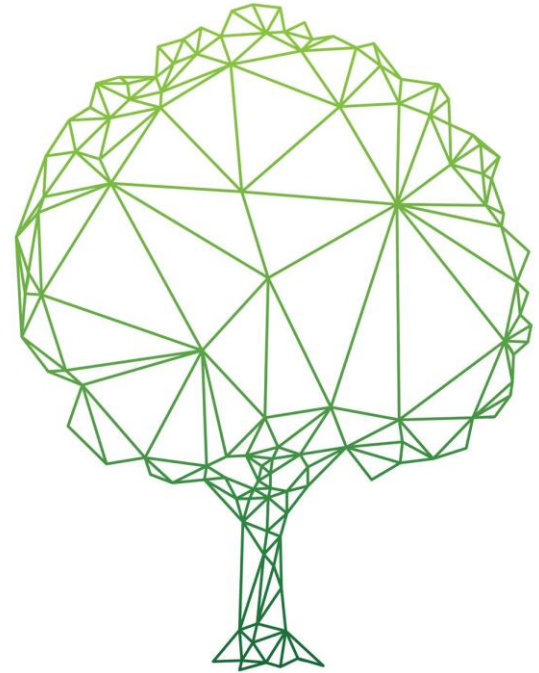**DETERMINING THE SHARED POPULATION BETWEEN SERVICE PROVIDERS**

How Tulsa Is Preserving Privacy and Sharing Data for *Social Good*

July 15, 2019

RESTORE HOPE
MINISTRIES

asemio

Background

Communities...
grapple with wicked social problems

And often see...
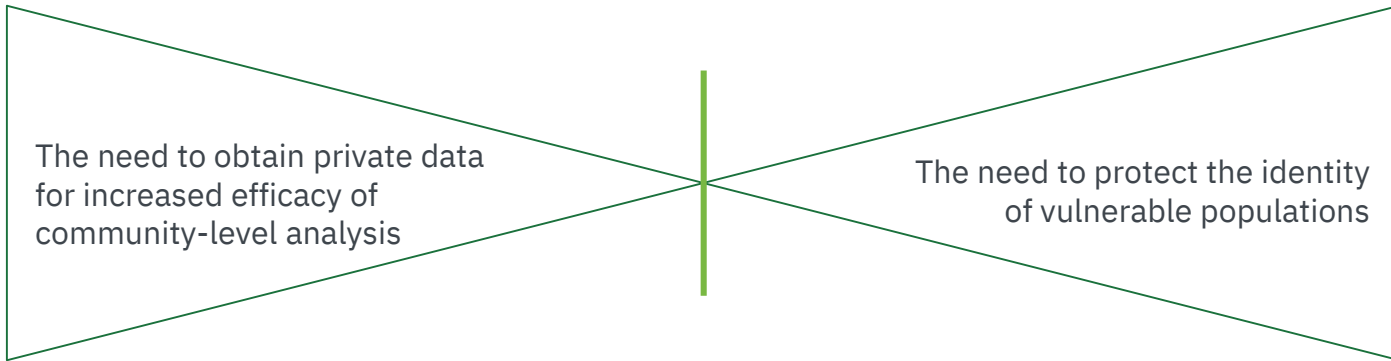*data as a panacea for achieving systems-level advancements*

Which results in a desire to...
increase interoperability, refine resource alignment, and streamline community services to improve outcomes

Problem Definition

The potential for access to data creates a tension between:

The need to obtain private data for increased efficacy of community-level analysis

The need to protect the identity of vulnerable populations

"

*There is a clear need to establish a model that can serve our communities better by enabling community analysis of integrated data **more quickly**, at a **lower cost**, and in a manner that **enhances both privacy and security protection for individuals contributing, and organizations** collecting, this sensitive data.*

Solution
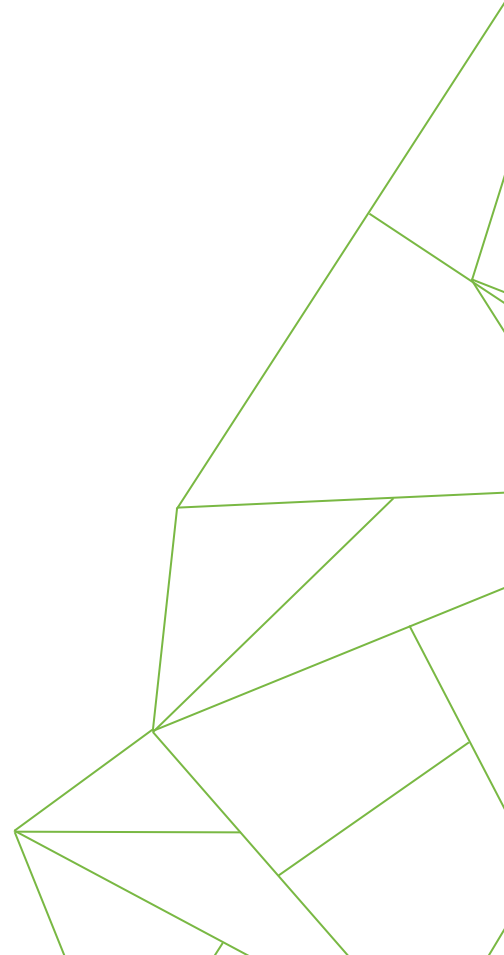
MPC Technology

WHAT IS IT?

*Secure multi-party computation (also known as secure computation, multi-party computation/MPC, or privacy-preserving computation) is a subfield of cryptography with the goal of creating methods for parties to **jointly compute a function over their inputs while keeping those inputs private**.*

(https://en.wikipedia.org/wiki/Secure_multi-party_computation)

MPC Technology

*"parties can **jointly compute a function over their inputs while keeping those inputs private**"*

HOW IS IT USED NOW?

- Subject of active research
    - DARPA (PROCEED Program)
        - https://www.darpa.mil/program/programming-computation-on-encrypted-data
    - Allegheny county (Demonstration project)
        - https://bipartisanpolicy.org/report/privacy-preserved-data-sharing-for-evidence-based-policy-decisions
    - Cybernetica (Sharemind)
        - https://sharemind.cyber.ee/

MPC Technology

*"parties can **jointly compute a function over their inputs while keeping those inputs private**"*

HOW CAN MPC TECHNOLOGY HELP?

- By providing faster access to broader data sets and more secure analysis techniques while improving personal privacy protections
  - Providers are not actually sharing PII, only (encrypted) versions of it that can't be used to reconstruct the original (plaintext) input.
  - Initial analysis suggests HIPAA / FERPA restrictions on sharing PHI should not apply.
  - No case law, but legal opinion is catching up. Europe is already there (c.f. Estonian case study with German legal opinion, to follow).
    - H.R.4479 - Student Right to Know Before You Go Act of 2017

Pilot
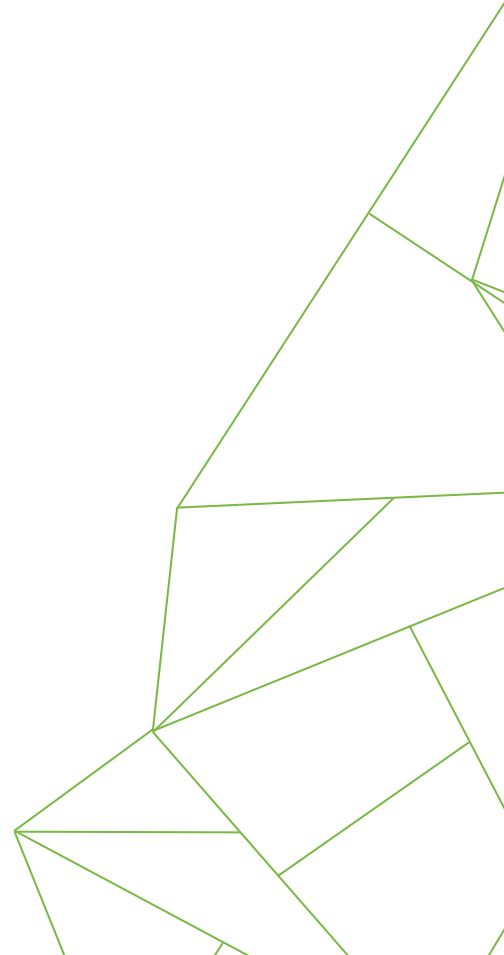
Pilot Question

What is the overlap of populations served by two disparate organizations?

- What does the answer to this question provide?
  - Information that can inform internal and external decision-making and next steps
    - A flashlight for additional analysis opportunities
    - To have a way to test and measure the implications of certain implementations

Pilot Process: Discovery

- Finding the question to ask
  - Collaborative working session to come up with case examples of beneficial insights that could be derived using MPC
- DSA's
  - Expedited agreements
  - Third party computation reduces redundant efforts
  - The legal necessity of DSA's may be minimized according to legal reviews

Pilot Process: ETL

Extraction

- Data specifications: Turning a question into a measurable hypothesis
    - There exists a sub-population of children who have received services from an AssistOK organization but who are not enrolled in CAP Tulsa's early childhood education program.

Partner	Sites	Extraction Parameters
CAP	CAP Tulsa	Children under the age of 5 who were enrolled in CAP during the time period of 2/1/2017 - 1/31/2019.
AssistOK	Restore Hope Ministries	Children under the age of 5 (i.e. born on or after 2/1/2012) who sought services at an AssistOK location during the time period of 2/1/2017 - 1/31/2019.
	Owasso Community Resources	
	Neighbors Along the Lines	

Pilot Process: ETL

Transformation
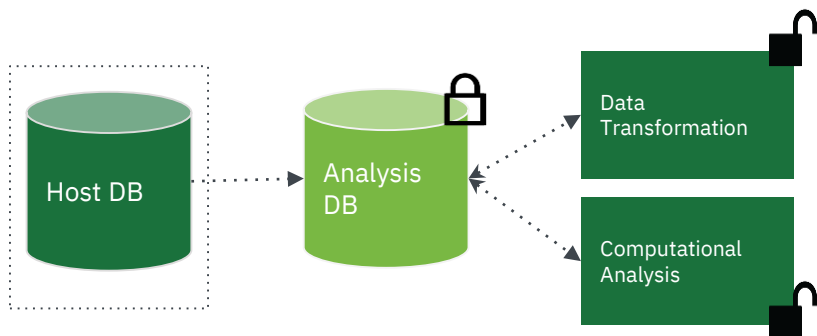- Text cleaning
- Arranging data
- Deduplication
- Review

Loading into analytics tools
- Traditional plaintext analysis
- MPC platform (Sharemind)

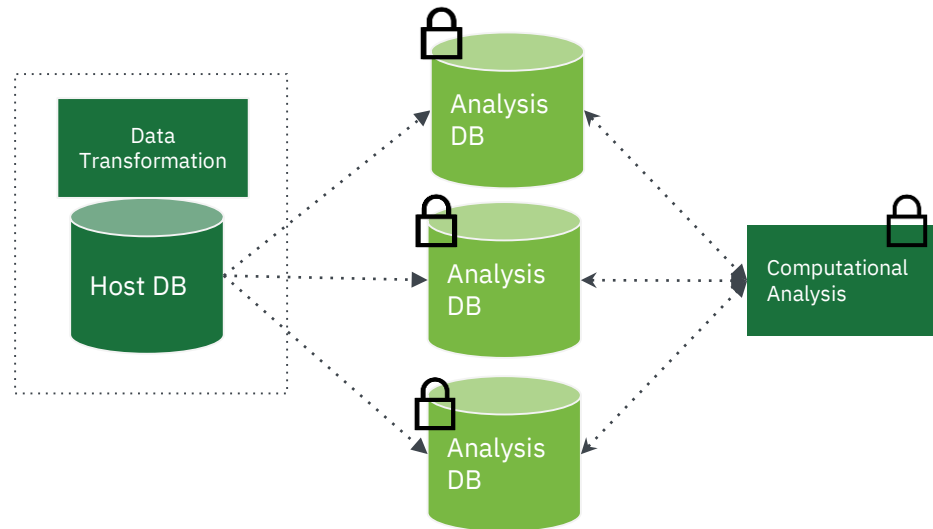# Process: Comparison

4

Non-Privacy Preserving Computation

Host DB → Analysis DB → Data Transformation / Computational Analysis

Privacy Preserving Computation

Data Transformation / Host DB → Analysis DB / Analysis DB / Analysis DB ↔ Computational Analysis

Process: Comparison

Traditional and MPC comparison

Partner	Site	Post ETL Count of Unique Children 4 and under	Traditional Plaintext Shared Overlap Count	MPC (Sharemind) Shared Overlap Count
AssistOK	Restore Hope Ministries'	147	13	13
	Owasso Community Resources	249	8	8
	Neighbors Along the Lines	700	44	44
CAP	CAP Tulsa	4133	65	65

Portal: Community Analytics Mapping Portal
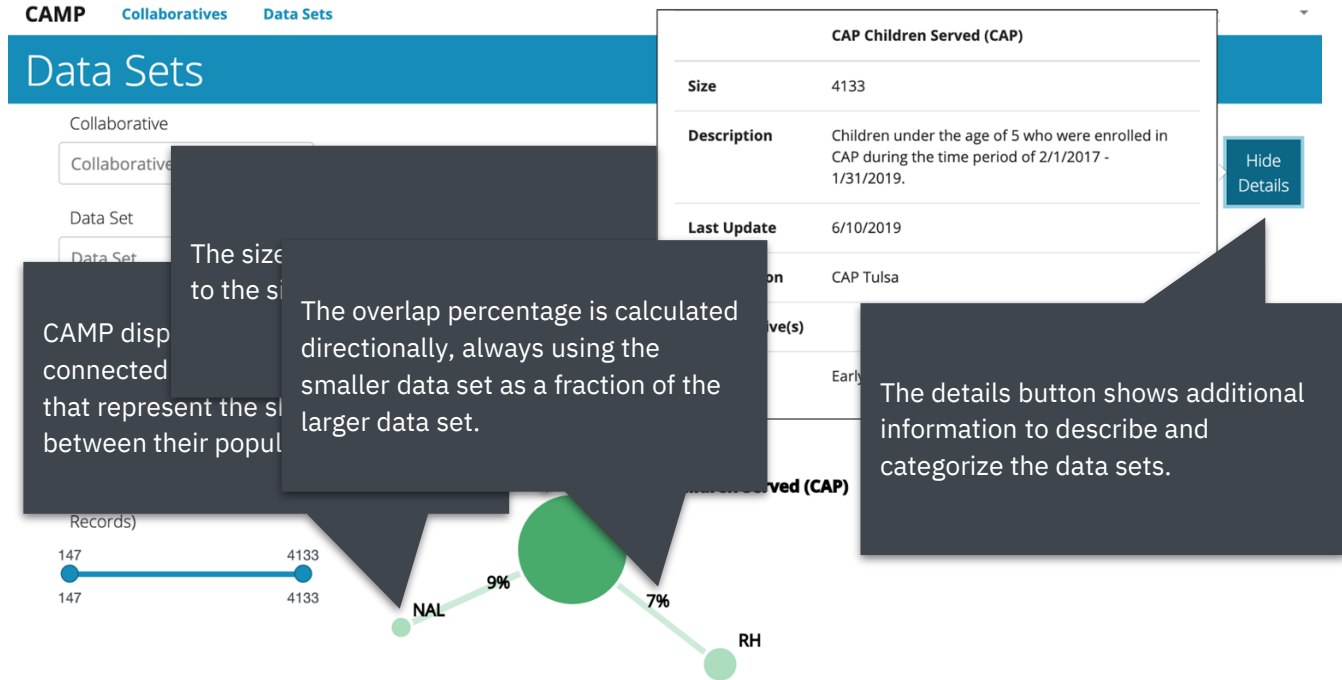
Impact

- New collaborative efforts between project stakeholders
- Igniting interest among community stakeholders
- Additional investments

Insights

- Lessons learned

  - You can never start *too* small
    - Find success first with a small group of trusted partners who are willing to try innovative approaches to better understand the populations they serve

  - Low context = low value
    - The technology is important but the resulting impact to the populations being analyzed should be front and center throughout the process

  - There are no magic bullets
    - Truly combating wicked social problems will require finding the nexus between both individual care coordination and population research and evaluation

Project Background

Restore Hope Ministries was funded by the DASH CIC-START program, which supports short-term activities that help local collaborations take meaningful steps toward planning or implementing multi-sector data systems.

Through DASH CIC-START, Restore Hope Ministries worked with Asemio to apply analytics technology to analyze the overlap between individuals who require basic needs assistance (e.g. rent, food, utilities, etc.) and those whose children attend early childhood centers.

Asemio developed this presentation and its corresponding white paper to share lessons learned from their use of innovative technology that allows for analysis of personally identifiable information while preserving client privacy.

Acknowledgements