

Improved access to research data in the field of vocational education and training BIBB's Research Data Centre

HOLGER ALDA

► **BIBB launched a Research Data Centre (BIBB-FDZ) in 2008. The FDZ processes the microdata from BIBB research projects and makes the respective data sets available to external researchers and scholars for their own analysis. As a constitutive part of the German Data Infrastructure coordinated by the German Data Forum (RatSWD), BIBB-FDZ helps to reduce the workload and costs involved in conducting data surveys in the field of vocational education and training (VET) by assisting external researchers in the use of existing data.**

IMPLEMENTING THE BIBB-FDZ INFRASTRUCTURE: PRIORITIES DURING THE SET-UP PHASE

Setting up a research service such as BIBB-FDZ requires a special infrastructure, including data documentation standards and regulations for data access for external researchers in line with the German data protection laws. Working within a joint development project, the Federal Ministry of Education and Research (BMBF) and BIBB supported the establishment of the BIBB-FDZ between 2008 and 2009. Some of the main areas of work during the set-up phase were:

- establishing an FDZ data archive and a secure IT-environment for FDZ studies and FDZ data sets;

- clarifying data protection rules for transmitting project information and (weakly or formally anonymised) data sets (e. g. Scientific Use Files, SUF) to external researchers, scholars, and institutions;
- laying down standards for data documentation;
- setting up a web site;
- allocating unique identifiers for data sets (digital object identifiers) and standardising the citation of BIBB-FDZ data sets, and
- informing and advising interested parties and users on the possible uses of BIBB-FDZ data sets for their own research.

These individual work segments were coordinated in such a way that transparent, standardised access procedures, sets of data documentation and other working tools (e. g. the questionnaire for each data set) are now in place for BIBB's microdata. The respective procedures and standards apply to all data sets in the FDZ.

The recent work of the BIBB-FDZ has focused on filling the established structures (cf. Figure) with more data sets, such as the BIBB Transition Survey 2006 on the level of individuals or the BIBB-Cost-Benefit-Survey 2007 on the level of firms.

AVAILABLE DATA SETS

An initial examination during the implementation of BIBB-FDZ showed that BIBB has generated more than 200 data sets since 2000; in addition, new sets are continuously generated by current BIBB research projects. Of course, not all of these data sets are of equal interest for secondary analyses. For this reason, the first data sets that BIBB-FDZ prepared for external usage were those with a proven demand from external researchers. Furthermore, emphasis is placed on the preparation of data covering the entire range of BIBB's research topics as comprehensively as possible. To this end, the FDZ developed documentation systems that take account of the heterogeneousness of BIBB's microdata without becoming arbitrary or too idiosyncratic.

This was achieved by dividing data sets in the field of VET into five classes based on the stations a (German) individual typically passes through in the process of acquiring education, skills and qualifications: general schooling, the first threshold, initial vocational training, the second threshold, and working life. The introduction of a further differentiation – between data on the level of firms and individuals (persons) – produces a ten cell-matrix filled with appropriate data sets from BIBB's research activities (cf. Figure).

By now BIBB-FDZ offers at least one data set in each cell. All data sets include a full set of documentation consisting of a data/methods manual, questionnaire, test data and other working tools. In total, 21 single data sets are currently accessible to external researchers.

HOW BIBB-FDZ'S DATA SETS ARE USED

Since 2008, BIBB-FDZ has analysed enquiry statistics in order to obtain information about the use of its data sets. In the following, results of these analyses will be summarised and commented upon, covering the time period from April 2008 to the end of 2010.

In this period, BIBB-FDZ answered 468 enquiries referring to BIBB microdata. Of these, 392 (approx. 84 %) were external requests, while 76 came from inside BIBB. The requests cover the entire range of BIBB-FDZ's services and products. As expected, most of the data-related enquiries pertained to accessing data, the terms of use and the like. However, BIBB-FDZ also answered methodology-related enquiries (regarding certain regression techniques, for example) and requests concerning specific topical VET-related content (e. g. the effect of continuing vocational training activities on unemployment risks). The range of research topics extends from studies on the potential for creativity in the workplace all the way to potential earnings and opportunities for wage increases against the background of technical and organisational change in the workplace environment.

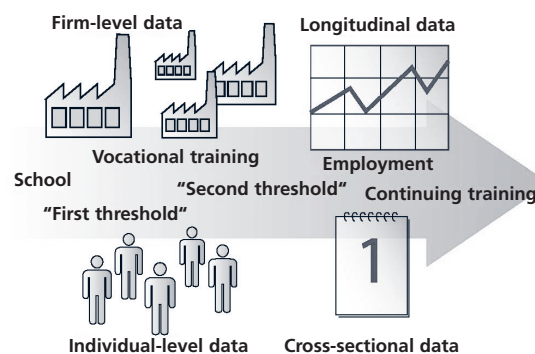
Looking at individual BIBB data sets, the BIBB/BAuA Employment Survey 2006 and the preceding surveys (BIBB/IAB surveys in 1998, 1991, 1985, 1979) currently account for the greatest number of enquiries by far. Approximately 42 per cent of all external requests pertain to the use of this data set or lead to its use. Another six per cent revolve around the BIBB/IAB surveys. About 18 per cent of all requests do not concern a specific data set and can thus be classified as pertaining entirely to content-related or methodological consultation services.

The number of enquiries received is not identical with the number of formal contracts with data users. The BIBB-FDZ signed a total of 107 use contracts with 234 persons for the different types of data access (SUF, remote data access, use of on-site work stations for guest researchers at BIBB-FDZ). The vast majority of these contracts relate to the BIBB/BAuA Employment Survey 2006.

Based on this information, the question of whether there is an actual need for an FDZ at BIBB can be answered with an unconditional "yes." This is also made clear by a glance at the institutions using BIBB data: they range from numerous national and international universities to well-known social and economic research institutes. The table illustrates how productive this use is, showing which kinds of papers external researchers produce with the BIBB-FDZ's data sets.

The table shows that the output of external researchers using BIBB data covers the entire spectrum of academic publications. Given that BIBB-FDZ has only established its infrastructure and prepared a significant number of BIBB data sets for external usage since the beginning of 2009, and seeing that it takes some time until scholarly articles are published – in reviewed journals, in particular, – the current output is highly encouraging and promises to increase significantly in the near future.

Figure The structure of BIBB-FDZ data sets



Source: www.bibb.de/de/50124.htm

Table External academic output generated by use of BIBB-FDZ data in 2009/2010

Kind of publication	Number	Percentage
SSCI journal	2	5.6
Other reviewed journal	5	13.8
Other journal	3	8.3
Monograph	2	5.6
Contribution to edited collection	2	5.6
Working paper	6	16.7
Research note	3	8.3
Manuscript	2	5.6
Research report	4	11.1
Work/manufacturing report	5	13.8
Other academic paper (e.g. master's thesis)	2	5.6
Total	36	100

OUTLOOK

The establishment of regular FDZ operations always requires a certain amount of work and investment in (IT) infrastructure. The analysis of the BIBB-FDZ enquiry statistics shows that within the first two and a half years since the FDZ's foundation, these efforts and costs have been offset by the substantial benefit yielded for VET research. The German Science Council and the German Data Forum have recommended the establishment and operation of a research data centre in all institutions generating data of potential interest to external researchers. By setting up its FDZ, BIBB has put these recommendations into practice and reached a good interim level in providing data-related services to the VET research community. A visible expression of this assessment is the accreditation of BIBB-FDZ as a part of the German data infrastructure coordinated by RatSWD. It is therefore BIBB's firm aim to continue to operate the FDZ at the current level. Interested parties can check the Research Data Centre's website at any time (<http://www.bibb.de/en/50113.htm>),¹ where they will find detailed and up-to-date information about the services and support offered by BIBB-FDZ. ■

¹ Note that, currently, information on our website is mainly in German. However, further English content will be added throughout 2011/2012.