Database Plan Development

by Samira Ghadyani

Under the supervision of Dr. Kathy Weaver

INFM737 Information Management Capstone Experience
The University of Maryland – College Park
Spring, 2016

PALS - Partnership for Action Learning in Sustainability
An initiative of the National Center for Smart Growth

Gerrit Knaap, NCSG Executive Director
Uri Avin, PALS Director

Contents

Introduction

Howard County would like to develop and populate a single holistic database that includes all the information on commercial buildings and occupancies for use by County agencies in various initiatives like inspections, licensing, and billing. Currently, there are many different County agencies that use information regarding commercial businesses, their occupancy and structure characteristics. For example, the County's Fire and Rescue service needs this information to monitor fire sprinkler alarm systems in safety inspections, and for distributing safety information.

Further, a central database would be useful and beneficial for the agencies listed below based on their needs and purposes:
- Department of Health
- Department of Police
- Department of Citizen Services
- Health Department
- Economic Development
- Department of Planning and Zoning
- Department of Licensing, Inspections and Permits
- Department of Finance
- Bureau of Utilities (part of the Department of Public Works)
- Office of Emergency Management
- Office of Consumer Affairs

The challenge in creating a holistic database is that there is no one County or State source that can provide current information on all aspects of the data needed, either about the structure of business that they are in, or about the business themselves. Thus, the project's ultimate goal is to develop a central database for the County, populating it with more accessible data, and creating an ongoing data management system to keep the database updated.

This project's database plan should include comprehensive sources listing, which can be used as a guide to build the central database, starting with identifying data sources and developing a management plan for ongoing data management. This document includes research about available data sources in agencies, including their approaches to acquiring and maintaining it, research about external data sources, and the advantages and disadvantages of including the data sources in the central database.

Further, a preliminary "Entity Relationship Diagram (ERD)," is suggested to clarify the attributes and the relationship between them to have a better understanding of the required data fields.

Through this research, a final solution is provided that meets the agencies' requirements to build a holistic central database.

Internal Data Sources

Department of Fire and Rescue Services
This agency uses the dataset from Department of Assessment and Taxation (SDAT), which is comprehensive and incorporates information on ownership and address, valuation, and basic information about the land and structure(s) associated with a given place. The SDAT dataset includes the following data fields, which are required for the central database: Address, Year Built, Number of Stories, Square Footage, Owner Name, Owner Address, Assessment Value, Tax Exempt, Occupancy Percentage, Tax ID.
Note: SDAT"S data quality, specifically for "Number of Stories" and "Owner information," is questionable because there is a significant amount of missing data in these fields.

Office of Fire Marshals (OFM)
OFM is part of Department of Fire and Rescue Services and they use the Mobile Eyes software to store their data. Although Mobile Eyes stores many of the required data fields, it is not a reliable and updated source. This system was populated many years ago, and only has updated information on the occupancies that OFM has been able to inspect. The data fields stored in Mobile Eyes are Address, Business Name, Year Built, Number of Stories, Square Footage, Hazardous Material, Property Manger Info, Owner Name and Tenant Info, Business Name.

Howard County Economic Development Authority (HCEDA)
EDA subscribes to an external data source, called "CoStar," which they find useful in tracking variables such as vacancy rates, absorption, and asking rents, as well as a source of information specific to buildings such as square footage, percent leased, and leasing or property management contact information. The data fields are Address, Business Name, Number of Stories, Square Footage, Number of Employees, Property Manger Info, Owner Info, Tenant Info, Assessment Value, Tax Exempt, Tax ID, Occupancy Percentage.

Office of Emergency Management (OEM)
The County's OEM has access to the Maryland Online Tier II Reporting System (MOTTRS), which is implemented by the Maryland Department of Environment for facilities to report their Tier II Emergency and Hazardous Chemical Inventory Reports. These reports, required by Section 312 of the federal Emergency Planning and Community Right-to-Know Act (EPCRA), require covered facilities to provide detailed reports about the hazardous materials they use or store on site. The reports are submitted to the State Emergency Response Commission (SERC), the Local Emergency Planning Committee (LEPC) with jurisdiction over the facility, and the local fire department that would respond

to any emergency at the facility. OEM has access to the following data through MOTTRS: Address, Hazardous Material Name, Hazardous Material Amount.

Note: Although the hazardous material was a required data field at the beginning of the project, it changed to an optional data field due to high sensitivity of this data.

Department of Inspections, Licenses and Permits (DILP)

DILP uses Address Points, State Department of Labor, Licensing, and Regulation (DLLR—for license and permit purposes), and SDAT databases. They also store information about commercial buildings permits and licenses that are manually entered into their system; the Accela software, used to build a data warehouse that integrates the mentioned data sources. DILP does not have direct access to DLLR. However, they access the DLLR website regularly and as a resource when they need to confirm a license or license professional's current status (http://www.dllr.state.md.us/pq/). Further, DLLR sends an updated licensed professional data file every day, and DILP manually imports the latest data file into Accela on an as-needed basis. They track these data fields: Address and GIS coordinates, Business Type, Year Built, Number of Stories, Square Footage, Property Management Name (No contact info), Owner Name (No contact info), Tenant Name (No contact info), Business Name, Assessment Value, and Tax ID.

Department of Planning and Zoning (DPZ)

DPZ has access to a DLLR database for employment purposes. They only have 2nd quarter 2015 DLLR data, received from the Baltimore Metropolitan Council (BMC) as a GIS database file. BMC has an agreement with DLLR to receive the Quarterly Census of Employment and Wages (QCEW) data for all counties in the Baltimore region (Howard, Baltimore County, Baltimore City, Harford, Carroll, Anne Arundel) for regional cooperative forecasting projections. BMC received the data from DLLR, decoded it, and then provided it to DPZ in GIS format. DPZ does not have direct access to the database but they, along with representatives of other counties, have signed a non-disclosure agreement. DPZ tracks Address, Business Type, Number of Stories, and Business Name.

Department of Citizen Services (DCS)

DCS focuses almost entirely on services and transactions, based on people as opposed to properties. They capture very limited property related information, and only then as a marker for the location of people that they might serve.

Office of Consumer Affairs

This agency tracks information about vendors or companies with which customers have experienced problems, recorded in a "Complaint Database." They track Address, Property Manger Company and Contact Info (very limited), Owner Name and Contact Info (very limited), Tenant and contact info (very limited), and Business Name.

Department of Finance
This agency uses SDAT public data recorded for financial purposes, particularly Address, Year Built, Owner Name, Number of Stories, Owner Address, Business Type, Square Footage, Assessment Value, and Tax Exempt.

Other Agencies
The County's Health Department tracks all food related information such restaurants and school food providers. The Police Department uses Computer Aided Dispatching (CAD) for 911.
The Department of Public Works tracks information about utilities, environmental information and trash collection mostly for residential buildings.

External Data Sources

CoStar: CoStar is a commercial real estate company that provides building-specific information through a proprietary database of commercial transactions. It includes subscription information, sample reports, and a user-friendly interface application. Costar tracks all the data fields that will be used in the central database. There are two clarifications about CoStar database:

- The class of the building used in CoStar database is determined by their research team, building characteristics, and the zoning information.
- Resident agent or attorney information is sometimes listed under the contacts information but it is not listed nearly as frequently as the owner and property manager contact information is listed.

Howard County Government's customized package will cost $1,085 per month for five user licenses, which consist of:

- Costar Property Professional: building data, owner contact info, property manager contact info, and assessment records
- Costar Tenant: tenants in each commercial building and contact info
- Access throughout the Baltimore market area.

Infogroup: Infogroup is another external data source that has never collaborated with Howard County, but they track reliable information about commercial buildings such as Address, Business Type (by industry code (SIC and NAICS), Year Built (the year a business was established appears in 11 percent of records; don't have data on the year a building was built), Number of Employees, Square Footage, Tenant Name and Contact Info.

There are some options for data updates through the 12-month license period:

- Six-month file update: $3,659 total price; the user receives the full file upon purchase, and then an updated file six months after the initial purchase.
- Quarterly file update: $4,391 total price; the user receives the full file upon purchase, and then

updated files three, six, and nine months after the initial purchase.

- Monthly updates: $5,122 total price; the user receives the full file upon purchase, and updated files on a monthly basis throughout the 12-month license term.

Infogroup has roughly 15,410 business records/locations. The retail price is 19 cents per record, or approximately $2,927 (data counts change on a monthly basis as the database is updated). The price of 19 cents per record is based on a one-time data purchase of the approximate 15,410 records. At this price, users are licensed to use the data for one year. There are no updates made to the dataset by Infogroup.

	SDAT	DLLR	CoStar	Infogroup	Accela	MOTTRS
Address	✓	✓	✓	✓	✓	✓
Business Type		✓	✓	✓	✓	
Year Built	✓		✓	?	✓	
No. of Stories	?		✓		?	
No. of Employees		✓	✓	✓		
Square Footage	✓		✓	✓	✓	
Hazardous Material						✓
Property Manager Name and Contact Info			✓		?	
Owner Name and Contact Info	?		✓	✓	?	
Tenant Name and Contact Info			✓	✓	?	
Business Name		✓	✓	✓	✓	
Assessment Value	✓		✓		✓	
Tax Exempt (Y/N)	✓		✓			
Occupancy Percentage	✓		✓			
TAX ID	✓		✓	?	✓	

Table 1: Data Field List with Accessible Data Source
(A question mark indicates that the source is either unreliable or includes only part of the data)

Entity Relationship Diagram (ERD)

A preliminary ERD has been created for the central database to clarify the data fields and the relationships among them. As the ERD shows, there are four entities: Properties, Property Manger, Owner, and Tenant. Further, the ERD represents the attributes of each table with the interrelationships and logical structure of the database. Obviously, a more advanced and detailed ERD should be created if the database is going to be built and maintained internally.
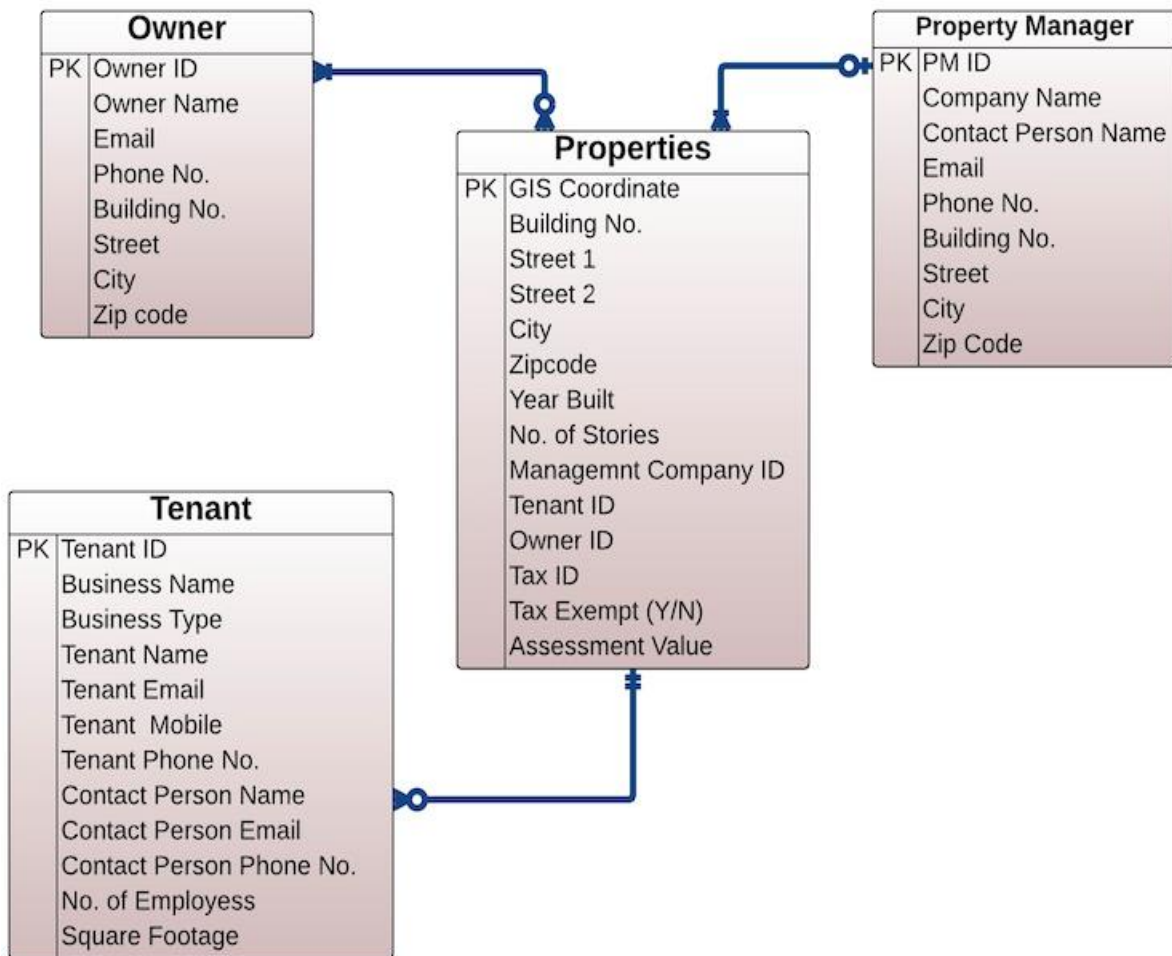

Diagram 1: Preliminary Entity Relationship

Conclusion

Based on the research and provided information, CoStar is the best source to use. It tracks all required data fields, and contains updated and reliable data. Its user-friendly interface is another advantage, enabling agencies to access comprehensive information in different layers and providing extra information such as pictures of the buildings.

Subscribing to CoStar will save Howard County time and money by using an existing and reliable database, as opposed to trying to create an internal database. The data is updated every 30 days, which makes it timely and reliable. Additionally, most CoStar customers believe it is a great return on their investment and have continued using the product for many years. Local users include the Mayor and City Council of Baltimore, Maryland Port Administration, Harford County Government, Maryland State Department of Assessments & Taxation, and Montgomery County Fire and Rescue Services.

On the other hand, "high cost" and "limited number of licensed user access to County agencies" are two major drawbacks of using CoStar.

An alternative solution would be integrating multiple government open source datasets such as SDAT and DLLR. However, these data sources do not contain all the required data fields such as property management companies, their contact information, or tenants' names and their contact information. Furthermore, the data quality of available data sources is questionable. There is a lot of missing data in SDAT database, especially in the number of stories, year built, and owners' names and contact Information. Building an internal database using unreliable data sources does not justify the extra money and time for employees.

Recommendation

The analysis and conclusion of this report are based on the limited information that was accessible due to data confidentiality policies.

Thus, further research is highly recommended into the Department of Labor, Licenses and Regulations database. DLLR may have updated and reliable data for the data fields not currently used in the Department of Planning and Zoning. Since DLLR is an open source dataset provided by the government, it will be valuable to investigate more about this dataset.