Big Ideas Submission

Big Idea Proposed Name	Big Data Based **Government Economic Monitoring** and Targeted Action (GEM) – Illustrating with the city of Newark (From Surveillance Capitalism to Cuddling Government)
What problem(s) does this Big Idea aim to solve?	The US Government has six major welfare programs (Temporary Assistance for Needy Families program, Medicaid, Child's Health Insurance Program, Food Stamps, Supplemental Security Program, Earned Income Tax Credit, and Housing Assistant program) aimed at improving citizens' well-being. These programs are similar to efforts in many other countries. They are usually aimed at goals such as decreasing poverty treating the sick, supporting the elderly, etc. These programs are of great importance but are often seen as extremely inefficient and targeted at wide social pathologies inadequate for narrow problems of the needy. This proposal aims to use modern information technology to directly guide government action toward serving individuals and family units that are in need. Additionally, the methodology monitors the evolution of a case as it progresses toward the resolution of its identified pathologies. This methodology also allows for the rapid creation and/or dissolution of actions / programs to deal with emerging needs such as opiate addiction, temporary local unemployment, child neglect, partner abuse, temporary unemployment, family distress, etc. Extant government policies are often based on census data. Big Data based Government Economic Monitoring and Targeted Action (GEM) will enable the close-to-timely-reality tracking of microeconomic status (income, expenses, demand, etc.) in real-time, based on big data collected from multiple data sources. For instance, the Consumer Price Index (CPI), a basis of estimating consumer expenses, is calculated using a survey to determine the commodities included in a typical market basket. Instead of relying on a market basket provided by an outdated survey, GEM will identify the demand based on existing consumers for each commodity. This will be determined using big data provided on each individual or household through extant accessible e-commerce purchase data. Notably, we will be able to capture microeconomic

	activities in greater detail. For example, measuring spending on entertainment as a portion of total income. Starting from a granular household unit (or possibly individual) up to an aggregated statewide level, GEM will provide microeconomic information to the state government as well as to the general public. Such information could be used to help increase social welfare. Further, we will provide a continuous monitoring system for a variety of microeconomic status variables. Thus, the information provided by GEM will enhance the government's ability to track microeconomic status indicators leading to better policy adoption, quicker identification of the peoples' needs, and increased support for people using these services. Traditional measures are very blunt and government action (e.g. food stamps) is coarse and wasteful as a result. This approach allows very targeted government actions, which protected by privacy enhancing measures (PEM), will decrease human suffering at the point of action, allow for recovery, and increase economic output. Furthermore, the approach will allow the action on social pathologies that are today too infrequent for government action. Although this discussion talks about income and expense data the sources of open and limited exogenous data are innumerous. For example car registration (can detect car sales), Zillow (can detect real estate sale and distress), cab rides in New York City, triangulation on wireless towers in addition to GPS data for location, pictures from satellites, progressively information from self-driving cars, from home ovens and refrigerators, etc. Each social pathologies may have a different set of markers. Big data has generated serious concerns pertaining to both invasion of privacy and its use as a tool for totalitarian repression. This methodology will provide privacy by organizing data into separate action elements such as information provisioning, algorithmic use, diagnostic and government actions. These elements would not be traceable due to different forms of normalization, encryption and data protection.
Goal(s)	Development of methodologies designed to: 1. Provide real-time microeconomic status information. • From granular details of microeconomic activities to aggregated state level measures. • Sustainable, timely, and relevant information gathering. 2. Provide a decision support system for the government to use for better policy adoption.

	• Continuous monitoring system of microeconomic statistics. • Direct action leads for individual needs related to social pathologies. 3. Privacy management and individual household need identification. 4. Sectioning data and the creation of encrypted indices for groups, families and individuals. 5. Continuous monitoring of implemented social actions and changes. 6. Bottom up analytics and aggregate case findings on social events such as: the opioid addiction crisis, food shortages, family abuse, alcohol addiction, job loss, functional dysfunction, technological displacement, job incompetence, mental dysfunction, etc.
Does this project have a target completion date?	N/A This is an effort to totally revamp the identification, implementation and evaluation of social actions instituted by the government. This unique approach uses many measures that adhere to a privacy protection approach yet enable analysis on an individual level in order to maximize effectiveness. This approach results in a continuing effort to understand individual/household stress-points. This information in turn empowers the government to act democratically to improve the lives of citizens. As a consequence of the localized nature of this project there are no holistic completion dates. There will however, be a set of discrete sub-projects with milestones, and deliverable deadlines. The multitude of data sources already in use in the private domain will allow for constant updates of social status. This provides a base of information for state intervention. We expect that after one year some of the work will be actionable by the government entity (e.g. the City of Newark) while further research and testing continues.
Please describe any work that is already under way.	The Continuous Auditing and Reporting Lab (CARLAB) has led a variety of projects which utilize exogenous data (big data) to provide insight and solutions to private sector firms and regulators globally. A highly trained group, the CARLAB is continuously gathering exogenous data from a variety of sources including the dark web. CARLAB members utilize data analytic skills to dismantle big data into useful information. As a result, the group has already developed numerous tools tailored for data mining. These tools are highly applicable to the GEM project.

	Additionally, during a number of miscellaneous projects, software robots were developed to capture social media data and other exogenous data for a designated purpose. These robots can also redirect data in order to validate data strings, create dashboards, or continuously monitor performance. Such robots can be used to facilitate the processes in the GEM project.
List your collaborators, e.g. faculty, staff, departments, institutes, centers, schools	Rutgers Business School Rutgers Accounting Research Center / Continuous Auditing & Reporting Lab (principal investigators) Potential collaborators: City of Newark Rutgers Law School School of Public Affairs and Administration Rutgers NJ Medical School School of Public Health School of Nursing Program in Criminal Justice
Interdisciplinary collaboration between departments, schools, colleges, and even chancellor-led units is a fundamental element of Big Ideas. Please provide additional details about the collaboration that will take place (or is taking place) for this project.	GEM captures a variety of available data. To process data into useful information for further analysis, the project needs to acquire domain knowledge. For example, we currently have a partial module of GEM capturing drug trafficking data from the dark web. Furthermore, we are in conjunction with the auditor general of the state of New York, collecting data on cleanliness of the city of New York. To proceed with further analysis, we need to collaborate with different academic organizations within Rutgers to garner competencies in areas such as medical and legal. Additionally, we will continuously need to look carefully at the legal and technical aspects of the work to protect privacy. In this way GEM requires interdisciplinary collaboration between different departments and schools in the Rutgers community. The results of GEM are designed to help public administrators come up with the new public policy or retool extant social welfare programs (Temporary Assistance for Needy Families program, Medicaid, Child's Health Insurance Program, Food Stamps, Supplemental Security Program, Earned Income Tax Credit, and Housing Assistant program).

What are the benefits of this collaboration and how will it advance knowledge? How might this Big Idea be transformational beyond a single school or unit?	The benefits of this project will be far reaching. Since the information gathered by GEM includes various microeconomic items, it can prove useful in a variety of different fields. For example, GEM will capture information on the demand for certain goods in specific region. This information could help aid both economics and marketing researchers. Additionally, the criminal data collected could be used by research programs in the field of criminal justice and law schools. Medical / health data may help on the identification of endemics and deep learning evaluation of the data structures may help in some health diagnostics and treatment.
In what ways might this Big Idea support and advance the Rutgers Strategic Plan?	Rutgers supports research to better future society. However, such data has been utilized by private companies to generate higher profit. We feel that what has been called in the literature "surveillance capitalism" can be transformed into "cuddling government." The main goal of GEM is to enhance the social welfare. This can be done by ensuring that no particular group monopolizes the information gathered by GEM for their own interest.
Why is Rutgers uniquely positioned to house this Big Idea?	CARLab is the preeminent leader in Accounting Information Systems research and has been rated #1 in the world over the last 28 years by Brigham Young University ranking. The faculty and graduate students are both devoted, and highly knowledgeable about the topic of data analytics. A large number of prior studies for corporate and government organizations have provided a wealth of relevant and readily transferable technologies.
How will this Big Idea help Rutgers fulfill its mission?	GEM contributes to every one of Rutgers' threefold missions. First, GEM will provide a great microeconomic data source to researchers that will illuminate a host of future research opportunities. Additionally, the construction of such system will facilitate research that caters to the current needs of New Jersey citizens. Second, GEM will utilize numerous sets of cutting-edge techniques including machine learning and data analytics. The construction of such systems has been dominated by a limited number of private firms who possess enough resources and related technologies. However, as the leading group in Accounting Information Systems, we are highly confident that the CARLab will be able to build a robust data system at a nationwide level for the benefit of society. Finally, GEM will help the public service sector, including state government, provide better support to its citizens. Information provided by GEM could be utilized as a type of decision support system for the state government.

What might be this Big Idea's impact on the university at large, or society as a whole? How might this Big Idea change the world?	Continuous data analytics based on many sources of available data enables GEM to have statewide microeconomic census in real-time basis. Such methodology could be highly applicable not only in New Jersey, or the United States, but in every nation. Every government could use such a microeconomic monitoring system for better policy adoption. GEM could be the starting point for building such Big Census systems. These programs could be segmented into direct action programs in needy communities to address social pathologies such as hunger, opioid addiction, family abuse, elderly abandonment, family disintegration, hunger, etc. See how it works This proposal aims at using the community of Newark, if adequate interest and commitment can be obtained, as the illustration of direct data-based action, and privacy preserving methodologies. The Rutgers Newark campus has extensive links and work with the extended Newark government and community. This project will not only make the impact on the City of Newark, but also will make the impact on the States, the Nation, and the world. To protect privacy, we introduce a database which is engineered based on blockchain technology. The system will transfer every information into a non-interpretable nor trackable form to link any individual who owns the data. The data will be stored in a private blockchain operated by a limited number of government agencies without allowing any outsider access to the data source. For instance, if a government officer queries the system, the government officer will get a limited answer where highly personal information is disguised. Other information which the officer is prohibited to access will be in a hashed form which will appear meaningless. Followed are the partial examples of certain modules of our privacy preserving system. Also, we would like to explain the impact by using one pilot case: In 2016, Newark, the largest city in New Jersey recorded at least 12,960 children to be in extreme poverty. Also, one-third of New Jersey's residents could not afford basic necessities of food. Poverty leads children to be in hunger, which is ironic since New Jersey is the fourth wealthiest state in the nation. Our goal for this example is to illustrate how GEM could help Food Stamp distribution to provide direct support for children who will be starving at this point as well. GEM first collects the data near the neighborhood related to poverty and hunger. For example, to find out poverty level in

	the neighborhood, we access to publicly available credit history, property report, forfeiture penalty reports, and any other necessary data. Based on this information, we conduct series of analysis to determine the neighborhood economic status while preserving personal privacy. In the other hand, we utilize the regional amazon purchase history, nearby grocery store sales data, and previous food stamp distributed history to find the neighborhood grocery purchase pattern. During data analysis, we heavily secure the data by using our proposed privacy-preserved blockchain database. Every input information provided by different institutions will be transformed into a hashed form where the data is stored in a blockchain which is visible only to permitted agencies. The security system will replace all the personal information into hashed indexes which changes every millisecond. This information will be highly secured in a private blockchain, however, even if this information be breached by any chance, there will be no mean to trace back the hashed index to individuals. As well as considering the neighborhood activities, GEM will be mapping the school districts where direct state actions are needed. On top of these results, the system suggests the area where the food stamp should have to be distributed. The system tracks the food stamp distribution history while keep monitoring to capture children in hunger in the neighborhood, it calculates the best timing for subsequent distribution. Therefore, the state could provide the food stamps to the children without any distribution gap. Moreover, when a state receives a request from the resident for food stamps, the system could provide evidence material which could assist the request review. Notably, the system will not utilize or provide any personal information which violates privacy.

Roughly, what level of investment will these additional resources require?	
Human Resources	10 Ph.D. Students over 4 years = 2,000,000 Faculty stipends 30K* 4* 4 = 480,000 Administrative support 200K*4= 800,000 Miscellaneous 200k*4=800,000 Promotion 200k*4=800,000 Total – 4,880,000
Capital Resources	Computer databases, storage, and resources: $2,000,000

Other Resources	Resources for studies at others Rutgers faculties 4 * 4* 200,000 =$3,200,000

Initial Team

Miklos A. Vasarhelyi, Principal Investigator (miklosv@rutgers.edu)

Alex Kogan

Dan Palmon

Farrokh Langdana

Arion Cheong

Jamie Freiman

Andrea Rozario

Xinxin Wang

Divya Anantharaman

William Glasgal (Volcker Alliance)

Ricardo Lopes Cardoso (FGV, Brazil)