



MIDWEST ARCHITECTURE COMMUNITY COLLABORATION 2020

NOVEMBER 5, 2020

**MACC 2020: Adaptable architecture:
building resilience in a time of change**

MACC MISSION

- The Midwest Architecture Community Collaboration's (MACC) purpose is to bring all domains of architecture together to share information and techniques of interest to all of us. It is our shared belief that through collaboration, we can better understand and promote the significance of architecture to business success.



INTEGRATED DATA, GRAPHS, AND LOCAL GOVERNMENT: EARLY LESSONS

ERIK ERICKSON, CHIEF DATA OFFICER, HENNEPIN COUNTY

CONTEXT: HENNEPIN COUNTY



- Home to 1.2 million residents
- ~8,900 employees (in 2019)
- \$2.1bn annual operating budget
- Organized into 7 lines of business
 - Disparity Reduction
 - Public Works
 - Public Safety
 - Human Services
 - Operations
 - Health
- Hard to estimate total residents served every month: HS serves ~350,000 per month, DOCCR (within public safety) serves ~24,000 per day)

THE PROBLEM

- Context
 - Lines of business evolved to support core local government functions
 - Internal processes, policies, data practices, etc. developed in support of each function
 - Local government is not unique, but perhaps subject to more public scrutiny and transparency requirements
- The business problem
 - We have limited understanding of how county services intersect, overlap, and/or influence one another.
 - This limited understanding results in less efficient and less effective county services (or so we believe)
- How can we leverage our existing data to support a more holistic and effective delivery of county services while maintaining data privacy and security?



The CDO's core mission:

How can we leverage county data to support the effective delivery of county services and the more efficient use of taxpayer dollars?

KEY CHARACTERISTICS OF A POTENTIAL SOLUTION

- Flexibility
 - Incremental & Iterative
- Speed
- Utility
 - Known use cases
 - Unknown use cases
- Ease of use
- All of this in the context of our data privacy landscape



INITIAL FORAYS INTO GRAPH

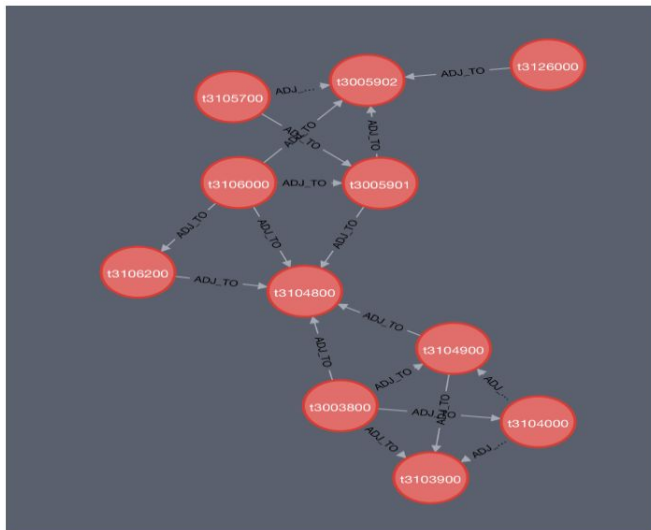
- Internal pilot with CosmosDB →
- Achieved functioning graph, but very complex to load and query

- Carlson Analytics Lab →
- Illustrated potential of integrated data via graph
 - Querying neighborhood characteristics
 - Analyzing SNAP “market penetration” (i.e. to what extent was SNAP participation aligned with food insecurity estimates)
 - Optimizing community resource locations

CARLSON ANALYTICS LAB

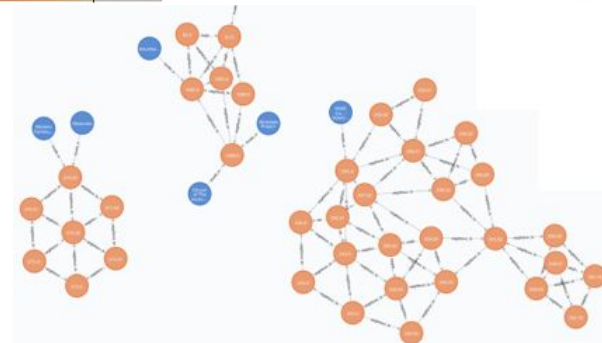
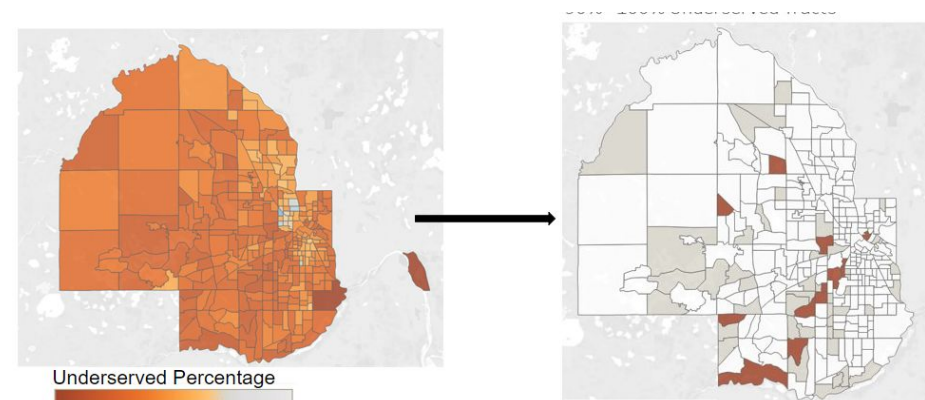
Simplifying complex queries

- Identifying adjacent census tracts with similar characteristics



Identifying potentially undeserved areas

- Is the county meeting residents needs related to food insecurity?



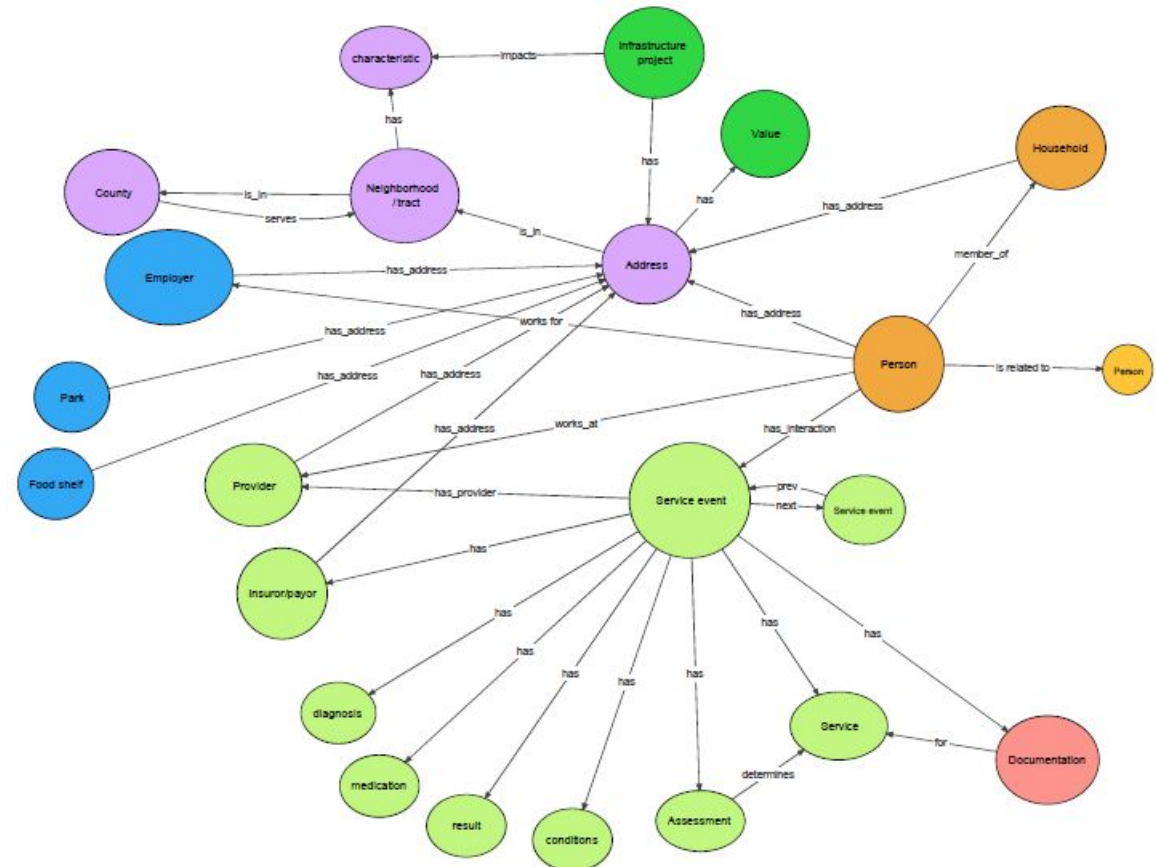
INTERNAL PROOF-OF-CONCEPTS

- Currently engaged in an internal proof-of-concept
 - Builds off initial forays
- Testing a few key characteristics
 - How flexible are graph DBs?
 - How easy are graphs to implement and use?
 - How quickly can we illustrate value?



FLEXIBILITY

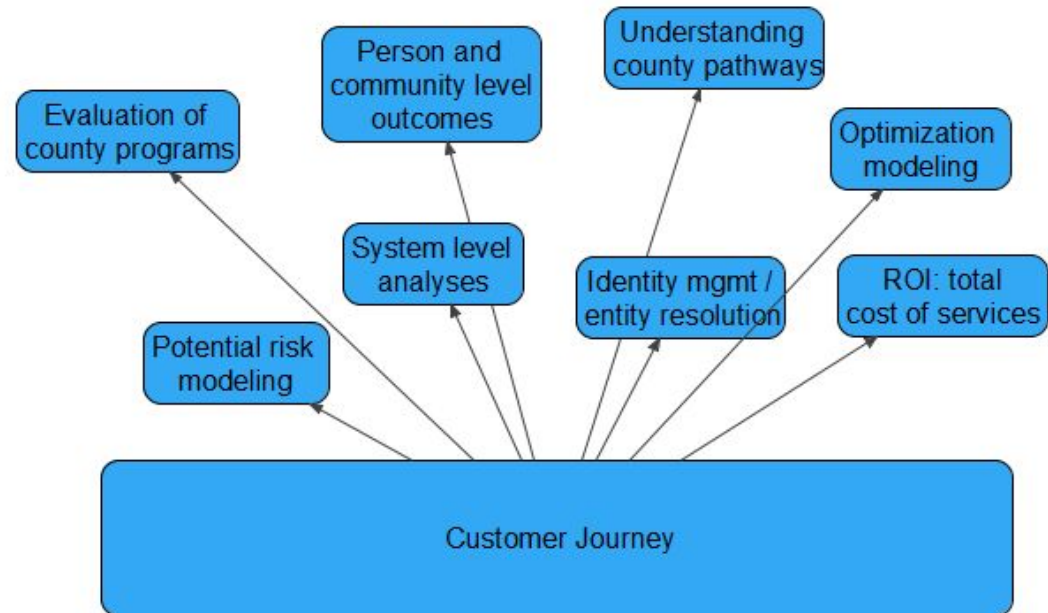
- Data integration
 - Able to ingest and meaningfully link a variety of data at different granularities.
 - Person
 - Household
 - Property
 - Business / organization
 - Community (census tract, zip code, city, county)
 - Still testing “the Jenga problem”
 - Starting with a “good enough” model and iterating has worked thus far



An early logical model

FLEXIBILITY

- Business use cases
 - Framing our primary use case around the customer journey or citizen experience has helped us link otherwise disparate lines of analysis
- Balancing the need to develop for known uses & maintaining flexibility for future (and often unknown uses)



EASE OF USE



- Biggest challenge to date has been standing up graph DBS in our Azure environment
 - the county still lives in a relational world and is in transition to the cloud
- Graph query languages have thus far proven relatively easy for analysts to learn
 - Have not been able to venture into advanced graph analytics yet
 - Requires a definite shift in thinking: from joins to traversals
- Data engineering talent and skill a concern
- Haven't had to make a major revision to the model... yet

ILLUSTRATING VALUE

- Aggressively working to illustrate value as part of our internal proof-of-concept
 - Currently looking at a 1-month time to delivery from initial stand-up of graph to first prototype report
- Customer journey has been an effective tool
 - Especially as means to illustrate what we could do with a graph DB
 - Drastic reduction in time for internal evaluations of county programs (“the 3-year study” □ the 3-month study)
 - Promising use cases around identity resolution and clustering
 - Early efforts suggest we can incorporate person level and community level data (opportunity atlas)
- Non-technical aspects of integrating data (governance, data privacy and security, etc.) have proven much more challenging



QUESTIONS?