

# Rutgers AICPA Data Analytics Research Initiative (RADAR)

## April In-Person Meeting

April 30, 2018

### Attendees:

Brian Miller	Won No
Dan Balla	Kari Lee
Mary Grace Davenport	Feigi Huang
Alessandro Peirano	Nuriddin Tojiboyev
Mike Leonardson	Qiao Li
Denise Motta	Tiffany Chiu
Muthu Raj	Abdulrahman Alrefai
Michael Wynen	Soo Hyun Cho
Brian Collins	Ben Yoon
Kristine Hasenstab	Jamie Freiman
Ami Beers	Brittany Kelley
Amy Pawlicki	
Dorothy McQuilken	<b><i>Via Teleconference</i></b>
Miklos Vasarhelyi	Kelly Hnatt
Jason Guthrie	Susan Davis
Vanessa Teitelbaum	Nicole Deschamps
Brian Wolohan	

### Meeting Minutes:

- The purpose of this meeting was to discuss the status of the process mining and MADS projects, and to determine the next steps and path to completion.
- The board approved the January 10<sup>th</sup> and April 4<sup>th</sup> meeting minutes for posting to the RADAR [website](#).
- **Process Mining Project Update:** The meeting began with an update from the process mining research team.
  - Process mining is a technique that uses event logs (i.e. chronological records of computer system activities) to analyze business processes. The objective of this project is to leverage process mining techniques to help form a view of the effectiveness of internal controls.
  - The research team has developed a risk assessment framework that:
    - Identifies variants (i.e. a single path followed by one or more process instances with identical routings) in the data and classifies them as either “acceptable” or “notable”.

- Prioritizes the subcategories of “acceptable” and “notable” variants based on risk scores and a materiality threshold (similar to the MADS framework discussed below).
- The research team is working with a procure-to-pay data set that includes both event log and transactional data. The team has identified variants and has classified them as either “acceptable” or “notable”. For each sub category identified, they have applied a risk score (i.e. “less acceptable”, “notable”, “moderately notable” or “highly notable”) and have prioritized the final output based on the risk score and a materiality threshold. When analyzing the data set, the team looked for missing activities, activities that may not be in the correct order and redundant activities.
- The board discussed the results of the analysis and the methods used to score and prioritize the final output. The board will regroup and determine what the next steps will be.
- **MADS Project update:** The MADS framework begins with a full population of data and, by applying filters and data analytics, narrows the population down to an output of “notable” items (i.e. riskier or items more likely to contain a misstatement) that may require additional analysis. The research team presented four instances of how the current MADS framework can fit within different cycles. The MADS framework was applied to the following cycles: (1) Expenditures, (2) Revenue, (3) Payroll and (4) General Ledger (Journal Entry testing).
  - **Expenditure Example:**
    - The research team is working with a procure-to-pay data set for this example. The team has applied specific filters (based on the data set and feedback from the board) to the data set covering the existence audit assertion. After applying the step 1 filters, the team analyzed the output and applied additional sub filters to further sift through data. The final population of “notable items” were prioritized based on a scoring criteria determined by the research team.
    - The team has begun to evaluate the framework and compare it against traditional sampling techniques using three different methods.
  - **Revenue Example:**
    - This team is working with a data set covering the revenue cycle. Similar to the expenditure team, the researchers have developed specific filters to sift through the population of data. The team applied three of the four filters related to areas deemed higher risk, and is still working on the fourth due to computing issues. They have identified their population of “notable items” and, in some cases, have applied additional sub-filters to narrow down the population. The team is determining whether to apply the filters to the data for areas deemed moderate or low risk.
  - **Payroll Example:**
    - This team is working with a payroll data set. The research team has also developed filters based on the data set and feedback from the board. Similar to the other teams, filters and sub filters have been applied to the data set in order to narrow down the data to a population of “notable” items. The team prioritized the final output based on dollar amount and the number of filters violated. The next step is to determine whether or not to continue with this

example, as it was suggested that it may be difficult to apply the MADS framework to the payroll area. If it is decided that the team should move forward they will continue their evaluation of the framework by comparing it against traditional sampling techniques.

- ***General Ledger (i.e. Journal Entry) Example:***
  - The team is working with a data set and is focused, more specifically, on one business unit. The team has developed a number of filters, and has begun applying them to the data. Similar to the other teams, the researchers are determining ways to prioritize or weight the final output of “notable items” and evaluate the framework.
- Overall, each example illustrated how the MADS framework can be applied to different areas. It was noted that the framework can be refined based on the situation. That is, in some cases sub-filters were applied to the output from the first set of filters. This does not always have to be the case. In some instances, it may be better to apply a different data analytic technique (e.g. visualizations, clustering analysis, etc.) to help analyze the population further.
- The board discussed the results from each of the examples, and agreed that additional testing should be done to help validate the filters that have been developed. Due to the limitations of the current data sets, it was noted that it would be helpful to apply the frameworks to new data sets or different cycles.

**Next Steps:**

- The board will regroup and determine ways to help “test” the frameworks (e.g. obtain additional data set, run pilots, identify additional cycles to apply the framework to etc.) and move the project to completion.