BTS:

Building Time-Series Dataset: Empowering Large-Scale Building Analytics

Arian Prabowo

Xiachong Lin

Imran Razzak

Hao Xue

Emily W. Yap

Matthew Amos

Flora D. Salim

A NeurIPS Dataset and Benchmark 2024 poster.
Official repo: https://github.com/cruiseresearchgroup/DIEF_BTS/

This research is supported by the NSW Government through the CSIRO's NSW Digital Infrastructure Energy Flexibility (DIEF) project, funded under the Net Zero Plan Stage 1: 2020-2030.
This project is also funded by the Reliable Affordable Clean Energy for 2030 (RACE for 2030) Cooperative Research Centre.
This research was undertaken with the assistance of resources from the National Computational Infrastructure (NCI Australia), an NCRIS enabled capability supported by the Australian Government.

Video by Amit from Pexels

BTS
Building Time-Series Dataset

- 3 Buildings

Anonymized buildings in Australia.
Pictures for illustrative purposes only.
Photo by Simone Hutsch on Unsplash

BTS

Building Time-Series Dataset

- 3 Buildings
- 3 Years

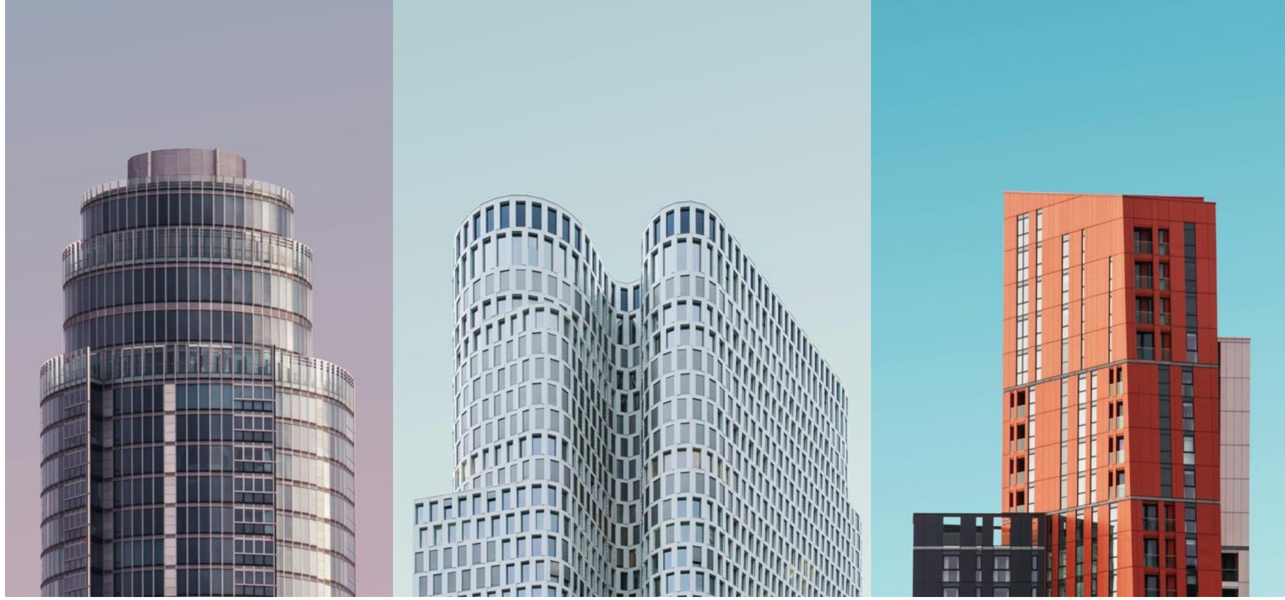
2021 **3 Years** 2024

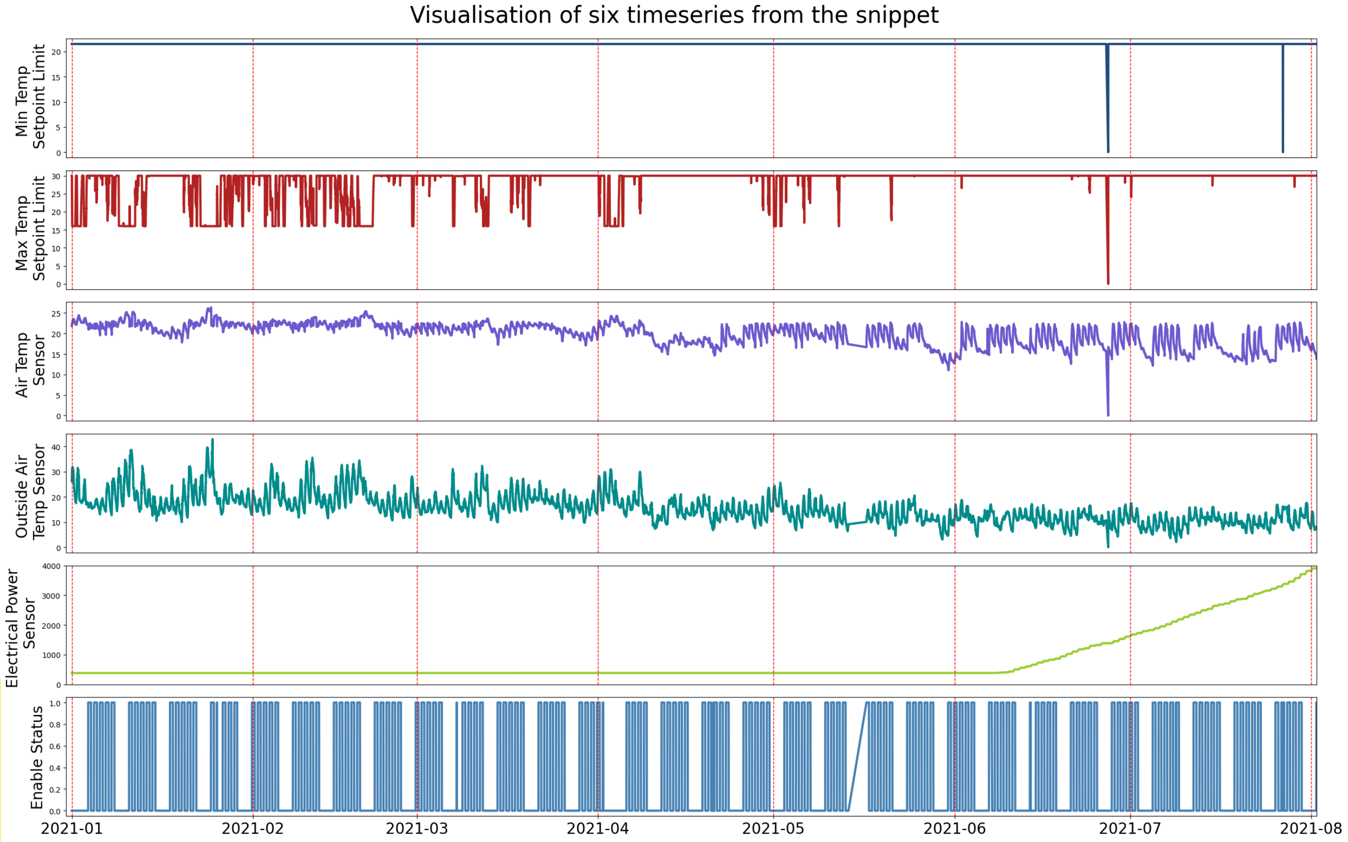
UNSW SYDNEY

BTS

Building Time-Series Dataset

- 3 Buildings
- 3 Years
- >15 000 Time-series
- >300 Classes





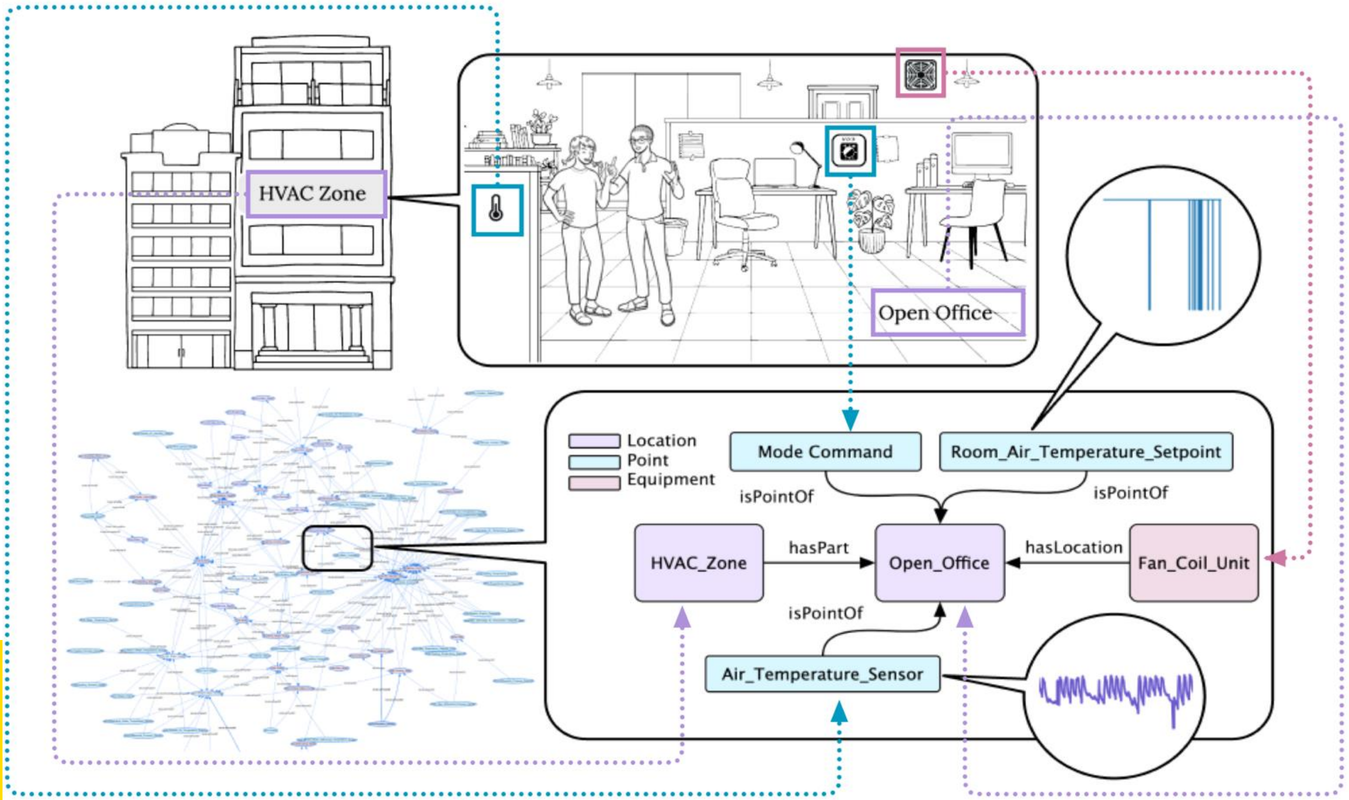
BTS

Building Time-Series Dataset

- 3 Buildings
- 3 Years
- >15 000 Time-series
- >300 Classes
- Knowledge Graph

The knowledge graph describes the "**physical, logical and virtual assets** in buildings and the **relationships** between them."

We use the Brick schema:



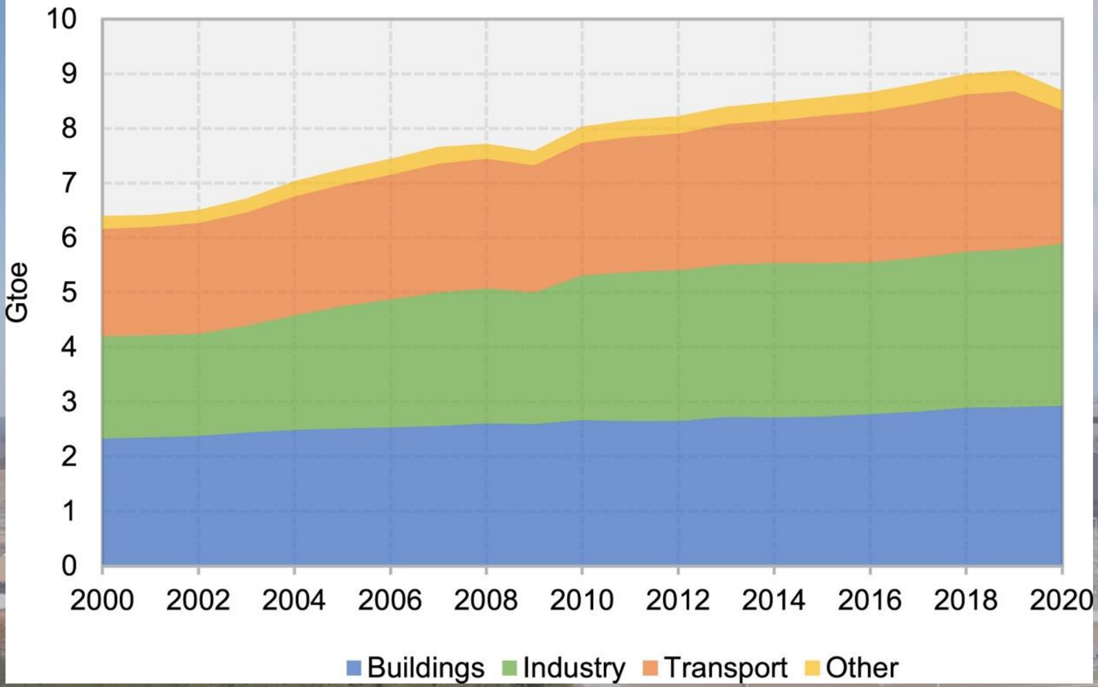
Why Building Analytics?

Video by Kelly from Pexels

Why Building Analytics?

A **third** of global energy consumption.

(Gtoe: Giga tonnes of oil equivalent)

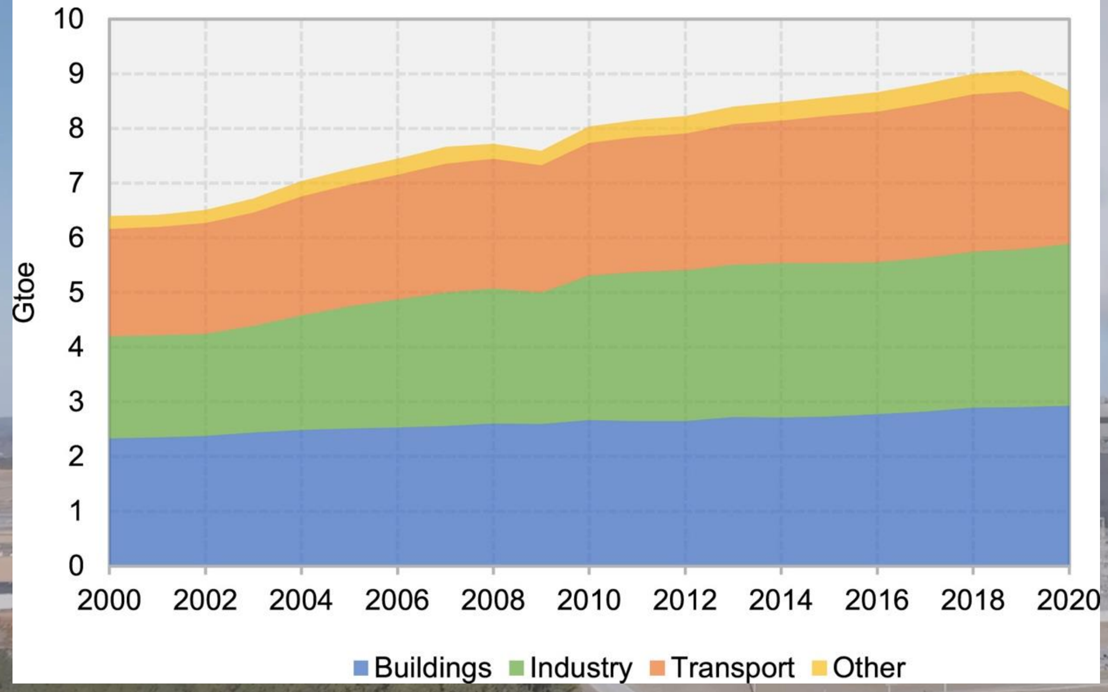


M. Gonz.lez-Torres, L. P.rez-Lombard, J. F. Coronel, I. R. Maestre, and D. Yan. A review on buildings energy information: Trends, end-uses, fuels and drivers. Energy Reports, 8:626–637, 2022.

Why Building Analytics?

A **third** of global energy consumption.

(Gtoe: Giga tonnes of oil equivalent)



A **quarter** of CO_2 emissions.

Table 1. Share of direct and indirect CO_2 emissions by sector in 2019.

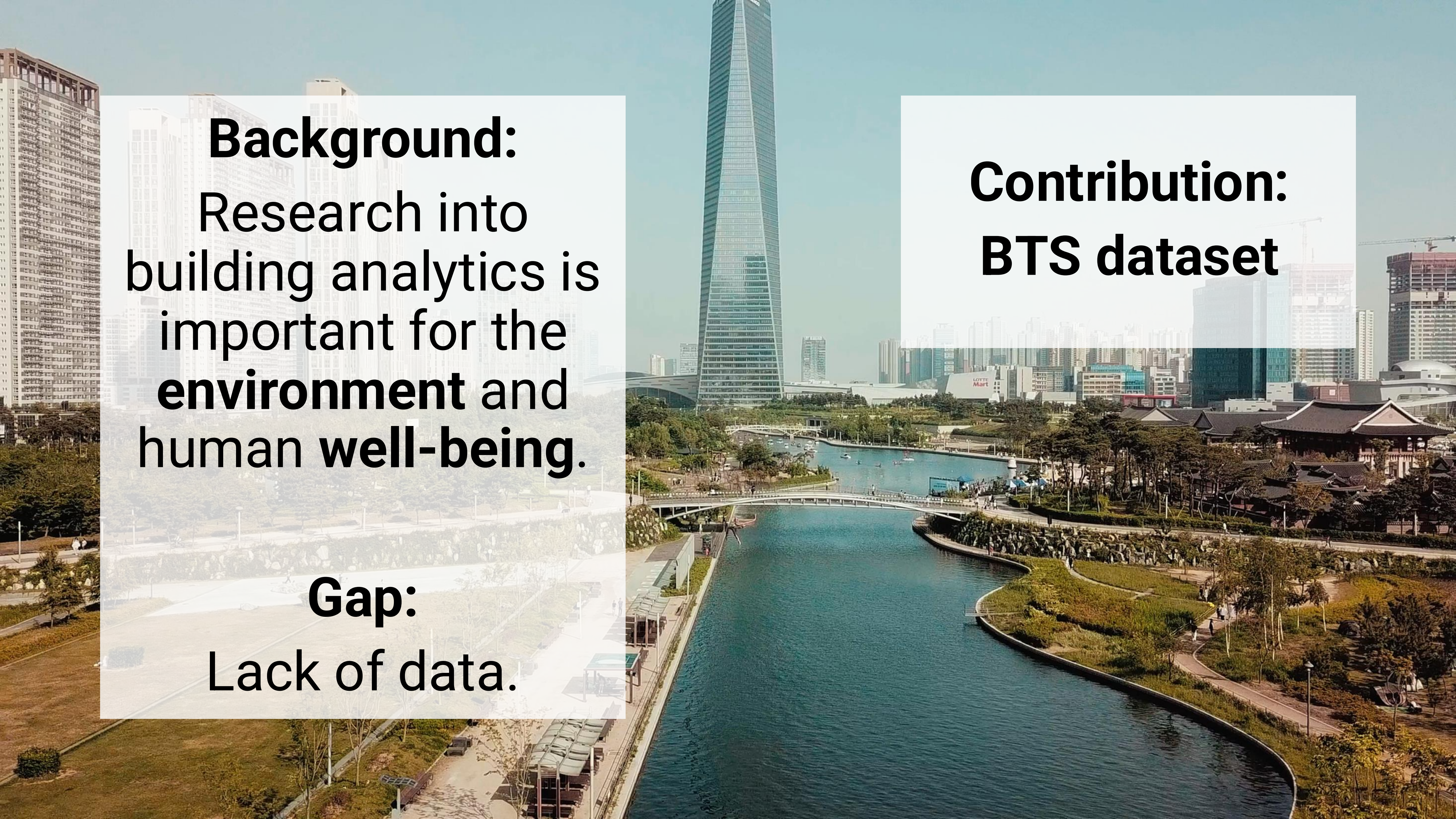
Sector	*Direct*	*Indirect*	*Total*
Industry	19%	19%	38%
Buildings	9%	19%	28%
Transport	25%	3%	28%
Other	2%	4%	6%

M. Gonz.lez-Torres, L. P.rez-Lombard, J. F. Coronel, I. R. Maestre, and D. Yan. A review on buildings energy information: Trends, end-uses, fuels and drivers. Energy Reports, 8:626–637, 2022.

Why Building Analytics?

As we spend more time in buildings, they will have increasing influence over our physical and mental wellbeing.

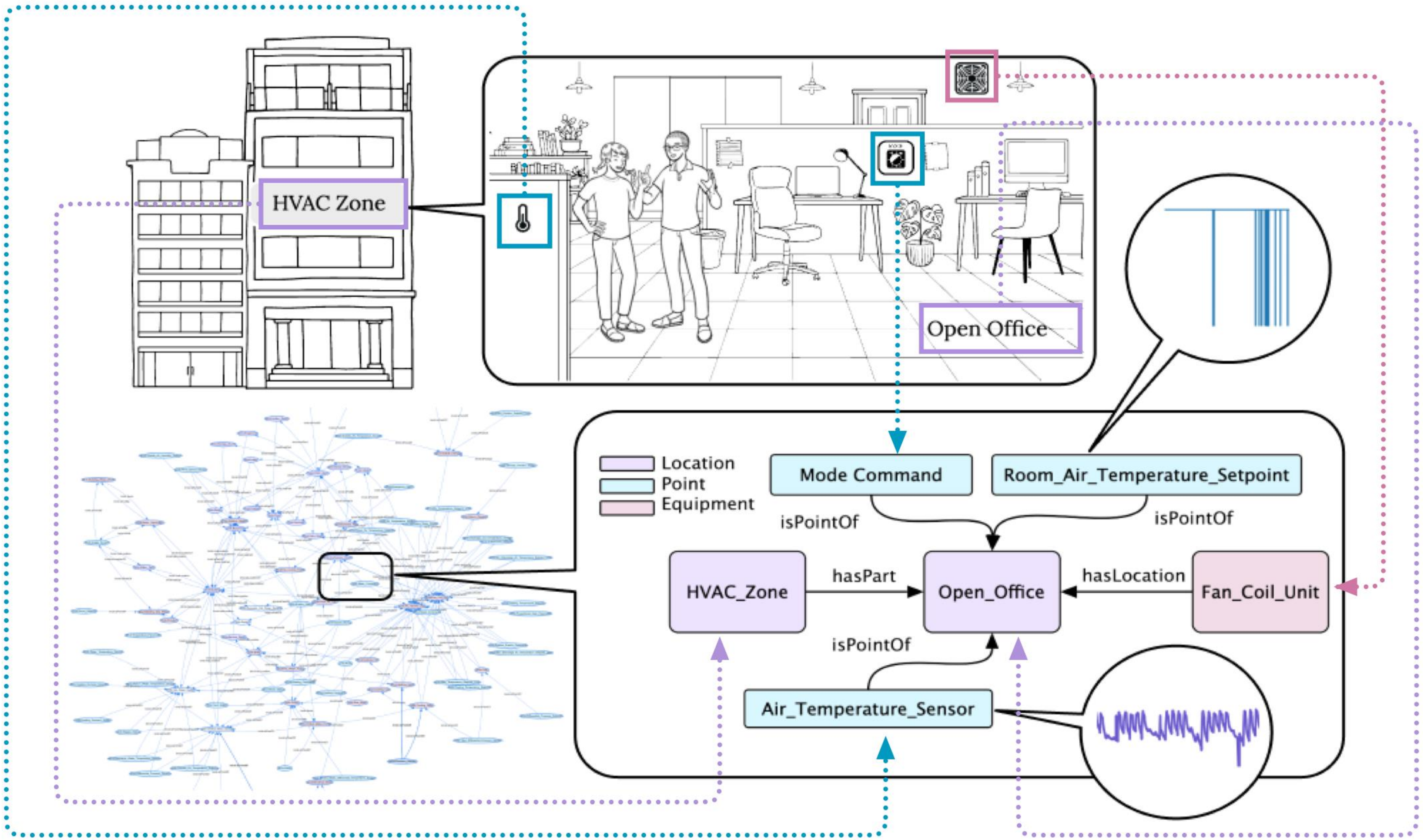
Video by fauxels from Pexels


Background:
Research into building analytics is important for the **environment** and human **well-being**.
Gap:
Lack of data.
Contribution:
BTS dataset
LOTTE Mart

BTS has 2 main components:

1. Multivariate **timeseries** from real-world building operations

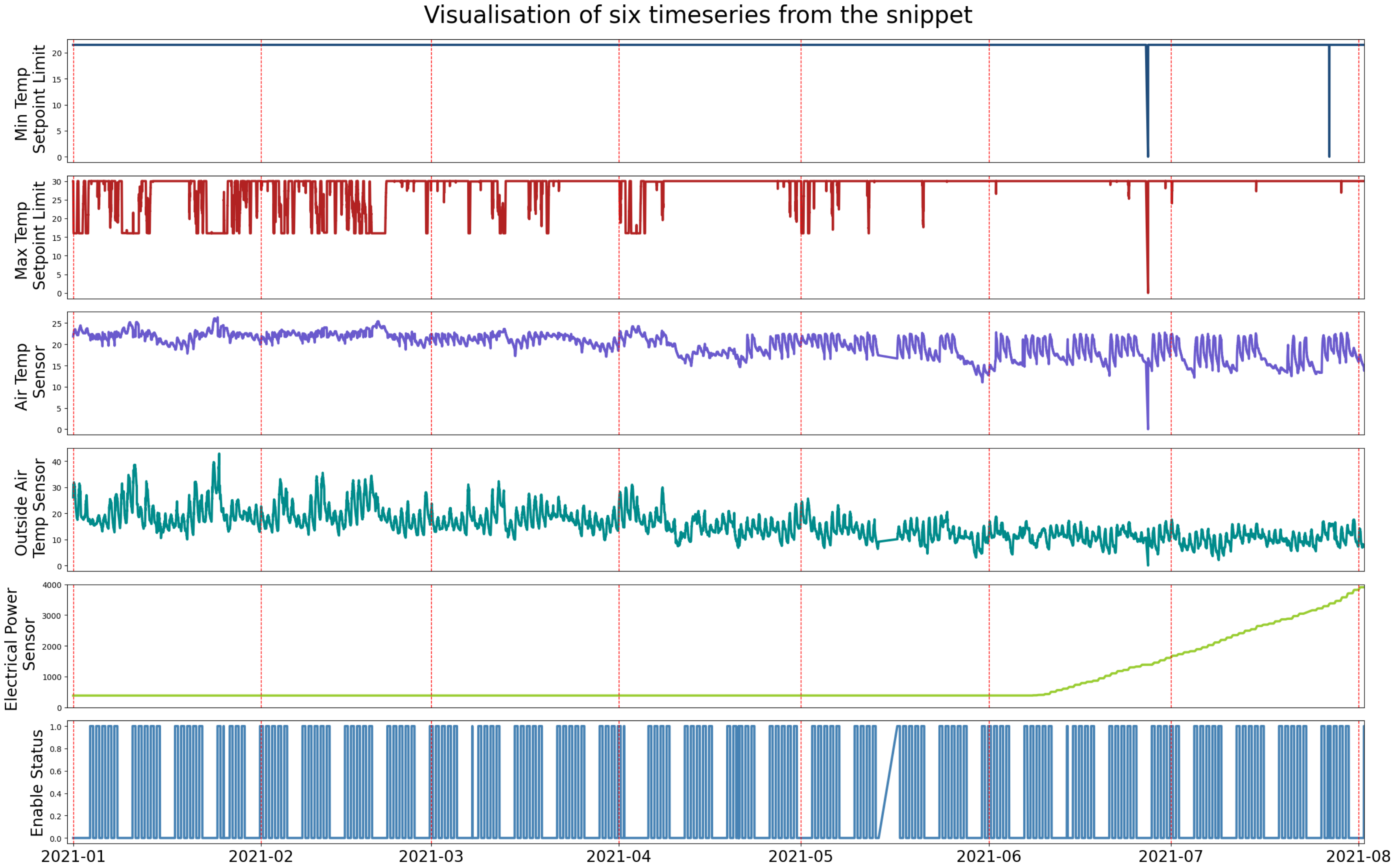
2. Building meta-data in the form of a **knowledge graph**



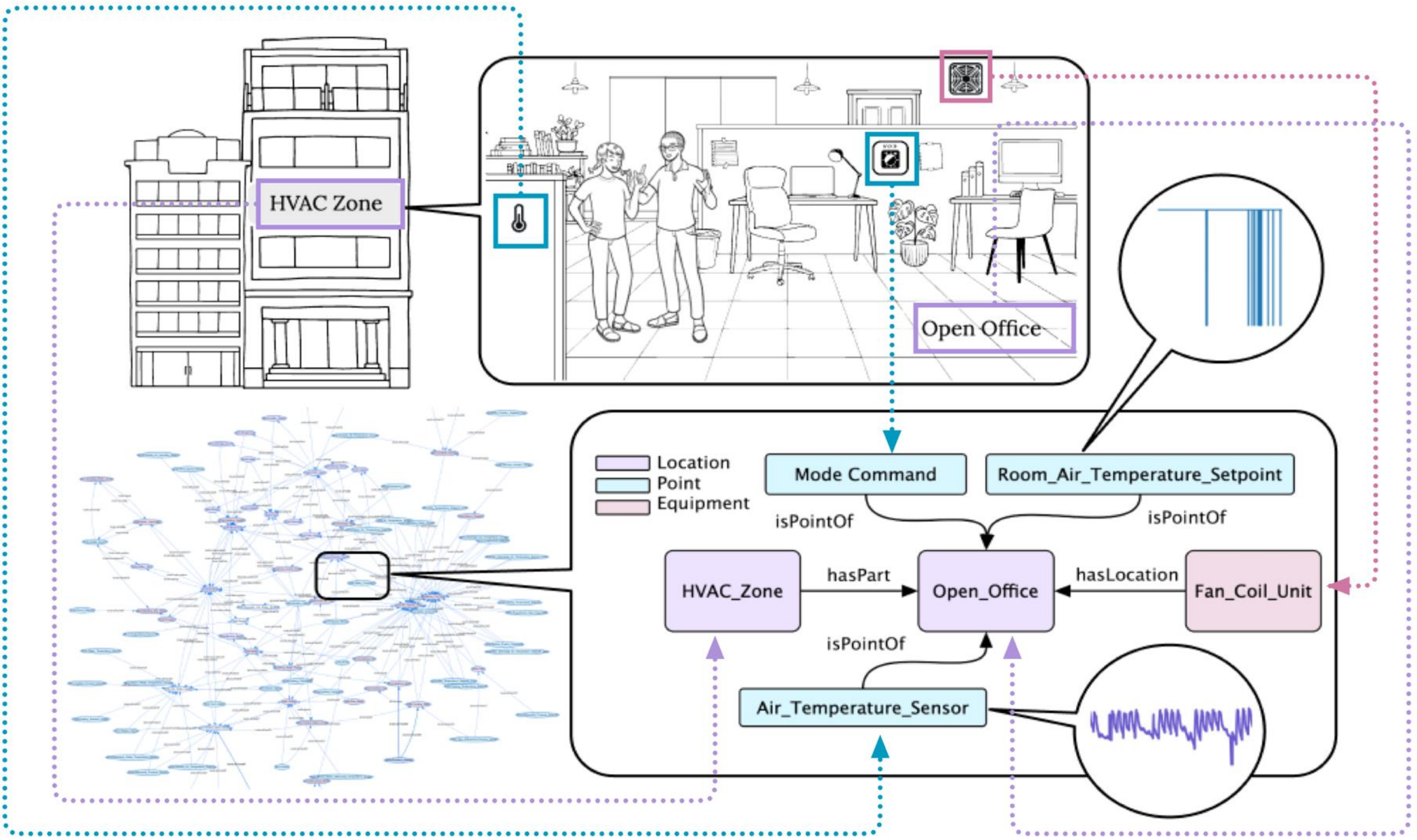
Timeseries: Diverse class and distribution

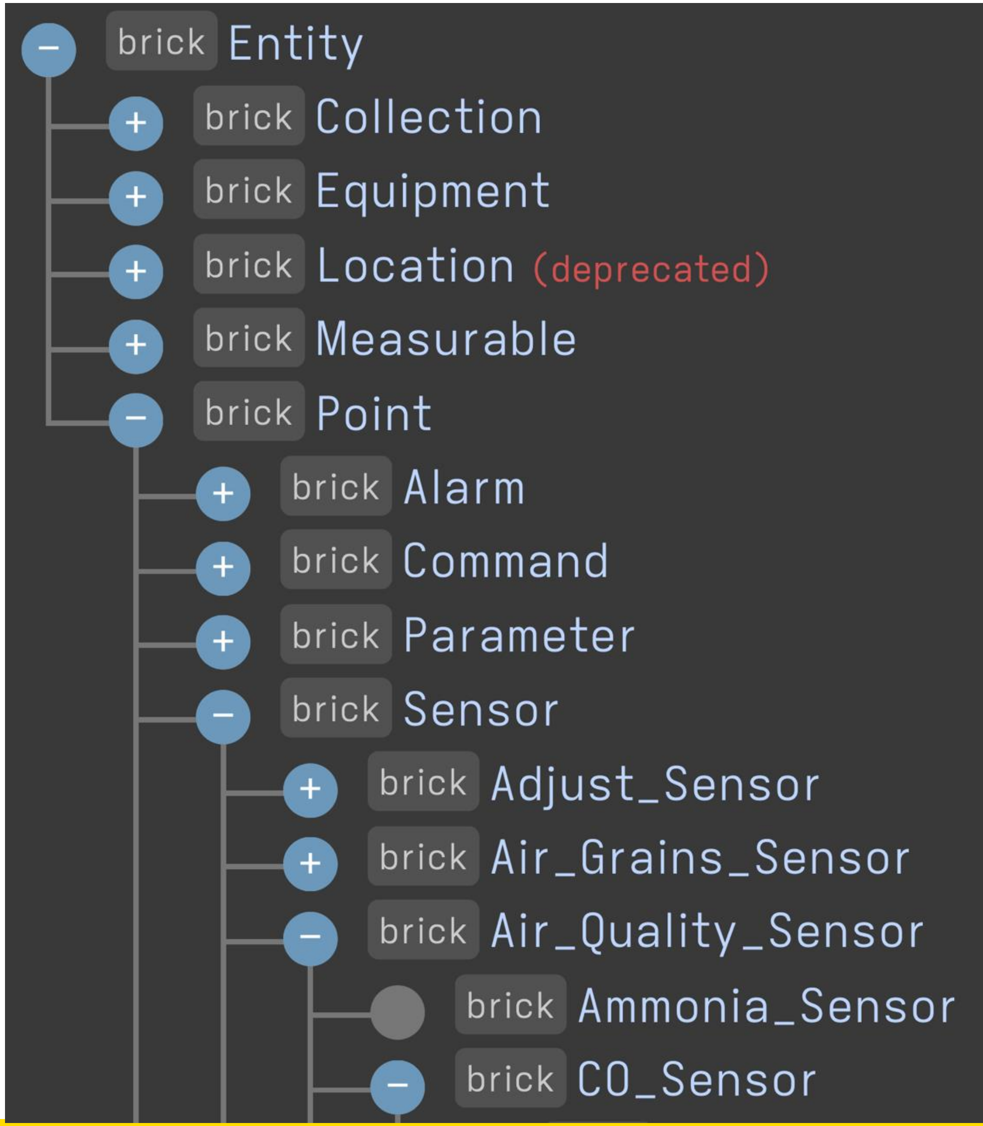
Most other timeseries dataset only have a single class with similar distributions family e.g., traffic, stock market, EEG, IMU.

BTS has few hundreds unique classes.



Metadata Knowledge Graph using a standardized schema





3 new Buildings, enabling: inter-building generalisation

- Transfer learning
- Domain adaptation
- Distribution shifts
- Few & Zero shots

	Count	(Unique)	LBNL59		BTS_A		BTS_B		BTS_C	
Top Level	Collection		0	(0)	4	(2)	2	(2)	8	(1)
	Equipment		59	(3)	547	(24)	159	(25)	963	(41)
	Location		73	(3)	481	(9)	68	(17)	381	(26)
	Point		230	(11)	8374	(126)	851	(57)	10440	(159)
	Timeseries		337		8349		851		5347	
Point Subclass	Alarm		0	(0)	798	(16)	5	(2)	109	(8)
	Command		0	(0)	363	(6)	97	(5)	785	(13)
	Parameter		0	(0)	79	(6)	36	(2)	935	(17)
	Sensor		144	(8)	4396	(56)	266	(25)	4062	(68)
	Setpoint		86	(3)	772	(26)	232	(16)	1629	(41)
	Status		0	(0)	1628	(17)	110	(6)	2187	(19)
	Location		Berkeley, USA		Undisclosed locations in Australia					
	Start Date		01-01-2018		01-01-2021		01-01-2021		23-06-2021	
	End Date		31-12-2020		31-12-2023		31-12-2023		18-01-2024	
	Duration (Days)		1094		1094		1094		939	
	Size Zipped (GB)		0.26		8.48		1.31		8.98	

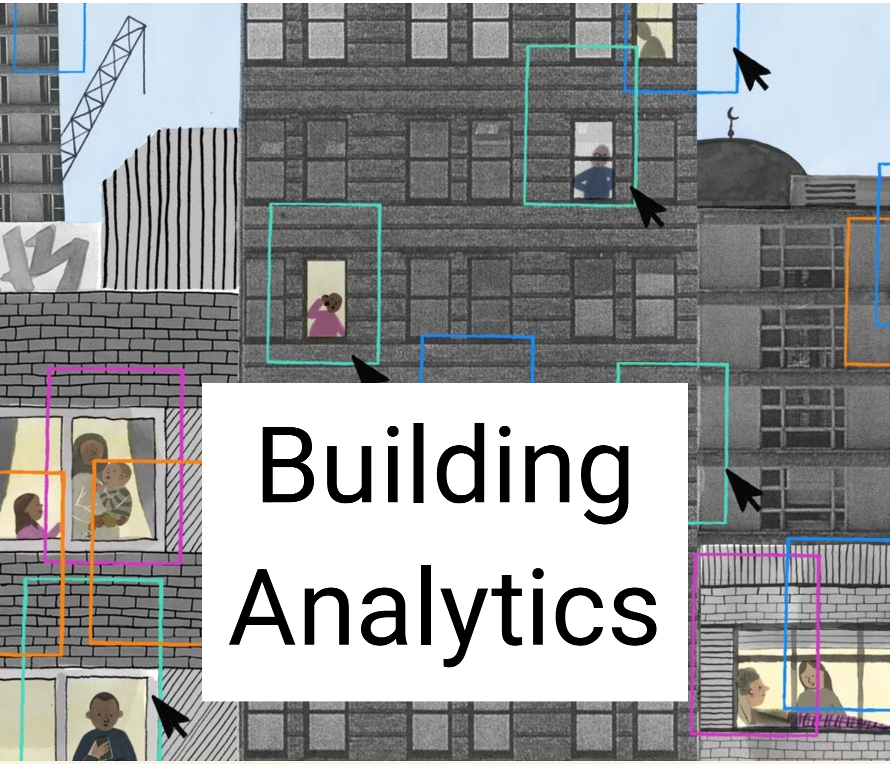
Why use BTS?

1

2

3

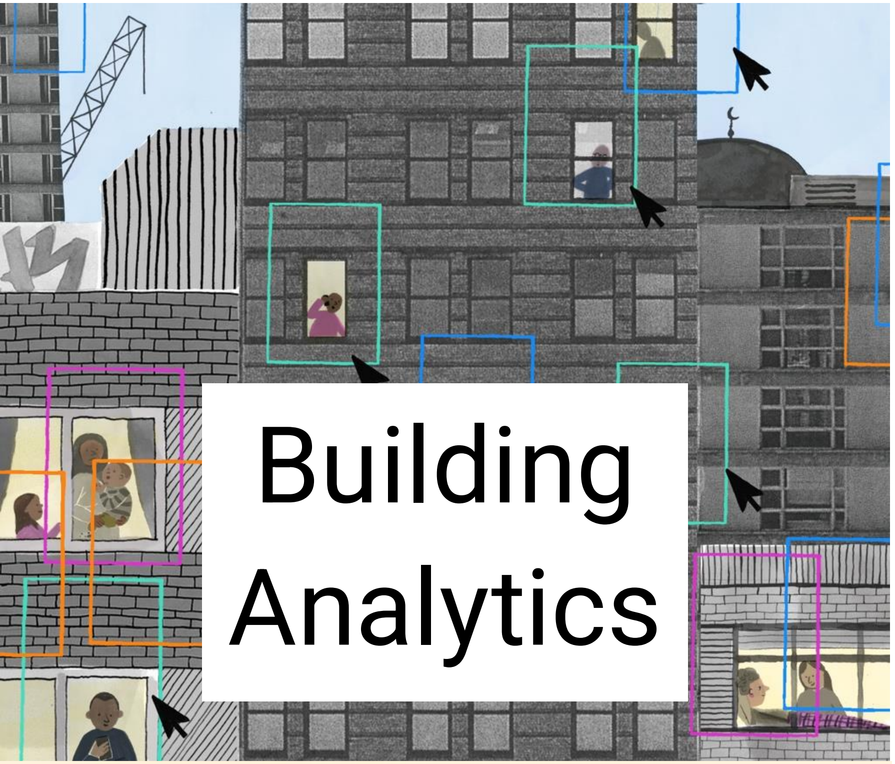
Why use BTS?

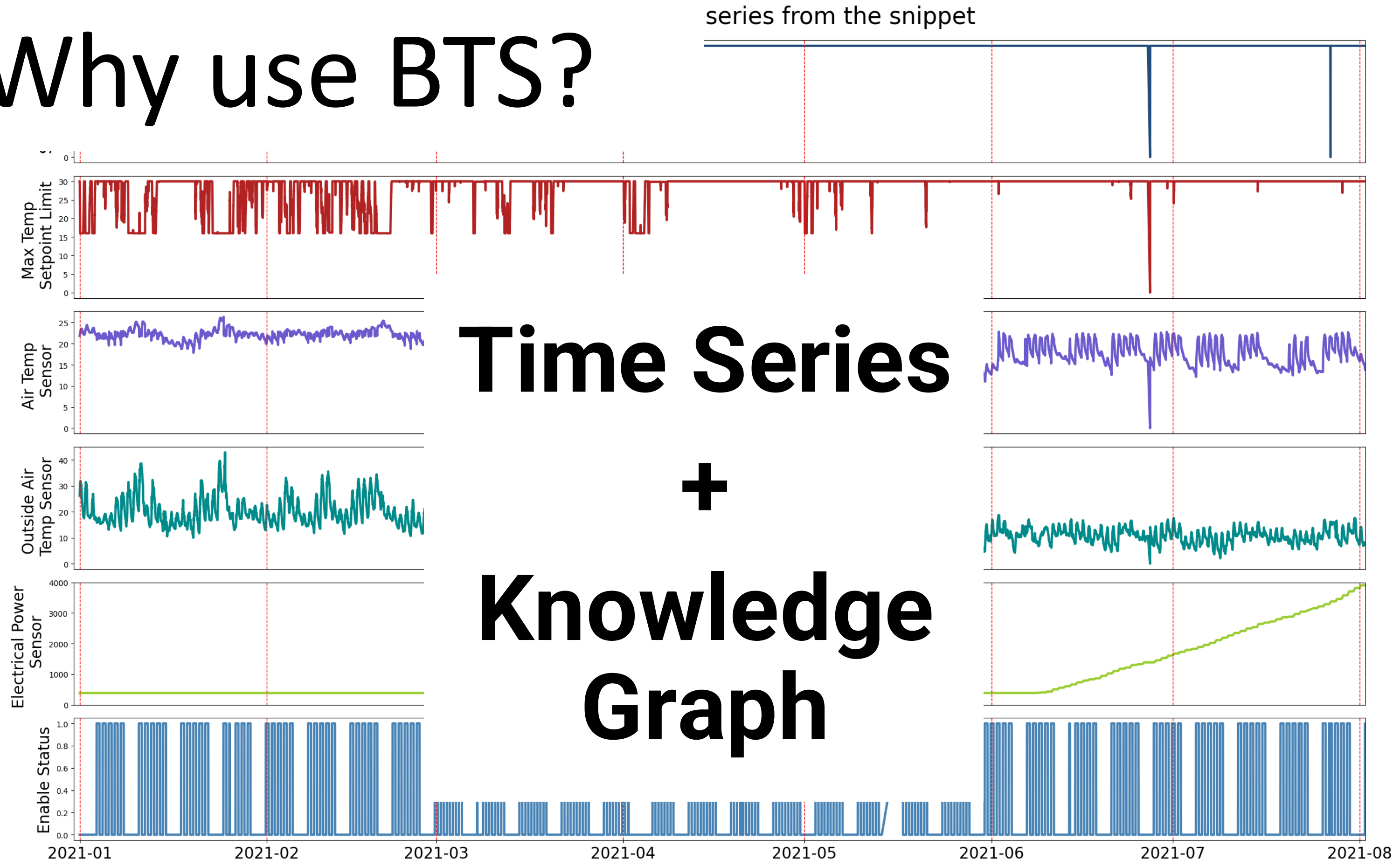


Example tasks:

- Forecasting
- Anomaly Detection
- **AI Chat for buildings:** To enables building managers to get data-driven insights via seamless interface like AI chat.

Why use BTS?

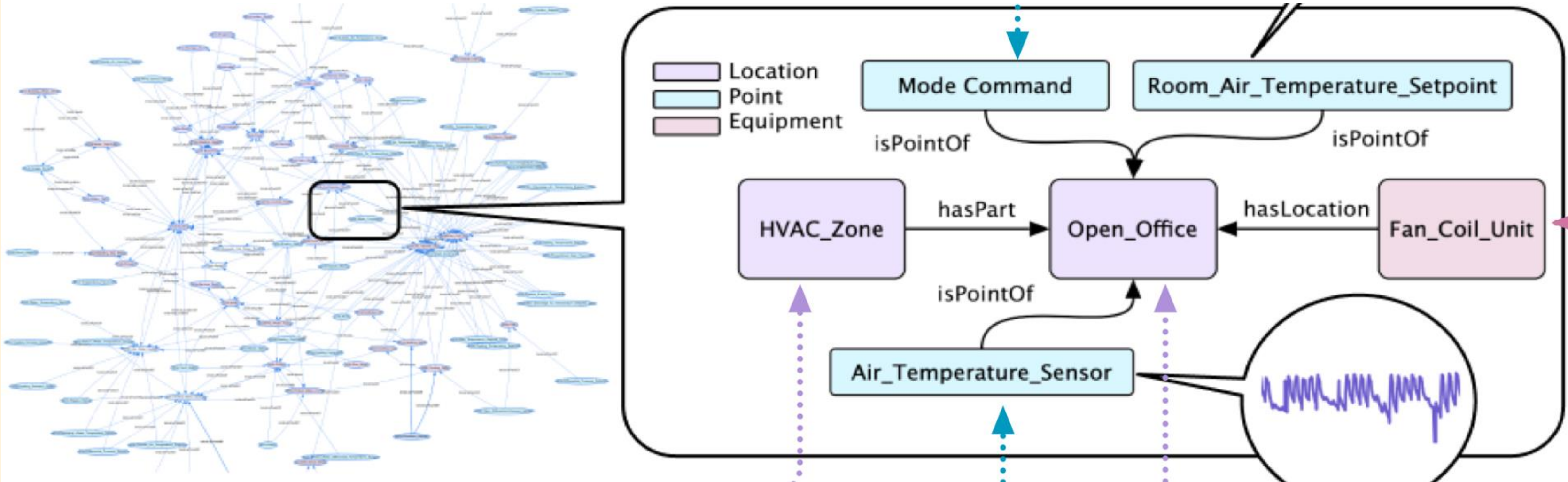




Time Series

+

Knowledge Graph

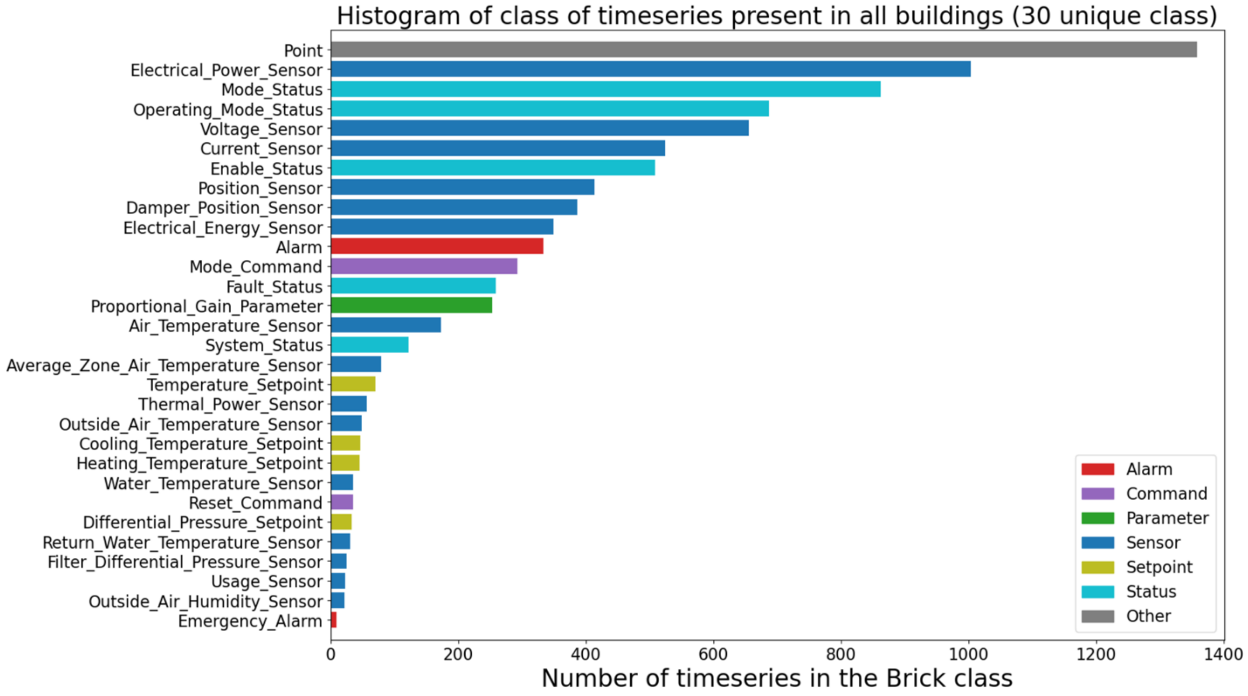


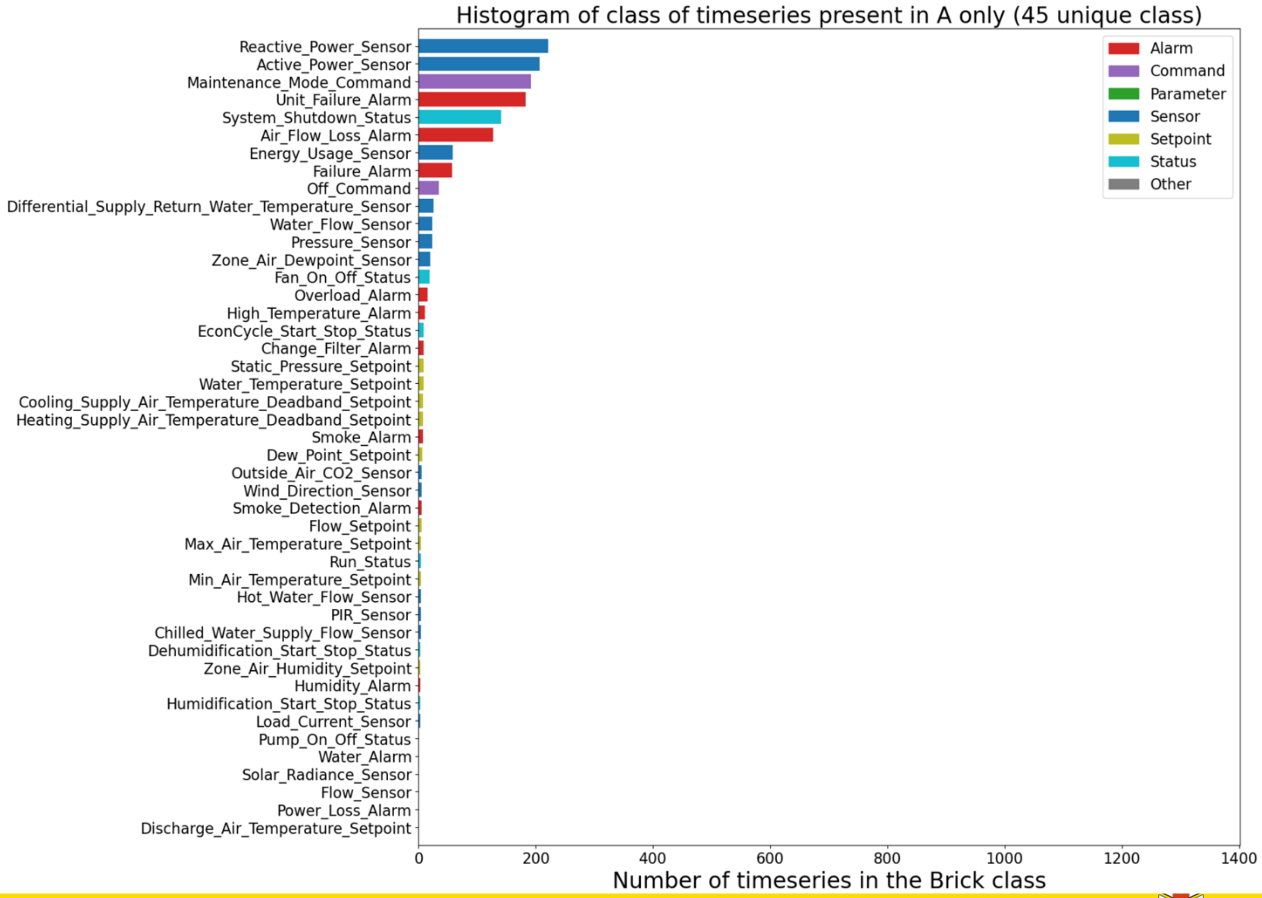
Why use BTS?

- **Unbalanced, Long-tailed**
  - Class: Some sensors (e.g., temperature sensor), are very common, while other sensors (e.g., dewpoint sensors) are very rare.
  - Data: The value for Alarm is zero most of the time.
- **Distribution shifts between buildings**
  - Few-shots, zero-shots, transfer learning, domain adaptation.
- **Irregular**
  - time step
  - any-variates

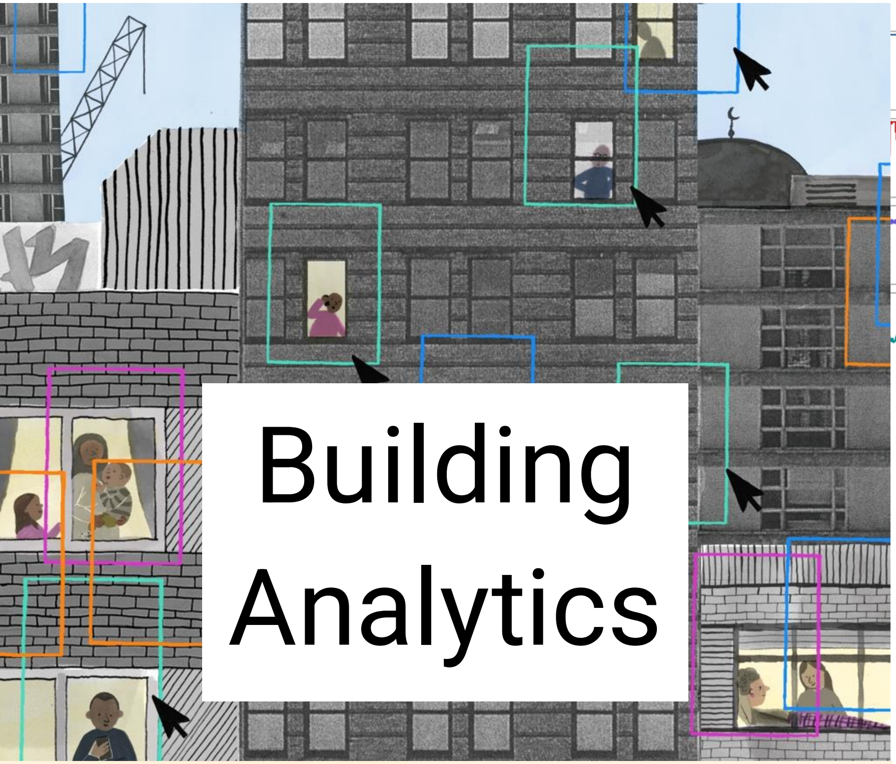
Electric_Power_Sensor have many instances.

But 45 unique classes only exist in Building A

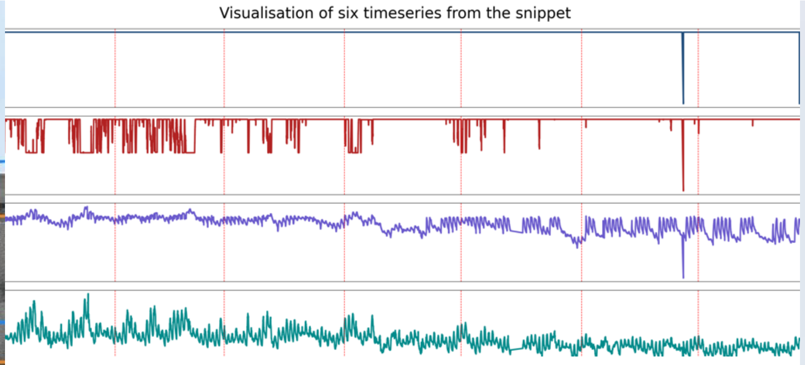




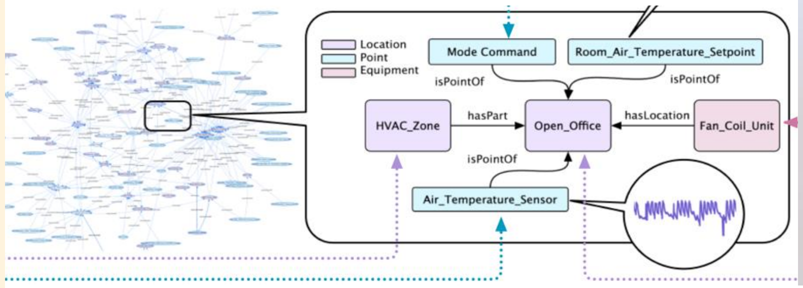
Why use BTS?



Building Analytics



Time Series
+
Knowledge
Graph



Complexities
of
real-world
data
in the wild

Two Benchmarks

Multi-label timeseries classification

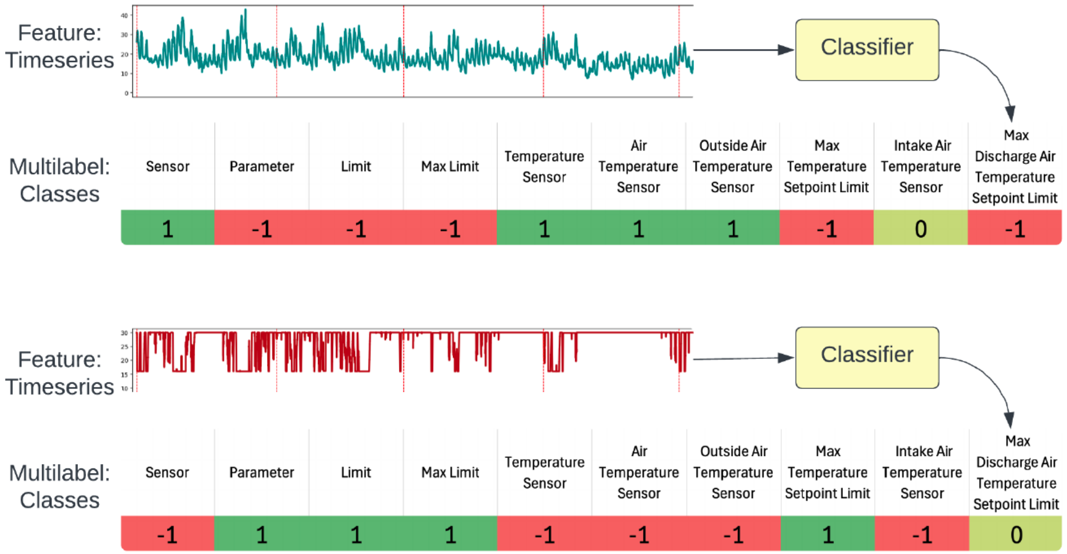


Figure 3: Visualisation of the multi-label timeseries classification task.

Zero-shot Forecasting

Table 5: Benchmark results on the zero-shot forecasting task. The columns refer to the training set, whereas the row represents the testing set.

		BTS-A		BTS-B		BTS-C	
		MAE	SMAPE	MAE	SMAPE	MAE	SMAPE
Previous Day Persistence		0.5377	48.1539	0.4976	43.2985	0.5458	45.7014
Previous Week Persistence		0.6190	57.2713	0.5918	51.3867	0.6499	58.1922
BTS-A	DLinear	N/A		0.4324	35.9846	0.4262	36.2734
	PatchTST	N/A		0.3748	29.2570	0.3712	29.5552
	Informer	N/A		0.5968	49.2217	0.5920	51.9745
	iTransformer	N/A		0.4026	31.1924	0.3842	30.1102
BTS-B	DLinear	0.4940	41.2264	N/A		0.4206	35.3121
	PatchTST	0.4575	36.7689	N/A		0.3711	29.2135
	Informer	0.5233	45.9279	N/A		0.4592	39.7068
	iTransformer	0.4783	37.5907	N/A		0.3901	29.9940
BTS-C	DLinear	0.4858	40.7421	0.4158	34.1473	N/A	
	PatchTST	0.4542	36.9451	0.3723	28.9325	N/A	
	Informer	0.5213	46.6112	0.4602	39.7162	N/A	
	iTransformer	0.4859	39.5158	0.4262	32.6550	N/A	

UNSW SYDNEY

BTS:
Building Time-Series Dataset

Thank you

Official repo:

https://github.com/cruiseresearchgroup/DIEF_BTS/

We will launch a competition using this dataset.