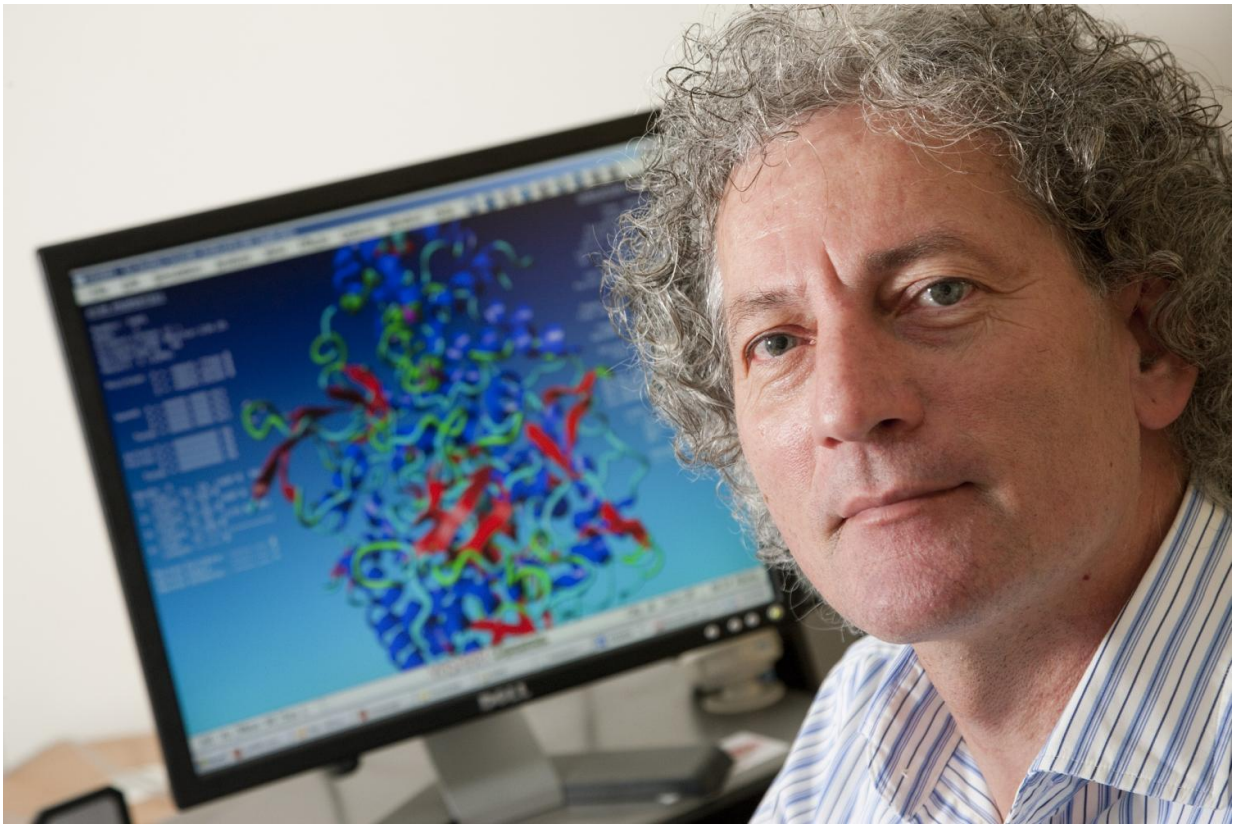
Big data analytics—Nostradamus of the 21st century

November 30 2016



Griffith acting head of School of ICT Bela Stantic

With much of the debate as to whether Donald Trump or Hillary Clinton would win the election taking place online, people blogging, tweeting or updating social media with their thoughts on the topic provided data

researchers with a rich source of information about what people were thinking and feeling about the election race.

Associate Prof Stantic was so confident in the result that he publicly announced his prediction even for the known swing states - and his calculations for all swing states were right.

"My algorithms showed clearly to me that based on past patterns and sentiment in social media that Trump, by November 8, would take over the lead, despite only having a 10 per cent chance to win according to all polls at that time," he said.

"In a public address on big data the day before I even correctly identified all main states that Trump would win (including Florida, North Carolina and Pennsylvania). Someone in the audience quickly checked online and said according to polls Hilary was 84 per cent favourite.

"I answered that people are likely to be more honest when telling friends rather than answering polls. It is scary how accurate prediction can be done by analysing social media."

When Griffith's Big Data and Smart Analytics Lab analysed comments on Twitter towards the end of July, it predicted that if the US Presidential election had been held at that time, Trump would have been the winner over Clinton. Those results were shared at the time in an article on The Conversation.

The same lab using the same method predicted and announced in a public lecture a week before the Australian federal election that the Coalition would win over the ALP.

Over the past several years, presidential elections have served as great testbeds in social media, big data, and analytics, which can go into great

detail on how campaigns use this information to find out more about voters."Such analytics can provide much more accurate information than telephone polling, especially in a day and age where people have caller ID and don't have landlines," Associate Prof Stantic says.

"This is why the polls leading up to the election had such inconsistent results."

"The amount of data that all of us generate is truly staggering, and it is continuing to grow. This publicly available data is secret treasure of information if we know how to discover it."

Associate Professor Stantic says big data analytics is a discipline faced with the challenge of managing the sheer volume of data and turning it into something useful.

"It makes predictions about the future based on the patterns of the past, and finds relationships buried in the data that no one has noticed."

Similar predictions about the environmental changes of Great Barrier Reef based on 'Human Sensors' and Gold Coast visitor satisfaction have been done on projects funded by National Environmental Science Program and City of Gold Coast.

"To further improve predictive power of Big Data analytics there is a need for smarter and faster algorithms to perform deep learning on the large volumes of data drawn from diverse, and we are working on it."

Provided by Griffith University