Report

Archived data set (according to the current DDI standard)

Author(s):
Frei, Andreas; Axhausen, Kay W.

Publication Date:
2009

Permanent Link:
https://doi.org/10.3929/ethz-a-005864349

Rights / License:
In Copyright - Non-Commercial Use Permitted
Deliverable D9

 Archived data set (according to then current DDI standard)

 Contract-No. TREN/06/FP6TR/S07.66711/038682-KITE
Work Package Team
IVT
ETH
CH – 8093 Zürich
Switzerland
Tel. +41 (1) 633 4102
Fax +41 (1) 633 1057
Email frei@ivt.baug.ethz.ch
URL: www.ivt.ethz.ch
Andreas Frei
Kay W Axhausen
### Document History:

<table>
<thead>
<tr>
<th>Date</th>
<th>Author</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>15/06/2009</td>
<td>Andreas Frei</td>
<td>Initial revision</td>
</tr>
<tr>
<td>18/06/2009</td>
<td>Tobias Kuhnimhof</td>
<td>1&lt;sup&gt;st&lt;/sup&gt; revision</td>
</tr>
<tr>
<td>18/06/2009</td>
<td>Andreas Frei</td>
<td>2&lt;sup&gt;nd&lt;/sup&gt; revision</td>
</tr>
</tbody>
</table>
Table of contents

Abbreviations .................................................................................................................................. 5
Summary ............................................................................................................................................. 6
1  Introduction and background ....................................................................................................... 7
   1.1  NESSTAR ............................................................................................................................... 7
   1.2  ETHDA .................................................................................................................................... 8
2  Archived dataset .......................................................................................................................... 9
   2.1  Description of the dataset ...................................................................................................... 9
   2.2  Structure of the dataset .......................................................................................................... 9
3  How to access and use the archived KITE dataset ...................................................................... 11
   3.1  Introduction ............................................................................................................................ 11
   3.2  Authentication ...................................................................................................................... 11
   3.3  Using Nesstar WebView ......................................................................................................... 12
      3.3.1  Metadata ......................................................................................................................... 13
      3.3.2  Variable Description ....................................................................................................... 14
   3.4  Tabulation .............................................................................................................................. 15
   3.5  Analysis .................................................................................................................................. 15
   3.6  Additional Functionality ........................................................................................................ 15
   3.7  Download ............................................................................................................................... 15
      3.7.1  To download documentation ......................................................................................... 16
      3.7.2  Export PDF .................................................................................................................... 16
References ........................................................................................................................................... 17
Abbreviations

MEST – Methods for European Surveys of Travel Behaviour
INVERMO - Intermodale Vernetzung
DATELINE – Design and Application of a Survey for European Long-Distance Trips Based on a International Network of Expertise
KITE - A Knowledge Base for Intermodal Passenger Travel in Europe
CATI - Computer Assisted Telephone Interview
SP - Stated Preference
IVT - Institut für Verkehrsplanung und Transportsysteme der ETH Zürich
ETHTDA - ETH Travel Data Archive
NESSTAR - Networked social sciences tools and resources
DDI - Data Documentation Initiative
ETHTDA - Eidgenössische Technische Hochschule Travel Data Archive
Summary

The KITE survey data is archived with the Nesstar software according to the DDI standard (see [http://www.icpsr.umich.edu/DDI/](http://www.icpsr.umich.edu/DDI/)). The data structure follows the survey structure and contains the following groups of variables: Household, Person, Journey roster, Vehicles, Regular trips, non-regular journeys and SP-Dataset.

Further it is described how to access the archived data and how to handle the WebView interface which allows browsing the data itself, the metadata and calculate and create first statistics of the data. The archived dataset can be downloaded in several data format, which makes it very easy to process and analyse the data the user’s known environment.

The dataset will be available for everybody, after requesting an account for the ETHTDA. The ETHTDA server allocates travel data since 2002 and will host the data also in the future.
1 Introduction and background

Work package 3 of the KITE project tested in a pilot survey a suitable survey methodology that intends to close remaining information gaps about long-distance travel behaviour. Long-distance travel is only in rare cases a part of daily mobility (e.g. salesmen or ambassadors). Therefore, such movements are reported with low frequencies in surveys of daily mobility. However national travel surveys (NTS) dedicated to daily mobility are the only source for long-distance travel in many European countries (e.g. in Denmark or Netherlands). The problem in these surveys is the difficulty of obtaining representative statistics of long-distance travel even with relatively big samples. Therefore in most national travel surveys modules are dedicated to long-distance travel with the exceptions mentioned above. The most common format in these surveys is that respondents are asked to report their long-distance journeys in addition to their mobility on a given day, but for a longer reporting period (e.g. Great Britain NTS (2002-2004), French NTS (1993-1994), Swiss Microcensus on Travel Behavior (2005), and Swedish RES (2005/2006)). A further, but less frequent source of data for long-distance travel are surveys exclusively dedicated to long-distance travel. But past surveys show very different amounts of long-distance travel demand for the same countries, which make that those numbers hard to trust (see Dateline dataset with its consistently lower value for the number of trips per person and year).

The KITE pilot surveys were carried out in Switzerland, the Czech Republic and Portugal.

A main objective of this pilot survey was to test a survey methodology in different countries in Europe and to asses the quality of information that can be obtained from it. In addition, a stated preference survey was designed to gather information about market potentials and user requirements.

The focus of this deliverable is to present how the dataset is archived and how it is possible in the future for other users to get access to the surveyed data including meta-data, which makes it possible to understand the data and perform desired analysis.

The software used to archive the data on the on the servers of the IVT allows sophisticated data browsing tools which is seamlessly integrated into the users' desktop. It is fully based on Web technology. The data structure used for the achieved data is built upon the use of interoperability standards, including the XML syntax and the DDI standard (see http://www.icpsr.umich.edu/DDI/ for further information).

A further feature of the archived is the ability to browse data interactively, thereby opening valuable data resources to a wider community. The generic data browser operates on the Web.

1.1 NESSTAR

The used software for the data archiving is called NESSTAR (see Norwegian Social Science Data Services, 2009). The Nesstar software suite is an integrated suite of
products aimed at facilitating the locating and use of socio-economic, and similarly structured, data. It has been developed by the UK Data Archive at the University of Essex and by the Norwegian Social Science Data Services in Bergen building upon the R&D of the EC funded NESSTAR and FASTER projects. It allows users to browse distributed data catalogues over the Web, examine detailed information about the data (metadata), carry out simple data analysis (e.g. tabulations and graphical displays) and then download data, in whole or part, in one of a number of popular formats. The system contains registration and authentication facilities to filter access to data as necessary and a suite of data publishing and server management tools.

1.2 ETHDA

ETHDA ([http://tda.ethz.ch/](http://tda.ethz.ch/)) is a virtual platform for the travel data, which is available at the Institute for Transport planning and Systems (IVT) since May 2002. ETHDA allows the end users to browse the archived travel data over the web and enables simple statistical analysis. ETHDA provides comprehensive information on the travel data that has been developed or analysed at the IVT. The information includes details of the corresponding travel survey from which the database was derived. ETHDA is one of the early travel data archives under Data Documentation Initiative (DDI) standards. The archive is build around the NESSTAR system.
2 Archived dataset

2.1 Description of the dataset

This KITE dataset consists of four parts: The first part is a dataset conducted with a CATI in Switzerland and Portugal and with a face-to-face interview in the Czech Republic contains the relevant sociodemographic information; the second part consists of data about the basic statistics of long-distance travel in terms of a long-distance journey roster (i.e. number of long-distance journeys and if they are regular or irregular ones), which were conducted again with a CATI in Switzerland and Portugal and respectively as a face-to-face interview in the Czech Republic; and the third part of the data is a detailed travel diary conducted with a CATI respectively a face-to-face interview with the relevant information on stage based level for intermodal long distance journeys. The fourth part is a follow up SP dataset, conducted as a postal survey, which contains stated preference data from mode and route choice experiments.

2.2 Structure of the dataset

The structure of the dataset follows the screening which was used. The first two parts of the survey were the same for the whole sample, whereas the last part is different for regular long-distance travelers, persons with non-regular long-distance journeys during the last 8 weeks and persons with no long-distance journeys during the past 8 weeks.

The structure follows the following structure:
Have you traveled in this period to a destination further away than 75 km?

- **YES**
  - How many of such journeys?
  - Do you travel regularly?
    - **YES**
      - Did you do some of these journeys not regularly?
        - **YES**
          - How many?
            - 1 irregular journey
            - up to 9 irregular journeys
              Next most recent journey
        - **NO**
    - **NO**

- **NO**
  - Last long-distance journey undertaken
  - END

Regular Trip data

Did you do some of these journeys not regularly?

- **YES**
  - How many?
    - 1 irregular journey
    - up to 9 irregular journeys
      Next most recent journey

- **NO**
  - 1 journey

non-regular journeys data

(up to 3 journeys)
3 How to access and use the archived KITE dataset

3.1 Introduction

This guide provides a brief introduction to Nesstar WebView and how to access and download the archived KITE dataset.

The WebView of Nesstar enables users to:

- View survey and tabular data
- Browse metadata
- Create cross-tabulations of variables
- Run correlation and regression analyses
- Produce graphical representations of data
- Create new variables
- Recode existing variables
- Download the data in 7 formats

3.2 Authentication

The ETHTDA is accessible with a normal Web browser by the address: http://tda.ethz.ch/. The start page is illustrated in Figure 1.

To access the Nesstar WebView a user account with a password is needed. To get a user account, a request has to be sent to the Nesstar administrator.
With a password and a user name, the Nesstar web view can be accessed with the link in the upper menu of the ETHTDA’s start page (see Figure 1 for the illustration of the link).

If you followed the Nesstar WebView link, you will access the start page of the Nesstar WebView (see Figure 2).

**Figure 2** Start page of the Nesstar WebView

3.3 Using Nesstar WebView

The following information describes how to navigate around the Nesstar system. The data and metadata published to the ETHTDA Nesstar Server are available within WebView, the interface to the Nesstar Server.

Survey data, tabular data or other resources, e.g. reports, can be accessed by the ETHTDA Nesstar server. Published datasets and resources are associated with catalogues.

The left side of the screen contains the ‘browse list’. This contains a list of all catalogues and published resources available within the ETHTDA (see Figure 3 a)).

A represents a Nesstar catalogue in the browse list. - Click on to expand the catalogue. The Catalogues hold the different datasets which are published at the ETHTDA.

Clicking on a title will display the abstract text on the right side of the screen (see Figure 3 b)). Click on the dataset icon to display further information about that dataset. Clicking again on will open further levels of information.

The KITE data is archived under:
3.3.1 Metadata
Metadata relating to the chosen resource can be viewed in this section. Information selected in the browse list will display corresponding information in the main frame on the right, see Figure 4.

To view information about published data:
1. Click on a dataset title to view its associated abstract.
2. Select the dataset icon to open a list of directories associated with that dataset.
3. Click on Metadata and the following options will be:
   - Study Description
   - Data Files Description

3.3.2 Variable Description
All the variables contained in a dataset will be found in this folder. The KITE data variables are grouped into different sections:

- Household,
- Person,
- Journey roster,
- Vehicles,
- Regular trips,
- non-regular journeys and
- SP-Dataset.

Each section contains an ID variable which connects the different groups with each other.

The description section displays descriptive text about the data. Select a variable from the browse list to display information about that variable in the right frame. The data displays value listings, categories and the number of responses for the selected variable (see Figure 5). Note that income is country-specific and the Sysmiss here are the data from Portugal and Czech Republic.

Figure 5 ETHTDA Nesstar summary statistics of Income in Switzerland

<table>
<thead>
<tr>
<th>Values</th>
<th>Categories</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Below 2000 CHF</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>2001 to 4000 CHF</td>
<td>91</td>
</tr>
<tr>
<td>3</td>
<td>4001 to 6000 CHF</td>
<td>168</td>
</tr>
<tr>
<td>4</td>
<td>6001 to 8000 CHF</td>
<td>189</td>
</tr>
<tr>
<td>5</td>
<td>8001 to 10000 CHF</td>
<td>113</td>
</tr>
<tr>
<td>6</td>
<td>10001 to 12000 CHF</td>
<td>62</td>
</tr>
<tr>
<td>7</td>
<td>12001 to 14000 CHF</td>
<td>33</td>
</tr>
<tr>
<td>8</td>
<td>14000 CHF and more</td>
<td>0</td>
</tr>
<tr>
<td>9998</td>
<td>don't want to say</td>
<td>213</td>
</tr>
<tr>
<td>9999</td>
<td>don't know</td>
<td>102</td>
</tr>
<tr>
<td>Sysmiss</td>
<td></td>
<td>2531</td>
</tr>
</tbody>
</table>

Summary Statistics
- Valid cases 1011
- Missing cases 2531
This variable is numeric
3.4 Tabulation

The Tabulation and Analysis tabs will only be available once a dataset is opened. The active selection has a blue background and white font.

The option tabulation enables you to create a frequency distribution for one variable, or cross-tabulations for multiple variables.

To create a table containing one or more of the variables from a survey dataset:
1. Click on the Tabulation tab.
2. Select a variable from the browse list on the left. A list of variables is available once you open a survey by clicking on the study icon and selecting Variable Description.
3. Click on the variable name and select one of the four options that will then become available.
4. Repeat this process to select other variables that you wish to add to the table.

3.5 Analysis

This tab enables a correlation or regression analysis to be performed. Once selected following the instructions on the screen to add variables.

3.6 Additional Functionality

There are a number of icons in the top right-hand corner of the screen. Depending on the characteristics of the dataset that is currently open, certain icons will be available.

3.7 Download

To download a dataset (see Figure 6 for illustration):
1. Select a study from the browse list.
2. Click on the ‘Download’ icon.
3. Select the format of the dataset in the drop-down menu. The options are: SPSS, SPSS portable, NSDStat, Statistica, StatA, DIF, Dbase and SAS.
4. Click on ‘download’ and choose to save.
3.7.1 To download documentation
- Follow the steps above.
- Select the format of the documentation (html or xml)
- Click on ‘Download’.

3.7.2 Export PDF
Tables and charts can now be exported to PDF format. When the required information is displayed in the right hand frame, click on the ‘Export to pdf’ icon. A ‘File download’ window will then be displayed from which you can select to ‘Save’ the file to a selected directory.
References