	5	NA	NA
MA04	4	NA	NA
MA05	5	NA	4
MA06	4	NA	NA
MA07	2	NA	NA
MA08	5	NA	1
MA09	4	2	1
MA10	5	NA	3

MA Home Info (Heating)					
House	Secondary Heating	Secondary Cooling	Humidifier	Dehumidifier	Energy Type
MA01	Portable electric heater (basement), portable heater (office)	Ceiling fan (primary bedroom), standing fan (bedroom)	Portable humidifier	Dehumidifier (basement)	Fuel Oil
MA02	NA	NA	Portable humidifier (kid's room)	Dehumidifier (basement)	Fuel Oil
MA03	Baseboard electric heater (basement); wood burning stove (never used); electric portable heater (never used)	NA	Portable humidifier	Dehumidifier (basement)	Natural Gas
MA04	NA	NA	NA	NA	Bottled Gas
MA05	Portable electric heater (basement)	Ceiling fans (bedrooms and living/ dining room; used in summer and winter)	Portable humidifier	Dehumidifier (basement)	Natural Gas
MA06	NA	NA	NA	NA	Natural Gas
MA07	Portable electric heater (upstairs office)	Window box fans (upstairs); 2 window AC units (have but have not used in bedroom)	Portable humidifier	Dehumidifier (basement)	Natural Gas
MA08	Previously used portable electric heater in conditioned attic, have not used since having the roof insulated	Window AC unit (finished attic)	NA	Dehumidifier (basement)	Natural Gas
MA09	Gas fireplace (family room)	Ceiling fans (bedrooms), two window AC units (attic bedroom),	NA	NA	Natural Gas
MA10	Portable electric heater (in-law suite, upstairs office)	Ceiling fans	NA	Dehumidifier (basement)	Fuel Oil

A) The type and number of different sensors in each home, Occupancy, Temperature, Humidity, Door/Window contact sensors, and SmartPlugs that measure the usage of power throughout the home.

B) Physical information about the home such as its age, number of floors, elevation, and floorspace.

C) Data about how many full time, and part time occupants each home has, as well as any additional pets.

D) Information about additional methods of climate control in each home not already detailed.

Four tables, their contents is specified in section A-D.

A) Number of Sensors in MA Homes

B) MA Home Info

C) Occupant Info

D) MA Home Info (Heating)

Acknowledgements

NSF REU-Site: Research Experience for Undergraduates site to create "Pathways Opening World Energy Resources" (REU - POWER)

This work was supported by the National Science Foundation Grant ##2150417