

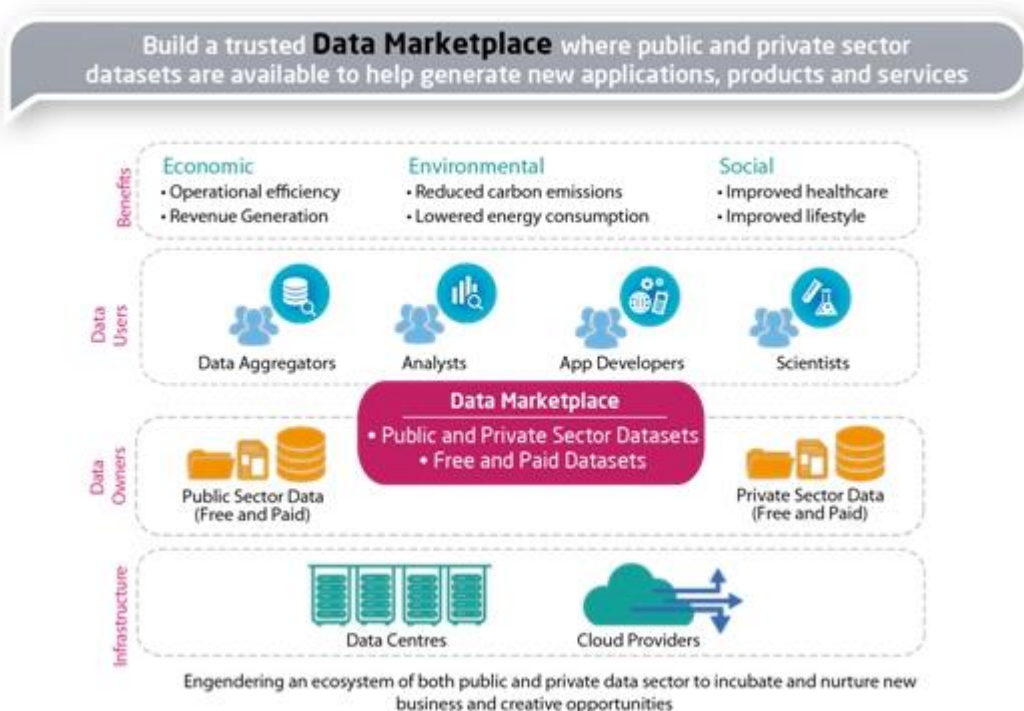
## Data Marketplace

Described as the “New Oil” of the 21<sup>st</sup> century, data has the potential to shape economies around the world. Just as how oil has driven much of our economic developments, the value of data is derived from its potentially limitless applications.

A healthy economy achieves a balance between the supply and demand of goods and services. Likewise, a well-balanced data-driven economy relies on the availability and the use of data to generate economic value. The challenge of striking such a balance remains. On the demand side, data users find it hard to discover datasets as the datasets are not tagged. Searching online typically only brings users to a webpage. Even if datasets are found, users struggle to use the data meaningfully as information such as data source, quality, version and terms of use is usually not available. When presented with multiple datasets, users have a problem comparing and choosing the dataset to use. “Meshing” multiple datasets also requires significant time and resources. Sometimes, individual programmes have to be written to access customised and disparate Application Programming Interfaces (“**APIs**”).

On the supply side, data owners face challenges making data available and accessible for monetisation. They need help to make their datasets discoverable, and require advice on the data format to adopt for easy access.

We want to create a data marketplace, comprising both public and private sector datasets, to address the obstacles faced by both providers and users of data. This will differentiate us from other data marketplaces in creating an effective data-driven economy. The data marketplace can help data owners monetise their data with data licensing fees. Such features can help them realise new revenue streams and further encourage them to share data. This helps the users of data, such as innovators and businesses, generate a myriad of new applications, products and services, thereby realising economic, environmental and social benefits.



To realise this, we propose the following measures:

- **Standardised Guidelines for Application Interfaces**

To allow datasets to be easily accessed and integrated, we need to implement industry-wide guidelines and best practice standards to overcome the problem of disparate APIs. One possibility is to encourage the adoption of a common technical standard to enable developers to access and mesh datasets without having to write specific APIs.

- **Assessing the Risk of Re-Identification Systems**

We need a system that measures the risk of de-identified data being re-identified. This will help data owners to de-identify personal identification information while retaining data value. With this, data owners may be more willing to share data without compromising the confidential information of individuals.

- **Registry to Enable Easy Access and Discovery of Data**

To help people discover and access data more easily, we propose setting up a registry of datasets built on the use of open source software. In addition, to nurture an environment of trust in the consumption and provision of datasets, users will be required to register before accessing or purchasing datasets.

- **Data Certification to Instil Confidence in Data Use**

We propose to implement a self-assess certification scheme where data owners answer a free online questionnaire. Based on the results, the dataset will be accorded the appropriate ratings. The questionnaire guides data owners on the practical steps of providing data for effective use. There will also be information to help data users understand the use of data. Such a certification system helps both data owners and users have a common understanding of the data, thereby fostering trust.

- **Community Participation to Develop a Data Economy**

Whether we succeed in building a vibrant data economy depends on the degree of community participation. The community must be aware of the value of data and how data sharing can benefit them. We can convey this information by sharing success stories and organising activities such as a Festival of Ideas. To cultivate a data-literate community of developers and users, we can start a series of accredited courses and educational programmes. Platforms such as data-themed competitions and “hackathons” will also engage industry experts, business leaders and data owners, and spur them to discover new innovative ideas.

### **Opportunities for the ICM Sectors**

A data marketplace with a high degree of community participation can generate significant outcomes. Firstly, a new sector called “info-mediaries” may emerge. Info-mediaries are agents who help organisations manage, control and monetise their data. Secondly, data analytics can help businesses plan their marketing strategies. Insights drawn from data can reveal valuable information on the lifestyles and purchasing patterns of their customers. This information can help in creating tailored strategies to increase outreach and growth. Thirdly, the demand for data-related jobs, in areas such as data analytics, data science, business intelligence and data-oriented software development, will increase.