

Washtenaw Equity Partnership (WEP) Data Subcommittee

June 6, 2022, 11:30am-1:00pm

Attendees:

MaryAnn Sarosi (co-chair)
Eli Savit (co-chair)
Grady Bridges
Linda Rexer
Meghan O'Neil
Jeff Rose
Angela Tripp
Linh Song
Trevor Bechtel
Alex Roth (Vera)

- Not enough people for quorum so wait to vote on minutes.

Mapping analysis discussion

- 20 agencies responded to the survey (~70% response rate.) Word document created by Alex of initial analysis of survey responses with potential next steps of what to do.
- Survey results shared in folder with group by Alex. 2 more agencies responded and spreadsheet with all responses are in the google folder.
- Still waiting for replies from 6 agencies. Alex asked how much follow up do we want to do to get people to complete the survey?
- Co-chair starts discussion about mapping subgroup survey responses
 - Follow-up – can we ask new chief about Saline Police Department? Do people have connections with other agencies that haven't responded? Vera to send email to agencies that haven't responded and cc co-chairs.
 - Discussion re: unique identifiers – are there resources for data normalization that could connect records for people across system? Python code for doing this. CJARS also may have some code for doing this. JDW does some data normalization but not sure about unique IDs. Is this issue something the subcommittee needs to come up with an answer for or is it something that we could recommend be done as part of a planning process for data integration. Maybe focus should be more on why unique identifiers are needed in order to integrate data rather than try to determine the specifics of how it should happen. Any group implementing the data warehouse would need to have the right stakeholders in place to work on implementing something like this.
 - Tracking demographic information – anything that can be done at the source to consolidate race/ethnicity data would be much better than trying to combine that later.
- Co-chair suggested reaching out to subcommittee members who have connections to agencies, so those contacts fill out surveys.
 - Survey doesn't take too long, can note estimated length in emails/communication to contacts so they can have an idea and participate.

- Different contacts for each of the district courts – got responses from some courts but not others.
- Ann Arbor city attorney’s office – might be easier to interview them because didn’t get reply about survey.
 - Co-chair will email contact in Ann Arbor government who can try to get a reply instead of having to do a FOIA request as they initially stated.
- 14B – subcommittee member will reach out to them.
 - May have the same data as 14A because the data system is the same.
 - Copying person who completed survey for 14A would be helpful.
- Survey responses re: whether and how agencies use unique person identifiers
 - Split responses, 6 agencies don’t use them and 7 do.
 - No consistency on type of unique identifier used. MDOC has their own identifier they use for jail booking, others use name, DOB, social security numbers, state ID, or some combination of those.
 - Should look into this more, especially if want to have some kind of data warehouse or cross-systems integrated data.
 - Subcommittee member: has been using statistical analysis and code to analyze a data and has been able to create unique identifiers with 95-96% consistency/accuracy.
 - A data warehouse could use data normalization to ensure correct data is being collected and associated with a unique identifier (e.g. when someone is entered in multiple systems but with their name spelled differently).
 - JDW: uses data normalization and standardization, but the Member was unsure if the JDW has any unique identifiers or tracking numbers etc.
- Subcommittee member: Need to be careful about how sensitive data that could be traced back to an individual is stored and presented. Have to make sure not to create publicly accessible information that would allow anyone to identify an individual, especially when charges didn’t result in conviction
- Co-chair: should consider the comments and concerns brought up here in the recommendations. For example, if we decide to recommend a data warehouse, can also recommend that the planning process addresses these types of concerns and questions.
- Subcommittee member: could use a subject number to consolidate data on same individual without using unique identifiers that could identify the actual person.
 - Data norming techniques could potentially identify people across different sources while keeping them anonymous.
- Subcommittee member: can use records to identify the same individual if listed multiple times or with different spelling of their name. Can configure a system that labels them with a random number, don’t need the name anymore.
 - Can use unique identifiers to do work moving forward, random number will not convey sensitive information.
- Subcommittee member: In the data subcommittee’s report, we can explain why anonymized unique identifiers will help us toward our goal of identifying where disparities are in the systems or where they get compounded.
 - There should be standards and consistency in language around this.
 - Need to get to basics about why to support the project.

- Providing guidance and clarity on what ought to be recommended and why.
- Subcommittee member: Implementing data standards can take time across disparate systems, need to be able to deal with it and think strategically for purposes of the project.
- Subcommittee member: approach may apply to more than one recommendation – taking time to implement.
- Future meetings: the data subcommittee will be discussing and adopting recommendations in September and will see draft recommendations in August.
- Co-chair- At the August meeting, the co-chairs will present an outline of the structure of the final report: executive summary, background, methodology, etc.
- Discussion of how agencies tracked demographic information.
 - Everyone tracks race, ethnicity, gender, and age except agency, which may have forgotten to check box for race/ethnicity in survey.
 - There is a lot of inconsistency about where agencies are getting their demographic information from (self-reporting, criminal history, officer perception, etc.) Not much information about sources of demographic information and how they prioritize sources of information.
 - Figuring out where demographic information should come from and developing a standard way of recording and reporting that information could be part of the planning process for a data warehouse.
 - Correlation and consolidation at the source will be more accurate for data collection.
- Differentiating between data collection and data possession.
 - E.g., courts may pull information from outside systems, so they don't collect it but have it/have access to it.
- Consistency – 7 agencies reported they use CLEMIS to collect criminal legal data, others used a variety of commercial/internal systems.
 - There were comments asking how difficult it would be to collect and integrate data from different agencies.
 - A member noted that the CLEMIS system is more friendly to utilize than older systems.
 - Another member reminded the group that data is only as good as due diligence of people who collected and entered it.
 - Any data system may have the same individual entered multiple times so data normalization is critical to consolidate information about single individuals.
 - Question: who has tracking data on cases the prosecutor's office has declined? Where that can be analyzed and examined?
 - Working on this now in prosecutor's office but not every case or charge is sent to their office.
 - Can be incidents law enforcement responds to that don't get to prosecutor's office.
 - Missing information about cases where no warrant request is submitted, or where warrant requests go to local prosecutors instead of the county prosecutor's office.
- Data storage and management

- 12 responses about where data is stored. Two agencies said they store data both externally and internally. Five agencies reported they store data internally and another 5 reported they store externally.
- No detailed information from most about where data is stored externally, except for Prosecutor's Office, where databases managed by Washtenaw County IT department.
- Next steps:
 - Email with draft language for following up and scheduling 1:1 interviews for those who responded to the survey.
 - Hold in depth conversations – 30 min to follow up and getting deeper into questions asked.
 - Alex/Vera will have follow-up conversations and drill into how data is collected/what it means for them. Look for places to clarify with participants.
 - Need to think of questions about process or database management.
 - Need to ask what we will do with the information collected. Needs to be a formal planning process that would involve all stakeholders that would flesh out more complicated issues of integration and the details.
 - What kind of information do we want to know – point to features of a data warehouse the subcommittee potentially wants.
 - Survey has highlighted problems that need to be discussed in a formal planning process.
 - Need to have a planning process, ability of individual agencies to look at historical data and their responses are limited (many responded they were unable to do that).
- Initial scan – seek responses from non-respondents and take Sandhya's approach for Phase 2.
- Subcommittee member: A lot of unknowns that need to be resolved, need to get out of the theoretical and go into specific fields that need to be collected, to be tactical and look at nuts and bolts of how it needs to be deployed. Those steps will be addressed in a data warehouse planning process.
- Co-chair: Our goal is to provide roadmap to look at issues and what needs to be addressed in a planning process, doing education with select institutions. Talk to them in real time about what the subcommittee is doing and what subcommittee is thinking of proposing so they can educate other institutions that potentially can be part of this too.
 - Question: Who is going to organize it or who can we talk to?
 - Vera/Mapping subgroup can interview people on the list.
 - Interviews – less than 30 minutes to discuss questions that Sandhya proposed.
 - Adjust questions based on who Vera interviews.
- Subcommittee member: Next meeting is July 13, and questions need to be formalized and invitations out soon to get things lined up.
 - Follow up concluded by July if goals are to have other planning processes sorted out by then as well.
 - Member suggested asking other subcommittees how ability to have a warehouse help their goals or work they're doing, touching base about what impact the ability to understand or track data would have on other subcommittee goals.
- Alex: May not make sense to ask other subcommittees now, as most probably aren't thinking about data at this point, but maybe can ask this later.

- Subcommittee member: Maybe in the future can be a question from co-chairs in a working group meeting.
- Co-chair: discusses excel spreadsheet that was shared with subcommittee – looked at different integrated data systems in different counties.
 - Two that have been around for a while – Allegheny, PA, and Multnomah, OR.
 - Allegheny – lots of substance from the interview, trying to figure out how to make it a narrative to share with whole subcommittee.
 - Allegheny County began its data warehouse in 2000 and Multnomah County began in 1996. Both had initial discussions with coalition of people, decided to move forward and had to bring together relevant stakeholders. Both planning processes included policy makers, IT management staff and public participating as well.
 - We would need a planning process in Washtenaw because we don't have all representatives at the table on this data subcommittee.
 - The excel sheet presented to the subcommittee is a draft. The co-chair asks all to look at it and noted that they are planning to refine it.
 - Data warehouses for Allegheny and Multnomah were developed in stages. Allegheny started with DHS then built out to justice systems then schools. Over 22 years, built up and out. Multnomah started with a few agencies then built out.
 - Need to discuss question of what kind of things we want to see in data warehouse, need a lot of input from the mapping folks, need discussion on what things to flag for the planning phase
 - Homework assignment: co-chair asked the subcommittee members to look at the spreadsheet describing the counties' data warehouse efforts and send any edits back.
 - Subcommittee member mentions the culture of openness and accountability in the counties to make sure the data warehouse is sustained.
- Subcommittee member: We should note that these county efforts have been in place for a long time and learning from them is very important. The state of the art has shifted in 20 years so how we approach this will also contrast what's in place in Allegheny/Multnomah.
- Both sites have said they got a lot of funding – federal, state, foundation – because of they had an integrated data system in place.
- Meeting has quorum by the end. Vote to approve last meeting's minutes.
 - Minutes unanimously approved. Minutes adopted.
- Co-chair: May have to add a meeting or add an additional 30 minutes to another meeting if we are pressed for time before our recommendations are due to the Working Group.
- Some subcommittee members say they won't be able to make next meeting.
 - Co-chair will check with group about moving the date of the next meeting because it's vacation season.
- Meeting adjourns.