



BEDA Accelerator

February 12, 2019





UBID Matching Washington, DC

Kevin Keene, Nora Wang, Mark Borkum
Pacific Northwest National Laboratory



PNNL is operated by Battelle for the U.S. Department of Energy

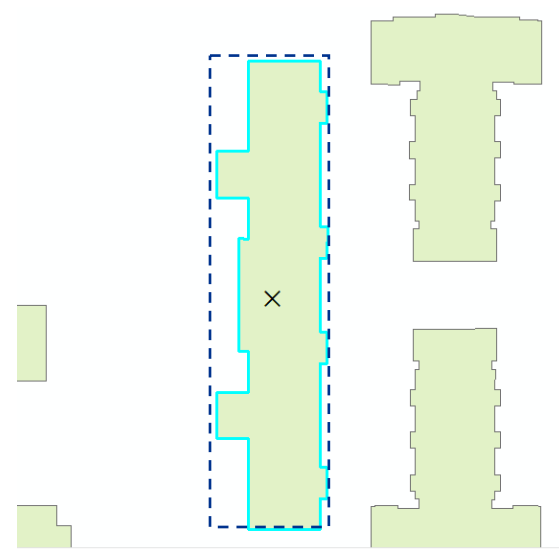


Datasets

- Open Data Footprints
 - http://opendata.dc.gov/datasets/a657b34942564aa8b06f293cb0934cbd_1
 - 163,467 entries
 - No local ID (“GIS_ID” field empty)
- Energy Benchmarking 2017
 - <https://doee.dc.gov/publication/2016-building-benchmarking-dataset>
 - 1,846 entries
 - pid, dc_real_pid, and pm_pid are local IDs, pid only with no duplicates and value for every entry
- Other datasets used for analysis:
 - Street Centerlines
 - <http://opendata.dc.gov/datasets/street-centerlines>
 - Address Points
 - <http://opendata.dc.gov/datasets/address-points>

UBID Matching Process

- 0 invalid geometries from footprints and benchmarking
 - All features were compatible
- 1,846 benchmarking points in total
 - 1,608 benchmarking points (UBID₀) matched to footprint UBIDs
 - 191 duplicate UBID₀ created (2 or more UBID₀ with same geolocation)
 - 238 UBID₀ not matched



Benchmarking Point

PID: PM05823132

87C4VXJM+452-0-0-0-0

Footprint (Bounding Box)

87C4VXJM+456-29-5-29-6

UBID₀ is used for intermediate matching step between points and polygons

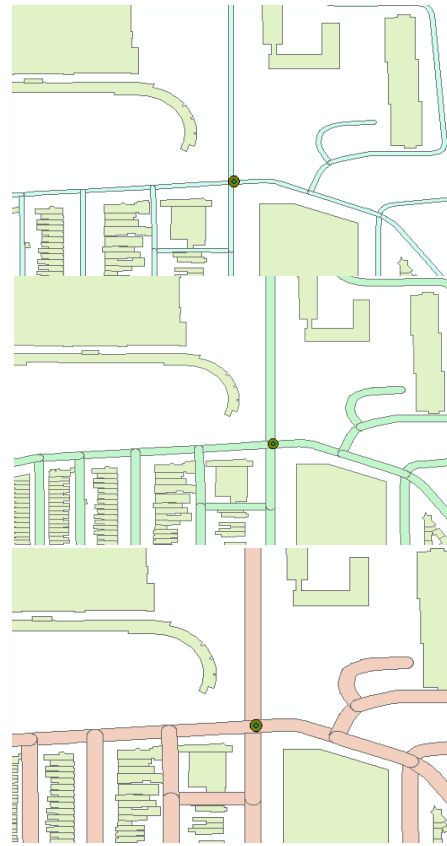
MATCH

Summary

Overview:

- Unmatched
 - Street
 - Property with multiple buildings
- Matched
 - Verified incorrect
 - ✓ Match not totally representative
 - Verified correct
 - ✓ Address/local ID/parcel/etc. correspond
 - Unverified
 - ✓ Unverified but correct
 - Address/etc. don't correspond or not available
 - ✓ Unverified incorrect
 - (Unknown cases)

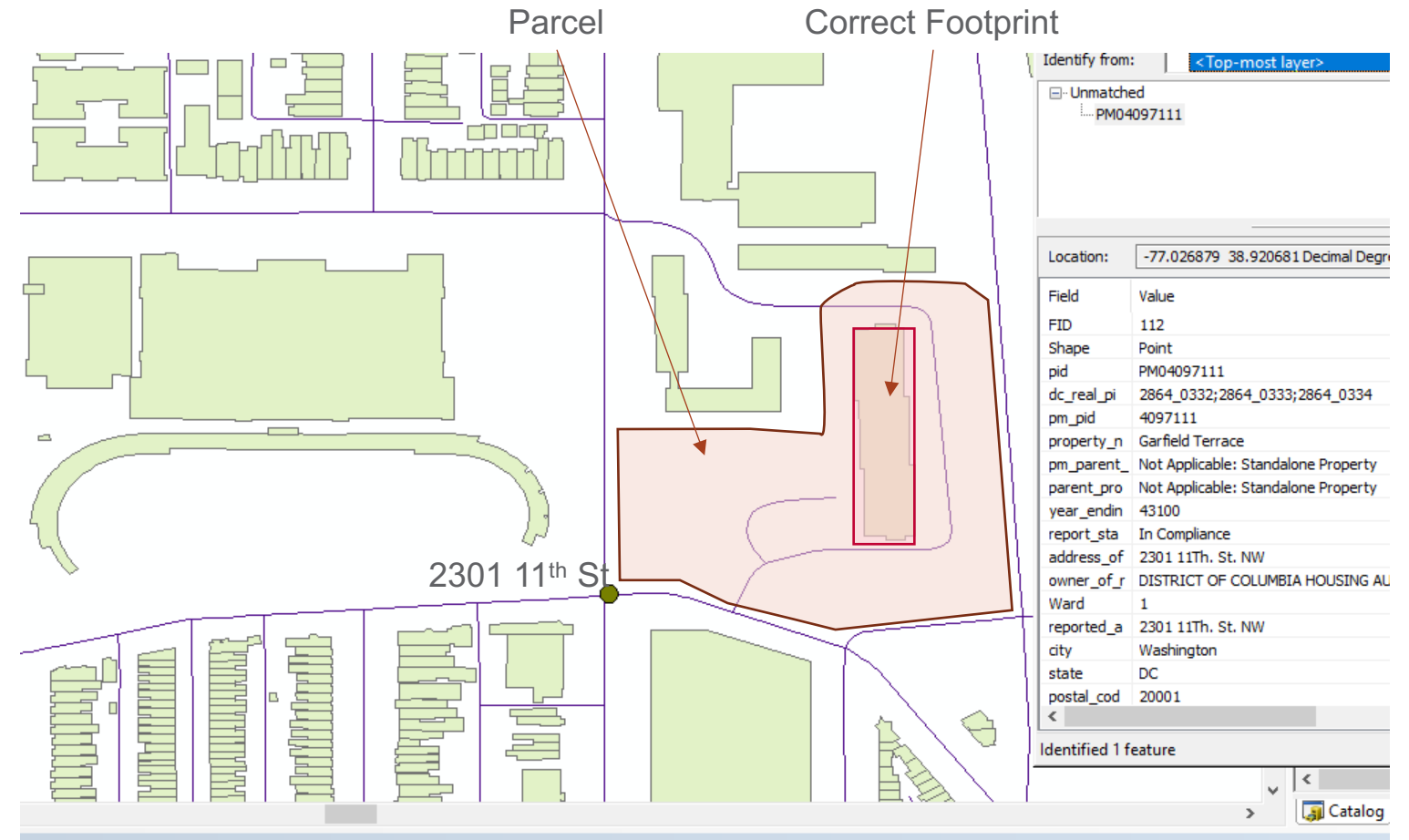
Unmatched: UBID₀ in Street



- 5ft buffer: 32 UBID₀
- 10 ft buffer: 66 UBID₀
- 15ft buffer: 71 UBID₀

Note: streets are specified as alley, avenue, etc – can make different buffers size for each (buffer sizes above just a proof of concept)

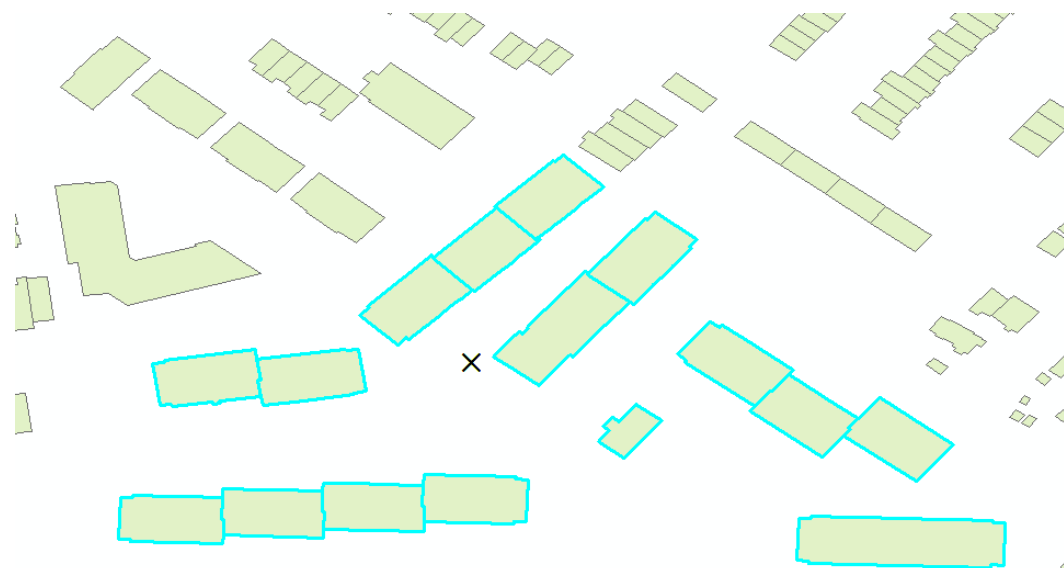
- Detect using street shapefile and buffer zone



Recommended Solution: Use address matching to match majority of cases, manual matching for what remains of the ~70 cases

Unmatched: One UBID₀ represents multiple buildings

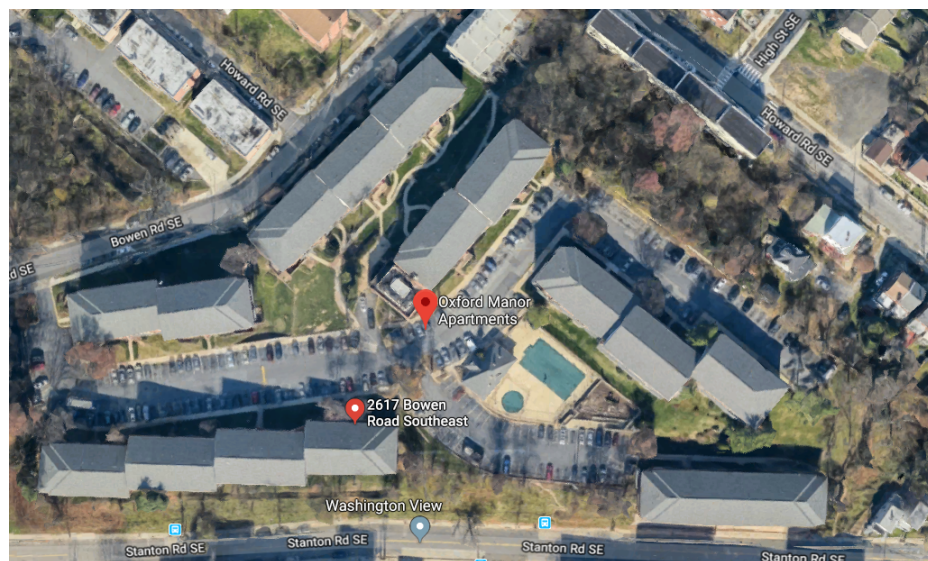
(About 140 instances)



- Detection: all unmatched cases that aren't in street (and have multiple buildings on property)

Example:

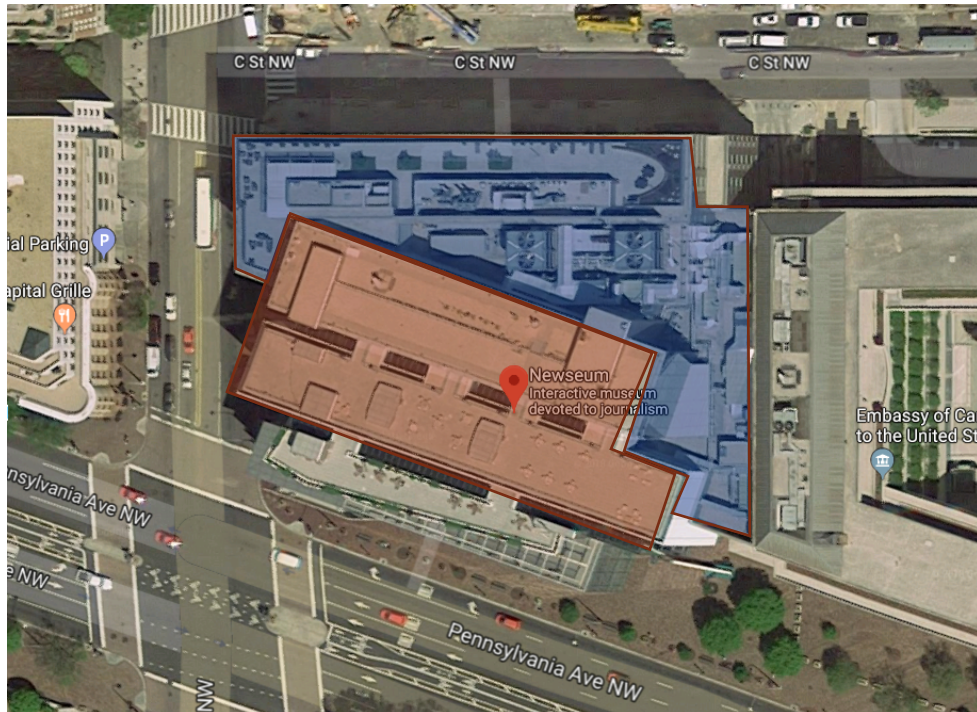
- Estimated area (with area map tool): 77k * 4 floors = ~308k sqft
- Reported area = 280k sqft
- Conclusion: Benchmarking data represents all buildings in this multi-family housing unit but didn't match because fell outside bounding box of all footprints



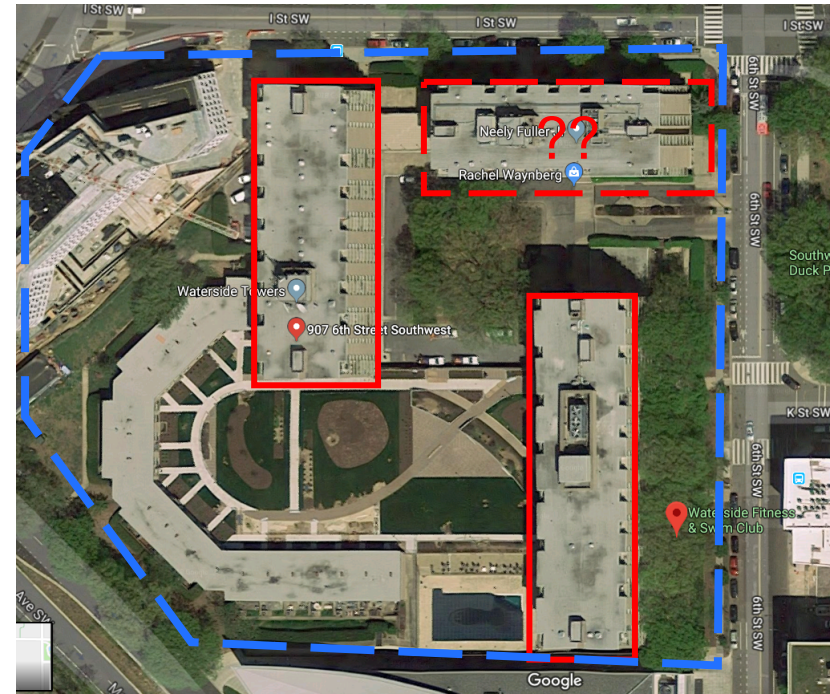
Recommended Solution: If parcel footprints are available, match to corresponding parcel footprint UBID (parent-child)

If parcel footprint isn't available, can use lot ID from address data to detect all footprints with matching lot ID (more data processing involved)C

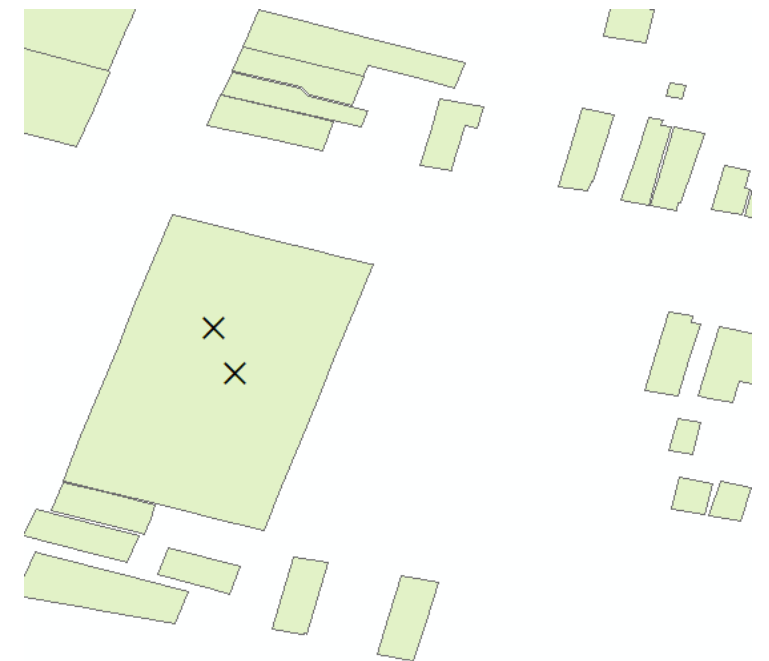
Verified Incorrect: Matched but not totally representative



UBID₀ represents one space in a building footprint



UBID₀ represents multiple buildings, but not all buildings, on a property



Multiple UBID₀ represents various spaces in a building footprint

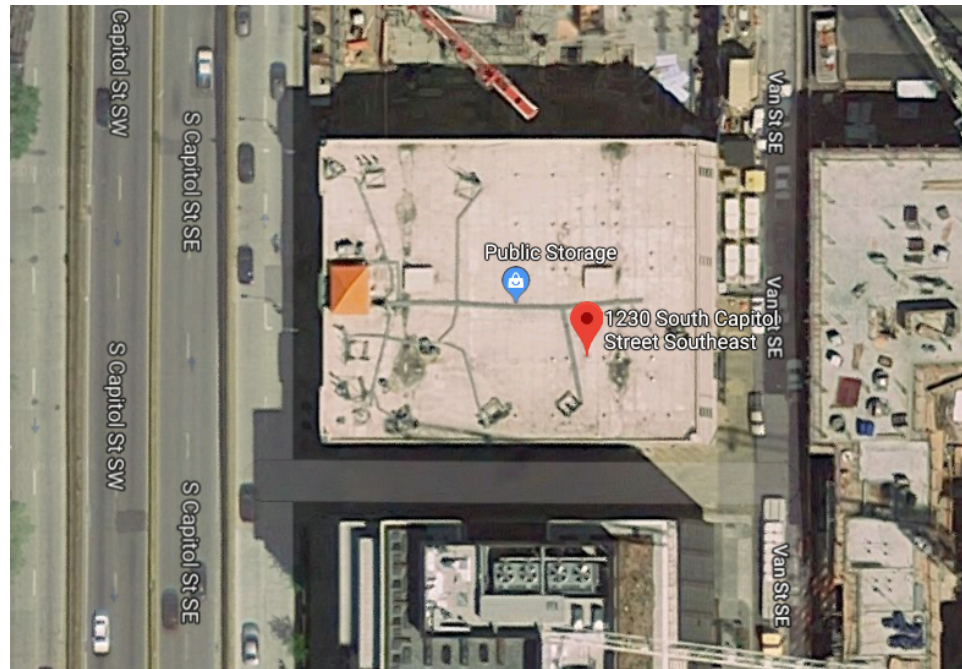
Recommended Solution: No way to determine which buildings or spaces correspond to benchmark, even manually
Match to the corresponding footprint or property and flag as not totally representative (if able to detect)

Verified Correct and Unverified Correct

Address Point	Benchmark Address (Reported + Tax)	
AA	BV	BX
FULLADDRESS	address_of	reported_a
3600 GEORGIA AVENUE NW	3600 GEORGIA AVE NW	3600 Georgia Ave NW
1475 EUCLID STREET NW	1475 EUCLID ST NW	1475 Euclid Street N.W.
1515 OGDEN STREET NW	1515 OGDEN ST NW	1515 Ogden Street NW
3900 16TH STREET NW	3900 16TH ST NW	3900 16th Street NW
7420 GEORGIA AVENUE NW	7420 GEORGIA AVENUE NW	7420 GEORGIA AVENUE NW
3109 14TH STREET NW	1390 KENYON ST NW	1390 Kenyon St NW
7820 GEORGIA AVENUE NW	7820 GEORGIA AV NW	7820 Georgia Ave
3239 14TH STREET NW	1375 KENYON ST NW	1375 Kenyon St NW
3025 ONTARIO ROAD NW	3025 ONTARIO RD NW	3025 Ontario Road N.W.
1712 NEWTON STREET NW	3322 MOUNT PLEASANT ST NW	3322 Mount Pleasant Street NW
2400 6TH STREET NW	2400 6TH ST NW	2400 6th St. NW
2600 6TH STREET NW	2400 6TH ST NW	100 6th and Trumbul St.
611 EDGEWOOD STREET NE	0611 EDGEWOOD ST NE	611 Edgewood Street Northeast
6925 WILLOW STREET NW	6856 EASTERN AV NW	6856 EASTERN AV NW
3228 GEORGIA AVENUE NW	3232 GEORGIA AVE NW	3232 Georgia Avenue Northwest
1631 EUCLID STREET NW	1631 EUCLID ST NW	1631 EUCLID ST NW
1629 COLUMBIA ROAD NW	1629 COLUMBIA RD NW	1629 Columbia Road
3336 14TH STREET NW	3300 14TH ST NW	3300 14TH ST NW
4501 KANSAS AVENUE NW	4501 KANSAS AVE	4501 Kansas Ave NW
915 GALLATIN STREET NW	915 GALLATIN STREET NW	915 GALLATIN STREET NW
2455 6TH STREET NW	2440 6TH ST NW	2455 6th Street NW

- Some addresses will match well and signify a true positive (with high certainty)
- Other addresses don't match, but could still be unverified correct or be unverified incorrect
 - Not easily distinguishable

Unverified Incorrect: Incorrect Location

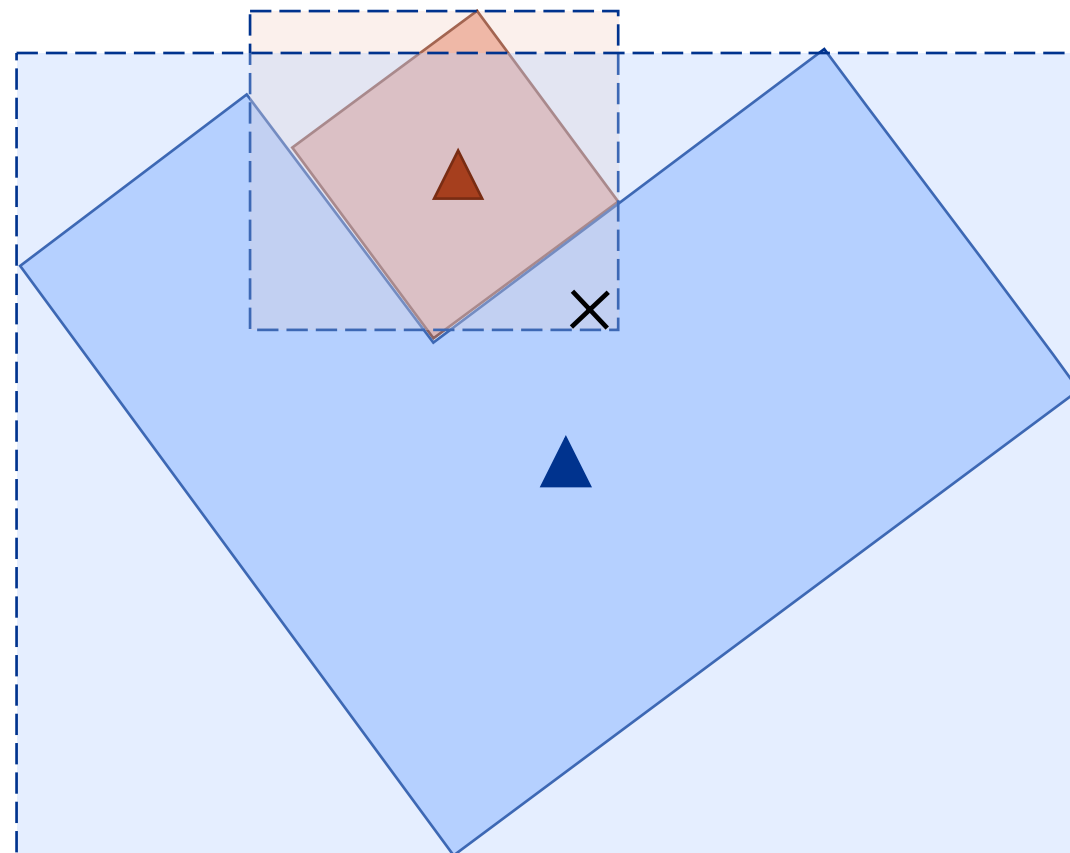


- Example:
 - Listed address: 1263-1265A A 1st St SE
 - Address corresponding to location: 1230 S Capitol SE (a few blocks away)
- Conclusion: Incorrect coordinates entered for second address
- This was caught by chance, unclear how many instances exist

	A	I	L	O	P	Q	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AL	AP
1	pid	address_of_record	reported_address	postal	year	primary_p	tax_rec	reported	energy	site_eu	weather	source	weather	total_g	total_g	water	water	electric	natural	UBID	
5	PM05932679	1263-1265 1ST ST SE	1263-1265A A 1ST ST SE	20003	2015	Hotel	118944	118944	85	54.2	54.2	131.9	131.9	589.2	5	3065.4		1249441	21889.93	87C4VXGR+6HF-0-0-0-0	

Unverified Incorrect: Incorrectly Matched

Hypothetical Example:



- If benchmarking geolocation isn't well aligned with center of the bounding box of the footprint, there is a chance it could be falsely matched to neighboring footprints

- × Benchmarking Location
- Footprint
- Footprint Bounding Box
- ▲ Footprint Bounding Box Centroid

Currently not detected, but further UBID analysis could help detect this:

- Compare 1st and 2nd centroid distance
- Large bounding box to footprint ratio
- Large distance between UBID₀ and centroid

Manually inspect results and fix

Recommended Solutions – Long Term

- Future benchmarking data:
 - Ask reporters to quickly verify address and footprint visually
 - For entries that have multiple buildings or subspaces, in order to be compliant, need to have way to determine which spaces it corresponds to
 - ✓ Have field with critical data like: number of buildings, addresses of all buildings, does it represent entire building Y/N, etc.

Thank you



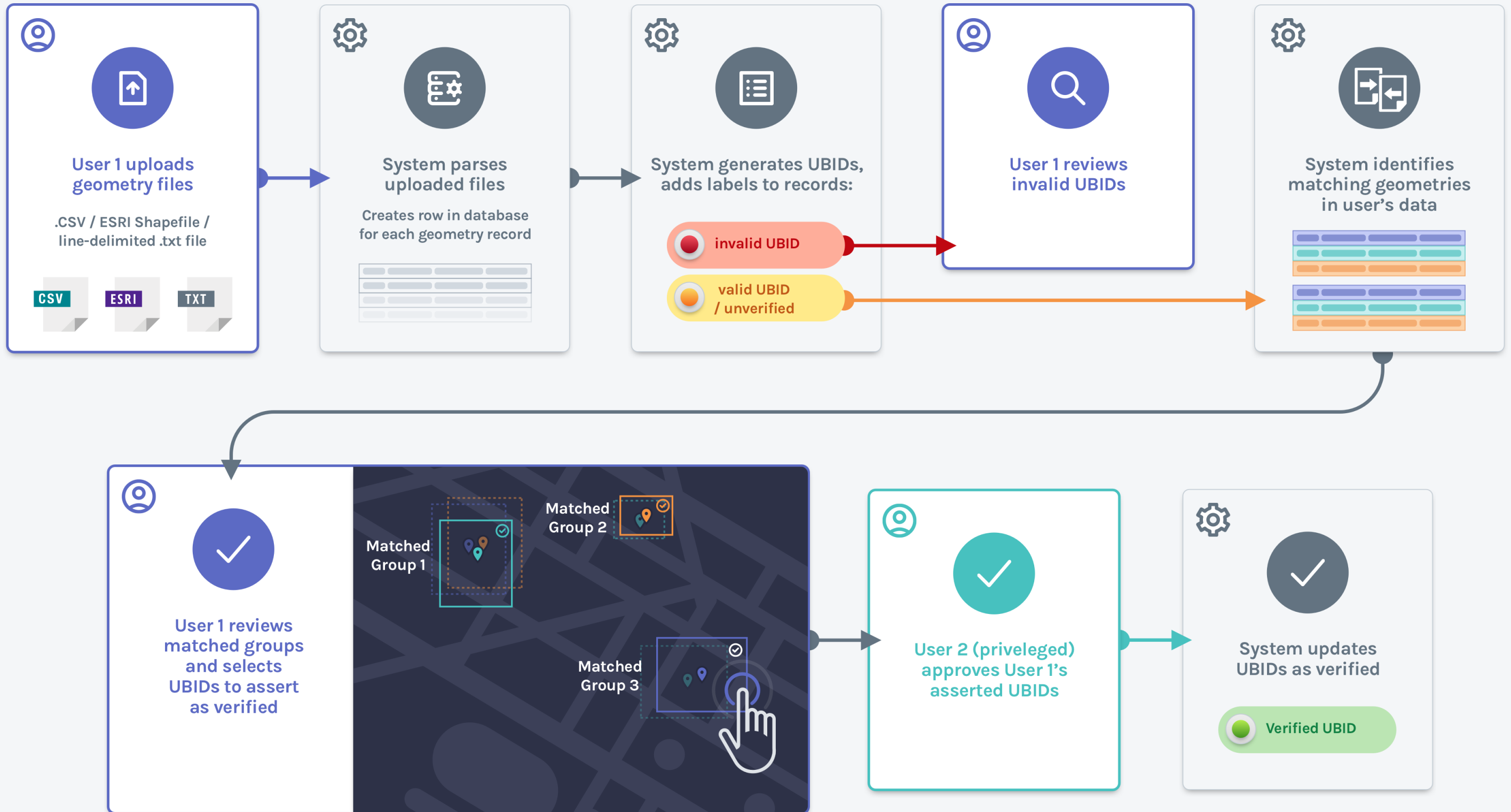


How USGBC is using UBIDs

Chris Pyke



Authoritative UBID Process





Thank you!

Next call:

March 12th, 2019

