CQA

James Stegeman
President – CEO

CostQuest Associates, Inc
1430 E. McMillan St.
Cincinnati, OH 45233

Jstegeman@costquest.com
513-662-2124 x104

November 16, 2018

Ms. Marlene H. Dortch
Secretary
Federal Communications Commission
445 12th Street SW
Portals II, Room TW-A325
Washington, DC 20554

RE: Modernizing the FCC Form 477 Data Program, WC Docket No. 11-10; Connect America Fund, WC-10-90

Dear Ms. Dortch:

On November 15, 2018, James Stegeman, Mark Guttman and Jesse Amundsen of CostQuest Associates, Inc. (CostQuest) met by telephone with the following Commission Staff: Steve Rosenberg, Roger Woock, Kirk Burgee and Ying Ke. The purpose of the meeting was to discuss data, methods and procedures needed to support increased granularity in 477 submission of address or locational information and how that information can be used to identify broadband gaps.

As described in the attached slide presentation, CostQuest is supportive of methods described by AT&T –the Cooperative Address-Based Broadband Deployment Database[1]. But if the end-goal of data submissions is to understand structure level broadband availability, additional data sources and methods will need to be included. Address data collection will be an important attribute of any broadband availability data source, but it won't be the only source.

In the meeting, CostQuest described differences in broadband analysis in urban and rural areas. Examples of alternative approaches and data sources were reviewed in the context of their utility to provide information to identify and measure broadband availability. CostQuest believes that there will not be an off the shelf data solution to provide structure locations to identify broadband availability and gaps. Rather, gaining a granular understanding of the broadband deployment gap will involve weaving together complimentary data sources and methods into a locational fabric. The notion of a geospatial fabric [2]has been used in other situations where partnerships among data developers and providers is necessary. CostQuest described and showed examples of data products that may be useful. This includes imagery, parcel and georeferenced address points.

[1] See Letter of Ola Oyefusi, Director Federal Regulatory, AT&T Services Inc., to Marlene H. Dortch, Secretary, FCC, WC Docket No. 10-90, 11-10 (filed October 12, 2018)

[2] See EPA's discussion of Hydrologic Fabric, https://www.epa.gov/waterdata/weaving-national-hydrologic-geospatial-fabric visited 11/14/2018.

CQA

James Stegeman
President – CEO

CostQuest Associates, Inc
1430 E. McMillan St.
Cincinnati, OH 45233

Jstegeman@costquest.com
513-662-2124 x104

CostQuest appreciates the opportunity to work with the Commission and contribute information to help identify the data sources, tools and methods to identify broadband gaps. As required by the Commission's rules, this *ex parte* record is now filed in the above referenced dockets.

Sincerely,

/s/
James W. Stegeman,
President / CEO CostQuest Associates, Inc.

Attachment: CQA Location Fabric.pptx

cc: Steve Rosenberg
Roger Woock
Kirk Burgee
Ying Ke

Location Fabric

CostQuest Associates

Nov. 2018

Today's discussion outline

- A more granular view of broadband deployment
- Review current state and potential future states
- Developing the Address Catalog
- Integrating the Address Catalog with locations
    - Ideal is to have near 100% of locations where homes and businesses are located - a "Location Fabric"
- Review of the current data sources
    - Review Sussex, DE (Buildings, Addresses, Parcels, Roads)
- Start with a proof of concept

Challenges of ubiquitous broadband

Urban Challenges

- Building qualification
- Untangling ownership
- Cost of construction
- Zoning / permitting impediments

Rural Challenges

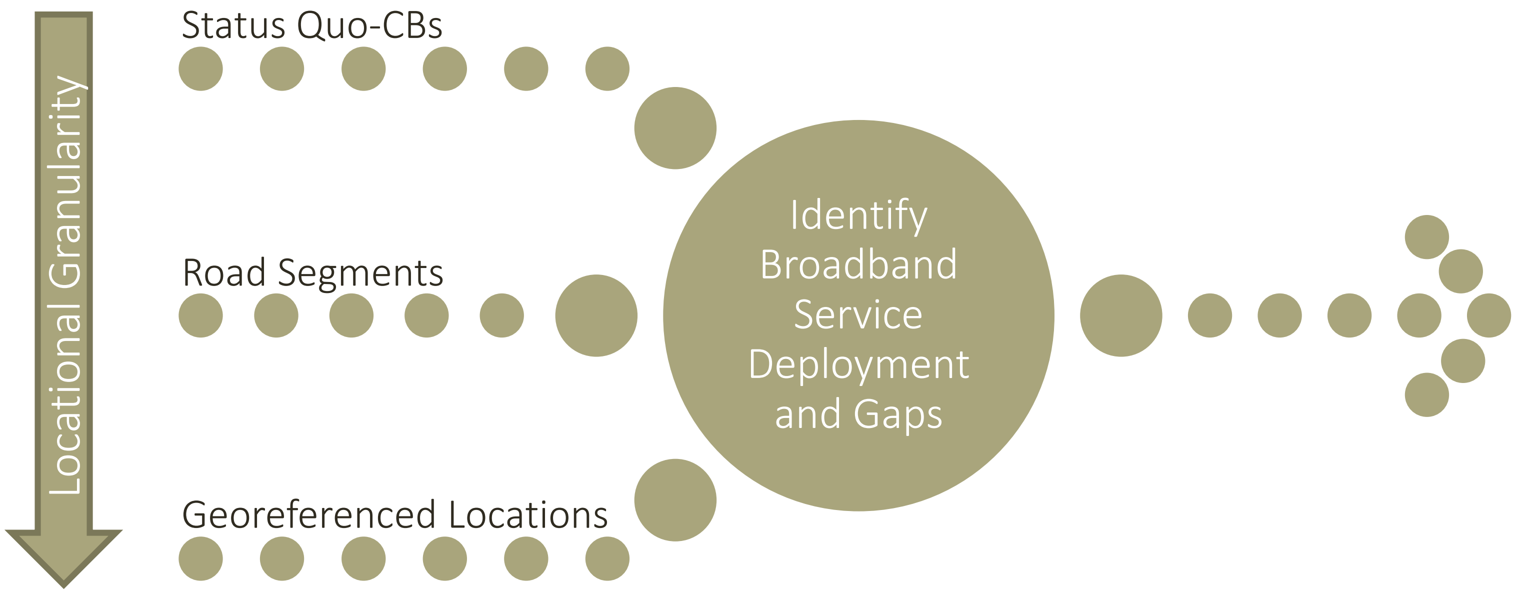
- Provider identification
- Market density / take rate
- Network age and quality
- Cost of construction

The current Form 477 approach provides a good understanding of broadband deployment, but if policy needs are favoring a more granular understanding of the broadband gap, the data collection will need to move, as well.

The broadband coverage issue

- We agree with the intent of AT&T's stated issues in its October 12, 2018 filed Ex-parte in regard to *"Modernizing the FCC Form 477 data Program, WC Docket No. 11-10: Connect America Fund, WC 10-90"*
- That is, to better understand and support the deployment of broadband in unserved areas, it will be beneficial to have more granular information about the location of homes and businesses to identify, first, where broadband is available and, second, to understand where broadband needs to be deployed
  - With this more granular information, parties would be able to more accurately
    - Identify gap locations
    - Measure the cost to serve those locations
    - Monitor the progress toward a goal of ubiquitous broadband
- However, we believe there are some adjustments that could improve the AT&T suggested approach

Improving the granularity



Finer locational granularity supports new approaches to understand, measure, and address the broadband gap

Form 477 Collection: Current State to Potential Future State

Current 477 state

- Uses Census Blocks to identify where broadband can or could be provided for fixed home and business locations
  - Requires identification of any part of census areas where existing service can be provided
  - Can create a one served-all-served issue
    - Carriers are not required to be able to serve all locations in the Census Block

Potential 477 road segment approach

- US Census provides TIGER road segments
- Road segments integrate with other Census products and are attributed so you can understand
  - Type of road
  - Type of block (urban / rural)
  - Address range
  - Demographics
- ISPs could identify the segments that they can or could serve upon request
- Advantage is a more granular data source that diminishes the one-served / all-served concern for a block
- Disadvantage is carriers may not link their outside plant information to road data sources…a geospatial process must be developed
- AND, whether a road is considered served or not does not necessarily inform the measurement of whether residential and business locations are unserved or not
  - UNLESS, you know where the residential and business locations are and if that location is covered

Potential 477 georeferenced location approach

- Create a national georeferenced location dataset – “Location Fabric”
- ISPs could identify the service locations they serve
  - Or the road segments upon which homes and/or businesses are located AND are entirely served
- Advantage is the ability to gain a granular identification of the served/unserved issue, assuming the ISPs file correctly and that a Location Fabric can be made available
- Disadvantage is three-fold
  - A national georeferenced location dataset - “Location Fabric” - does not currently exist
  - ISPs currently do not have a uniform basis to provide georeferenced service locations
  - Locations change, the data set is a snapshot in time and will need to be maintained

… Developing the address catalog

AT&T's cooperative address approach

AT&T's suggested cooperative address-based broadband deployment database

1. ISPs provide addresses of current and former customers
2. Validate address form, de-duplicate, standard geocode, exception report
3. Open crowdsource for corrections, additions and removals
4. ISPs use the cleaned database to identify where they serve. ISPs note where there is fallout

AT&T's cooperative address approach

- If location level reporting granularity is required, the cooperative address approach will need to be augmented
  - Address level data presents issues and by itself is only part of the answer
    - Issues:
      - Addresses need to be standardized, de-duplicated
      - Addresses from providers will not likely represent a full census of locations
      - Addresses do not provide a location until georeferenced – and today the georeferencing occurs through geocoders, which are:
        - Not consistent between vendors
        - Rely on data sets that are incomplete, not focused on rural America, and not necessarily built for the georeferencing of potential broadband demand locations
      - Addresses
        - Are not necessarily unique
        - Change
        - Are not assigned to every building; may or may not correspond to Census enumerated housing units
      - Validity
        - Which addresses are valid is subjective...validity may differ among service providers and FCC
        - Is an address without a structure valid (never occupied but serviceable)?
        - Is an address without a marketable customer valid (not occupied)?
  - Crowdsourcing (i.e., human intervention) needs to be structured, with a level of expected accuracy

…Integrating the address catalog with locations

A single method or source won’t tackle all the challenges

- A successful approach for location level granularity requires
  - Uniformity
  - Acknowledgement that error free is not economically achievable
    - With geographic data, error is scale dependent
      - Accuracy in maps of large areas is different than accuracy in small areas
  - Acknowledgement that a 100% complete dataset is not achievable
    - In part, what is the control against which the completeness is measured?
  - Input from users and service providers
  - Development of an open, self-correcting, learning process
    - “Wikipedia”-Like / Curation
  - A structure and set of methods that will improve over time
  - Geospatial ‘fabrics’ are used when multiple data sources, methods and parties come together to address a common challenge
    - These collaborative structures have and do work
      - EPA: Hydrography < https://www.epa.gov/waterdata/weaving-national-hydrologic-geospatial-fabric >
      - Australia: Hydrography < http://www.bom.gov.au/water/geofabric/>

The Location Fabric

- Identifying and remedying the lack of broadband at a granular level will require information to inform what an address means
  - Addresses are important, but by themselves, they are not objective locations
  - And, what about new addresses, addresses never served, addresses in areas with no providers?
- Rather than a view of competing concepts of a Location Fabric and Address Catalog, the two are complimentary requirements to achieve the goal
  - Georeferenced locations in the Location Fabric are needed to provide a granular view of broadband areas and gaps
  - Addresses are needed to give the locations a frame of reference and bridge to other systems
- To create the "Location Fabric", multiple data sources, scoring routines, and a managed visual review plan are required
  - Data sources include: Parcels, georeferenced addresses, georeferenced building locations, roads, and more
  - Scoring provides a level of certainty
  - Managed Visual Review process can be used in areas of uncertainty

The importance of the Managed Visual Review

- While scoring routines against various forms of data can guarantee some certainty, there will be areas of the country that need a review by a human to provide certainty
- Our Managed Visual Review method is a compliment to AT&T's suggested crowd source approach
  - Managed Visual Review is a process of using various managed human resources (including crowd labor) to visually inspect, and/or review specified data
    - Can be used in areas of uncertainty and provide an acceptable quality level
    - Can be used to test overall quality
    - Can be used to form the basis of machine learning
- Caveat: while Managed Visual Review is critical to addressing uncertainty, it needs to be weighed against potential cost
  - Higher certainty pushes the need for greater review
  - Components of potential cost:
    - ~21 million: 400M*400M grids in the U.S. with any type of road
    - ~16 million: 400M*400M grids in the U.S. with likely home or business location
    - ?? Number of uncertain grids to review
    - ~$1.25 - $2.50: Cost to apply Managed Visual Review of a single grid
- Note: any manual human modification needs to result in traceable actions

Review of Potential Data Sources

Location identification: Data sources

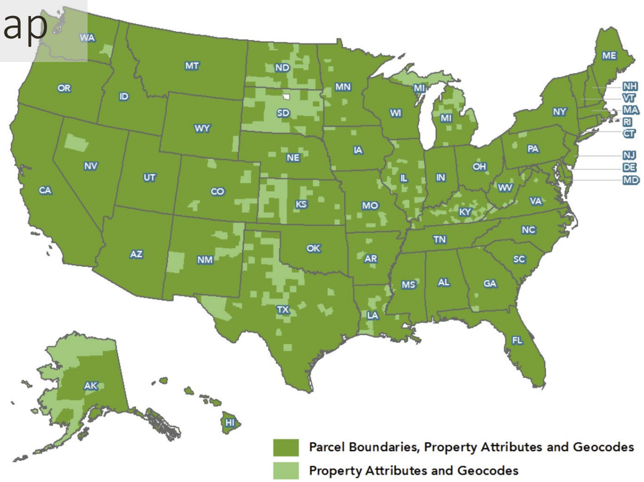
- Data Sources
  - Parcels
    - Public: collection is at the county level, some free
    - Third-party: nearly complete (151M plus)
  - Imagery / Commercial Rooftop sources
    - Public: Microsoft dataset (125M rooftops)
    - Third-party sources
  - Roads
    - Public: Tiger
    - Third-party sources
  - Georeferenced address datasets
    - Public: OpenAddress (179M to date)
    - Third-party sources

Location identification: Data sources

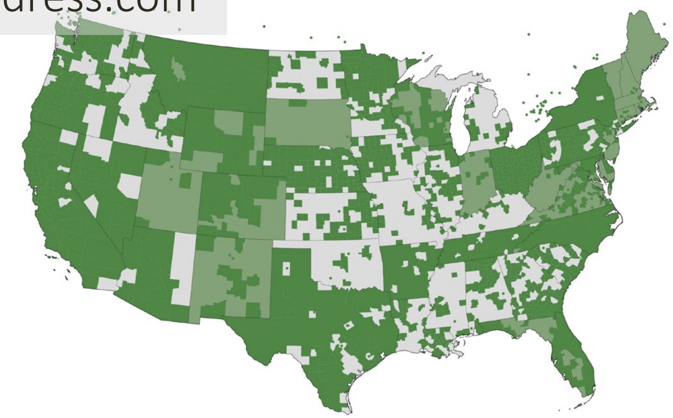
State		Open Address					MicroSoft Building		Parcels		CQA Estimate
State		Coverage Group	Land Area Covered	Population Covered	Addresses Captured	Estimated Total Addresses	Count	Addresses per Building	Count	Addresses Per Parcel	Locations
Alabama	AL	Substantial	43.30%	59.00%	2,357,777	3,996,232	2,460,404	1.62			2,148,041
Alaska	AK	Substantial	12.60%	75.70%	293,620	387,873	110,746	3.50			308,282
Arizona	AZ	Substantial	72.50%	93.80%	3,488,800	3,719,403	2,555,395	1.46			2,730,135
Arkansas	AR	Substantial	89.70%	92.80%	1,602,496	1,726,828	1,508,657	1.14			1,307,168
California	CA	Complete	92.50%	98.90%	15,884,000	16,060,667	10,988,525	1.46			12,350,590
Colorado	CO	Complete	100%	100%	5,209,386	5,209,386	2,080,808	2.50			1,995,157
Connecticut	CT	Complete	100%	100%	1,404,863	1,404,863	1,190,229	1.18			1,294,892
Delaware	DE	Complete	100%	100%	522,264	522,264	345,907	1.51	424,904	1.23	408,260
District of Columbia	DC	Complete	100%	100%	294,385	294,385	58,329	5.05			177,458
Florida	FL	Complete	100%	100%	23,062,056	23,062,056	6,903,772	3.34			8,018,225
Georgia	GA	Substantial	59.40%	58.80%	4,262,188	7,248,619	3,873,560	1.87			4,086,988
Hawaii	HI	Substantial	37.10%	86.30%	468,218	542,547	252,891	2.15			438,713
Idaho	ID	Substantial	33.50%	46.90%	853,508	1,819,846	883,594	2.06			720,183
Illinois	IL	Substantial	48.70%	74.90%	5,130,608	6,849,944	4,855,794	1.41			4,781,642
Indiana	IN	Complete	100%	100%	4,367,720	4,367,720	3,268,325	1.34			2,839,840
Iowa	IA	Substantial	56.70%	53.30%	877,268	1,645,906	2,035,688	0.81			1,344,742
Kansas	KS	Substantial	41.40%	60.30%	1,170,895	1,941,783	1,596,495	1.22			1,239,104
Kentucky	KY	Substantial	13.90%	26.30%	1,145,922	4,357,118	2,384,214	1.83			1,914,057
Louisiana	LA	Substantial	42.50%	53.80%	1,936,594	3,599,617	2,057,368	1.75			1,944,069
Maine	ME	Complete	100%	100%	689,751	689,751	752,054	0.92			698,111
Maryland	MD	Complete	100%	100%	4,577,960	4,577,960	1,622,849	2.82			2,169,515
Massachusetts	MA	Complete	100%	100%	3,928,544	3,928,544	2,033,018	1.93			2,230,378
Michigan	MI	Substantial	17.80%	49.90%	3,248,546	6,510,112	4,900,472	1.33			4,406,476
Minnesota	MN	Substantial	71.10%	67.30%	2,213,122	3,288,443	2,815,784	1.17			2,291,975
Mississippi	MS	Substantial	16.60%	25.40%	743,061	2,925,437	1,495,864	1.96			1,302,429
Missouri	MO	Substantial	15.50%	39.40%	2,550,051	6,472,211	3,141,265	2.06			2,695,246
Montana	MT	Complete	100%	100%	1,537,224	1,537,224	762,288	2.02			510,093
Nebraska	NE	Substantial	87.70%	77.20%	1,057,719	1,370,102	1,158,081	1.18			804,153
Nevada	NV	Complete	78.30%	99%	2,108,731	2,130,031	932,025	2.29			1,038,072
New Hampshire	NH	Complete	100%	100%	623,604	623,604	563,487	1.11			571,210
New Jersey	NJ	Complete	100%	100%	3,873,240	3,873,240	2,480,332	1.56			3,160,915
New Mexico	NM	Complete	100%	100%	1,710,964	1,710,964	1,011,373	1.69			909,564
New York	NY	Substantial	96.60%	69.80%	7,234,129	10,364,082	4,844,438	2.14			6,222,525
North Carolina	NC	Complete	100%	100%	11,311,905	11,311,905	4,561,262	2.48			4,338,363
North Dakota	ND	Substantial	33.40%	50.00%	323,095	646,190	559,161	1.16			356,999
Ohio	OH	Complete	100%	100%	6,597,750	6,597,750	5,449,419	1.21			4,836,669
Oklahoma	OK	Substantial	8.40%	29.30%	1,284,766	4,384,867	2,091,131	2.10			1,644,493
Oregon	OR	Substantial	80.80%	87.00%	3,253,750	3,739,943	1,809,555	2.07			1,670,400
Pennsylvania	PA	Substantial	76.50%	71.60%	4,463,248	6,233,587	4,850,273	1.29			5,424,598
Rhode Island	RI	Complete	100%	100%	484,380	484,380	366,779	1.32			394,319
South Carolina	SC	Substantial	63.30%	69.40%	2,118,924	3,053,205	2,180,513	1.40			2,183,304
South Dakota	SD	Complete	100%	100%	401,268	401,268	649,737	0.62			385,193
Tennessee	TN	Substantial	92.50%	89.60%	4,248,600	4,741,741	3,002,503	1.58			2,867,235
Texas	TX	Substantial	89.40%	89.50%	15,992,499	17,868,714	9,891,540	1.81			9,878,662
Utah	UT	Complete	100%	100%	1,349,895	1,349,895	1,004,734	1.34			1,014,132
Vermont	VT	Complete	100%	100%	343,596	343,596	345,911	0.99			316,559
Virginia	VA	Complete	100%	100%	6,472,704	6,472,704	3,057,019	2.12			3,248,310
Washington	WA	Substantial	68.80%	84.60%	3,349,836	3,959,617	2,993,361	1.32			2,814,839
West Virginia	WV	Complete	100%	100%	1,216,107	1,216,107	1,020,031	1.19			882,161
Wisconsin	WI	Complete	100%	100%	5,362,094	5,362,094	3,054,452	1.76			2,528,952
Wyoming	WY	Complete	100%	100%	610,160	610,160	380,772	1.60			278,913
					179,613,791	**217,536,484**	**125,192,184**	**1.74**	**152,000,000**		**124,122,309**

Location identification: Data sources

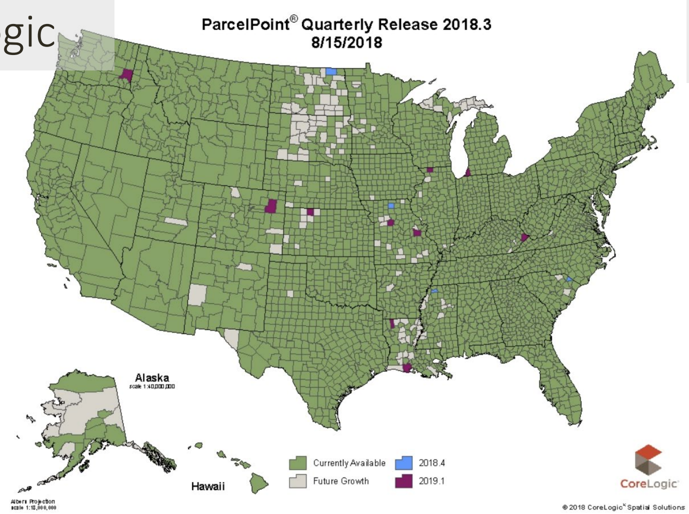
Parcels:
Digital Map



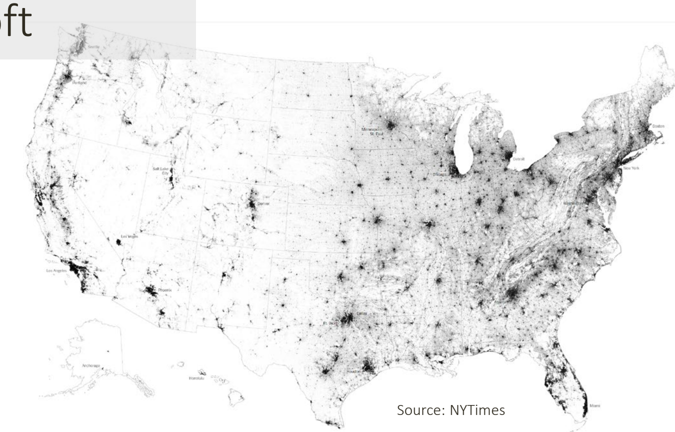
Addresses:
OpenAddress.com



Parcels:
Core Logic



Rooftops:
Microsoft



Sussex County, DE

Sussex County, DE

Building Structure	Few
Parcel	Yes
Address	Multiple per Building / Parcel
Road	Yes

Sussex County, DE

Building Structure	Multiple per Parcel
Parcel	Yes
Address	Multiple per Building
Road	Yes

Sussex County, DE

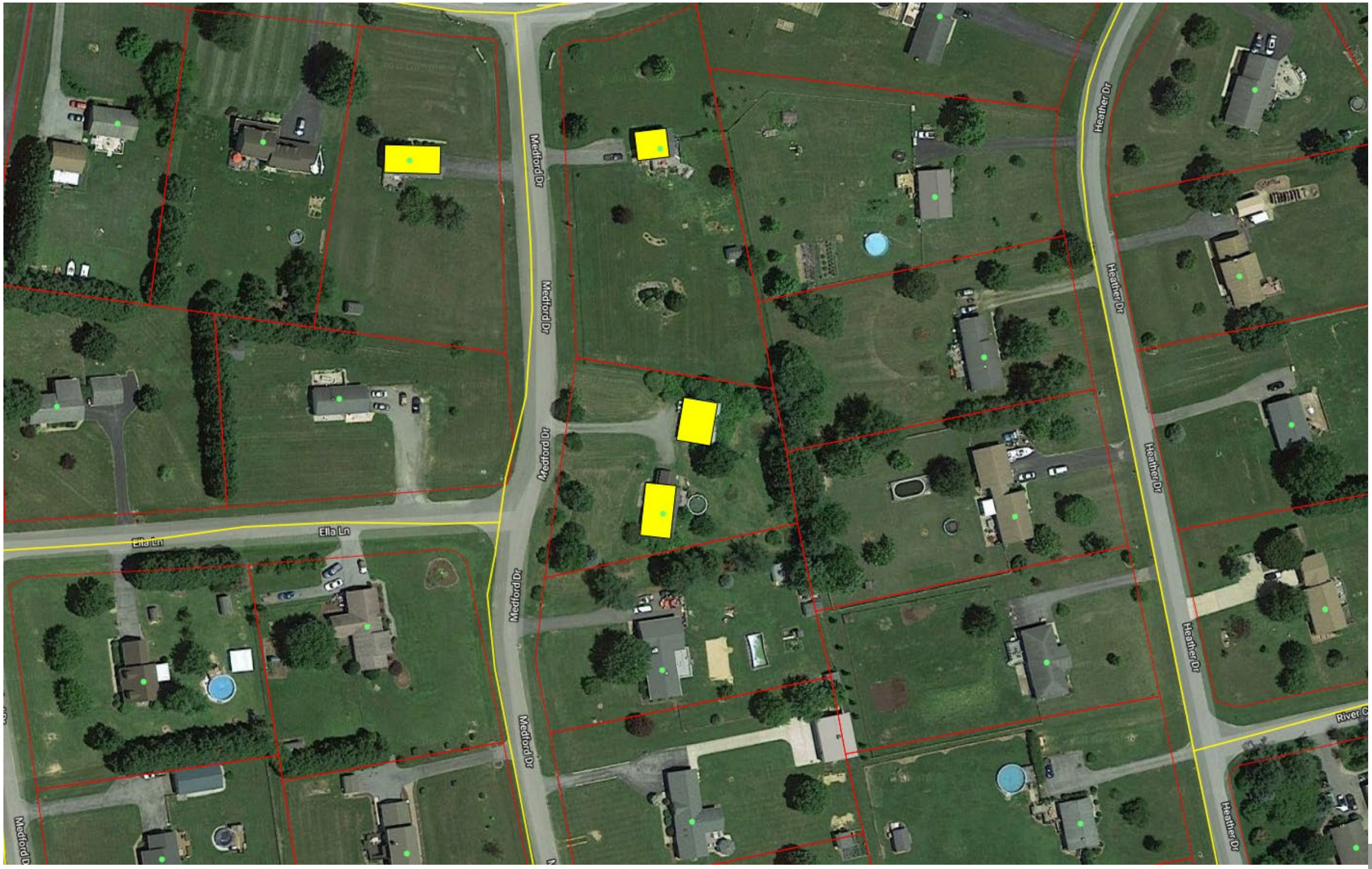
Building Structure	No
Parcel	Yes
Address	Multiple per Parcel
Road	Yes

Sussex County, DE

Building Structure	Yes
Parcel	Multiple per Building
Address	Multiple per Building
Road	Yes

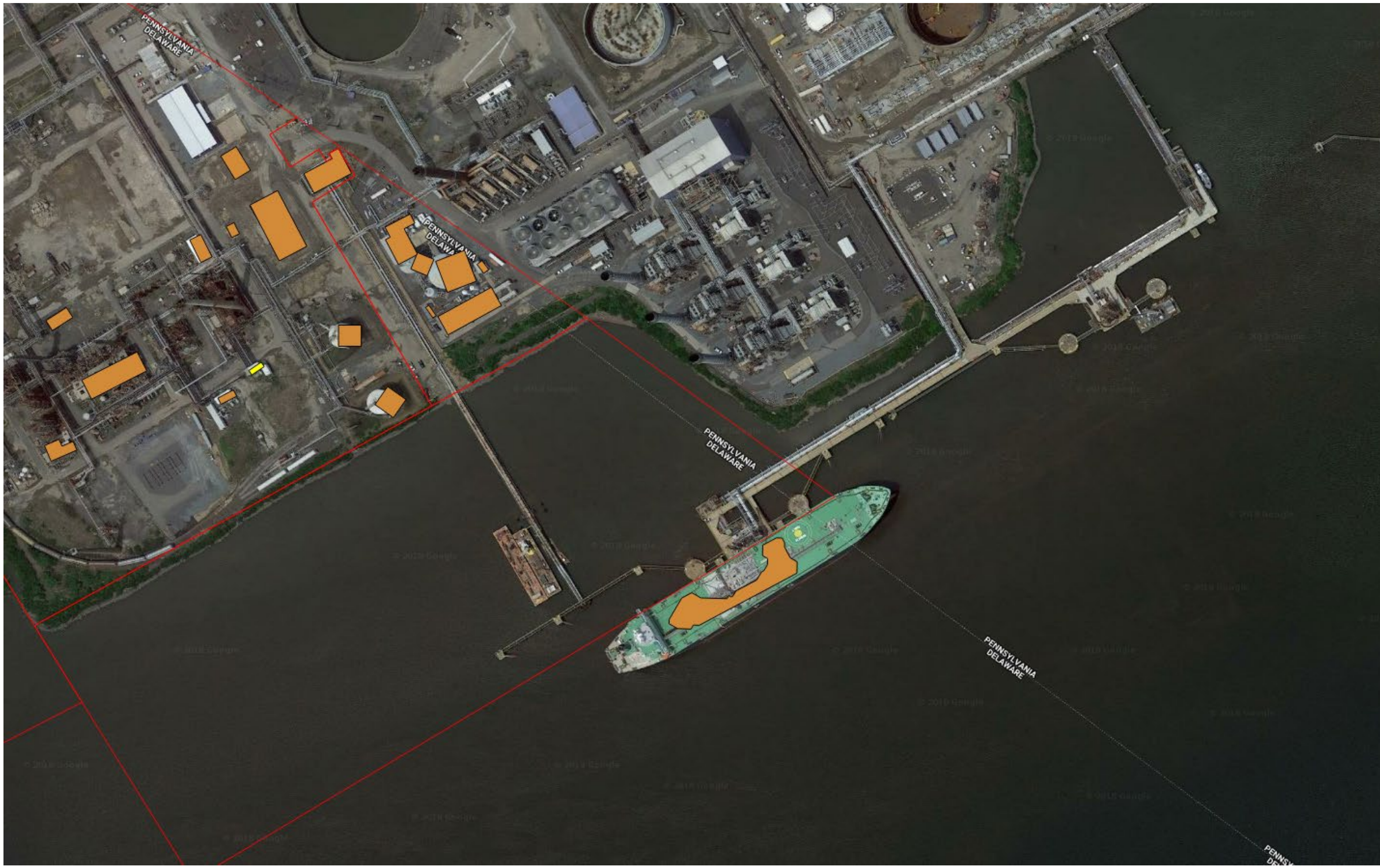
Sussex County, DE

Building Structure	Yes
Parcel	Yes
Address	Mostly
Road	Non-Addressable Segment



Sussex County, DE

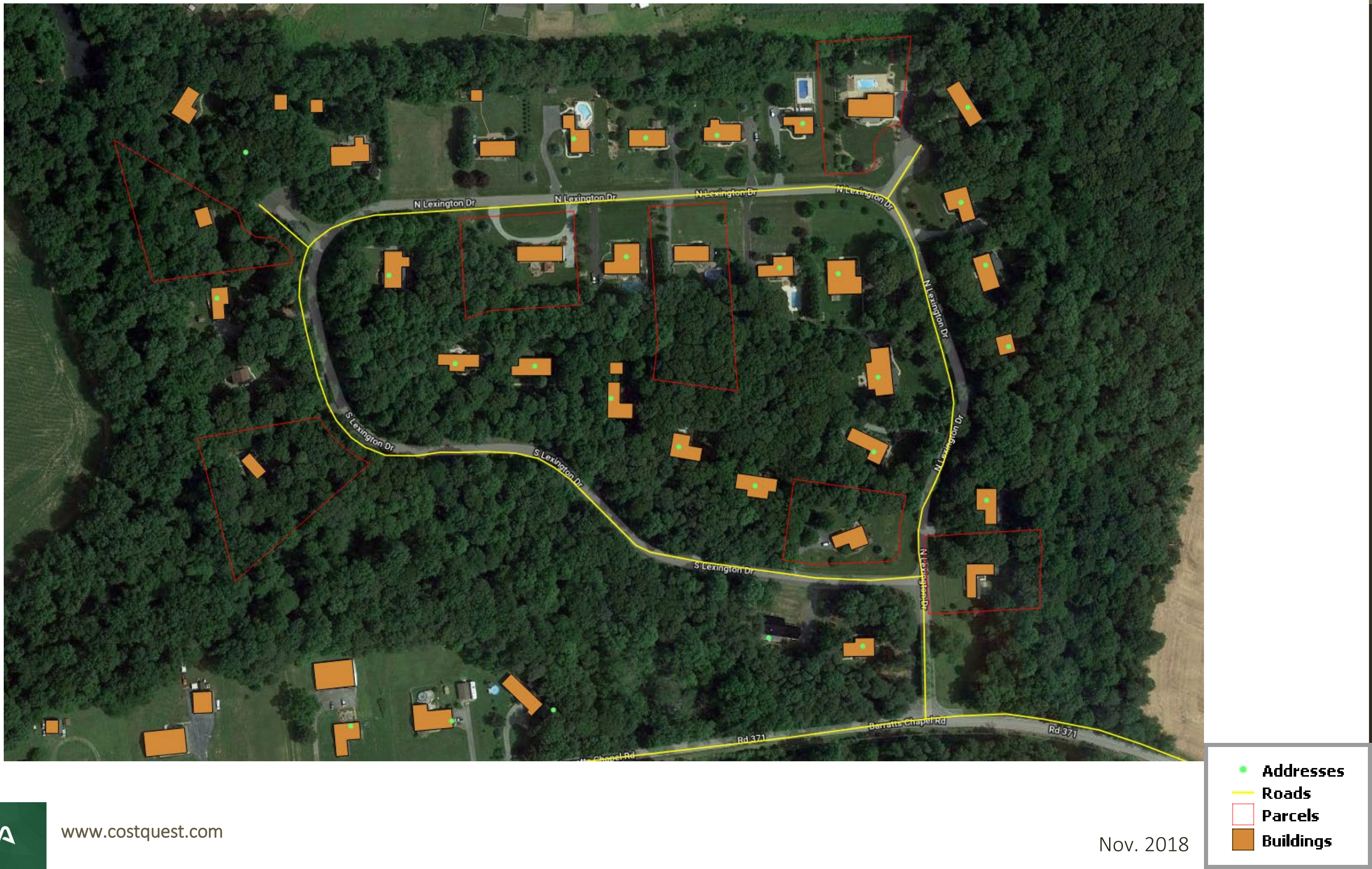
Building Structure	Multiple per Parcel
Parcel	Yes
Address	No
Road	No



Addresses
Roads
Parcels
Buildings

Sussex County, DE

Building Structure	Yes
Parcel	Yes
Address	No
Road	Yes



DE summary

Parcels are control source

Delaware

Parcel Summary				
How many buildings are on the parcel?	How many address points are on the parcel?	How many building footprints touch the parcel?	How many parcels?	
building_category	**address_category**	**footprint_category**	**count**	
Single	Single	Multiple	25,287	6.0%
Empty	Empty	Empty	47,255	11.1%
Single	Empty	Multiple	397	0.1%
Empty	Single	Empty	65,812	15.5%
Empty	Single	Single	39,315	9.3%
Single	Multiple	Multiple	836	0.2%
Multiple	Empty	Multiple	798	0.2%
Multiple	Multiple	Multiple	5,756	1.4%
Empty	Single	Multiple	3,493	0.8%
Empty	Multiple	Single	769	0.2%
Empty	Empty	Single	2,803	0.7%
Empty	Multiple	Multiple	117	0.0%
Single	Empty	Single	2,134	0.5%
Single	Multiple	Single	3,438	0.8%
Multiple	Single	Multiple	26,388	6.2%
Empty	Multiple	Empty	1,379	0.3%
Empty	Empty	Multiple	437	0.1%
Single	Single	Single	198,490	46.7%
			424,904	

DE summary

Buildings are control source

Delaware

Building/Footprint Summary						
Based on a centroid intersection with a 2018 TiGER block	Centroid is within 750m of a "road" (MTFCC=S*)	Is the segment that the centroid is linear referenced to addressable in TiGEr 2018?	How many parcels touch the building footprint?	How many address points fall within the building footprint?	How many buildings?	
urban_rural	**on_road**	**addressable**	**parcel_category**	**address_category**	**count**	
?	1	0	Empty	Empty	5	0.0%
?	1	0	Single	Empty	2	0.0%
?	1	1	Empty	Empty	9	0.0%
?	1	1	Single	Empty	6	0.0%
?	1	1	Single	Single	1	0.0%
R	0	0	Empty	Empty	4	0.0%
R	0	0	Multiple	Empty	2	0.0%
R	0	0	Single	Empty	48	0.0%
R	0	0	Single	Single	8	0.0%
R	1	0	Empty	Empty	26	0.0%
R	1	0	Empty	Single	3	0.0%
R	1	0	Multiple	Empty	434	0.1%
R	1	0	Multiple	Multiple	37	0.0%
R	1	0	Multiple	Single	263	0.1%
R	1	0	Single	Empty	6,718	1.9%
R	1	0	Single	Multiple	125	0.0%
R	1	0	Single	Single	3,625	1.0%
R	1	1	Empty	Empty	19	0.0%
R	1	1	Empty	Single	2	0.0%
R	1	1	Multiple	Empty	3,377	1.0%
R	1	1	Multiple	Multiple	238	0.1%
R	1	1	Multiple	Single	3,555	1.0%
R	1	1	Single	Empty	36,957	10.7%
R	1	1	Single	Multiple	609	0.2%
R	1	1	Single	Single	40,297	11.6%
U	0	0	Empty	Empty	8	0.0%
U	0	0	Multiple	Empty	1	0.0%
U	0	0	Single	Empty	25	0.0%
U	1	0	Empty	Empty	35	0.0%
U	1	0	Empty	Multiple	1	0.0%
U	1	0	Empty	Single	6	0.0%
U	1	0	Multiple	Empty	1,013	0.3%
U	1	0	Multiple	Multiple	671	0.2%
U	1	0	Multiple	Single	1,318	0.4%
U	1	0	Single	Empty	6,655	1.9%
U	1	0	Single	Multiple	1,609	0.5%
U	1	0	Single	Single	8,457	2.4%
U	1	1	Empty	Empty	21	0.0%
U	1	1	Empty	Single	4	0.0%
U	1	1	Multiple	Empty	10,620	3.1%
U	1	1	Multiple	Multiple	9,613	2.8%
U	1	1	Multiple	Single	18,561	5.4%
U	1	1	Single	Empty	57,689	16.7%
U	1	1	Single	Multiple	5,676	1.6%
U	1	1	Single	Single	127,554	36.9%
					345,907	

Parcel data

- Parcel datasets provide other potential "scoring" data:
  - LandUse, Value, OwnerName, etc..
- The following is an illustrative sample of the LandUse data for a single category

EXEMPT, GOVERNMENT AND HISTORICAL

- CEMETERY (EXEMPT)
- CHARITABLE ORGANIZATION, FRATERNAL
- CORRECTIONAL FACILITY, JAILS, PRISONS, INSANE ASYLUM
- EMERGENCY (POLICE; FIRE; RESCUE; SHELTERS, ANIMAL SHELTER)
- GOVT. ADMINISTRATIVE OFFICE (FEDERAL; STATE; LOCAL; COURT HOUSE)
- HOMES (RETIRED; HANDICAP, REST; CONVALESCENT; NURSING)
- HOSPITAL-PUBLIC
- INSTITUTIONAL (GENERAL)
- MEDICAL CLINIC
- PAROCHIAL SCHOOL, PRIVATE SCHOOL
- POST OFFICE
- PRIVATE UTILITY (ELECTRIC; WATER; GAS; ETC.)
- PUBLIC SCHOOL (ADMINISTRATION; CAMPUS; DORMS; INSTRUCTION)
- RELIGIOUS, CHURCH, WORSHIP (SYNAGOGUE, TEMPLE, PARSONAGE)
- WELFARE, SOCIAL SERVICE, LOW INCOME HOUSING (EXEMPT)

Source: Digital Map parcel data

…the path forward

The path forward

- Start with a Pilot project
  - Identify what defines a location
  - Identify potential data sources
  - Identify “scoring” approach to identify an accepted location
  - Identify the crowd approach
    - What needs to be verified, what level of accuracy, …
  - Determine the level of beginning quality and the target