Energy Analytics: Identifying Anomalies in Campus Energy Usage Data

Miguel Alvarez, Dominic Enriquez, Henry Koelling,
Minh Pham, Ben Preiss, Nacho Rodriguez-Cortes

Comps 2022
Carleton College

Overview

- Introduction
- Energy Analytics 4
- Conclusion

Overview

- **Introduction**
  - Motivation
  - Key Terms
  - Previous Work
- Energy Analytics 4
- Conclusion

Introduction

- Motivation
- Key Terms
- Previous Work

Introduction

- **Motivation**
  - Goal
  - Impact
- Key Terms
- Previous Work

Goal

- **Task**: Develop a web-tool for identifying anomalous data in Carleton's energy system
  - Long term goal of EA Comps
- Importance & relevance of the project
  - Carleton's 2011 Climate Action Plan
    - Net zero GHG emissions by 2050
    - "Energy audits, Green IT" (2010-2025)
  - Financial consequences
    - Wasted energy costs
    - Money that isn't going 'green'

Impact

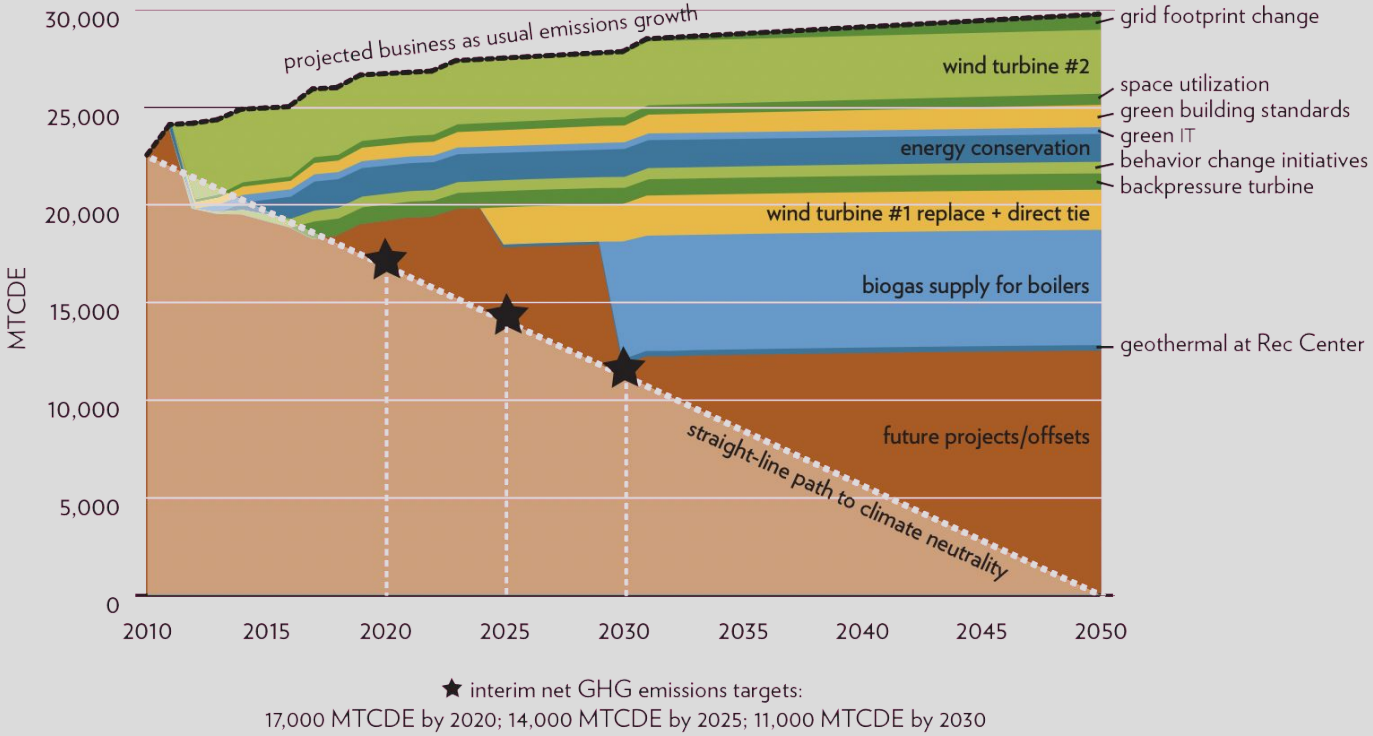


Figure 1: Timeline for GHG emission targets & reductions

Introduction

- Motivation
- **Key Terms**
  - Energy Analytics
  - Points
- Previous Work

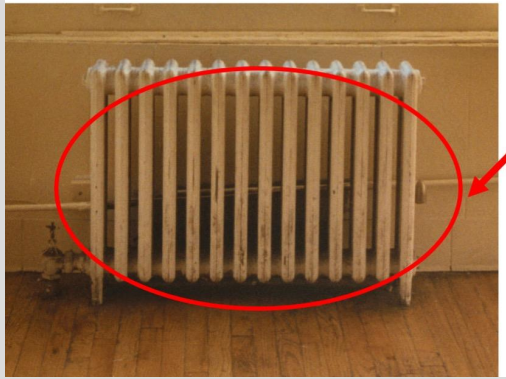
Energy Analytics: definition

- “The process of collecting electrical data and applying sophisticated analytical software and algorithms to deliver insights around consumption and time of use reductions”
- Key factors in *Energy Analytics*
  - Energy used
  - Cost
  - Detecting anomalies (eg, room is too cold)
  - Automation

Points - Devices

Thermostat

VAV Box



Radiators

Air handling unit

Points

- What are they?

Points

- What are they?
    - Room Temperature

Points

- What are they?
  - Room Temperature
  - Room Set Temperature

Points

- What are they?
  - Room Temperature
  - Room Set Temperature
  - Air Vent Percentage Open

Points

- What are they?
  - Room Temperature
  - Room Set Temperature
  - Air Vent Percentage Open
  - Wind Speed

Points

- What are they?
  - Room Temperature
  - Room Set Temperature
  - Air Vent Percentage Open
  - Wind Speed
  - Wind Direction

Points

- What are they?
  - Current Room Temperature
  - Temperature Set Point
  - Air Vent Percentage Open
  - Wind Speed
  - Wind Direction
- Tons of points!

Points

- Naming

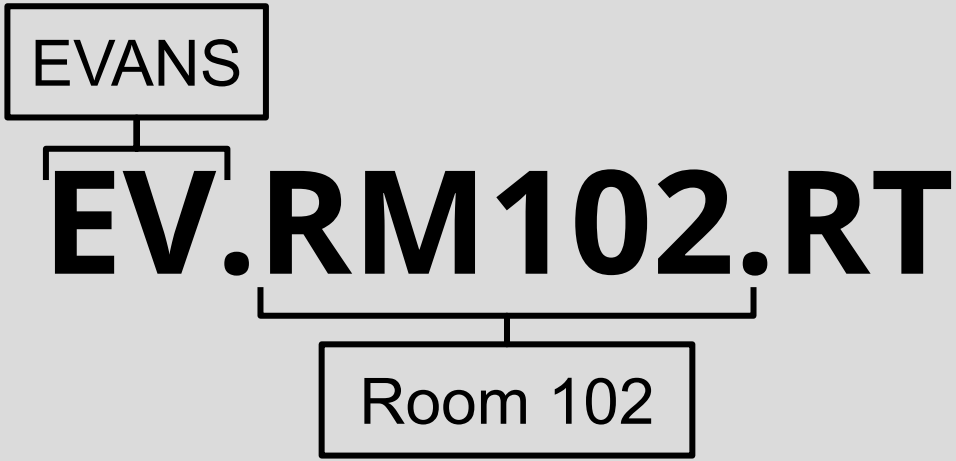
EV.RM102.RT

Points

- Naming

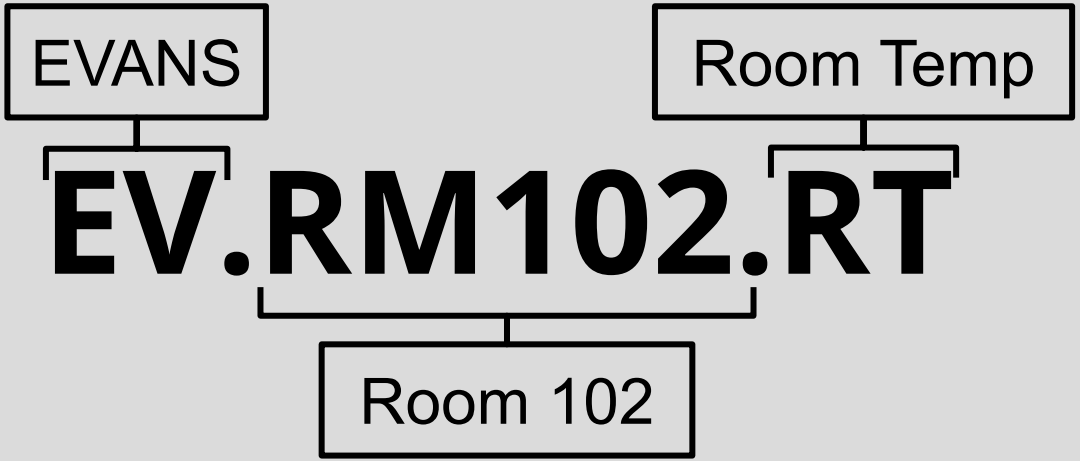
Points

- Naming



Points

- Naming

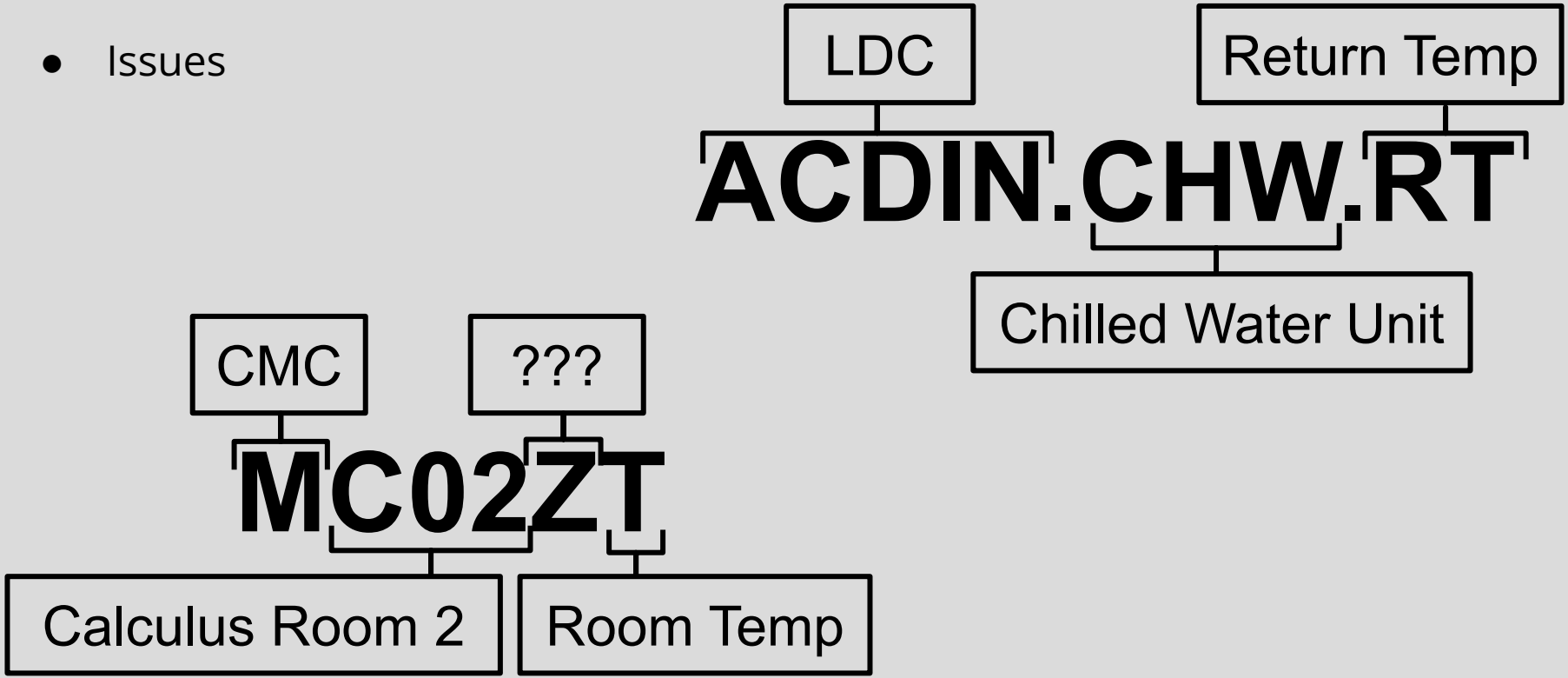


Points

- Issues

Points

- Issues



Introduction

- Motivation
- Key Terms
- **Previous Work**
  - Value pipeline
  - API
  - Point graph
  - Search feature
  - Machine learning algorithms

Previous Work

- **Value Pipeline**
  - Importer script to collect data from points
  - Rube Goldberg machine of automated scripts
  - Data goes straight to us
  - No disrupting production server

Previous Work

- Value Pipeline
- **API**
  - API endpoints to search database

GET Point IDs

http://energycomps.its.carleton.edu/api/points/ids?search=<search_request>

Returns the point_id's that follow the specified search parameters

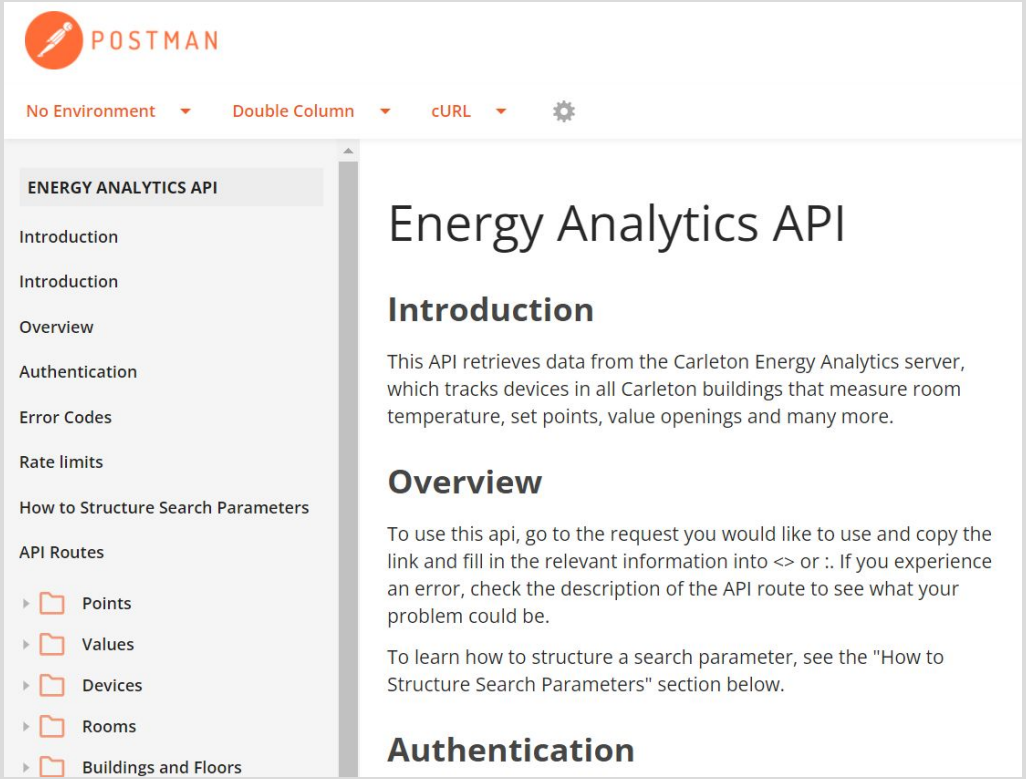
GET Point by ID

http://energycomps.its.carleton.edu/api/point/:id

Gets a point by its point_id

Previous Work

- Value Pipeline
- **API**
  - API endpoints to search database
  - Postman documentation



Previous Work

- Value Pipeline
- API
- **Point Graph**

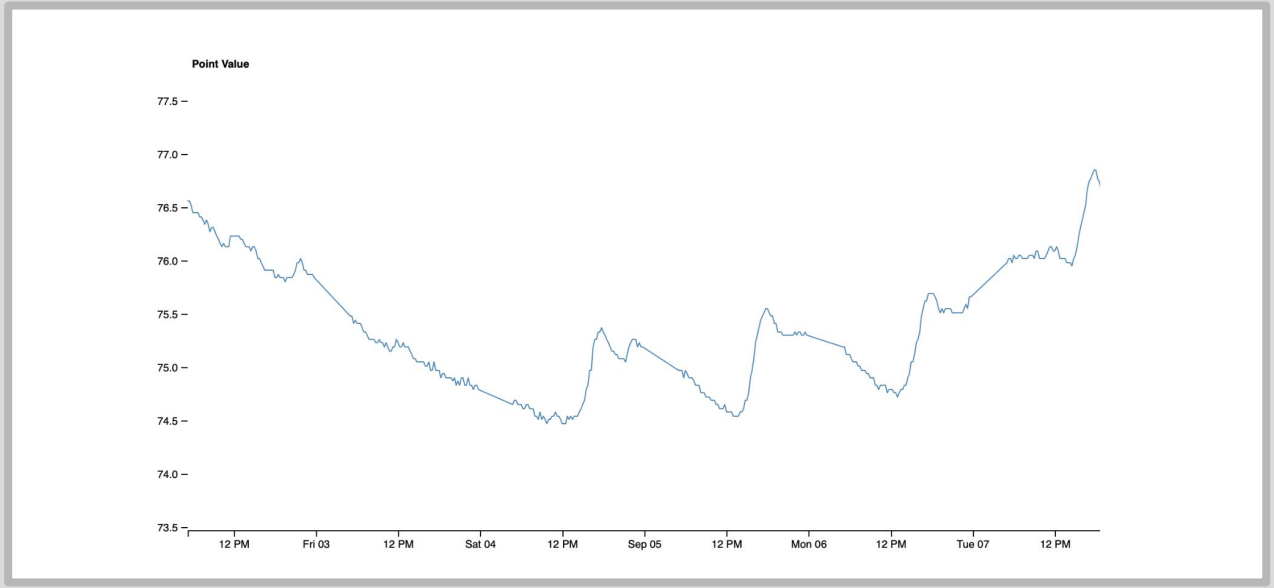
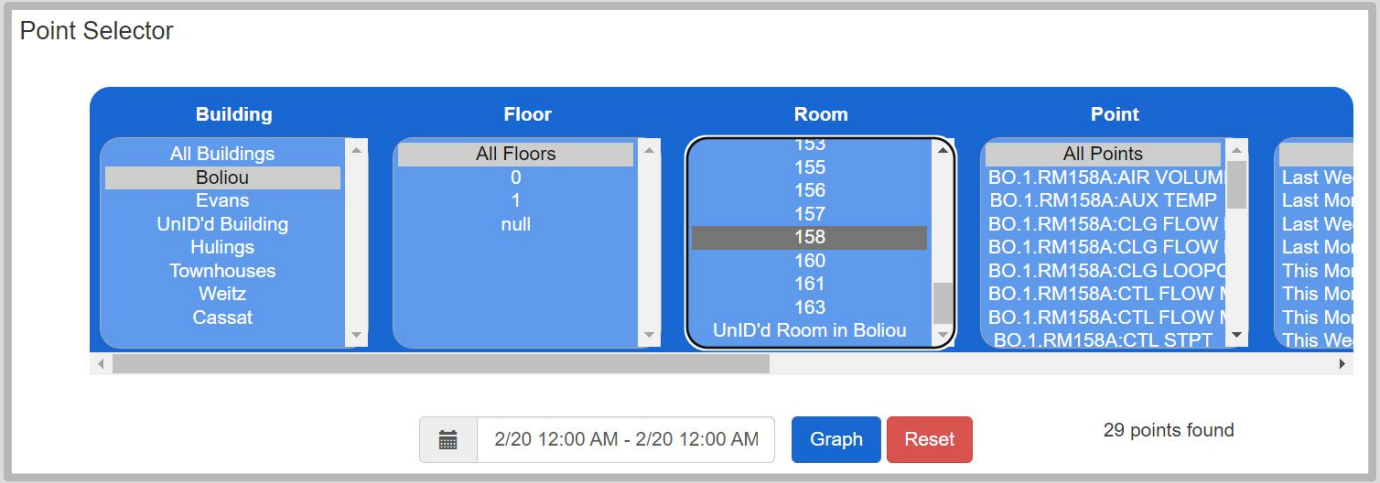


Figure 1: Evans 106 room temperature from Sept 1st to Sept 8th, 2021

Previous Work

- Value Pipeline
- API
- Point Graph
- **Search Feature**



Previous Work

- Value Pipeline
- API
- Point Graph
- Search feature
- **Machine learning algorithms**
  - To find energy usage problems, devices problems, etc.
  - K-Means
  - STL or Seasonal-Trend decomposition using LOESS
  - Isolation Forests

Energy Analytics: at Carleton

- Reduce costs by:
  - Improving energy efficiency
  - Not relying on a third-party to process our data
- Try to be even more environmental friendly
- Carleton’s climate goals from 2030 to 2050 (and their current reliance on “unknown tech”) – Climate action plan (2030 to 2050), Utility master plan (done last year)

Overview

- Introduction
- Energy Analytics 4
- Conclusion

Overview

- Introduction
- **Energy Analytics 4**
  - Energy data
  - Data Management
  - Anomaly algorithms
  - Heuristics approach
  - API
  - Website interface
- Conclusion

Energy Analytics 4

- Energy data

Energy Analytics 4

- Energy data
- Data Management

Energy Analytics 4

- Energy data
- Data Management
- Anomaly algorithms

Energy Analytics 4

- Energy data
- Data Management
- Anomaly algorithms
- Heuristics approach

Energy Analytics 4

- Energy data
- Data Management
- Anomaly algorithms
- Heuristics approach
- API

Energy Analytics 4

- Energy data
- Data Management
- Anomaly algorithms
- Heuristics approach
- API
- Website interface

Energy Analytics 4

- **Energy data**
- Data Management
- Anomaly algorithms
- Heuristics approach
- API
- Website interface

Energy Analytics 4

- **Energy data**
  - What we have
  - Difficulties
- Data Management
- Anomaly algorithms
- Heuristics approach
- API
- Website interface

Carleton's Energy Data

- What we have now:

Carleton's Energy Data

- What we have now:
  - Each building's energy use - hour to hour

Carleton's Energy Data

- What we have now:
  - Each building's energy use - hour to hour
    - Can be useful for large issues

Carleton's Energy Data

- What we have now:
  - Each building's energy use - hour to hour
    - Can be useful for large issues
  - Whole building data visualizations

Carleton's Energy Data

- What we have now:
  - Each building's energy use - hour to hour
    - Can be useful for large issues
  - Whole building data visualizations
- What's hard to do:

Carleton's Energy Data

- What we have now:
  - Each building's energy use - hour to hour
    - Can be useful for large issues
  - Whole building data visualizations
- What's hard to do:
  - Compare buildings side-by-side

Carleton's Energy Data

- What we have now:
  - Each building's energy use - hour to hour
    - Can be useful for large issues
  - Whole building data visualizations
- What's hard to do:
  - Compare buildings side-by-side
  - Identify single rooms with problems

Carleton's Energy Data

- What we have now:
  - Each building's energy use - hour to hour
    - Can be useful for large issues
  - Whole building data visualizations
- What's hard to do:
  - Compare buildings side-by-side
  - Identify single rooms with problems
  - Automate identification of problems

Carleton's Energy Data

- Examples of problems:
  - Finding open windows in the winter

Carleton's Energy Data

- Examples of problems:
  - Finding open windows in the winter
  - Locating a stuck vent

Carleton's Energy Data

- Examples of problems:
  - Finding open windows in the winter
  - Locating a stuck vent
  - Discovering that a floor is leaking energy

Energy Analytics 4

- Energy data
- **Data Management**
  - ALC and Siemens
  - Database
  - Compression and backup
- Anomaly algorithms
- Heuristics approach
- API
- Website interface

Data Management

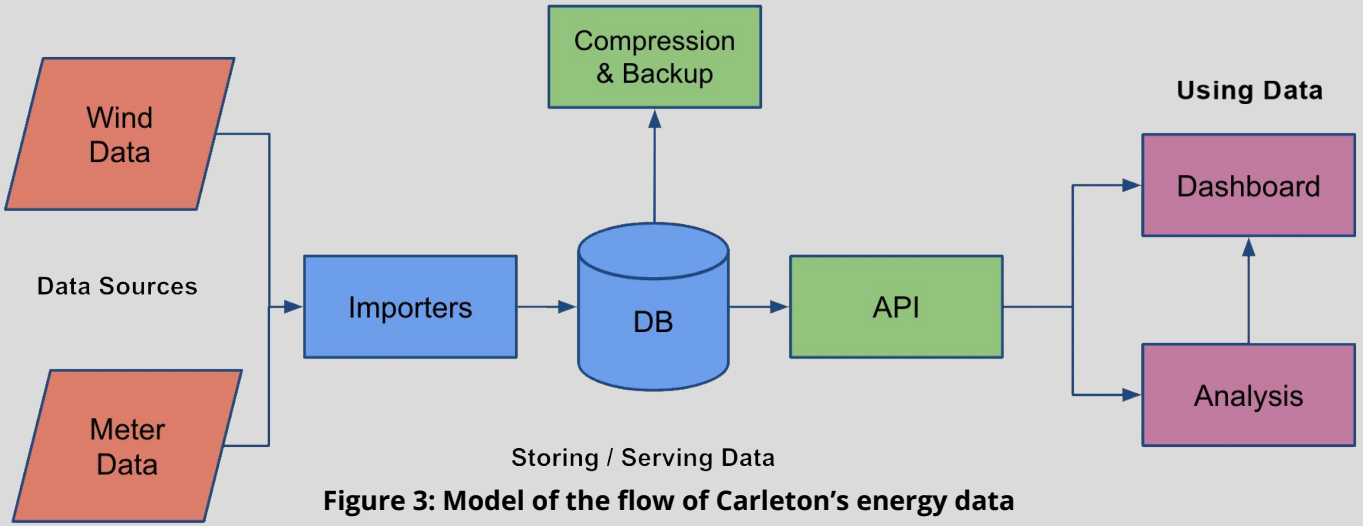


Figure 3: Model of the flow of Carleton's energy data

Data Management

- ALC & Siemens
  - Value pipeline to database

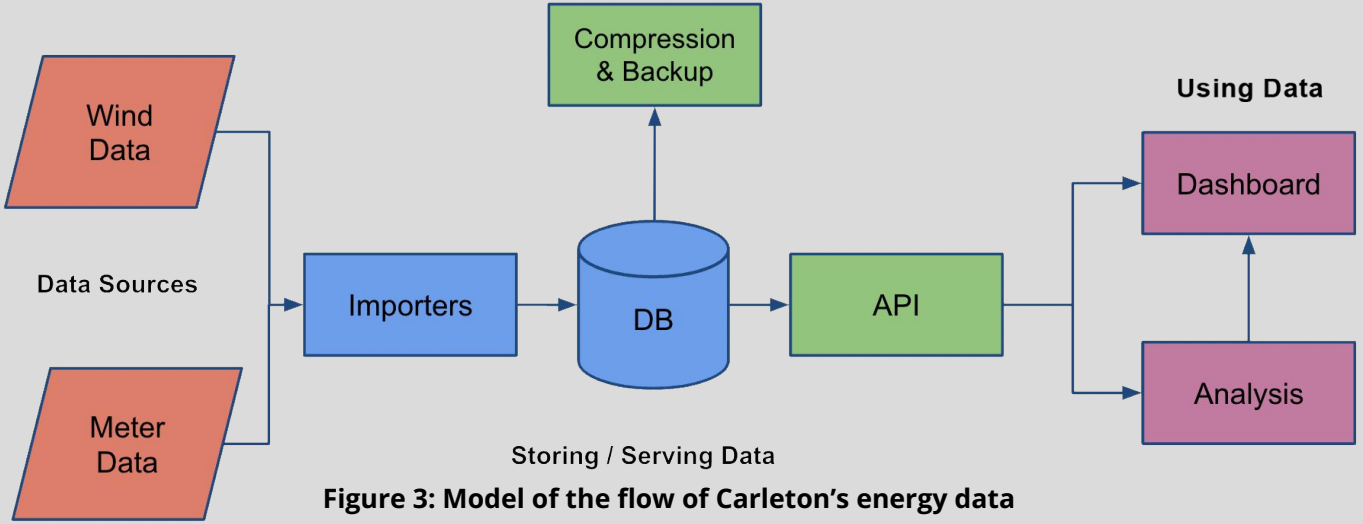


Figure 3: Model of the flow of Carleton's energy data

Data Management

- ALC & Siemens
  - Value pipeline to database
  - Difficulties with ALC data

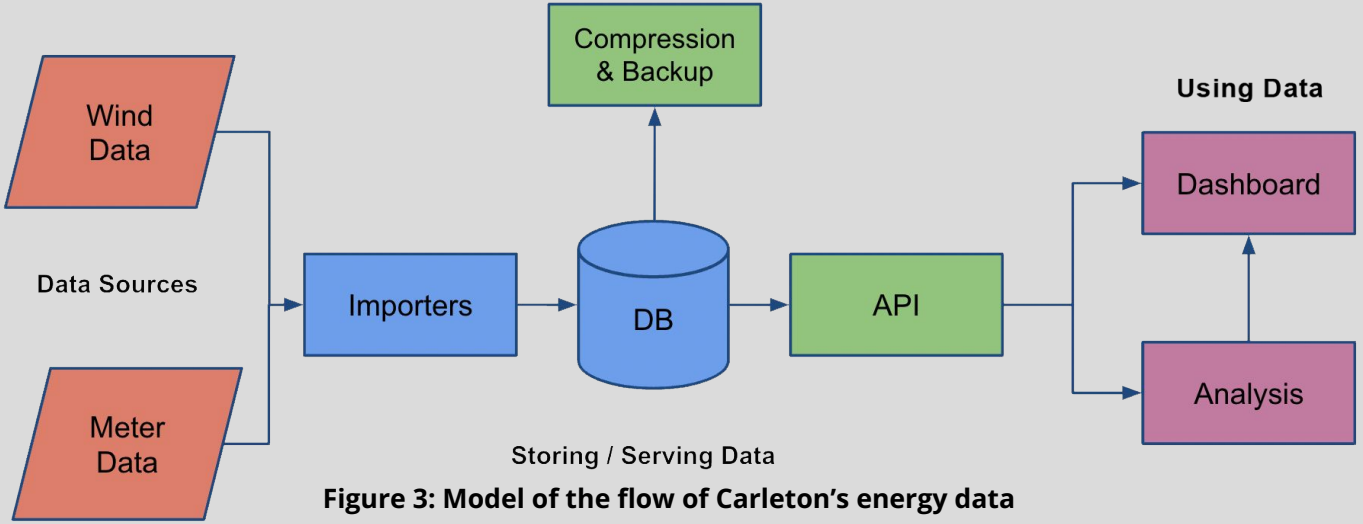


Figure 3: Model of the flow of Carleton's energy data

Data Management

- ALC & Siemens
  - Value pipeline to database
  - Difficulties with ALC data
- Creating an automated compression & backup system for the database
  - .CSV file format clean-up

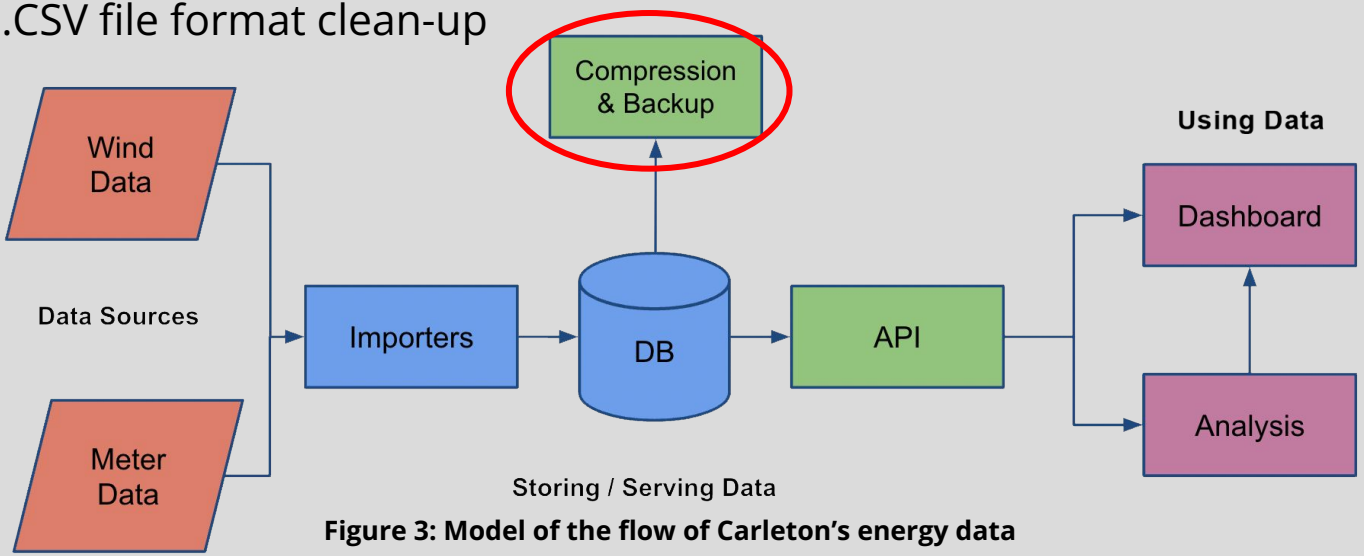
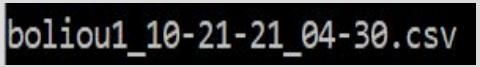


Figure 3: Model of the flow of Carleton's energy data

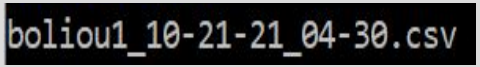
Data Management

- Data was originally all dumped into one place with no organization



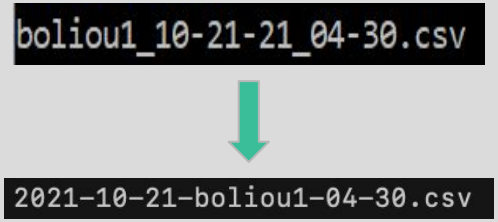
Data Management

- Data was originally all dumped into one place with no organization
    - Naming conventions made the data impossible to sort



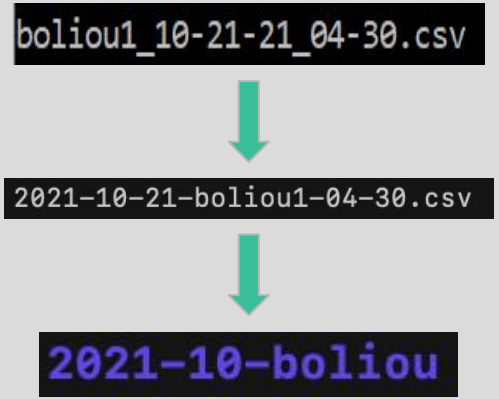
Data Management

- Data was originally all dumped into one place with no organization
    - Naming conventions made the data impossible to sort



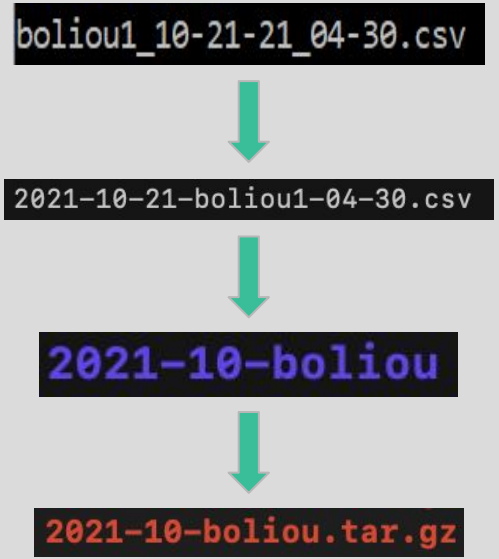
Data Management

- Data was originally all dumped into one place with no organization
  - Naming conventions made the data impossible to sort
- Next group by buildings



Data Management

- Data was originally all dumped into one place with no organization
  - Naming conventions made the data impossible to sort
- Next group by buildings
- Finally, combine into one file and compress



Data Management

- Created 3 bins:

Data Management

- Created 3 bins:
    - Failed to import

Data Management

- Created 3 bins:
  - Failed to import
  - Ready to archive

Data Management

- Created 3 bins:
  - Failed to import
  - Ready to archive
  - Uploads

Data Management

- Created 3 bins:
  - Failed to import
  - Ready to archive
  - Uploads
- Movement through bins automated with a cronjob

Data Management

- Created 3 bins:
  - Failed to import
  - Ready to archive
  - Uploads
- Movement through bins automated with a cronjob
  - Runs bash scripts daily

Data Management

- Created 3 bins:
  - Failed to import
  - Ready to archive
  - Uploads
- Movement through bins automated with a cronjob
  - Runs bash scripts daily
  - Successful processing entailed tar and zip

Data Management

- Created 3 bins:
  - Failed to import
  - Ready to archive
  - Uploads
- Movement through bins automated with a cronjob
  - Runs bash scripts daily
  - Successful processing entailed tar and zip
    - After a month’s collection of data

Energy Analytics 4

- Energy data
- Data Management
- **Anomaly algorithms**
  - K-means
  - STL
- Heuristics approach
- API
- Website interface

Anomaly Algorithms

- K-Means & STL (Seasonal Trend Decomposition with Loess)

Anomaly Algorithms

- K-Means & STL (Seasonal Trend Decomposition with Loess)
  - Two anomaly algorithms previously investigated by EA Comps groups

Anomaly Algorithms

- K-Means & STL (Seasonal Trend Decomposition with Loess)
  - Two anomaly algorithms previously investigated by EA Comps groups
  - K-Means good at analyzing multivariate data

Anomaly Algorithms

- K-Means & STL (Seasonal Trend Decomposition with Loess)
  - Two anomaly algorithms previously investigated by EA Comps groups
  - K-Means good at analyzing multivariate data
  - STL especially sensitive to identifying anomalies associated with seasonal changes

Anomaly Algorithms

- K-means not what we needed
  - Conflated high positive rate
  - Computationally expensive
  - Results not as interpretable

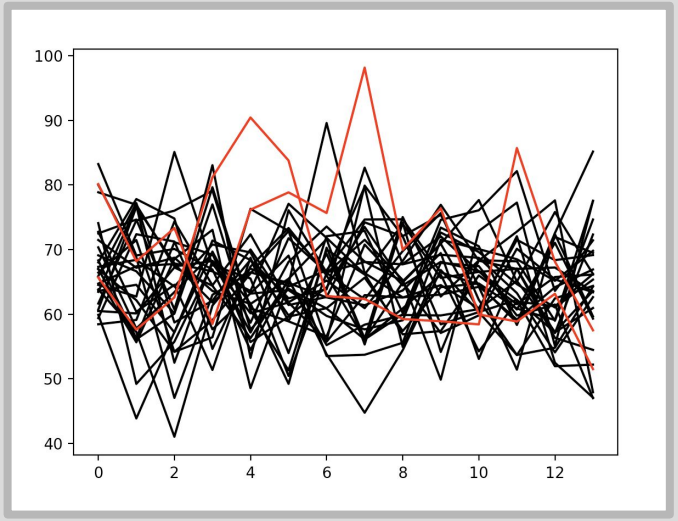


Figure 1: K-means anomaly detection algorithm run on test data

Anomaly Algorithms

- K-means not what we needed
  - Conflated high positive rate
  - Computationally expensive
  - Results not as interpretable
- STL
  - Very slow

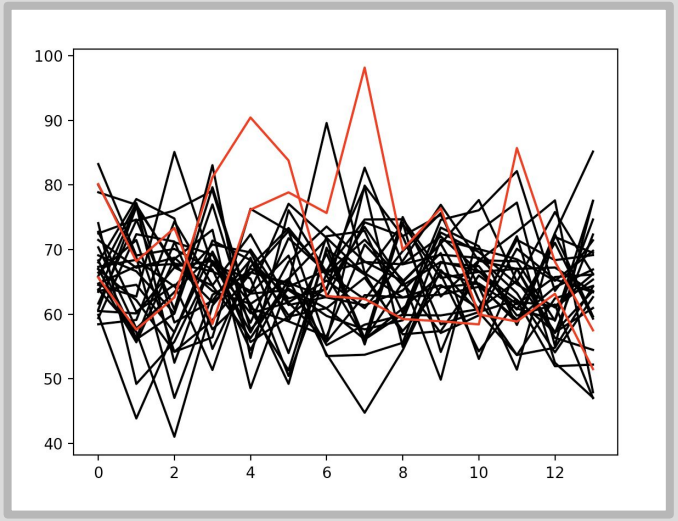


Figure 1: K-means anomaly detection algorithm run on test data

Anomaly Algorithms

- K-means not what we needed
  - Conflated high positive rate
  - Computationally expensive
  - Results not as interpretable
- STL
  - Very slow

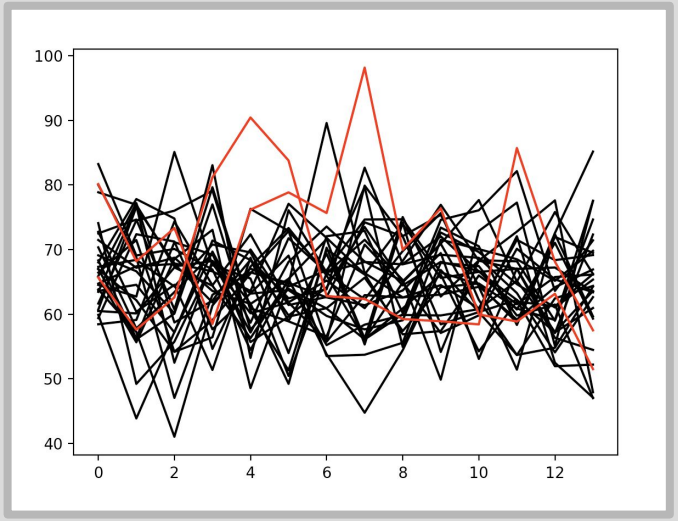


Figure 1: K-means anomaly detection algorithm run on test data

- Promising in the long term, and merits further research
  - Decided to focus on tools that allowed for higher interactivity

Energy Analytics 4

- Energy data
- Data Management
- Anomaly algorithms
- **Heuristics approach**
  - Idea
  - Example
- API
- Website interface

Heuristics Approach

- Short term over long term

Heuristics Approach

- Short term over long term
- Finds data points in range of parameters for *combination of points* selected by user

Heuristics Approach

- Short term over long term
- Finds data points in range of parameters for *combination of points* selected by user
- UI catered towards facility staff
  - Intent to make identifying broken points easier

Heuristics Approach

- Short term over long term
- Finds data points in range of parameters for *combination of points* selected by user
- UI catered towards facility staff
  - Intent to make identifying broken points easier
- Interpretable presentation of identified points

Heuristics Approach

Example: Using *Room Temperature* and *Damper Percentage*

- High RT, low DP
    - Stuck vent

Heuristics Approach

Example: Using *Room Temperature* and *Damper Percentage*

- High RT, low DP
  - Stuck vent
- Low RT, high DP
  - Stuck vent

Energy Analytics 4

- Energy data
- Data Management
- Anomaly algorithms
- Heuristics approach
- **API**
  - Explanation
  - Implementation
- Website interface

API

- Previous comps groups developed an API that allowed for the extraction of values from one specific point over time

API

- Previous comps groups developed an API that allowed for the extraction of values from one specific point over time
  - Highly granular, very targeted querying of the database

API

- Previous comps groups developed an API that allowed for the extraction of values from one specific point over time
  - Highly granular, very targeted querying of the database
  - Perhaps *too* targeted

API

- Previous comps groups developed an API that allowed for the extraction of values from one specific point over time
  - Highly granular, very targeted querying of the database
  - Perhaps *too* targeted

In practice:

- More than any one particular point value, it was *the combination of point values within a given room* that signaled to facilities that a problem had arose

API

- Previous comps groups developed an API that allowed for the extraction of values from one specific point over time
  - Highly granular, very targeted querying of the database
  - Perhaps *too* targeted

In practice:

- More than any one particular point value, it was *the combination of point values within a given room* that signaled to facilities that a problem had arose
  - Zoomed out API endpoints

API

- Previous comps groups developed an API that allowed for the extraction of values from one specific point over time
  - Highly granular, very targeted querying of the database
  - Perhaps *too* targeted

In practice:

- More than any one particular point value, it was *the combination of point values within a given room* that signaled to facilities that a problem had arose
  - Zoomed out API endpoints
    - Room level data

API

- Previous comps groups developed an API that allowed for the extraction of values from one specific point over time
  - Highly granular, very targeted querying of the database
  - Perhaps *too* targeted

In practice:

- More than any one particular point value, it was *the combination of point values within a given room* that signaled to facilities that a problem had arose
  - Zoomed out API endpoints
    - Room level data
    - Multiple point value readings at the same time

API

/anomalies/vent-and-temp

Parameters: start time, end time, vent, temp

API

/anomalies/vent-and-temp

Parameters: start time, end time, vent, temp

Big idea: “Show me all of the rooms on campus within a given time frame, where the temperature is greater than a particular threshold, but the vent isn’t opening wide enough to try and cool it down”

API

/anomalies/vent-and-temp

Parameters: start time, end time, vent, temp

Big idea: “Show me all of the rooms on campus within a given time frame, where the temperature is greater than a particular threshold, but the vent isn’t opening wide enough to try and cool it down”

Context: Facilities and directory of sustainability have identified this type of “vent/temp mismatch” as a common problem, one they’d like to be able to spot quickly

API

/anomalies/vent-and-temp?start_time=1636478378
&end_time=1636564778&vent=30&temp=75

API

/anomalies/vent-and-temp?start_time=1636478378
&end_time=1636564778&vent=30&temp=75

“Show me all of the rooms that had a temperature greater than 75 and a vent angle of less than 30, between November 9th and November 10th, 2021”

API

/anomalies/vent-and-temp?start_time=1636478378&end_time=1636564778&vent=30&temp=75

“Show me all of the rooms that had a temperature greater than 75 and a vent angle of less than 30, between November 9th and November 10th, 2021”


```
{
  "Boliou 140": {
    "damper_name": "BO.1.RM140:DMPR COMD",
    "temp_name": "BO.1.RM140:ROOM TEMP",
    "values": {
      "temp": [
              75.75, 76.25, 76.25, 76.50, 76.50, 76.50, 76.00, 75.50,
              75.25, 75.25, 75.25, 75.25, 75.25, 75.25, 75.25, 75.25,
              75.25, 75.25, 75.25, 75.25, 75.25, 75.25, 75.25, 75.75,
              75.25, 75.25, 75.25, 75.25, 75.25, 75.25, 75.75, 76.25,
              76.25
              ],
      "timestamp": [
              1636479000, 1636479900, 1636480800, 1636481700, 1636482600,
              1636483500, 1636484400, 1636485300, 1636533000, 1636533900,
              1636534800, 1636535700, 1636536600, 1636537500, 1636538400,
              1636539300, 1636540200, 1636541100, 1636542000, 1636542900,
              1636543800, 1636544700, 1636545600, 1636546500, 1636547400,
              1636548300, 1636549200, 1636550100, 1636551000, 1636561800,
              1636562700, 1636563600, 1636564500
              ],
      "vent": [
              0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
              0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
              ]
    }
  }
}
```


Figure 4: JSON data returned by API heuristic query

Energy Analytics 4

- Energy data
- Data Management
- Anomaly algorithms
- Heuristics approach
- API
- **Website interface**
  - Dashboard
  - Heuristics

Website Interface

- Dashboard
  -
- Heuristics
  -

Website Interface

- Dashboard
  - Seamless, automated display of most important information

- Heuristics
  -

Website Interface

- Dashboard
  - Seamless, automated display of most important information

- Heuristics
  - In depth interactivity and experimentation

Website Interface - Heuristics

- Heuristics

Website Interface - Heuristics

- Heuristics
  - Simple heuristics: Finding points easier and quicker than anomaly search algorithms

Website Interface - Heuristics

- Heuristics
  - Simple heuristics: Finding points easier and quicker than anomaly search algorithms
  - Allowing facilities to play with data

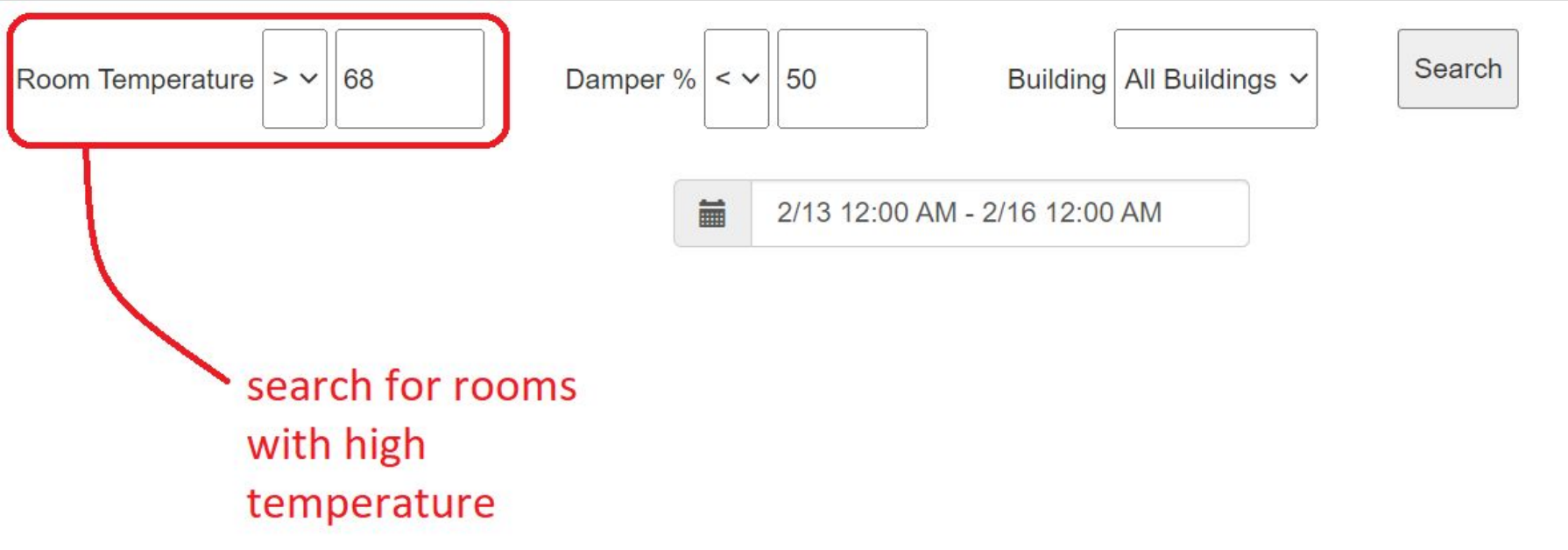
Website Interface - Heuristics

- Heuristics
  - Simple heuristics: Finding points easier and quicker than anomaly search algorithms
  - Allowing facilities to play with data
    - Finding open windows

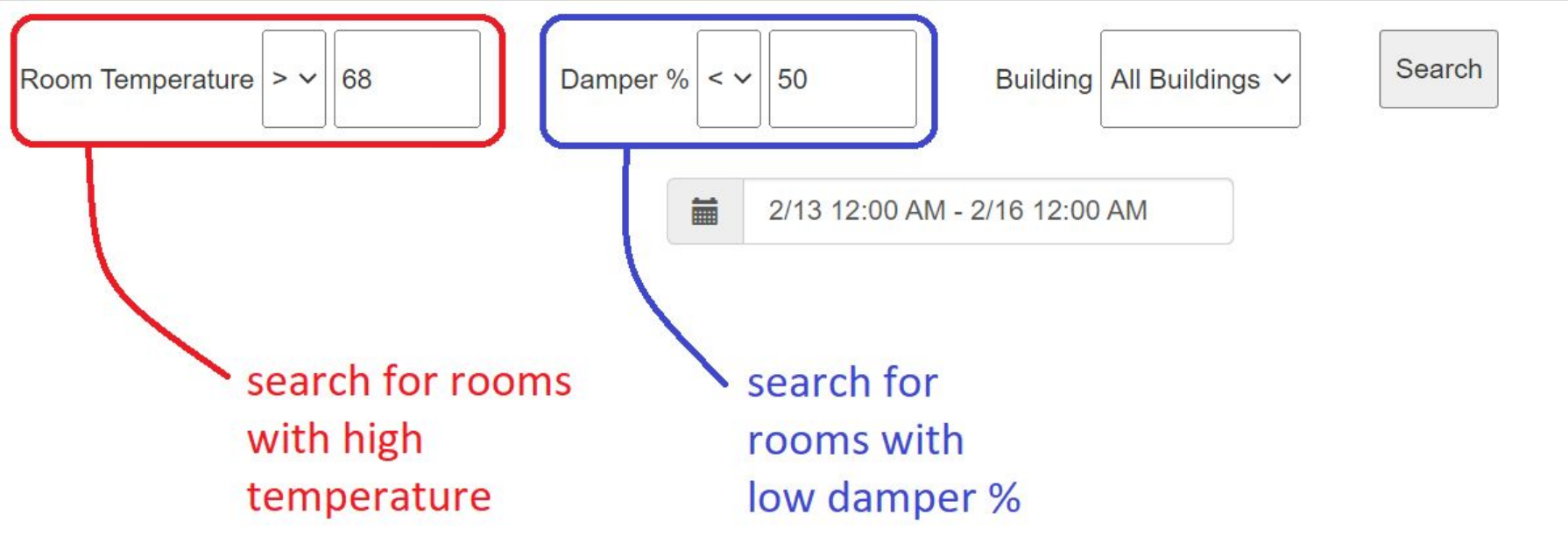
Website Interface - Heuristics

- Heuristics
  - Simple heuristics: Finding points easier and quicker than anomaly search algorithms
  - Allowing facilities to play with data
    - Finding open windows
    - Broken devices

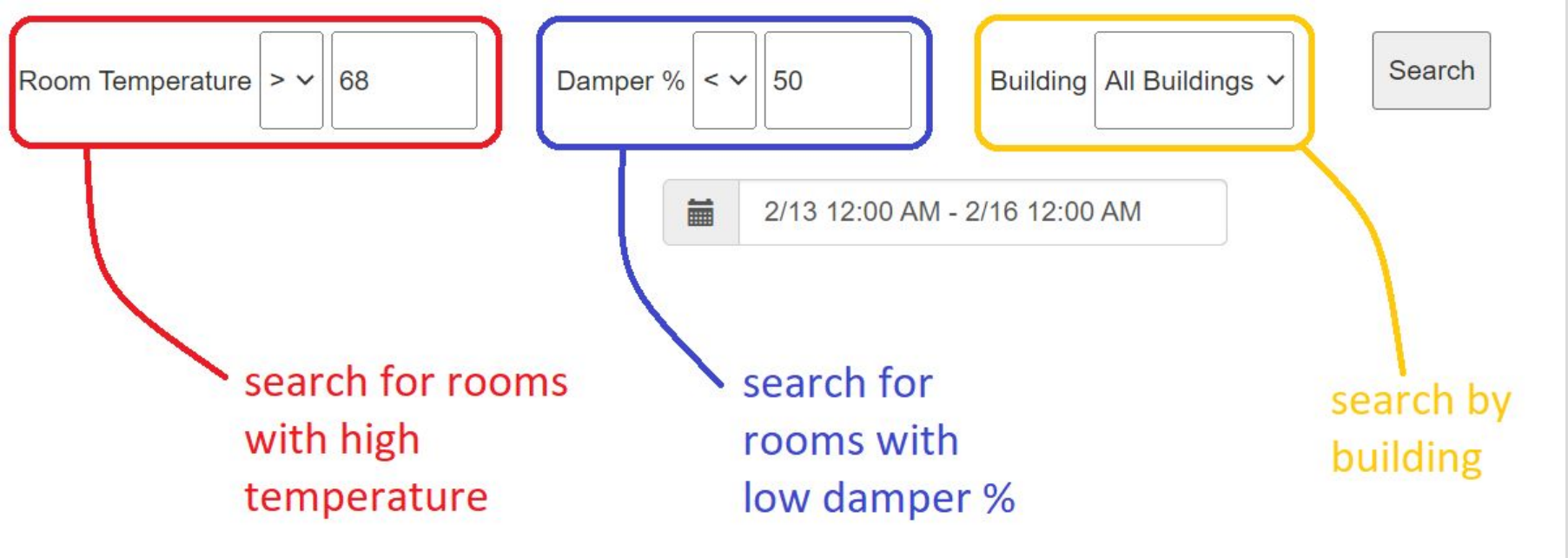
Room temperature & Damper %



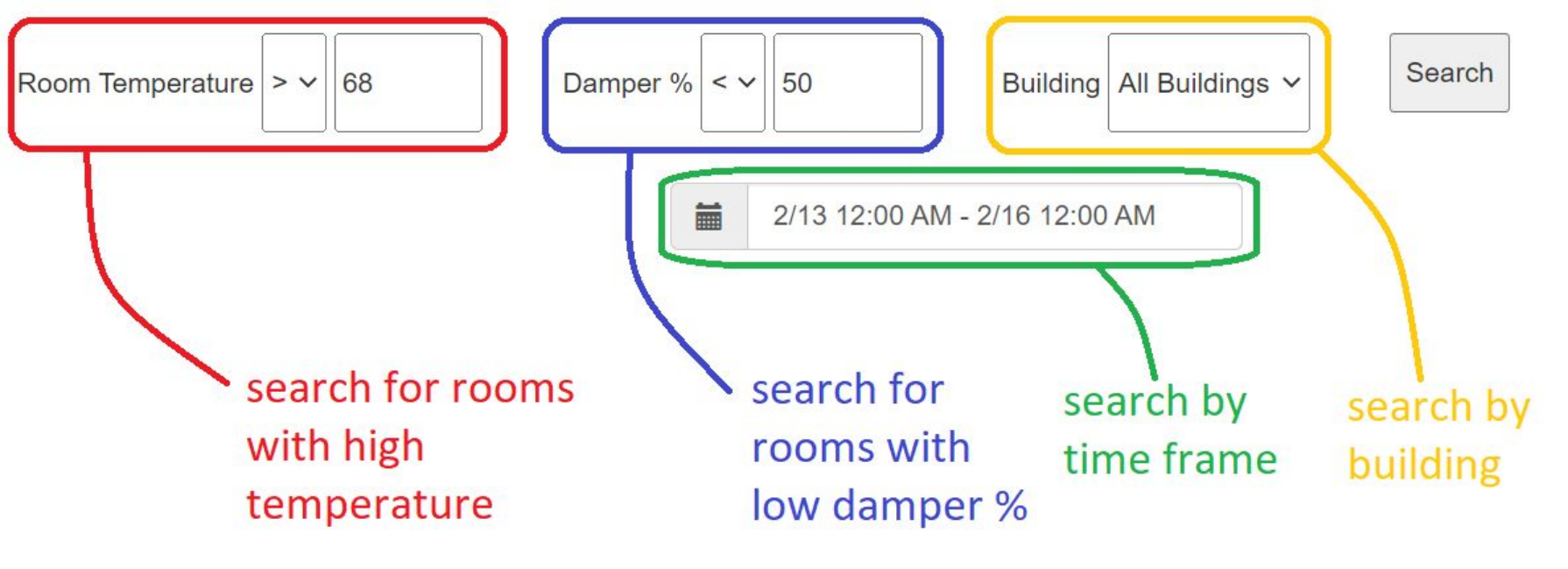
Room temperature & Damper %



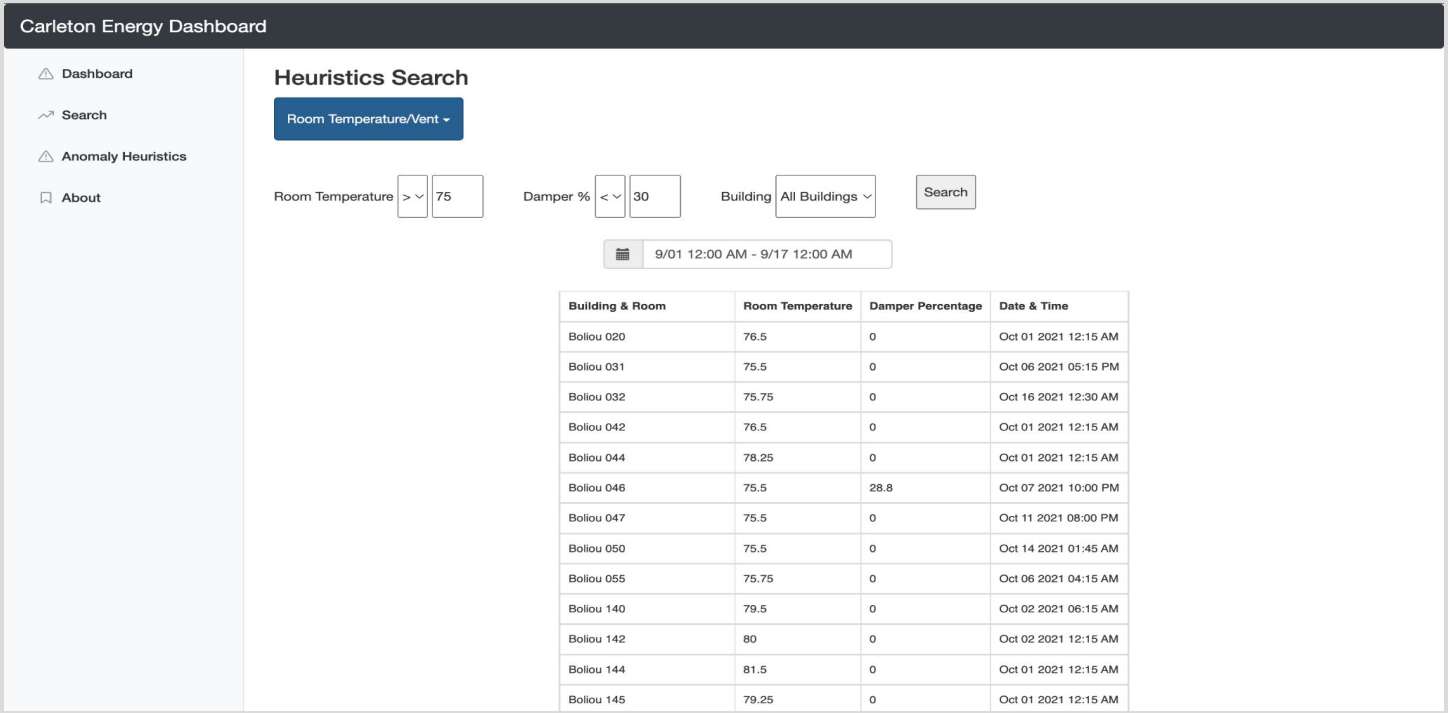
Room temperature & Damper %



Room temperature & Damper %

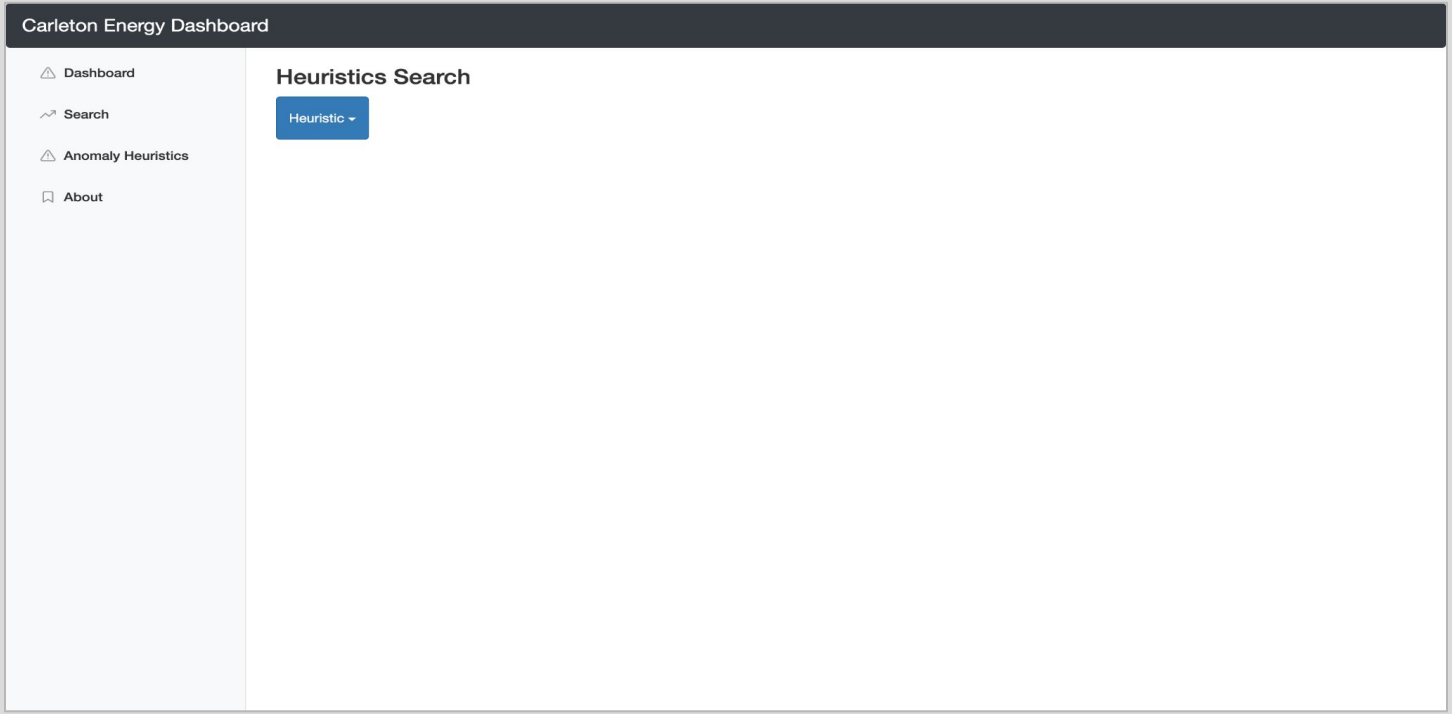


Room temperature & Damper %

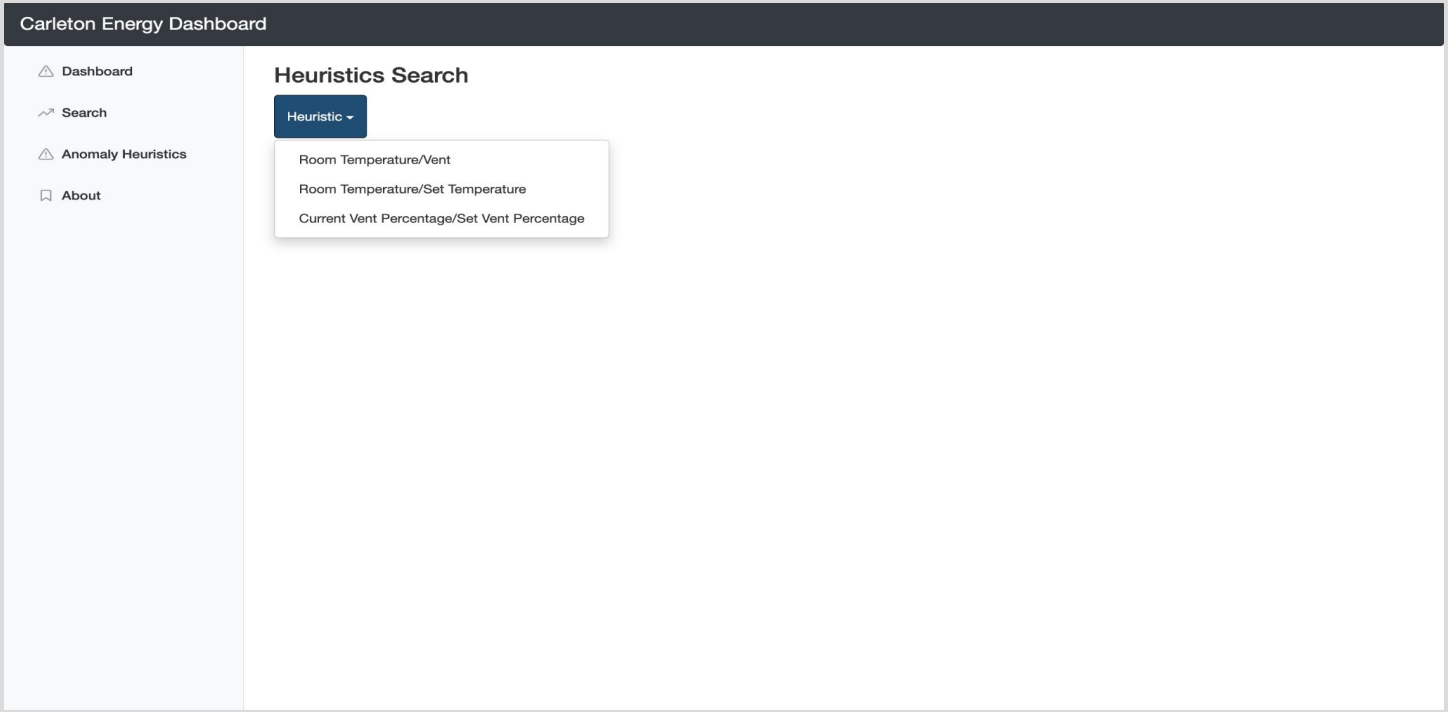


Building & Room	Room Temperature	Damper Percentage	Date & Time
Boliou 020	76.5	0	Oct 01 2021 12:15 AM
Boliou 031	75.5	0	Oct 06 2021 05:15 PM
Boliou 032	75.75	0	Oct 16 2021 12:30 AM
Boliou 042	76.5	0	Oct 01 2021 12:15 AM
Boliou 044	78.25	0	Oct 01 2021 12:15 AM
Boliou 046	75.5	28.8	Oct 07 2021 10:00 PM
Boliou 047	75.5	0	Oct 11 2021 08:00 PM
Boliou 050	75.5	0	Oct 14 2021 01:45 AM
Boliou 055	75.75	0	Oct 06 2021 04:15 AM
Boliou 140	79.5	0	Oct 02 2021 06:15 AM
Boliou 142	80	0	Oct 02 2021 12:15 AM
Boliou 144	81.5	0	Oct 01 2021 12:15 AM
Boliou 145	79.25	0	Oct 01 2021 12:15 AM

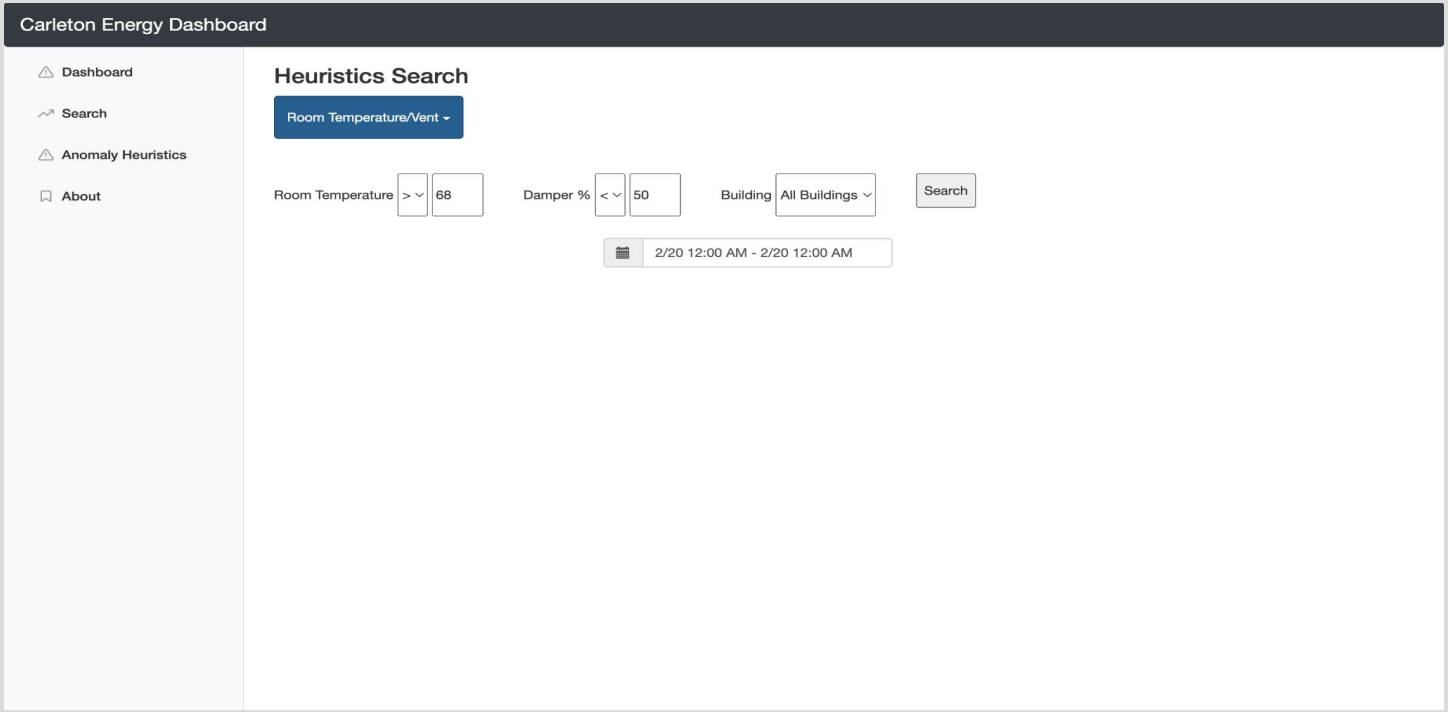
Heuristics: Functionality



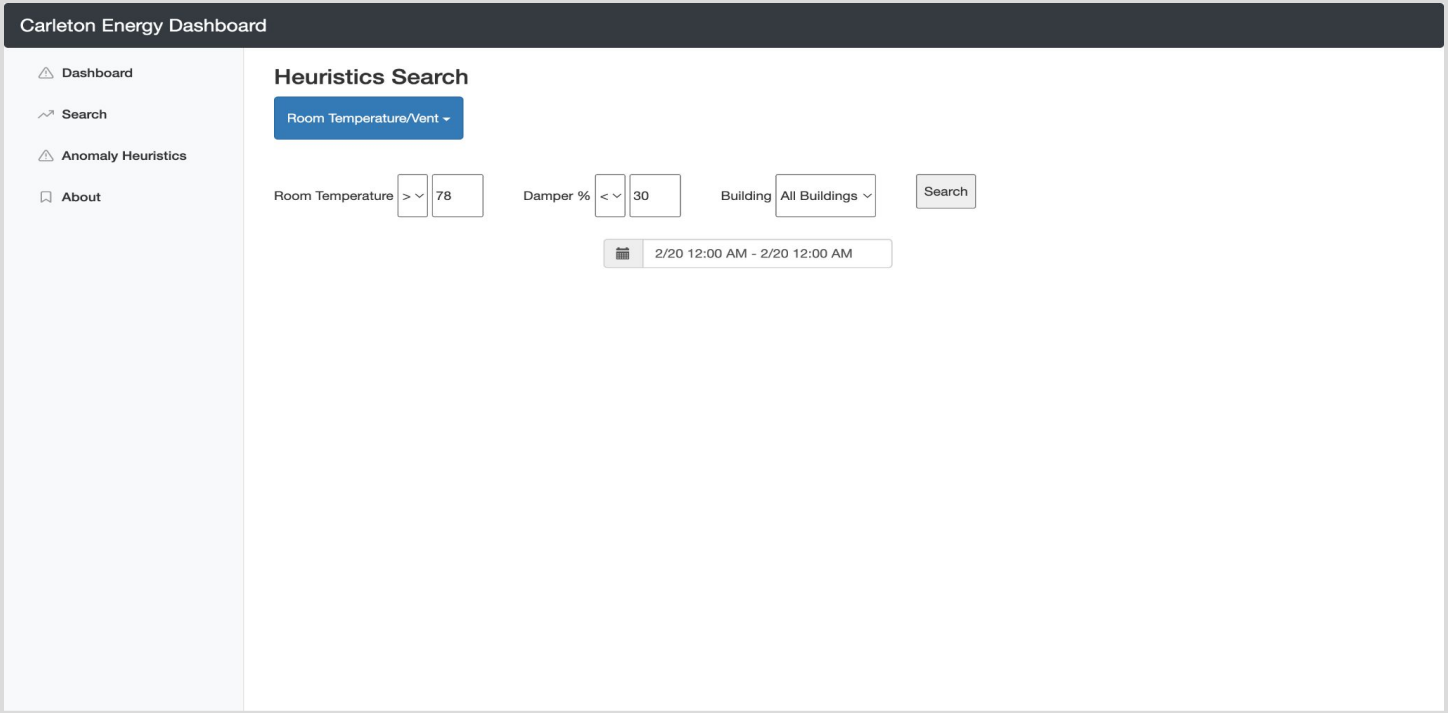
Heuristics: Functionality



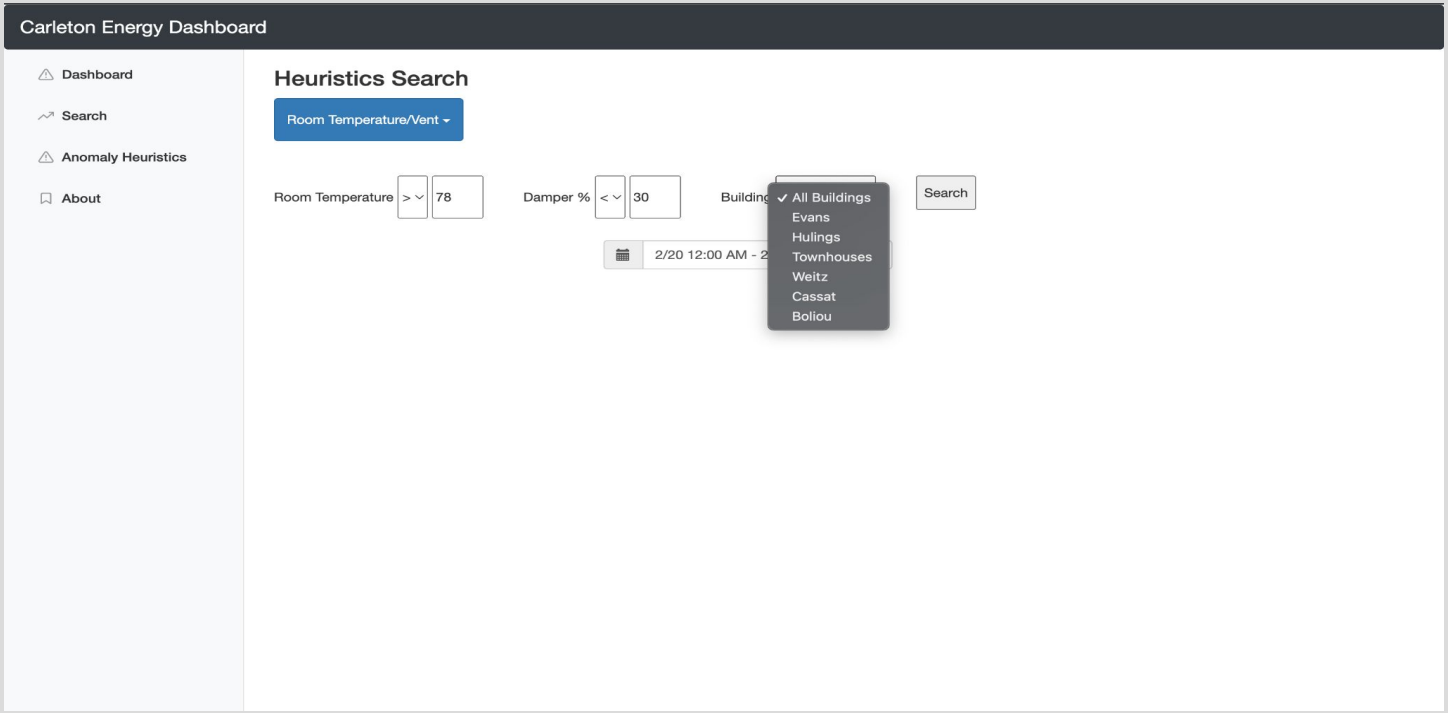
Heuristics: Functionality



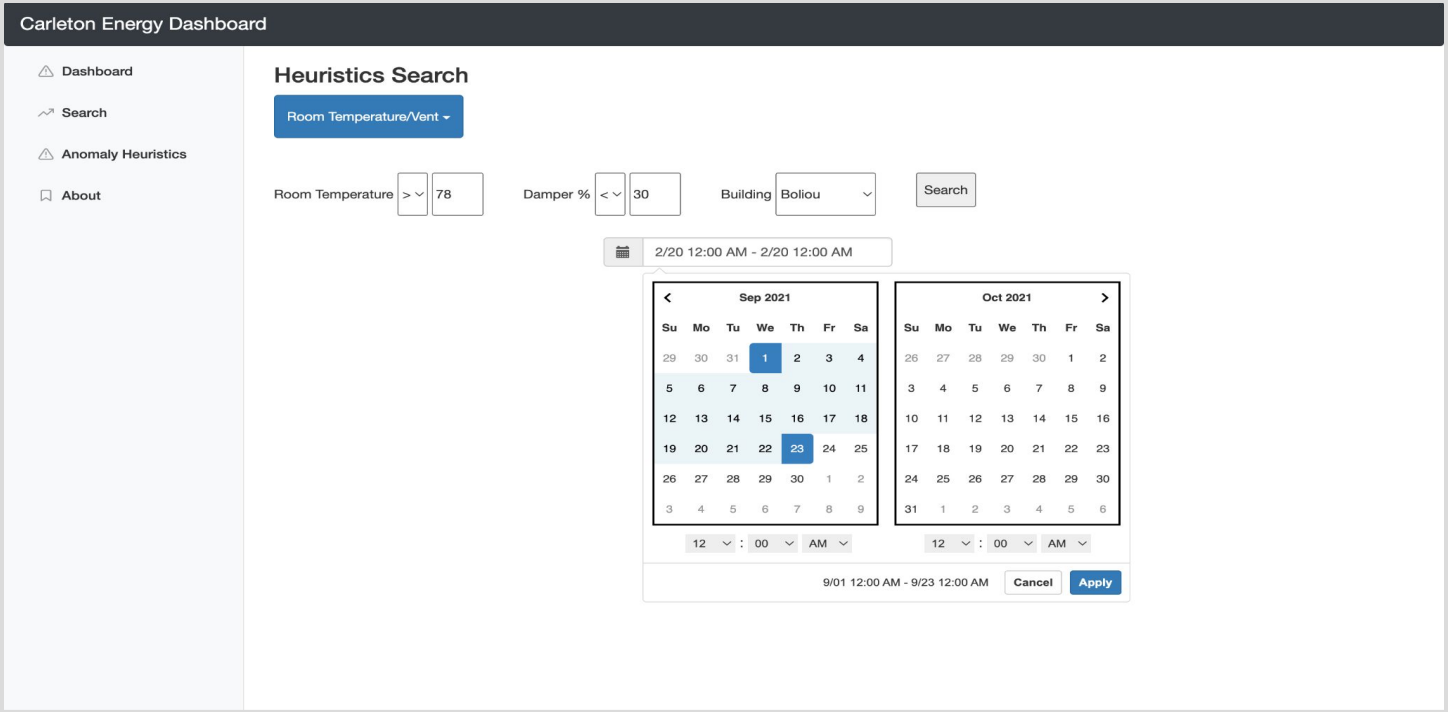
Heuristics: Functionality



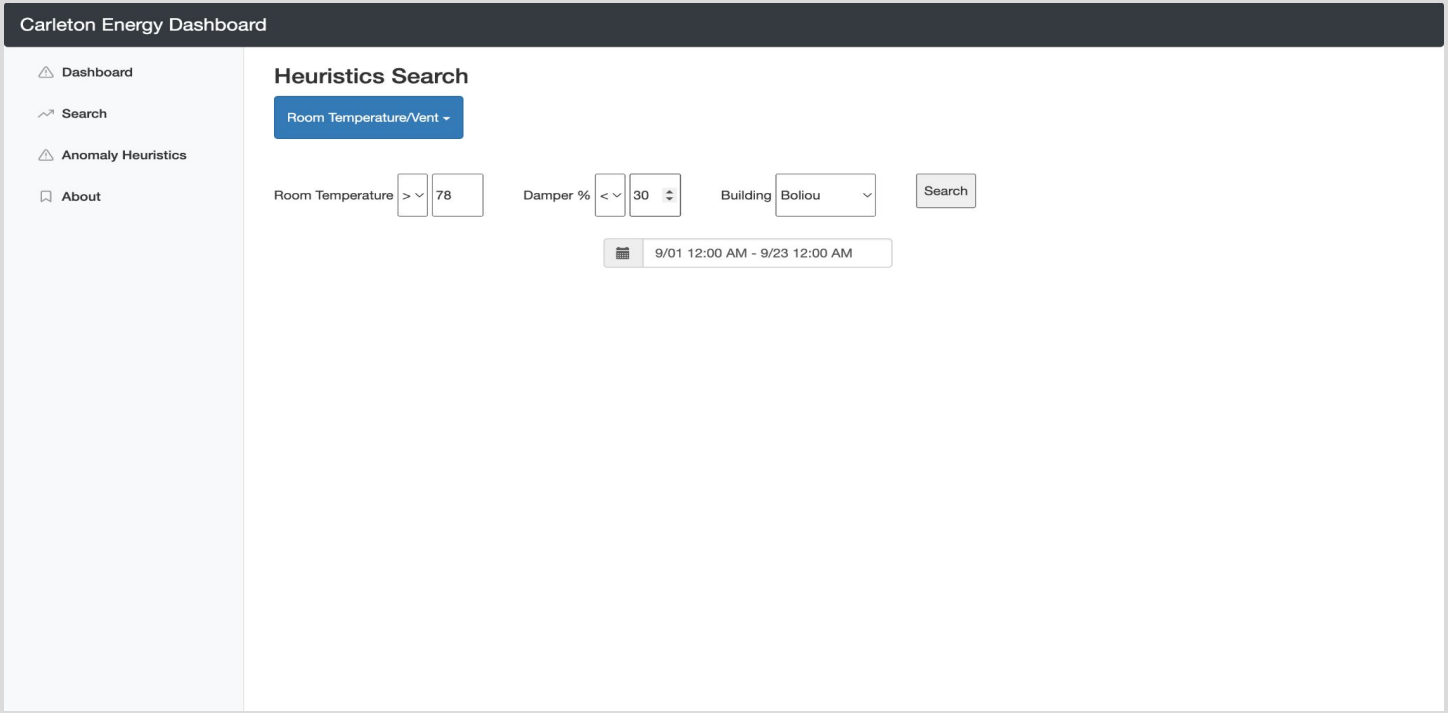
Heuristics: Functionality



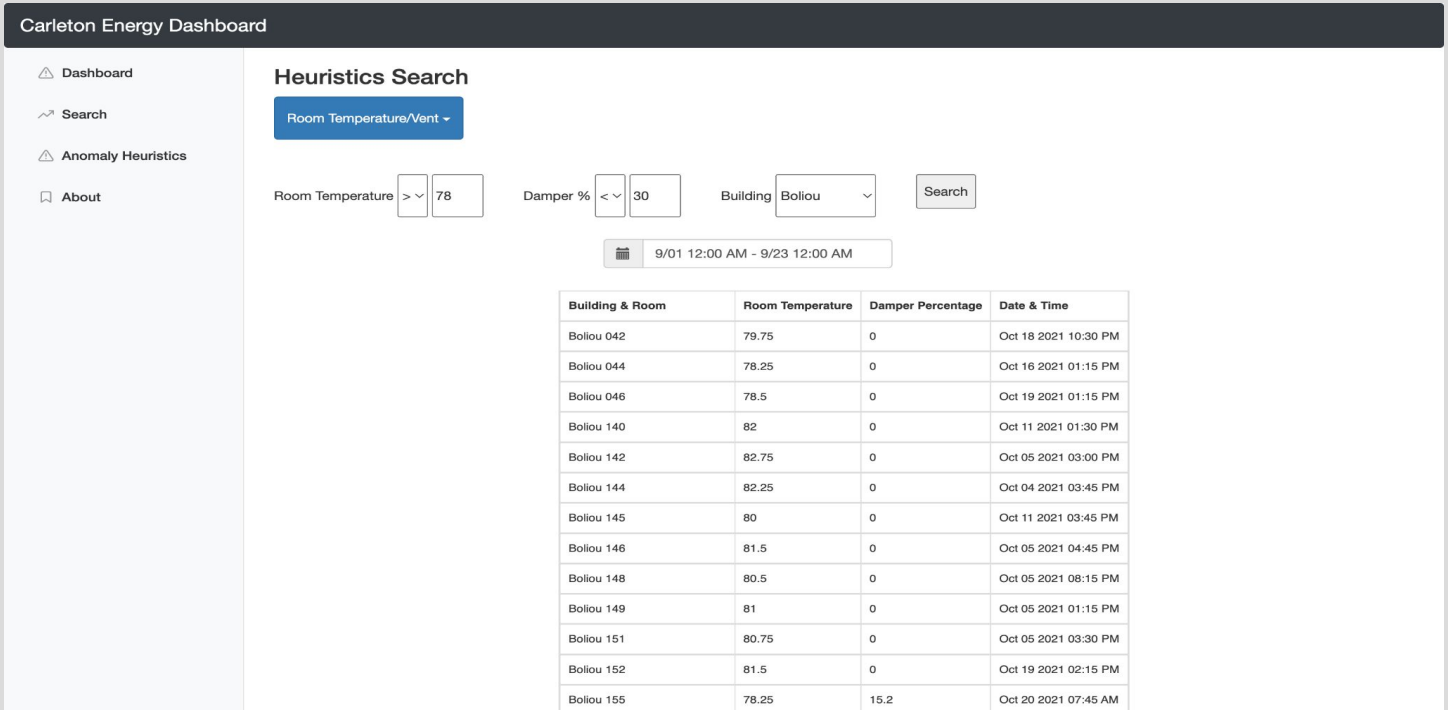
Heuristics: Functionality



Heuristics: Functionality



Heuristics: Functionality



Building & Room	Room Temperature	Damper Percentage	Date & Time
Boliou 042	79.75	0	Oct 18 2021 10:30 PM
Boliou 044	78.25	0	Oct 16 2021 01:15 PM
Boliou 046	78.5	0	Oct 19 2021 01:15 PM
Boliou 140	82	0	Oct 11 2021 01:30 PM
Boliou 142	82.75	0	Oct 05 2021 03:00 PM
Boliou 144	82.25	0	Oct 04 2021 03:45 PM
Boliou 145	80	0	Oct 11 2021 03:45 PM
Boliou 146	81.5	0	Oct 05 2021 04:45 PM
Boliou 148	80.5	0	Oct 05 2021 08:15 PM
Boliou 149	81	0	Oct 05 2021 01:15 PM
Boliou 151	80.75	0	Oct 05 2021 03:30 PM
Boliou 152	81.5	0	Oct 19 2021 02:15 PM
Boliou 155	78.25	15.2	Oct 20 2021 07:45 AM

Website Interface - Heuristics

- Functionality of heuristics page
  - Limitations of data and current heuristics
  - Potential of current heuristics
  - Future heuristics and their value

Website Interface - Dashboard

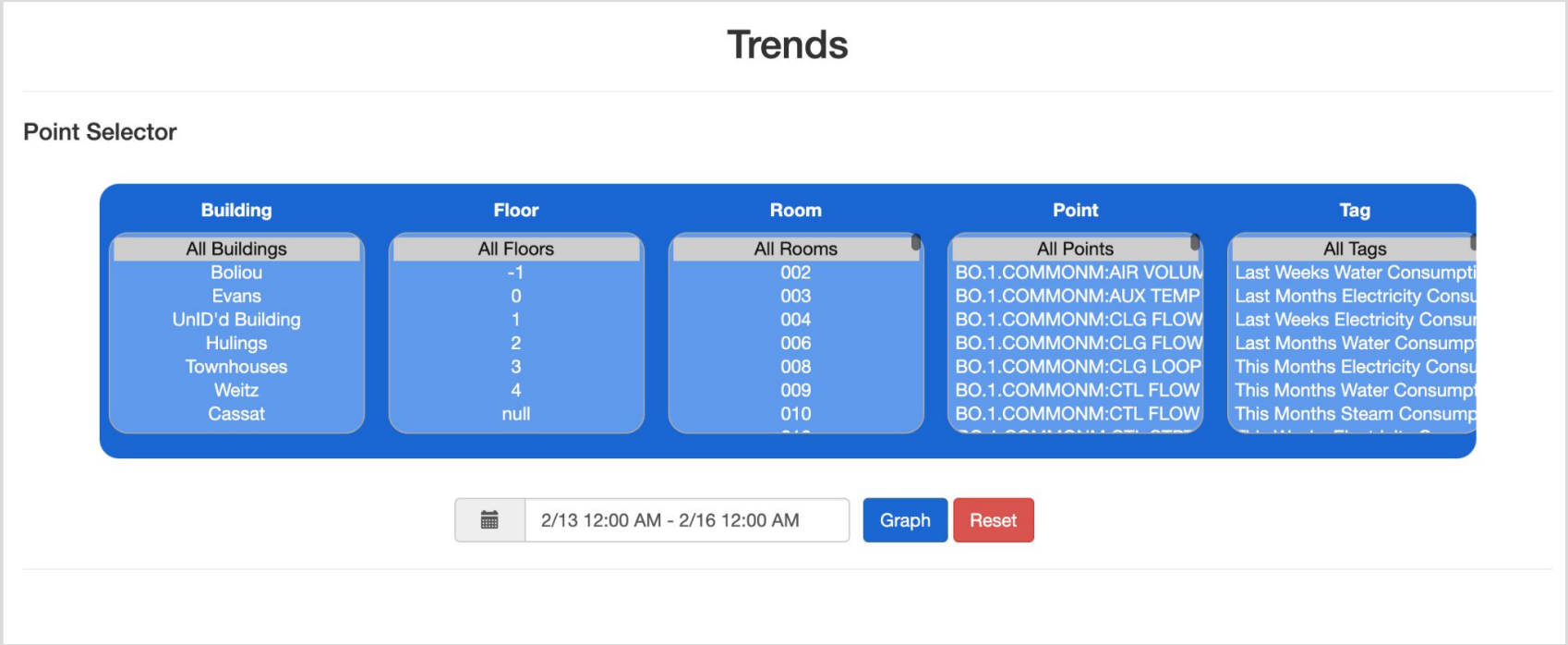


Figure 6: Point search tool

Website Interface - Dashboard

Potential Problems

Temp/Vent Mismatches (T > 75°, V < 30%)

Room	Date & Time	Room Temperature	Damper Percentage
Boliou 140	11:30 Nov 9	75.75°	0%
	11:15 Nov 10	76.25°	0%
Boliou 142	11:30 Nov 9	76.25°	0%
	11:15 Nov 10	76.75°	0%
Boliou 155	5:30 Nov 10	75.25°	10.4%
	9:45 Nov 10	75.5°	14.4%

Figure 7: Dashboard populated by potentially anomalous rooms

Overview

- Introduction
- Energy Analytics 4
- Conclusion

Overview

- Introduction
- Energy Analytics 4
- **Conclusion**
  - Challenges
  - Our work
  - Moving Forward
  - Acknowledgements

Conclusion

- Challenges
- Our work
- Moving Forward
- Acknowledgements

Conclusion

- **Challenges**
  - Legacy code
  - Data standardization & pipeline
  - API Documentation
  - Data Access
- Our work
- Moving Forward
- Acknowledgements

Challenges

- Legacy code
- Data standardization & pipeline
- API Documentation
- Data Access

Challenges

- **Legacy code**
- Data standardization & pipeline
- API Documentation
- Data Access

Legacy code

- Working but not fully documented
- Incomplete & didn’t work (STL)
- Irrelevant or with unclear purpose

Challenges

- Legacy code
- **Data standardization & pipeline**
- API Documentation
- Data Access

Data & pipeline

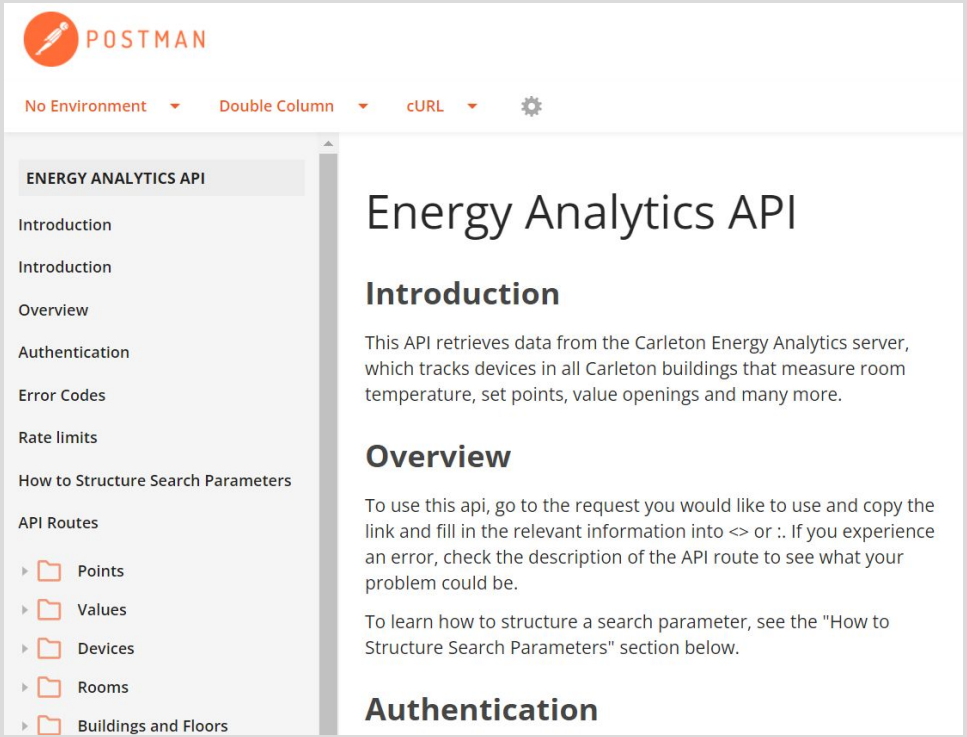
- Missing data
  - ALC, Wind, Windows upgrade
- Data integration
  - Siemens, ALC, Wind different formatting
- Ownership/Permissions
  - Automation didn’t go through, files couldn’t run, data unmovable

Challenges

- Legacy code
- Data standardization & pipeline
- **API Documentation**
- Data Access

API documentation

- No access to edit
- Incomplete sections



Challenges

- Legacy code
- Data standardization & pipeline
- API Documentation
- **Data Access**

Data access

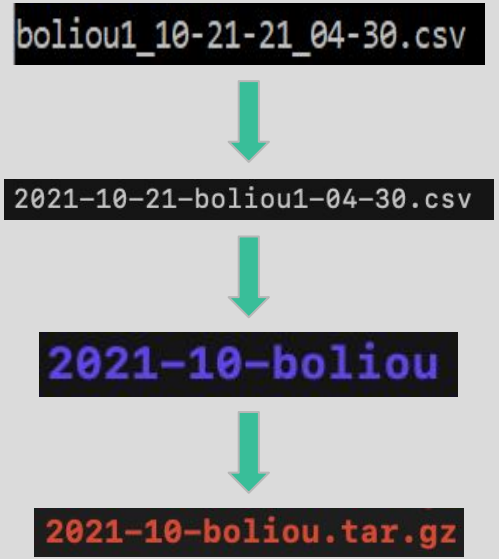
- No access to data from the points

Conclusion

- Challenges
- **Key takeaways**
  - Standardization and compression
  - User-friendly web interface
  - Dynamic heuristics concept
- Moving Forward
- Acknowledgements

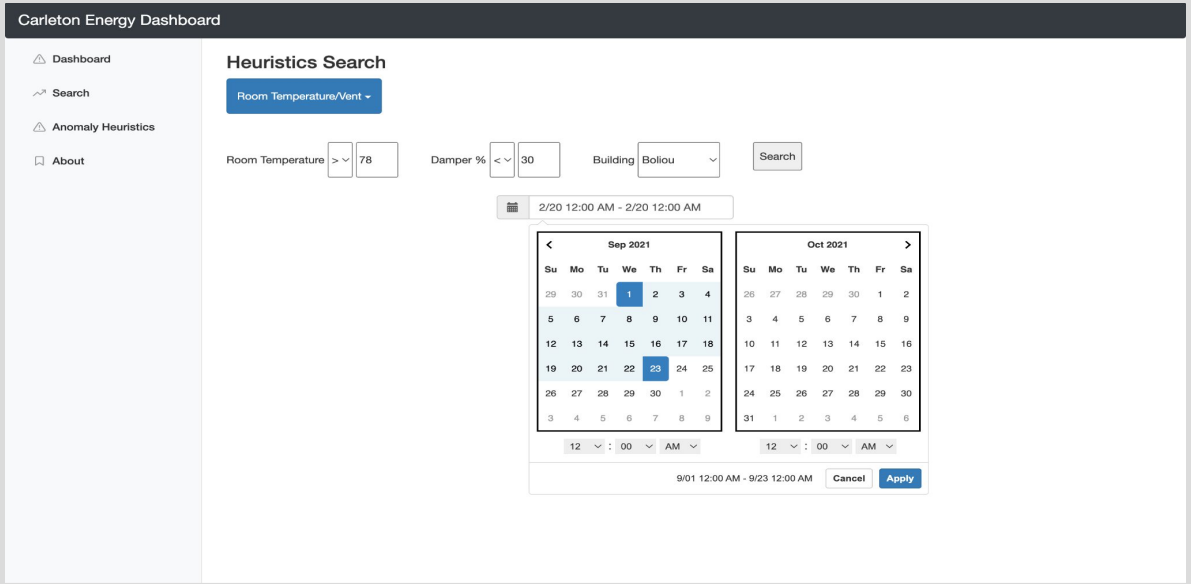
Key takeaways

- Data standardization and compression system



Key takeaways

- Data standardization and compression system
- User-friendly Dashboard and Heuristics page



Key takeaways

- Data standardization and compression system
- User-friendly Dashboard and Heuristics page
- Heuristics concept robust for further development
  - Room temp vs. set temp
  - Vent angle vs. set vent angle

Conclusion

- Challenges
- Key takeaways
- **Moving Forward**
  - More heuristics
  - Machine learning algorithms
  - Data integration
- Acknowledgements

Moving forward

- Further implementation of heuristics
  - Room temp vs. set temp
  - Vent angle vs. set vent angle

Moving forward

- Further implementation of heuristics
  - Room temp vs. set temp
  - Vent angle vs. set vent angle
- Machine learning algorithms
  - K-means
  - Isolation forests

Moving forward

- Further implementation of heuristics
  - Room temp vs. set temp
  - Vent angle vs. set vent angle
- Machine learning algorithms
  - K-means
  - Isolation forests
- Data integration
  - Siemens and ALC
  - Wind Turbine data

Conclusion

- Challenges
- Key takeaways
- Moving Forward
- **Acknowledgements**

Acknowledgements

- Jeff Ondich
- Martha Larson, Sustainability team
- Facilities Staff
- CS department
- Mike Tie

Questions