**Response from Victoria University on the
Academy of the Social Sciences in Australia (ASSA) discussion paper**

'*Doing social science in 2032: Charting national research infrastructure priorities*', a discussion paper that invites input into the Academy's *Decadal Plan for Social Science Research Infrastructure 2023–32*:

Victoria University (VU) would like to thank the Academy for the opportunity to provide feedback on the discussion paper, as a way to help inform the Academy's *Decadal Plan*.

Some initial observations on the discussion paper:

The language in the document has a tendency to be very STEM-related in its focus on *producing, accessing, analysing and managing data; generating new knowledge*. The Decadal Plan seemingly seeks to conform to STEM standards and economic parameters, which in effect, commodifies human experience. For example, there is a significant focus on collecting and managing 'big data' and the benefits of automated/computer-assisted analysis, but relatively minimal focus on the very important deep contextual, nuanced analysis that is unique to social sciences (i.e., the measurement of meaning). See for example the text on the ability to automate aspects of qualitative research, on page 26.

Thus, VU suggests that greater emphasis could be placed on the supportive role such automation could provide researchers. This, importantly, is not intended to reduce the involvement of the researcher, but free them to explore other areas of inquiry. In effect, this automation process, could/ should support the growth of the researcher's experiential knowledge, not (inadvertently) replace it.

As a small University, there is a concern that this model of national research infrastructure is likely to favour the larger, more resource-rich universities in terms of developing and managing 'assets'. VU is very supportive of reframing what social science 'infrastructure' is, and that the paper provides some useful direction to address skills/training gaps and specialised support for university social science R&D applications. Many social scientists struggle with accessing and manipulating large databases. Support staff with the requisite skills are essential to the process but sometimes viewed as an 'add on'. They are in fact an integral part of quantitative research teams and need funding support. VU would strongly endorse the comments in the Strategy about their importance.

VU would like to emphasise the importance of the shared infrastructure and the sharing of knowledge /skills across the sector and across the globe. Many of the issues Australian social scientists address are common to international social scientists, so collaboration to form common databases is highly valuable for those Australian social scientists studying these areas. Participation takes time and generally money to establish the collaborations. Given the tyranny of distance, the obstacles are greater for Australian researchers than for many other countries. This, therefore, needs funding support.

A good example of the benefits of such sharing and global collaboration is the Global Burden of Disease. This would probably be classified as a health database but for any health economist or other social scientists working in the field, it is the first port of call for any study to understand a

disease or evaluate its treatment. It is a collaboration of over 10,000 scientists, epidemiologists, health economists etc., world-wide, synthesizing nearly 300,000 data sources to cover 350 health outcomes and risk factors providing data for about 200 countries and some states within those countries. See: https://www.healthdata.org/research-analysis/gbd

Responses to the specific questions in the discussion paper:

Q1 How would you modify or augment our description of the current state of assets, systems, rules and skills and training?

Ethical considerations around consent for sharing of data is a serious issue that has not been addressed adequately in this discussion paper. Data sovereignty discussions are also crucial when working with Traditional Owners and the Aboriginal communities and it will be vital to ensure their engagement throughout the discussions around consent and sharing.

Q2 Can you provide specific examples of data-related challenges your research team faces, where shared infrastructure could significantly boost productivity or support your research aspirations?

Need for increased access to privately held collections. For example, in the area of literature and creative writing, these would include data held by organisations such as AustLit and Nielssen Bookscan. As more work is done on publishing inequity, this kind of data becomes more important.

There is a need for better and more consistent digitising of permanent and touring collections (in libraries, museums and self-managed archives).

Information about curriculum held by the Department of Education and Training could be made more accessible including access to the recommended curriculum lists.

Digitising, storing and translating narratives outside 'official' historical documents (such as oral histories and collected ephemera).

 The ability to share and receive knowledge internationally and improved translation functions – especially when incorporating languages that are spoken rather than written (such as Creoles and other patois),

Clear mapping of resources, drawing out connections, and have a simple resource to go to in order to determine where things can be found/explored. Improved discovery tools, and improved quality of information about how to access discovery tools, will assist.

In lieu of historical census records, researchers use historical Vital Registration data. Currently this has been digitalised primarily as a way of raising money by state based departments, catalogued by surname which the most relevant to family historians and indexed using different platforms which do not speak to each other.

Q3 Which needs can be met through improvements to existing assets, systems, rules or skills and training? Briefly describe the improvements required

Need for more discipline specific research skills - yes - not all social sciences are the same, and yet there is often a sameness about the research being undertaken. A lot of research in education is undertaken to provide evidence for what we already know, without focusing on impact. Because it can be hard to access large bodies of qualitative data a lot of small scale qualitative studies take

place - these are great - but there needs to be some sense of how they fit with everything else and how connections might be made. A lot of energy goes into the process (following strict research structures) and not enough into the thinking and connecting.

Q4 Which needs require that the sector advocates for new assets, systems, rules or training?

Briefly describe any new infrastructures you think are required, including where possible examples and any requirements for successful implementation (e.g., incentives, funding, partnerships)

A Program for keeping and making accessible census data is urgently needed.

Funding for social science research and the use of shared data in universities needs to be increased - people need time to think as well as research/ write. We need to break the formulaic writing that is evident in a lot of educational research.

Q8 Which needs require that the sector advocates for new assets, systems, rules or training? Briefly describe the required new infrastructures, including where possible, any requirements for successful implementation (e.g., incentives, funding, partnerships).

There would need to be very considerable ethics training for researchers concerning data sharing. Researchers also need legal training and clarification about who owns data and under what circumstances it can be shared. At the moment this varies from state to state.

If we are going down this path, we have to ensure that AI training begins with ethical considerations and that the sharing of data will in fact produce a result that is ethically acceptable to the participants, and not just those with vested interests. We certainly need very clear guidance from the NHMRC and AHEC.