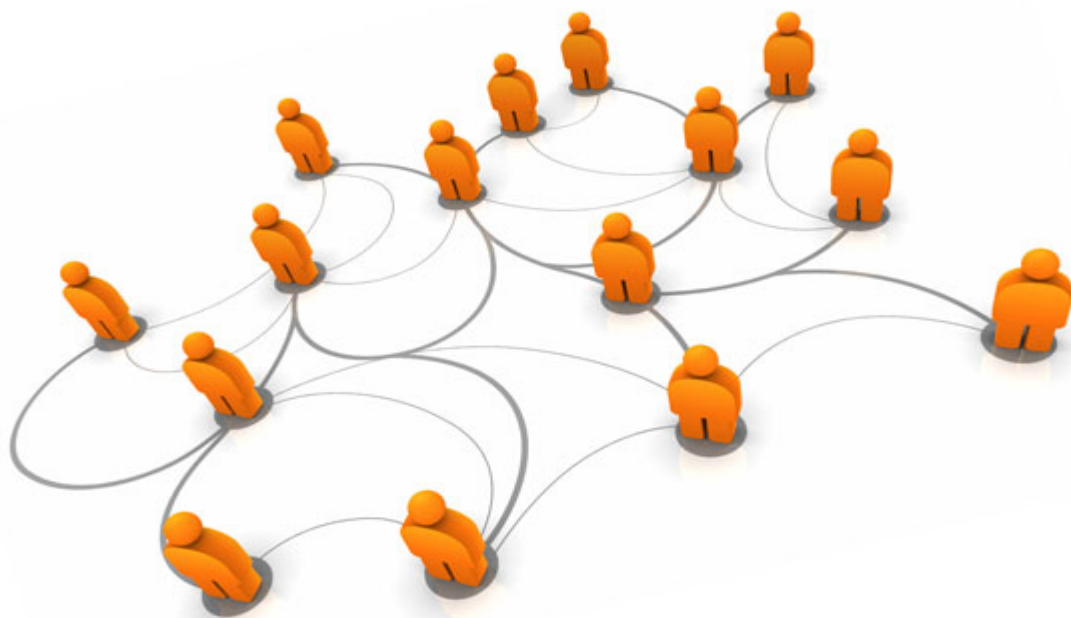




A network on
web-based data collection,
methodological challenges,
solutions and implementation



Supported by



Authors: Stephanie Steinmetz, Lars Kaczmarek, Pablo de Pedraza, Ulf-Dietrich Reips, Kea Tijdens, Katja Lozar Manfreda, Lilly Rowland, Francis Serrano, Marko Vidakovic, Carl Vogel, Ana Belchior, Jernej Berzelak, Silvia Biffignandi, Andreas Birgegard, Ernest Cachia, Mario Callegaro, Patrick J Camilleri, Gian Marco Campagnolo, Marta Cantijoch, Naoufel Cheikhrouhou, Daniela Constantin, Reuven Dar, Sophie David, Edith de Leeuw, Guy Doron, Enrique Fernandez-Macias, Niels Ole Finnemann, Muriel Foulonneau, Nicoletta Fornara, Marek Fuchs, Frederik Funke, Rachel Gibson, Solza Grceva, Gustav Haraldsen, Gudbjorg Jonsdottir, Martin Kahanec, Kathrin Kissau, Kirstine Kolsrud, Timo Lenzner, Laurent Lesnard, Helen Margetts, Yavor Markov, Goran Milas, Boris Mlacic, Liliana Mihaela Moga, Mihaela Neculita, Alina Irina Popescu, Suvi Ronkainen, Annette Scherpenzeel, Arto Selkala, Knut Kalgraff Skjak, Ana Slavec, Michele Ernst Staehli, Fanny Thorsdottir, Daniele Toninelli, Ravi Vatraru, Vasja Vehovar, Alberto Villacampa Gonzalez, Bernardo (Dov) Winer.

Acknowledgement: The authors would like to acknowledge networking support by the COST Action IS1004.



The word cloud is based on content from the network's website at <http://webdatanet.eu/>

What is in it for you?

Do you collect data via the Internet in your research? If you do, this European network is important for you. The network collects and combines experiences and research on the methodology of online data collection. It provides access to expertise that may be important in your research.

Come and join!

Contact details are available at the end of this document.



What is it?

WEBDATANET is a unique multidisciplinary European network bringing together more than 76 leading web-based data collection experts, (web) survey methodologists, psychologists, sociologists, linguists, economists, computer scientists, media researchers, internet scientists and public opinion researchers from 28 countries of the European Union and abroad.

By addressing methodological issues of web-based data collection (surveys, experiments, tests, non-reactive data collection, and mobile internet research) and fostering its scientific usage, WEBDATANET aims to contribute to the theoretical and empirical foundations of web-based data collection, stimulate its integration into the entire research process, and enhance the integrity and legitimacy of these new forms of data collection.

WEBDATANET is an Action of COST, which is an intergovernmental framework for European Cooperation in Science and Technology. COST is one of the longest-running European instruments supporting cooperation among scientists and researchers across Europe. It contributes to reducing the fragmentation in European research investments and opening the European Research Area to cooperation worldwide. It is a unique means for European researchers to jointly develop their ideas and new initiatives across all scientific disciplines through trans-European networking of nationally funded research activities. While COST does not fund research in and of itself, it does provide platforms for European scientists to collaborate on specific projects which are called *Actions*.

The need for collaboration

Throughout the last two decades, web-based data collection has increasingly become an important and indispensable instrument of current social research and the commercial survey industry. Results based on these data can influence political decision-making and public opin-

ion. Therefore, it is vital that no erroneous, distorted, vague or misleading conclusions are drawn from such data.

Since web-based data collection can offer unprecedented opportunities as to making data rapidly available with benefits in sample size, global country coverage, cost optimization, and access to rare groups, it is problematic that a coherent and systematic scientific methodology has not yet been developed. This is particularly surprising as Europe is home to a wealth of experts in a variety of fields related to web-based data collection.

In order to provide web-based data collection with the scientific validity that traditional research methods already enjoy and to tackle remaining methodological challenges WEBDATANET was established in June 2011 as a COST Action and will continue working on these issues through the COST Action until June 2015. It is anticipated that beyond 2015, the interactions of this scientific community will continue in the context of integrated projects and other strategic research initiatives.



WEBDATANET members who attended a meeting in Mannheim in 2012.

WEBDATANET defines itself as a unified research forum establishing a multidisciplinary network of web-based data collection experts, (web) survey methodologists, psychologists, sociologists, linguists, economists, computer scientists, media researchers, internet scientists and public opinion researchers who work together to accumulate and synthesize knowledge of methodological problems of web-based data collection. This includes surveying, experimenting, testing, non-reactive data collection, and mobile internet research.

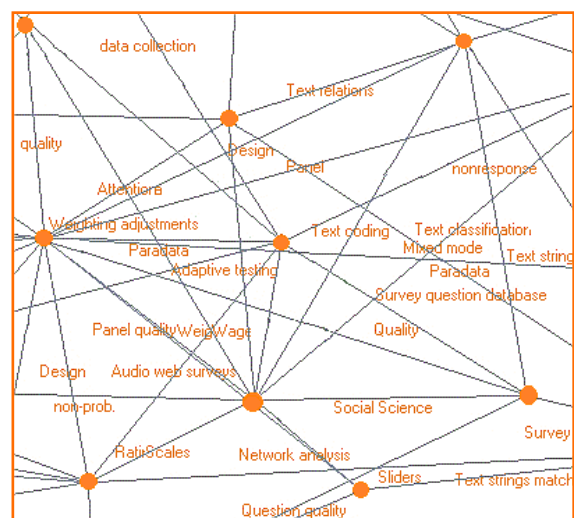
The aim of collaboration

The goals of WEBDATANET can be described as follows:

1. to provide the current loose network of researchers in web-based data collection with a platform to exchange experiences and discuss the impact of their research as well as new technological and methodological challenges and opportunities;
2. to develop a sound theoretical and empirical foundation for web survey methodology, testing, web-based experimenting, non-reactive data collection, and mobile internet research as well as for new emerging forms of web-based data collection;
3. to foster joint efforts to submit cross-national research proposals and to publish research findings in the form of articles in international scientific journals and books;
4. to stimulate early stage researchers to present and discuss their projects and establish contact with researchers abroad;
5. to provide a modern technological platform for continuous professional networking, where all researchers in this area will disseminate their materials, discuss issues, inform themselves, and communicate with users;
6. to disseminate the knowledge on web-based data collection to users, such as statistical offices, academics, policy-makers, the media, and to increase the awareness and usage of such data sources by enhancing its credibility.

An exercise organized and carried out by Working group 2 at a meeting of WEBDATANET demonstrates how various

research interests and possible collaborations were identified among new WEBDATANET members.



WEBDATANET members and part of a schematic illustration of the network showing members and their associations to others through different topics.

The scientific scope of collaboration

WEBDATANET provides researchers with a collaboration platform where current challenges and innovative ideas for web-based data collection can be exchanged, discussed and developed. Moreover, through its multidisciplinary framework it offers a unique availability of expertise.

In order to provide a framework for collaboration, the scientific scope of WEBDATANET is primarily divided into three broad research themes emerging as im-

portant areas in the field of web-based data collection. For each theme, one working group has been established. Each group can create task forces. A task force is a small group of researchers who work towards a specific goal, such as a special journal issue or a workshop. The output of each task force is closely linked to the goals as stated in the original WEB-DATANET proposal and the Memorandum of Understanding.

The three research themes are (1) Quality issues of web surveys, (2) Innovations in web-based data collection, and (3) Guidelines, codes, standards, implementation and dissemination.

Working group 1: Quality issues of web surveys

The purpose of working group 1 is to systematically investigate the principal weaknesses and methodological challenges of web-based data collection, in particular those of web surveys as a tool for gathering high-quality data. It analyses and compares errors in today's prevailing non-probability based web surveys from a cross-national perspective. In this context, existing post-survey adjustments, such as post-stratification, propensity score adjustment, calibration and imputation techniques are evaluated particularly according to the central question of improving the validity and representativeness of survey results. Special attention is paid to efforts to use small probability samples (e.g. telephone or face-to-face surveys) to calibrate a large (but less expensive) non-probability web sample. Commercial survey companies use this approach, but further foundational research on its scientific value needs to be done.

Furthermore, the working group explores the possibilities of conducting probability-based web surveys. It investigates factors and develops strategies to preserve low non-coverage and non-response as well as low measurement errors related to probability-based web and mixed mode surveys. The cost-error optimization of probability-based web surveys is also a high research priority. In this context, the working group also addresses the extent to which different web marketing strategies, such as attractive website content and the provision of incentives, might blur the distinction between probability and non-probability sampling by helping to attract large and heterogeneous groups of visitors.

Working group 2: Innovations in web-based data collection

Working group 2 concentrates on the opportunities and scientific challenges of recent and possible future innovations in web-based data collection. An example is paradata analysis (observation and use of participant behaviours during data collection – behaviours such as changes of answers, scrolling, and response times) which can inform the researcher of measurement problems. In addition, it explores how such techniques can mutually improve each other. For example, no report has yet been published on paradata from large-scale cross-cultural web surveys to inform the scientific community of the relative relevance and similarity of paradata interpretations (e.g. the difficulty or sensitivity of items). The group discusses and disseminates innovations and improvements in web-based data collection types that are already widely used in social sciences, such as web-based tests and experiments, non-reactive data collection (e.g. from so-

cial media and blogs), and mobile internet research.

Working group members have expertise and experience in a variety of web technology innovations that can contribute to methodological knowledge on web-based data collection, for example: text analyses from 'spiders', server log file analyses, visual web design (e.g. dynamic paging and multimedia) and its influences on measurement (survey question design, optimal routing scripts, search tree testing), software developments, innovative ways of recruiting participants, and other exciting innovations.

Working group 3: Guidelines, codes, standards, implementation and dissemination

This working group focuses on the practical implementation of web-based data collection methods within a broader framework of users. The aim of this working group is to develop guidelines, tools and sites for practical implementation as well as to adapt the developed innovations and findings of WEBDATANET to the substantial needs of different scientific fields.

Working group 3 evaluates existing guidelines of professional associations and researchers with regard to conducting web surveys and collecting web data. For the development of new guidelines and standards, specific case studies/scenarios are being developed and discussed, taking into account not only methodological aspects but also cost effectiveness and feasibility. It also explores the possibilities of outreach activities within Europe and beyond, and the group considers whether the experiences of the institutions that already imple-

ment web-based data collection can be applied in the institutions where such methodology has not yet been introduced. Furthermore, this group evaluates professional and ethics codes and standards in the light of opportunities created by new web-based data collection techniques.

Finally, this working group seeks to ensure the practical implementation of results, along with openness, flexibility, liaison and interaction with other research programmes as well as communication between web survey experts, other social scientists, official bodies, and media, communication, marketing and opinion researchers.

Examples of current task forces within the three working groups

Task Force Example 1: Challenges and pitfalls of measuring wages via web surveys – explorations on weighting

Online panels could offer a good source of data for integration with official statistics on wages. However, as the sub-population with internet access is quite specific, the driving question for the task force is whether online panels related to wages and working conditions are representative, and if not, *can* representativeness be achieved? The task force will analyse different existing probability and non-probability based web surveys focusing on labour markets (e.g. WageIndicator, LISS panel, and SSRI panel). Their characteristics are compared with reference data from official statistics to estimate the bias. For a selection of core variables adjustment models are applied, for example simple weighting and propensity score adjustment. Properties and theoretical advantages of the applied methods are discussed and evaluated particularly

according to the central question of improving the validity and representativeness of survey results.

Task Force Example 2: Criteria for the evaluation of questionnaire quality

The aim of this task force is to prepare a list of criteria that can be used to evaluate the quality of a web survey questionnaire. This list will, for example, be applied within the freely available online tool for implementing web surveys that is developed at the Centre for Social Informatics at the Faculty of Social Sciences, University of Ljubljana, Slovenia. A similar tool, the *questionnaire checker*, is available via the iScience Server currently hosted at Universidad de Deusto in Spain – thus, WEBDATANET brings together expertise from two groups. Both tools already include a set of criteria which gives an evaluation of the quality of a questionnaire and detected problems to alert to the researcher who is setting up a survey. The WEBDATANET experts in web survey methodology in this task force will comment and improve the set of existing criteria.

Task Force Example 3: Implementing web surveys to political opinion polls in Spain

The task force aims to increase the implementation and use of web surveys in political opinion polls conducted by Centro de Investigaciones Sociológicas in Spain (CIS). This is done by participating in the annual open competitive calls providing interested researchers the opportunity to present research proposals. So far, this task force has participated in one call. The proposal focused on studying the coverage error using CIS data sets. It was positively evaluated, and the resulting research project is now being conducted at the University of Salamanca. The idea is to partici-

pate in the following calls to propose new projects aiming, in the medium/long term, to pave the way for future implementation of web surveys at CIS.

Task Force Example 4: How are web surveys changing empirical (social science) research?

This task force deals with the changes in empirical social science research and research in other disciplines (e.g. health studies, engineering) due to the extensive use of new forms of web-based data collection. For different fields (e.g. sociology, political science, sociology of technology etc.) the prevalence and importance, particularly, of web surveys as a data collection mode (in comparison to other survey modes or other data collection methods) will be evaluated, and the question will be addressed whether these new forms of data collection methods have changed how social scientists and other disciplines are doing research.



WEBDATANET members who attended a meeting in Amsterdam in 2011.

The options for cooperation

Within the framework of WEBDATANET different forms of collaboration are supported. Among those are networking activities such as meetings, conferences, short term scientific exchanges (short term sci-

entific missions) and other forms of outreach activities.

MC and working group meetings

The Management Committee (MC) supervises and coordinates the implementation of the COST Action. It draws up detailed plans, distributes tasks and defines methods for the execution of the Action, and it makes executive decisions within the framework of the COST guidelines. The MC meets twice a year.

Working group meetings are intended to increase the collaboration among Action members with respect to specific goals and objectives of the Action. The meetings are normally scheduled before or after the MC meeting but can also be organized on demand by the specific working group.

Short term scientific missions

A short term scientific mission consists of individual research stays of WEBDATANET members. It aims at ensuring mobility and close collaboration among researchers.

In the first year of WEBDATANET six short term scientific missions have been carried out, mainly among WEBDATANET members who are early stage researchers. The topics of these short term scientific missions covered a broad area:

- WageIndicator, Labour Markets and WEBDATANET;
- Fostering interactions between working group 1 and 3: Online Master and Web survey tools;
- Measuring wages via online panels: review of existing panels, estimation problems and future challenges;

- Deepening interactions and collaborations on ‘Web Survey Methodology’;
- Large scale web-based data collection systems;
- Web-based registry for eating disorders.

Short term scientific missions are successful, particularly in support of early stage researchers in their scientific development, and also senior researchers profit from short visit stays in order to develop proposal(s) etc.

Interview with WEB-DATANET chair Dr. Pablo de Pedraza about his experience with the short term scientific mission (STSM) on ‘WageIndicator, WEBDATANET and Labour Markets’ at the IZA in Bonn, Germany.



Pablo, what was your project about?

I conducted research activities and dissemination activities. Regarding the former, I was working with Dr. Martin Guzi using WageIndicator data for life satisfaction research. Regarding the latter I gave a seminar explaining to IZA researchers the WageIndicator and WEB-DATANET.

Additionally, I had a meeting with Dr. Nikos Askitas focusing on how applied economics may benefit from web-based data and, therefore, their involvement in the network. We also talked about organizing a WageIndicator workshop in May 2013 at the IZA.

How did you benefit from the support of the WEBDATANET network?

I could use the WEBDATANET resources to pay part of the expenditures, the trip and accommodation. It was also useful for other network members to know my research interest in more detail.

What did you accomplish during your STSM?

I was able to make some good progress in the research activities with Dr. Martin Guzi, and established some new contacts with colleagues from IZA. I also benefitted from working in one of the most prestigious institutes for Labour research and participated in its activities.

Why would you recommend doing an STSM?

I find STSMs a very useful tool for the career development of early stage researchers, like myself. STSMs are also, in general, a very efficient networking tool. Everybody in the network is able to read STSMs proposals and reports which help to monitor the interactions that are functioning within WEBDATANET and, therefore fulfil one of our goals: develop a stable research network on the topic of web-based data collection for scientific research.

Do you have any tips or suggestions for other network members who are considering planning an STSM?

Some: define in advance the work to be done, in detail, by collaborating online with the host researcher; think on research and scientific goals but also in long run collaborations, networking and projects; think about other researchers in the network who may be working on similar topics.

Thank you for your time!

Training schools

The first WEBDATANET training school will focus on the practical implementation of web surveys. It will take place at the

University of Ljubljana (Slovenia) in Spring 2013.

The school is intended for (but not limited to) early stage researchers who would like to learn how to implement a web survey for their own research purposes in different fields. Trainees get additional specific theoretical and practical knowledge on how to conduct web surveys as a special type of survey research.

The course will be a mixture of lectures and lab sessions. Participants will be provided with the basic principles of implementing web surveys in order to collect survey data of adequate quality.

Workshops / Conferences

The next Internet Survey Methodology Workshop will be jointly organized by WEBDATANET/COST and the University of Ljubljana. The workshop will take place from November 14–16, 2012 in Ljubljana (Slovenia).

The aim of the workshop is to bring together survey methodologists to present and discuss their latest research on the methodology of web surveys. The topics will range from data quality in web surveys to innovations in this field. We expect up to 30 participants (world leading scientists in the field) to participate.

Being innovative through cooperation

In the context of increased competition among scientific disciplines and research areas regarding innovative contributions, WEBDATANET is innovative as it contributes towards the further development and improvement of a sound theoretical

underpinning of web-based data collection. It improves the efficiency of web-based data collection processes and their use in social, behavioural and media sciences. It provides the scientific community with a meeting and discussion forum of leading specialists who will consider how to maximize the potential benefits and minimize the possible drawbacks of these new approaches. In addition, it will strengthen the importance and visibility of web-based data collection methodology in the international arena by creating a sufficiently large and competitive scientific community of researchers. In addition, with the development of new cross-national joint research activities, the dissemination of results in the form of articles, books and reports as well as with the integration of the results into the educational programmes of the participating researchers and groups, the network will foster the further establishment of a new research stream within social science methodology.

But also from a societal perspective WEBDATANET is innovative as it aims to improve the integrity and legitimacy of web-based data collection. With such new forms of data collection, it is important that the results which contribute towards public opinion formation and executive decision-making are based on sound scientific underpinning. Thus, not only collectors of web-based data but also society in general will benefit from research conducted as part of this Action. At the same time, WEBDATANET activities which promote increased use of web surveys within Europe will benefit society by reducing the amount of public money spent on data collection.

Being productive through cooperation

The overview of recent publications of WEBDATANET members may serve as an illustration on how the network benefits from each individual and how the members benefit from participating and collaborating in the network.

Interested in joining?

WEBDATANET encourages people who are involved and interested in web-based data collection to join the Action. Participants of WEBDATANET are:

- senior/junior researchers and university students in the field of web-based data collection as well as academic research institutes involved in the coordination of European and worldwide surveys;
- commercial market research and public opinion agencies;
- governmental agencies and institutions carrying out web surveys;
- international and national public policy-makers;

Although there are specific rules and requirements on how to become a member of the Action, participation is rather easy to achieve. For more information about the network, please visit our homepage at

<http://webdatanet.eu/>

You can also contact our office directly:

Scientific and financial administrator
Mr. Francis Serrano

University of Salamanca, Spain
Email: francissh77@gmail.com

Action chair

Dr. Pablo de Pedraza
pablodepedraza@usal.es

Action vice chair

Dr. Stephanie Steinmetz
s.m.steinmetz@uva.nl

Working group 1 chair and co-chair

Dr. Katja Lozar Manfreda
Katja.Lozar@fdv.uni-lj.si

Dr. Gudbjörg Andrea Jonsdottir
gudbjorg@hi.is

Working group 2 chair

Prof. Dr. Ulf-Dietrich Reips
u.reips@ikerbasque.org

Working group 3 chair and co-chair

Prof. Dr. Silvia Biffignandi
silvia.biffignandi@unibg.it

Prof. Dr. Fanney Thorsdottir
fanneyt@hi.is

Short term scientific missions (STSM)

Prof. Dr. Ravi Vatrapu
vatrapu@cbs.dk

Financial rapporteurs

Dr. Lars Kaczmarek
lars.kaczmarek@gesis.org

Prof. Dr. Kea Tijdens
k.g.tijdens@uva.nl

**DC Rapporteur
(COST office)**

Ms. Ljiljana Kaliterna (HR)

**Science officer of the Action
(COST office)**

Dr. Giuseppe Lugano

**Administrative officer of the Action
(COST office)**

Mr. Leo Guilfoyle

Suggested citation: Steinmetz, S., Kaczmarek, L., Pedraza, P., Reips, U.-D., Tijdens, K., Lozar Manfreda, K., ... Winer, B. (2012). WEBDATANET: A network on web-based data collection, methodological challenges, solutions and implementation [Supplement]. *International Journal of Internet Science*, 7(1), 78–89.

