



Will the truth about MySejahtera please stand up?

By Ravee Suntheralingam April 2, 2022

- *Data, when analysed, can be used for all sorts of behavioural modelling*
- *Sceptical about claim data captured is only retained for a max 60 days*



MySejahtera was introduced in April 2020 to the Malaysian public. The Minister of Health then was Adham Baba, second from left. Khairy Jamaluddin (3rd from left) the current Minister of Health, was the Science, Technology and Innovation Minister then.

What exactly is the truth about MySejahtera? Malaysians are entitled to understand the contradictory positions taken by the Minister of Health and that of the Public Accounts Committee. After following the issue with great interest, even before the current flare up, these are my thoughts.

MySejahtera is not costly to develop. I doubt it cost more than US\$238,000 (RM1 million) for the initial three modules back in March 2020. This is not factoring in the running and management of such an app.

Today, there are commercially available off-the-shelf mobile banking applications that have more complex features than MySejahtera which can be developed and customized for less than RM1.2 million says the developer of such applications.

I studied the features of MySejahtera and did my own design of the technical and database architecture as well as required data flow. I am able to do this because I just happen to have these competencies after 23 years in the Information and Communication Technology sector. Apart from hands-on design, code cutting and implementation, a total of 10 years of my technology career was channelled towards managing the assessment of at least 1,000 research and development projects for its technical and commercial feasibility on behalf of the Ministry of Science, Technology and Innovation.

MySejahtera data structures can handle amongst others, an appointment and facility database, information and location updates, decision and status management as well as a user and identification code database. It has the capability to assess COVID risk, manage check-in by users, and identify hot spots as well as providing access to support services and a host of relevant information and statistics.

I am not concerned if the raw MySejahtera application without its data is sold or retained by the Government. It is the abuse of the analytics of recorded users and location data that is the concern. MySejahtera collects two types of data. The first type relates the personal information of the 24.5 million registered users such as name, identification, and address and vaccination status. The second type relates to the business and non-business premises visited by users of the MySejahtera application.

I do not think the Rakyat is aware that the data collected by MySejahtera, when analysed by an expert, can be used for all sorts of behavioural modelling of users as well as business and non-business premises. The statement by Datuk Seri Anwar Ibrahim regarding MySejahtera data being used for product and telemarketing does not reflect the true extent of the seriousness of the issue.

With so much data captured by MySejahtera, a data science analyst can predict many things about the behaviour of users as well as the business premises registered with MySejahtera. One example can be the estimation of revenue generated by business premises based on some form of stochastic model which can then be used by the Tax Department to compare with actual tax submissions. Another example is the creation of demographic models that can be used to refine race based economic policies.

I am also sceptical about the narrative that all scanned data captured by MySejahtera is only retained for a maximum of 60 days before it is deleted. How do we know this for a fact? How do we know if copies of such collected data have been made without the knowledge of Government? More importantly, has any form of predictive analytics

been applied to data retained within the 60 day window on an on-going basis and has the outcome of such analysed and consolidated data been retained and progressively refined over time?

There is no doubt in my mind that predictive analytics of MySejahtera data presents many spin-off commercial opportunities. Are Malaysians willing to allow the Government, or its agencies, or politically linked businesses, to exploit both their raw location and movement data as well as data that has been subject to the predictive analytics process?

I have three suggestions for the Government, to consider now that Malaysia enters the official endemic phase (from 1 April 2022):

1. Firstly, discontinue the mandatory scanning of QR codes using the MySejahtera application and restrict its use to recording vaccination status and public information only.
2. Secondly, permanent erase all raw scanned location data captured by MySejahtera from all databases as well as any outcome of analysed and consolidated data that was subjected to the predictive analytics process.
3. Thirdly, impose mandatory jail terms for anyone in possession of raw MySejahtera personal and location data as well as any derivative or outcome of data that has been analysed and consolidated using any predictive analytics tool.

There is no doubt that an application like MySejahtera is a useful tool to manage a crisis. However, this same application can become a double edged sword if data privacy is not honoured.