



The Computerworld Honors Program

Honoring those who use Information Technology to benefit society

Final Copy of Case Study

Status:

Laureate

Year:

2013

Organization Name:

United Nations Global Pulse

Organization URL:

<http://www.unglobalpulse.org/>

Project Name:

Big data & unemployment through the lens of social media

Please select the category in which you are submitting your entry

Emerging Technology

Please provide an overview of the nominated project. Describe the problem it was intended to solve, the technology or approach used, how it was innovative and any technical or other challenges that had to be overcome for successful implementation and adoption. (In 300 words or less.)

This project analyzed social media chatter and sentiment to identify trends related to unemployment increases, and inform policymakers of likely effects. UN Global Pulse is exploring ways that new types of digital data can complement official government statistics to track global development in real-time. Analyzing 500,000 blogs, forums and news sites, SAS® Social Media Analytics and SAS Text Miner examined two years of social media data from the US and Ireland for references to unemployment and coping mechanisms. SAS compared mood scores and conversation volume with official unemployment statistics to see if upticks in those topics were indicators of spikes in unemployment. The analysis revealed that increased chatter about cutting back on groceries, increasing use of public transportation and downgrading one's automobile could predict an unemployment spike. (Appendix 1) After a spike, surges in conversations about canceled vacations, reduced health care spending, and foreclosures/evictions shed light on lagging economic effects. This is invaluable for policymakers trying to mitigate negative effects of increased unemployment. By

analyzing sentiment, each unemployment reference received a "mood score" based on tone, such as "optimistic" or "depressed", and then sorted by themes including housing, transportation and finance. In the US, a rise in "hostile" or "depressed" mood occurred four months before the unemployment spike. Increases in "anxious" chatter in Ireland correlated with a spike five months later. Increased "confused" chatter preceded a spike by three months, while "confident" chatter decreased significantly two months out. A dashboard (Appendix 2) displayed trends such as mood change over time, and leading and lagging indicators of unemployment shocks. A large challenge was the data acquisition, filtering and validation process. After eliminating noise, sorting by topic, and determining categories, a sample was validated, manually, to ensure the quality and accuracy of the query and filtering processes.

When was this project implemented or last updated? (Please specify month and year.) Has it incorporated new technologies and/or other innovations since its initial deployment? (In 300 words or less.)

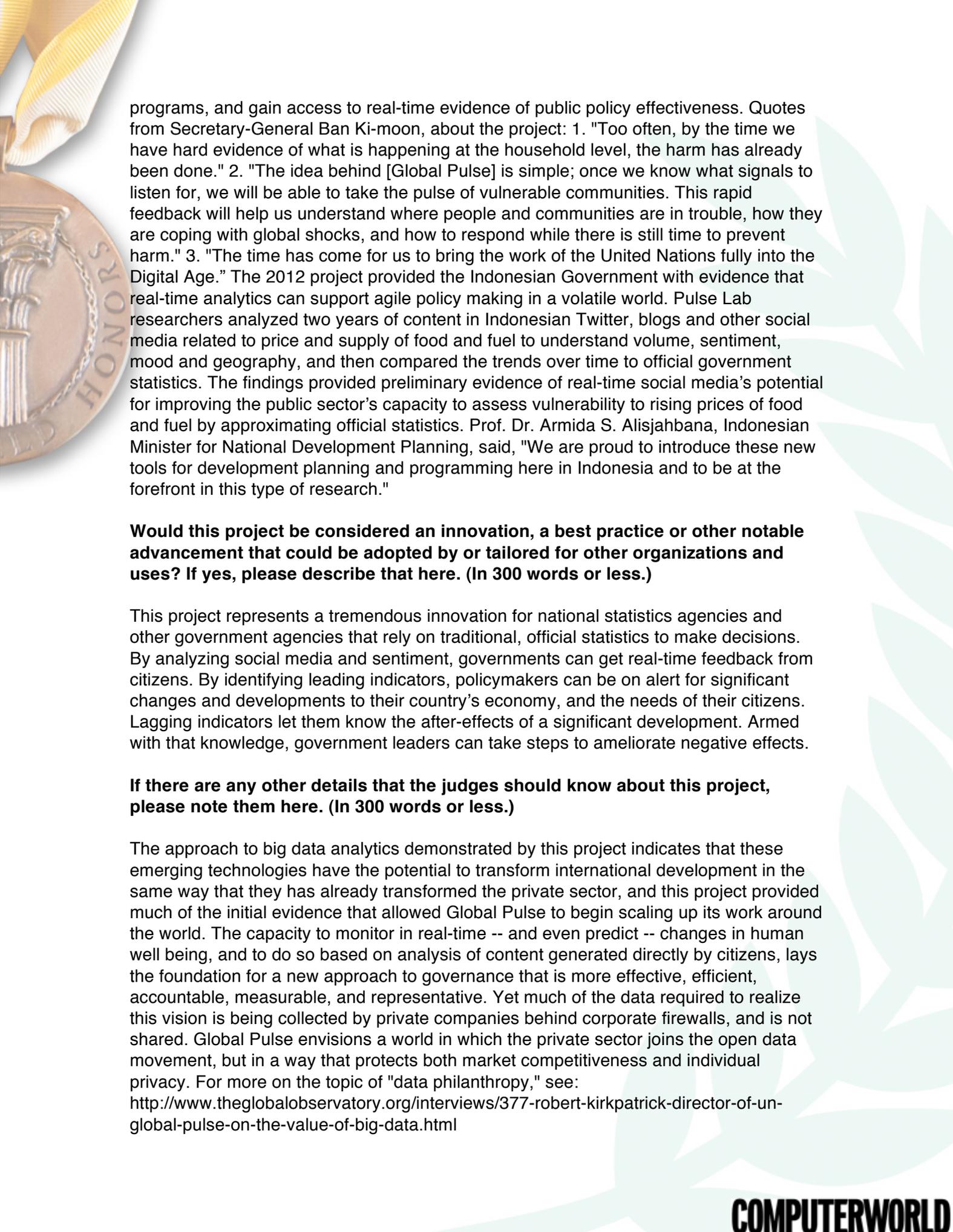
December 2011 was the conclusion of the project in the US and Ireland. In 2012, the same methodology was applied by Global Pulse in a similar project in Indonesia to explore how analysis of online conversation about food prices, fuel prices and unemployment could support faster, better targeting of national social protection programs. SAS and Global Pulse were able to: 1. Monitor more than 200,000 online Indonesian documents created every day. 2. Extract relevant discussion on rice, cooking oil, fuel and employment. 3. Capture sentiment and mood of social media content. 4. Detect location, price, availability, specific government programs. 5. Explore results with graphs, statistics, charts and demographic. As an example: on the issue of unemployment, several objectives of the program were about addressing questions such as: "Can we understand qualitative experiences and feelings around unemployment to complement official statistics?"; "Can online conversations provide an early indicator of impending job losses?"; "Can the data help policymakers enrich their understanding of how communities cope?" In this exercise, SAS Text Analytics was used to uncover key themes related to coping with economic stress, perform content categorization and text analytic model scoring to assign metric to mood state dimensions. Social media data was then correlated with macroeconomic and unemployment metrics. Longitudinal analytical models and algorithms were used to leading and lagging indicators of trends such as inflation and unemployment.

Is implementation of the project complete? If no, please describe the project's phases and which phase the project is now in. (In 300 words or less.)

The project is complete, but Global Pulse and SAS will continue to partner on similar projects in other countries.

Please provide at least one example of how the technology project has benefited a specific individual or organization. Feel free to include personal quotes from individuals who have directly benefited from the work. (In 300 words or less.)

The 2011 project demonstrated to the UN General Assembly that social media analytics represents an unprecedented opportunity for the public sector to strengthen its capacity for early warning, improve the design and implementation of international development

A gold medal with a ribbon is visible in the top left corner. The medal features a classical architectural design and the word "HONORS" is partially visible. A large, stylized green leaf graphic is on the right side of the page.

programs, and gain access to real-time evidence of public policy effectiveness. Quotes from Secretary-General Ban Ki-moon, about the project: 1. "Too often, by the time we have hard evidence of what is happening at the household level, the harm has already been done." 2. "The idea behind [Global Pulse] is simple; once we know what signals to listen for, we will be able to take the pulse of vulnerable communities. This rapid feedback will help us understand where people and communities are in trouble, how they are coping with global shocks, and how to respond while there is still time to prevent harm." 3. "The time has come for us to bring the work of the United Nations fully into the Digital Age." The 2012 project provided the Indonesian Government with evidence that real-time analytics can support agile policy making in a volatile world. Pulse Lab researchers analyzed two years of content in Indonesian Twitter, blogs and other social media related to price and supply of food and fuel to understand volume, sentiment, mood and geography, and then compared the trends over time to official government statistics. The findings provided preliminary evidence of real-time social media's potential for improving the public sector's capacity to assess vulnerability to rising prices of food and fuel by approximating official statistics. Prof. Dr. Armida S. Alisjahbana, Indonesian Minister for National Development Planning, said, "We are proud to introduce these new tools for development planning and programming here in Indonesia and to be at the forefront in this type of research."

Would this project be considered an innovation, a best practice or other notable advancement that could be adopted by or tailored for other organizations and uses? If yes, please describe that here. (In 300 words or less.)

This project represents a tremendous innovation for national statistics agencies and other government agencies that rely on traditional, official statistics to make decisions. By analyzing social media and sentiment, governments can get real-time feedback from citizens. By identifying leading indicators, policymakers can be on alert for significant changes and developments to their country's economy, and the needs of their citizens. Lagging indicators let them know the after-effects of a significant development. Armed with that knowledge, government leaders can take steps to ameliorate negative effects.

If there are any other details that the judges should know about this project, please note them here. (In 300 words or less.)

The approach to big data analytics demonstrated by this project indicates that these emerging technologies have the potential to transform international development in the same way that they has already transformed the private sector, and this project provided much of the initial evidence that allowed Global Pulse to begin scaling up its work around the world. The capacity to monitor in real-time -- and even predict -- changes in human well being, and to do so based on analysis of content generated directly by citizens, lays the foundation for a new approach to governance that is more effective, efficient, accountable, measurable, and representative. Yet much of the data required to realize this vision is being collected by private companies behind corporate firewalls, and is not shared. Global Pulse envisions a world in which the private sector joins the open data movement, but in a way that protects both market competitiveness and individual privacy. For more on the topic of "data philanthropy," see: <http://www.theglobalobservatory.org/interviews/377-robert-kirkpatrick-director-of-un-global-pulse-on-the-value-of-big-data.html>