Dirk Meusel, Peggy Göpfert, Nicole Wagner, Nora Merker, Wilhelm Kirch

Data Based Web Application for Public Health Relevant Research Data

Datenbank-basierte Web-Applikation für Public Health relevante Forschungsdaten

The demand of giving public account of research activities is met by German Public Health research by the compilation of empirical research data into so called Public Use Files and Public Documentation Files. Since research designs and deriving data sets in Public Health research can be described as heterogeneous as well as they frequently contain sensible data, disseminating these files via Internet requires special considerations concerning the computing architecture as well as software development. The workgroup Public Use Files has set up a pilot web application for testing the dissemination of Public Use Files via Internet. The web application is accessible at URL http://www.public-health.tu-dresden.de and introduced in the paper.

Keywords: Internet, Public Use Files, Research Practice


Stichworte: Internet, Public Use Files, Forschungspraxis

Introduction

German research in Public Health was funded with approximately 50 million Euro of public incomes. An issue, which faces new attention in Germany too, is giving a public account of research activities. Research teams in numerous European or North American countries meet this demand of accountability by publishing so called Public or Scientific Use Files. On the one hand, these files provide transparency of research practice as well as research results. On the other, they provide the opportunity of secondary analysis of the scientific data. Common examples of this practice are the Public Use Files of the Bundesgesundheitssurvey (Federal Health Survey), which are being carried out and provided by the Robert Koch Institute in Berlin. The compilation of Public Use Files furthermore implies the advantage of promoting research results without additional expenses as well as the opportunity of stimulating research in one's emphasised field of interest.
In accordance with the DGPH (German Association of Public Health) and the project's supporting agency, a workgroup PUF (Public Use Files) was founded for publishing data of Public Health Research in Germany. Members of this workgroup are, among others, the commission of Public Affairs of the DGPH and the German Coordinating Agency for Public Health in Freiburg. This workgroup chooses suitable research projects, designs the formal framework for the Public Use File creation of high quality, and preserves the interest of the particular research project groups. One of the main objectives within this PUF workgroup is the dissemination of Public Use Files via Internet.

For the task of disseminating Public Health relevant research data via Internet, the PUF workgroup has set up a pilot web application with a twofold objective. On one hand, the prospective PUF user shall be able to access information regarding the files. On the other hand, the research teams shall be able to keep the information presented in the PUF up-to-date. This paper examines the special characteristics of empirical data in Public Health science and their implications on the dissemination of Public Use Files via internet. Following that, the data based web application is introduced.

**Empirical data of Public Health Research and the Internet**

Archiving and disseminating empirical data derived from Public Health research is a complex and multifaceted undertaking. As mentioned elsewhere (Meusel/Göpfert/Kirch 2001), several reasons support the strategy to create public or scientific use files from empirical research data in Public Health science. Firstly, the data deriving from the original research project will not get lost after the research team finished work on the data and on the project. In the past, once the members of a research team left, undocumented data files confronted even experts in secondary analyses with the nearly unsolvable problem of uncovering the connection between variables in the data file and the corresponding items in the field instruments. Hence, the creation of a public use file is an efficient way of archiving the data and preserving the metadata about it. Secondly, utilising empirical data by scientific third party users promotes a more comparative approach to research in Public Health sciences in general. Research methods and results become more lucid and comprehensible. Finally, and most important, public use files provide the opportunity to achieve secondary data analysis without high monetary and time expenditures. They can be utilised for educational purposes in Public Health studies as well as for follow-up research. Furthermore, standardised public use files convey the scientific exchange between national and international research units and, for this reason, present an additional source to promote a project's results.

Empirical data deriving from Public Health research hold two special characteristics with serious implications for their dissemination via internet. Firstly, data sets are very heterogeneous. One end of this continuum presents epidemiological survey data that can be characterised by large samples sizes and an extensive set of quantifiable data variables. Objective of those studies is the population wide description of the examined phenomena. The other end symbolizes case studies with a qualitative approach. Most often, those studies survey interview data to be investigated by different forms of content analysis. Between these ends lay differing research and sample designs with a mixture