

SNSF Career Tracker Cohorts Data Management Plan

1 Data collection and documentation

1.1 What data will you collect, observe, generate or reuse?

This project produces questionnaire-based survey data and makes use of pre-existing data sources as described below. The volume of data for an individual survey wave including all files such as questionnaire data spreadsheets, administrative data spreadsheets, code of online questionnaires, or lists of standard classifications is around 200 MB. The main data used in this project are the following:

- Administrative data provided by the SNSF. The administrative data is made available to the project team in Excel format. It contains personal identification data (names, gender, year of birth, contact data) and administrative data related to the funding the respondent receives (scheme, discipline, grading of the application, funding decision etc.). The personal data will be used for the sole purposes of contacting the survey participants and linking the survey data with those elements of the administrative data that are relevant for data collection and analyses.
- Survey data collected from the participating applicants to SNSF career funding schemes using an online survey software. The raw survey data produced by the survey software is in Excel and CSV format. The project team will process the raw survey data in preparation for analysis. The resulting data will be stored in several formats such as plain text (CSV) and Stata format. The data processing includes linking the survey data to the administrative data where necessary for data analysis.
- Data related to the implementation of the online questionnaire. This mainly includes Excel files containing standard classifications (e.g. classifications of occupations, lists of disciplines, or lists of research institutions).

1.2 How will the data be collected, observed or generated?

We collect data by means of online surveys with applicants to SNSF career funding schemes at the postdoc level. The surveys are implemented using LimeSurvey, an open source survey software, and are hosted locally on a server within the subnet of the Department of Social Sciences in the IT network of the University of Bern. The survey data is then partially linked with the administrative data of the applicants (gender, year of birth, discipline etc.), which is provided by the SNSF via encrypted transfer. In order to link the data without using personal identifiers such as names (for reasons of privacy and independence), we use separate password-secured look-up tables, one on the person level and one on the application level. The former matches the SNSF person number and the person identification number used in the CTC project (CTC person ID), the latter matches the SNSF application number and the CTC application ID. The CTC IDs are random numbers consisting of either seven (application ID) or eight (person ID) digits and are different from the SNSF person and application numbers.

1.3 What documentation and metadata will you provide with the data?

For each survey, we provide a detailed documentation containing information on the development of the questionnaire (e.g. pretesting), the final questionnaire (including screenshots), the survey participants and the field work (e.g. characteristics of the survey population, response rate, duration of the interviews), the data preparation and data quality (e.g. data cleaning, classifications, definition of variables, nonresponse analysis), and the final data (codebook). This documentation will allow external users to understand and work with the data (see section 4). The documentation will be made publicly available at the project website. In addition, the documentation will be included in the FORS data archive.

2 Ethical, legal and security issues

2.1 How will ethical issues be addressed and handled?

The participants of the surveys are informed in advance by the SNSF about the aims of the Career Tracker Cohorts (CTC) study and the data collection procedure. The participants are assured that their data will be treated with strict confidentiality and that their participation will be unrelated to the evaluation procedure of their grant application. Furthermore, they are notified about the transfer of their contact information and the required administrative data to the project team, and the use of this data throughout the project. After receiving this information, participants are given the possibility to opt out of the study. In this case, their data is not transmitted to the project team and they are not invited to participate in any of the surveys. Participants are also given the possibility to retract their consent at any later point in time, in which case they will no longer be contacted for any future surveys.

The SNSF provides the project team with the data of all participants who did not opt out via encrypted transfer. As mentioned above, we use separately saved and password-protected look-up tables that contain the necessary IDs (SNSF person and application numbers and CTC person and application IDs) to link the survey data with the administrative data. Separate IDs are used in order to guarantee privacy and independence of the evaluation and administrative processes at the SNSF.

The ethics commission of the Faculty of Business, Economics and Social Sciences of the University of Bern approved the CTC project on 25 September 2018.

2.2 How will data access and security be managed?

Access to the project data is limited to the members of the project team (see also section 3.1). Files containing sensitive information (identifying information such as names and addresses) are exchanged through secure-transfer between the SNSF and the project team and kept in an encrypted and password-protected format. Moreover, all files that contain identification information (e.g. names, addresses) are secured by a password that is only known to the members of the project team.

2.3 How will you handle copyright and Intellectual Property Rights issues?

The SNSF is the owner of the data and the deliverables. The property rights are set out in the mandate contract between the Swiss National Science Foundation (SNSF) and the University of Bern acting through the Institute of Sociology and the Interdisciplinary Centre for Gender Studies ICFG.

3 Data storage and preservation

3.1 How will your data be stored and backed-up during the research?

The collected survey data, the administrative data from the SNSF and all other project relevant documents are stored in a project folder on the cloud-based SWITCHdrive. SWITCHdrive is a non-commercial cloud storage service for institutions of higher education in Switzerland. All data on SWITCHdrive is stored exclusively on servers that are part of the Swiss academic network. Access to this folder is limited to the members of the project team at the University of Bern.

Furthermore, personal information of the participants (e.g. names, addresses) are stored in a database on a server within the subnet of the Department of Social Sciences in the secured IT network of the University of Bern. The connection is established through VPN and it is ssl encrypted. Daily backups of the SWITCHdrive folder are stored on a server within the subnet of the Department of Social Sciences in the secured IT network of the University of Bern. In addition, daily backups of

the data stored on the servers of the University of Bern are performed. Access to these backups is limited to the members of the project team at the University of Bern.

3.2 What is your data preservation plan?

We will preserve all data and documents related to the project on SWITCHdrive or on a server within the subnet of the Department of Social Sciences, for as long as the project is conducted by us. Upon termination of the contract we will return or destroy, as requested by the SNSF, all information that is not publicly accessible (confidential information) without keeping any copy thereof. The datasets and documentation published on FORSbase will be available according to FORS's regulations.

4 Data sharing and reuse

4.1 How and where will the data be shared?

The collected data will be made available to the scientific community in the form of scientific use files. The scientific use files, together with the documentation and plain text dictionaries of variables and labels will be archived on FORSbase in plain text format (CSV) and Stata format.

The scientific use files will be made anonymous so that they will not contain any personal identifiers. The privacy of the respondents will be further protected by anonymization measures for the data as well as by procedural measures concerning access to and use of the data.

As with most social-science survey data, identification of at least some of the individuals may be possible through elaborate procedures that exploit covariate patterns. Such attempts to identify individuals are prohibited by the data protection contract that researchers willing to use the data have to sign and are thus illegal. In general, the number of individuals that could potentially be identified by such procedures is small because surveys are typically conducted with samples from large populations. This is different for the CTC study where no sample is drawn and all applicants are invited to take part in the study. The number of individuals that could potentially be identified may thus be substantial. To safeguard against misuse of the data and ensure that the privacy of the respondents is adequately protected we thus make the data de-facto anonymous and take procedural measures to regulate data access. De-facto anonymization is achieved by removing information from the data set that could potentially be used to identify individuals (such as, e.g., the name of the host institution). Detailed information on the steps taken to anonymize the data can be found in the data documentation that is provided with each scientific-use file (the documentation is made available on the project website at careertrackercohorts.ch and at FORSbase). The procedural measures controlling data access are as follows:

- 1) The scientific-use files will be archived on FORSbase. This ensures that only researchers and students affiliated with a university, a university of applied science, or a research institution for scientific research and/or education purposes can access the data.
- 2) Researchers requesting access to the scientific-use files have to submit a detailed description of the intended use of the data and the analyses and accept a user contract, before their request is approved. Use of the data will be limited to the purpose of the proposed project and the data will have to be destroyed after completing the project.
- 3) The SNSF controls access to the scientific-use files on a case-by-case basis. Data is only released upon approval by the SNSF after the contract has been signed (see point 4). A steering committee consisting of the SNSF internal Data Governance Group will evaluate the submitted requests (see also annex 1 for more details). The committee will base its decision on the quality and authenticity of the research proposal, on whether the CTC data is suitable to answer the research questions raised in the proposal, but not on the research question per se.
- 4) The applicant has to sign a physical contract with the SNSF before receiving access the scientific-use files (see annex 2 for the example contract).

4.2 Are there any necessary limitations to protect sensitive data?

The entire data sharing process (see section 4.1) complies with the Federal Act on Data Protection (FADP).

4.3 I will choose digital repositories that are conform to the FAIR Data Principles

yes

4.4 I will choose digital repositories maintained by a non-profit organization

yes

Annex 1: Access to scientific-use files of the SNSF-CTC Decision process

1 Steering committee

The steering committee that decides on whether a person gets access to the scientific-use files consists of the SNSF internal Data Governance Group. If the Data Governance Group is not up to take a decision, alternatively an ad-hoc group composed of one person from the legal services, one person from the data team and one person from strategy support will take the decision.

2 Evaluation criteria

Researchers requesting access to the data must submit an application containing a detailed description of the intended use of the data. Access to scientific-use files may be granted under the following conditions:

- The project pursues (a) legitimate research question(s) and the application demonstrates that the data will be used appropriately for scientific purposes or for teaching at an institution of higher education.
- The research question(s) described cannot be answered without the scientific-use files of the SNSF-CTC (justification of the data request).
- The application provides all the information necessary (goal of the research project, name(s) of the person(s) and affiliations of all the persons accessing the data, description of the measures taken to ensure data security at the organisational, technical, personnel and informational level).
- The project does not pursue any commercial purpose.

If the researchers intend to link the scientific-use files of the CTC with data from the Federal Statistical Office (FSO), this needs to be specified in the description of the intended use of the data. In addition, the following supplementary conditions apply:

- The purpose of the research needs to be clearly defined. Pursuing different research questions is not possible.
- The application demonstrates how the linkage with the FSO data is indispensable to the research project and that the data will be used appropriately for scientific purposes.
- The researchers have to be employed at a recognized research institution (e.g. university) or at an organization of the federal, cantonal or municipal administration.

If these additional conditions are met, the SNSF may confirm to the FSO that it approves of the data linkage intended by the researchers.

3 Process

All the members of the steering committee receive a copy of the access request. They decide together within a maximum of 4 weeks (20 working days) whether access can be granted or not. They send the person requesting access a justified decision. Afterwards and before the scientific-use files are made available the person requesting access has to sign a physical contract with the SNSF.

In case the researchers want to link the scientific-use files with data from the FSO and the SNSF approved this request, the SNSF sends the researchers a written confirmation that can be attached to the data linkage application to the FSO, as soon as the physical contract with the SNSF has been signed.

August 2024

Annex 2

www.snsf.ch
Wildhainweg 3, P.O. Box, CH-3001 Berne

careertracking@snf.ch

Contract

between

Swiss National Science Foundation (SNSF), Wildhainweg 3, 3001 Berne, represented by its legal department,

(hereinafter “the SNSF”)

and

the researchers of the research project “...”, represented by ...,

(hereinafter “the researchers”)

Article 1 Introduction, description and goal of the research project

¹ The persons involved, the research question and the goal of the research project are described in detail in the request for data access. The request for data access forms an integral part of the contract (Annex XX).

² The researchers request access to the scientific-use files of the SNSF Career Tracker Cohorts (SNSF-CTC) archived at FORSbase. The scientific-use files of the SNSF-CTC are indispensable for the conduct of the research described in their request for data access.

Article 2 Object

The researchers get access to the scientific-use files of the SNSF-CTC that are currently archived at FORSbase. New scientific-use files of the SNSF-CTC will be archived at FORSbase as they emerge. If the researchers wish to get access to additional/future scientific-use files they have to file a new request for data access.

Article 3 Right of use

The scientific-use files of the SNSF-CTC are to be used exclusively for the realisation of the research project described in the request for data access. Any other use or exploitation of the scientific-use files of the SNSF-CTC is prohibited. This also includes any other use or exploitation via third parties.

Article 4 Data protection, confidentiality and data destruction

¹ The researchers comply with the provision of the Federal Act on Data Protection (FADP, CC 235.1) and guarantee full protection and confidentiality of the data contained in the scientific-use files of the SNSF-CTC to which they have received access within the framework of the present contract, even after completion of the research project.

² The researchers guarantee that they will not initiate any research, contact-making or other measures aimed at identifying individuals concerned by the scientific-use files of the SNSF-CTC.

³ The researchers destroy the scientific-use files of the SNSF-CTC after the completion of the research project, without keeping any copy thereof. The researchers shall immediately confirm to the SNSF that they have destroyed the scientific-use files of the SNSF-CTC completely and irrevocably. If the research project has not been completed by the aforementioned date, the researchers must submit a written and justified request for extension to the SNSF at least one month before this date expires. The SNSF will decide on this request within four weeks.

Article 5 Prohibition of data linking

The researchers are prohibited from linking the scientific-use files of the SNSF-CTC with their own data, with further data from the SNSF or with data of third parties.

Article 6 Prohibition of data transfer to third parties

¹ Only the persons listed in the request for data access will process the scientific-use files of the SNSF-CTC. The undersigned ensures that these persons comply with all legal and contractual obligations and makes them aware of the confidentiality of the data. Persons having access to the scientific-use files have to be recorded in a separate list. The SNSF has to be informed about any changes to this list.

² The researchers are prohibited from passing on the scientific-use files of the SNSF-CTC to third parties or making them available to third parties in any other way.

Article 7 Right of publication

¹ The rights to the results of the research belong to the researchers.

² The results of the research are published in such a manner that particular individuals may not be identified. Prior to publication, the SNSF will receive a free copy of the publication for its information.

Article 8 Contract duration

¹ This contract comes into force upon signature by both parties and ends at the earliest at the end of the research. Article 6 paragraph 2 of this contract remains in force.

² The SNSF has the right to terminate the contract prematurely if it finds a breach of the other party's obligations.

Article 9 Amendments/additions to the contract

Amendments or additions to this contract must be made in writing and signed by both parties.

Article 10 Disputes

Both parties aim to settle disputes amicably. In the event that a dispute must be settled in the courts, the place of jurisdiction will be Berne, Switzerland. This agreement shall be governed by the substantial law of Switzerland.

This contract is drawn up and signed in two copies. The researchers and the SNSF each receive one copy.

Berne, dated _____

Location, dated _____

Swiss National Science Foundation:

The researchers:

For the Strategy Support division:
Name

For the researchers:
Name