SUMMARY

To perform its tasks properly, government regularly collects data on individuals, enterprises, institutions, markets, roads, nature, emissions, homes, and so on. If researchers were given better access to public data,[1] they could use it to generate new knowledge. Such knowledge could be useful for government policy, for education, or for other purposes in society. This advisory report discusses what is needed to promote the reuse of public data by researchers.

The reuse of public data by researchers serves a dual purpose. For academic research, public data is valuable because it often covers a large population and/or a long period of time; in addition, people do not usually drop out of populations covered by government records. For government, public data is a potential source of knowledge – knowledge that sheds light on the effect of its policies, and helps it to carry out complex tasks. An example of a complex task is the protection of such values as security, privacy, self-determination, solidarity and fair competition. We do not really know how data-driven government policy in a digital society impacts these values. Researchers can use public data to give policy firm scientific underpinnings while bearing in mind the broader context.

The Netherlands has access to solid statistical information; public authorities have been collecting it for a long time and their work is reliable. Nevertheless, when we compare the Netherlands with other countries, we see that it is not exploiting all the opportunities that this information offers. For example, the Scandinavian countries have national registers of diagnoses, drug use and so on that are easily accessible to researchers. The Netherlands has similar longitudinal datasets, but they are not

---

1 Not to be confused with 'publicly *accessible* data'.

necessarily made available to researchers. Government organisations often do not disclose data about persons and enterprises in order to protect privacy or trade secrets.

The national government, provincial and municipal authorities, other organisations that perform public tasks (such as housing corporations), Statistics Netherlands (CBS), but also researchers themselves can all contribute in their own way to improving the reuse of public data. The Committee on the Accessibility for Scientific Research of Data Administered in Public and Semi-public Record Systems argues that the national government bears primary responsibility for making data collected for public tasks permanently available.

The Committee makes the following recommendations.
a. National government: make inter-ministerial agreements on the procedure and conditions under which government administrators must actively make data available to researchers. One option would be to do so through the Data and Government Steering Committee, which reports to the Ministry of the Interior and Kingdom Relations. Appoint a national Chief Public Data Officer (CPDO), who will play a pivotal role in concluding these agreements. Let data nodes with national coverage facilitate those who provide and request public data. Make inter-ministerial agreements about this as well.
b. Government organisations: state in your annual report which public data has been collected in the reporting year, and how that data is made available for research purposes. Encourage reuse by making data Findable, Accessible, Interoperable and Reusable (FAIR). Appoint data stewards whose specific task is to make public data more accessible and reusable. In addition to expertise in the field of data management, the data stewards must also have a thorough knowledge of privacy issues and copyright law, for example. Do not charge researchers who wish to use the data more than the marginal cost of data provision itself, and make it clear in advance how the costs will be calculated. It is common practice for government organisations to charge researchers for providing them with public data. It is only fair that they should do so, but there is little clarity at the moment about the cost structure that they use.
c. Statistics Netherlands: wherever possible, facilitate access to and reuse of the data to which you yourself have access by virtue of your statistical duties. Statistics Netherlands and the Ministry of the Economic Affairs and Climate Policy: you are currently examining how to tighten up the legal framework in a way that gives private parties with a public task a firm legal basis for supplying their data to Statistics Netherlands. As you do so, be especially mindful of the importance of making public data available for research purposes.
d. Researchers: do not hesitate to approach government organisations with a research question, and ask them to make data from their records more accessible. It goes without saying that all parties must comply with privacy legislation and

prevent the disclosure of individual data. In publications, refer to the uniform resource identifier (URI) of your data sources. At the end of a study, provide the relevant government organisation with feedback on your research results, the contribution your study has made to scholarship, and the importance of the data the government organisation has provided as evidence for government policy. You should preferably also share this information with the Chief Public Data Officer.

The Committee in no way recommends that government organisations should collect more data for research purposes. Its recommendations concern the reuse of public data that has already been collected. As the advisory report reveals, there are already good examples of how research can be used to underpin government policy in many areas, such as the environment, labour markets and education. These examples can inspire greater efforts to reuse public data in research in other areas, something that serves the interests of both academic research and society.