

SUMMARY

Researchers select and construct 'big data' in the same way as they do 'regular' data. The difference between big data and regular data comes down to quantity and variety. This report addresses the question of what will happen to the relationship and interaction between data, knowledge and applied research now that 'regular' data are developing into 'big' data? What are the implications for academic research, especially those fields that work with personal data?

Big data offers the research community opportunities, but it also poses new challenges with respect to data accessibility, analysis and storage and raises various legal and ethical issues. Researchers increasingly need to call in experts to assist in using big data. Big data is also generating new questions concerning the reproducibility and validity of research, how far we can generalise from research results, and whether existing methods and statistical techniques remain useful.

Researchers who work with personal data should therefore develop cross-disciplinary and discipline-specific methods and techniques for dealing with big data, in accordance with the statutory frameworks and guidelines. That is what is needed to make full use of big data for the benefit of research while at the same time protecting research subjects, as well as researchers and their networks. Although research involving big data has greater affinity with inductive reasoning, theoretical underpinnings are precisely what is required. Big data may make it even more important to work from a theoretical basis than has traditionally been the case. It is also advisable for researchers who use personal data to work in teams with data specialists and legal and ethical experts, all of whom should be given full credit for their expert support. Another recommendation is that researchers should have easy

access to such expertise, preferably embedded in a sound, secure infrastructure within a research institute or faculty.

Current and future generations of researchers who work with personal data should be trained in using big data, a task in which all universities and university hospitals, the Association of Universities in the Netherlands (VSNU), and the Netherlands Federation of University Medical Centres (NFU) have a role to play. Training should address not only technical aspects but also such matters as data quality, ethics and privacy.

To ensure the integration of local data sources and infrastructure, and to facilitate cooperation within the research community and with public- and private-sector partners, an overarching infrastructure is required. When designing this new infrastructure, the Ministry of Education, Culture and Science, acting in consultation with the Netherlands Organisation for Scientific Research (NWO), SURF (collaborative ICT organisation for Dutch education and research) and VSNU, should focus on research that involves personal data and should join Open Science, Open Data and FAIR data initiatives, for example the international GO FAIR network and the Netherlands' Open Science Platform.

Big data offers new opportunities to enhance research in the Netherlands in many different areas of application. If the Netherlands takes advantage of these opportunities, it can boost its international reputation in big data research and attain a prominent position in academic research.