EQUIPEX
CALL FOR PROPOSALS

2010

Acronym

DIME-SHS

SCIENTIFIC SUBMISSION FORM B

Acronym of the project	**DIME-SHS**	
Titre du projet en français	Données, Infrastructure, Méthodes d'Enquêtes en Sciences humaines et sociales	
Project title in English	Data, Infrastructure, Methods of investigation in the social sciences and humanities	
Coordinator of the project	Name: **Laurent Lesnard** Institution: **Sciences Po** Laboratory: CDSP – Centre de données socio-politiques / Centre for socio-political data Unit number: UMS 828	
Requested funding	Tranche 1/Phase 1 **16 811 197 €**	Tranche 2/Phase 2 **552 407 €**
Disciplinary field	☐ health, well-being, nutrition and biotechnologies ☐ environmental urgency, eco-technologies ☐ information, communication and nanotechnologies ■ social sciences and humanities ☐ other disciplinary area	
Scientific areas	Methodology, survey, statistics, demography, economics, history, political science, sociology, information and communication	

Organization of the coordinating partner

Laboratory/Institution(s)	Unit number	Research organization
CDSP / Sciences Po	UMS 828	Sciences Po / CNRS

Affiliations des partenaires au projet/Organization of the partner(s)

Laboratory/Institution(s)	Unit number	Research organization
GENES		
SES / Ined		
CERLIS / Université Paris Descartes	UMR 8070	Université Paris Descartes/CNRS/Université Paris 3
Telecom ParisTech	UMR 5141	Telecom ParisTech/CNRS
GIS Réseau Quetelet		CNRS/EHESS/INED/ Université de Caen
Company	**Economic sector**	**Staff size**
EDF R&D	Energy	2000

1. SCIENTIFIC ENVIRONMENT AND POSITIONING OF THE EQUIPMENT PROJECT 5
2. TECHNICAL AND SCIENTIFIC DESCRIPTION OF THE ACTIVITIES 9
2.1. Originality and innovative features of the equipment project 9
2.2. Description of the project 11
2.2.1 Scientific programme 11
2.2.2 Structure and building of the equipment 16
2.2.3 Technical environment 18
3. DISSEMINATION AND EXPLOITATION OF RESULTS 20
4. PROJECT MANAGEMENT 23
4.1. Management 23
4.1.1 Relevant experience of the project coordinator 23
4.1.2 COORDINATION MODALITIES 24
4.2. Collaboration organization 25
4.2.1 Partners description, relevance and complementarity 25
4.2.2 Qualification, role and involvement of individual partners 30
5. FINANCIAL ASSESSMENT 40
6. APPENDICES 44
6.1. Literature review references 44
6.2. Partners' references 46
6.3. Estimate for the equipement 50
6.4. Commitment Letters 54

SUMMARY

The *Données, Infrastructures, Méthodes d'Enquêtes en Sciences humaines et sociales* (DIME-SHS) project aims to make up for the underdevelopment of French universities in data collection and in survey methodology, in order to position itself as an actor meeting the highest standards of international competition in the field. This infrastructure aims to provide France with an innovative intermediate structure for the collection, enrichment, and dissemination of data for qualitative and quantitative research in SHS. In this respect, DIME-SHS will complement the "Very Large Research Infrastructure" (Très Grande Infrastructure de Recherche) GIS PROGEDO – Réseau Quetelet and the CNRS's "Very Large Facility" (Très Grand Équipement) for the social and human sciences, Adonis. The originality of DIME-SHS, compared to existing structures in France, Europe, or the United States, lies in the fact that it considers these three stages of research protocols as a whole, and takes full advantage of the possibilities that internet and new technologies provide for the exploration and distribution of classic data and the creation of new types of data, as well as for the reduction of running costs. Eventually, this infrastructure aims to offer methodological resources with high scientific standards to the French research community, enabling researchers to carry out their own surveys and to reuse pre-existing data. It will also encourage the development of experimental research designs and methodological innovations and will improve the methodological training of French researchers.

DIME-SHS will be organised into three integrated components which are based on existing but separate projects and research teams. Those components will be developed to innovate in the gathering of quantitative data, the archiving and distribution of qualitative surveys, and the study of the web as a corpus.

The first component is inspired by American and European Survey Research Centres (SRC), which are in charge of collecting data using questionnaires bound for research. It will reinforce the partnerships between Universities and National Statistics Institutes (INSEE and INED). It will reduce the dependence of public research on private survey companies. On one hand, it will respect international methodological norms based on probability sampling and not on quotas (currently the only affordable method for French university researchers) and it will provide better control of the survey procedures. On the other hand, avoiding private intermediaries reduces the cost of collecting data. The SRC component will include a telephone call centre and a web panel. Each person, randomly selected, will receive a smartphone and a corresponding plan in exchange for participating in the web panel. This will resolve the issue of coverage for Internet-based surveys. This new and personal way of responding is expected to reduce the growing difficulty of non-response in surveys. These two complementary modes of data collection, by phone and by web, will allow experimental research designs like mixed modes studies. Such an innovative structure would place France at the forefront of international quantitative research in SHS.

The second module is meant to develop a qualitative data bank (BEQuali) and will raise France to the American and European level in the field. The BEQuali component will ensure the preservation of qualitative research history but will also establish it as a scientific field and allow it to progress on methodological aspects. To overcome the qual-quant divide, it will be integrated into international (CESSDA) and national (Réseau Quetelet) networks of resource centres to benefit from the previously accumulated experience in the area of quantitative methodologies". It will make a more complete use of the potentials of the web in order to promote the enrichment and the diffusion of qualitative data by implementing innovative tools for navigation and exploration. One of the expected benefits of this project is the improvement of the protocols and methods of qualitative researches, and their cumulative knowledge, notably through the promotion of secondary analysis, which remains scarce without the required infrastructure.

The third and last component of DIME-SHS aims to develop tools and methods related to the collection of data from the web, especially unstructured data from spontaneous online expression. The quantity, and above all the diversity, of the traces left by these social interactions represent new opportunities for the social sciences. Yet these abundant data might remain underexploited or misused without appropriate tools to investigate this field. Based on the experience of the Sciences Po médialab, the web data module will be the first infrastructure in

France to offer researchers a public service which collects spontaneous expression data from the web, thus reducing the dependence on the expensive and proprietary tools so far only developed by marketing companies.

The three components of the infrastructure are complementary and their activities and resources will be mutualised. Numerous research projects from various academic disciplines will certainly take advantage of the different and complementary methods available through DIME-SHS. The Social Sciences and Humanities are also in partnership with a host of other scientific disciplines and as a result, DIME-SHS will be able to contribute to the development of research in other areas (health, risk analysis, or environment for example) and work with other non-academic bodies: state services, local governments, NGOs and businesses.

Data collected and acquired by DIME-SHS will be disseminated for scientific purposes using international standards and innovative tools and interfaces. In pursuance of the open science research movement, tools, methods and results progressively accumulated are made available to researchers as well as policy makers and the whole civil society. This availability will benefit the public and private sectors by stimulating innovations. Other consequences expected are a better structuring of academic partnerships, and a greater collaboration amongst specialised researchers to overcome the quali-quanti divide. DIM-SHS will also contribute to better methodological teaching and as such will improve students' training and employability. Eventually, it will enable an improvement in scientific work along with a reduction of the costs of public funding to research through important economies of scale (numerous synergies, limitation of redundant tools and more extensive use of existing data).

The success of this infrastructure relies on the complementarity of the expertise brought by the partners of this project. Sciences Po has a long history of experience in political science surveys, either quantitative or qualitative. More recently, Sciences Po has started to expand this expertise to data documentation and dissemination with the CDSP, and to web data collection and analysis with the médialab. GENES will bring its expertise in conducting telephone surveys and designing surveys, especially in order to achieve comparability with national or international surveys. INED has a long tradition of excellence in conducting qualitative and quantitative investigations on sensitive subjects (sexuality, violence) or on difficult to reach populations. They have also participated in various comparative surveys. The University Paris Descartes has extensive experience in conducting empirical investigations in the social sciences, which recently took the shape of a Master specialised in this field. Telecom Paris Tech will contribute its unique dual knowledge in telecommunication engineering and the sociology of communication. EDF R&D will share its expertise in archiving, documenting, and disseminating qualitative surveys but also its competence in the quantitative analysis of qualitative materials. The project relies on the expertise of the GIS Réseau Quetelet to meet the highest standard in data dissemination and confidentiality.

Together, these partners have the expertise to build the innovative and cost-effective infrastructure for collecting, documenting, and disseminating data for the social sciences. On the scientific level, they have the necessary knowledge to invent and standardise new methods (extraction and cleaning of web data technologies, new ways to contact respondents and to explore qualitative data…). They have also the required level of experience and reputation to promote the new services offered by the infrastructure and to overcome the possible reluctance of adoption(epistemological, ethical or academic). The consortium will consult with the best legal experts to ensure the highest level of data confidentiality and to protect the intellectual property of collected data. The various resources they contribute to this project, combined with the required financial contributions, will make it possible to build a lasting infrastructure within 10 years. There is a high demand for the services that would be provided by DIME-SHS, not only in France, but also internationally where similar infrastructures (Knowledge Networks in the US; CentERdata in the Netherlands) exist, paving the way to cost effective and high-quality international comparisons.

1. SCIENTIFIC ENVIRONMENT AND POSITIONING OF THE EQUIPMENT PROJECT

As highlighted in the Social Sciences and Humanities alliance - SNRI report (SNRI , ten years after the Silberman report (Silberman, 1999) on the social sciences and their data, French universities still lag significantly behind in terms of collecting, distributing and capitalising on data and research, as well as teaching empirical research methods, either qualitative or quantitative. Methodological excellence, essential for international competition, is now more than ever acquired through the ability to develop innovative and experimental research designs. This means being able to conduct primary research and secondary analysis with full control of all the different stages involved. New areas of investigation that have emerged from the internet (notably the social web) must be added to studies using questionnaires and/or qualitative methods.

Collecting data using questionnaires

Although it was discussed in the Silberman report ten years ago, France is still behind in the way it deals with university studies using questionnaires. The reasons for this backwardness are well known, as are the consequences and solutions. In the absence of a French academic survey platform, data collection must be outsourced to private survey companies which present a number of drawbacks linked to their economic structure (Pages et al. 2006, Caveng 2009). The first of these is the high cost, resulting from the profit margins that are applied. The second is the fact that the quality standards of these companies are often too low for the most rigorous scientific research. This is partly due to the working conditions of the interviewers (short term contracts, no incentives to professionalisation, high turnover rates, see Duclos et al. 1996 p. 82-85 et Caveng 2007). The third is the lack of experience these companies have in the treatment of random samples or very large samples, and lack of a knowledge of the specific problems and issues in Socio-economic Sciences and Humanities (SHS) (most of the work in this sector is in market research and, for the companies that conduct public opinion surveys, this generally makes up 10 to 20% of their revenue). As a result, monitoring interviews is time or cost consuming.

Consequently, such studies are very expensive in France (for the RAPFI study with French people of North African, African and Turkish origin, close to 90% of the budget was spent on this). As a consequence, it is difficult to find funding for them. Moreover, the few studies conducted in spite of these elevated costs often do not meet international methodological norms which are based first and foremost on probability sampling and not on quotas, which is currently the only affordable method for French university researchers. Panel studies, which involve following a population over a period of time, are also rarely used in France. Lacking in training and generally paid per questionnaire, the survey interviewers working for the private survey companies frequently ignore the instructions and rules concerning the questionnaire procedure (especially probing and follow-up questions), in order to limit the amount of time spent on each interview (Caveng 2009). This weakens the methodological work of the survey designers. If the French context is not conducive to survey-based studies obeying international norms, it is even less conducive to research methods that by definition require conducting experiments according to strict protocols. Unable to develop its own surveys, the French academic community is struggling to compete on an international level in terms of empirical research and methodology.

The solutions to these problems are well known, and have been long practiced in North America and in Europe in the *Survey Research Centre* (SRC) framework. Present in almost every university that ranks highly in social sciences (Princeton and Berkeley for example), these university structures integrate the whole processing chain of a questionnaire based study. Moreover, they give researchers access to the means of conducting studies at a lower cost, and at the advanced methodological level which results from the expertise accumulated by the methodological research conducted by the SRC. For example, it is the SRC of the University of Michigan that conducts the *National Election Studies*, an academic study conducted every two years since 1948, which is considered by the international scientific community to be the benchmark both in terms of research questions and quality of methodology (with both pre- and post-election waves as well as panels that sometimes stretch over four years). In France, only the IEP in Grenoble has its own telephone survey centre, associated with a Masters in Opinion and Market Surveys (PROGIS). This structure is very small however (10 positions), and is not open to the national scientific community. INED is the only French institution, excluding INSEE, that has been able to conduct surveys at a higher methodological level and develop research into methodology. In the absence of its own means of data collection however, INED is dependent on INSEE and private survey companies.

Although France's backwardness on this issue is the same ten years after the Silberman report, the international context has significantly changed. Empirical quantitative research in SHS is in the process of being profoundly changed by the development of new information and communication technologies: online diffusion, new forms of communication and measurement such as mobile telephones, smartphones, tactile tablets (Ipad), GPS, or biomarker measurement tools (Raento 2009). These new tools offer the means of collecting data that is higher quality, and collecting it faster and for a lower cost than traditional survey techniques. The possibility of using visual aids (Toepel et al. 2009), videos, interactive applications, or geo-locators for respondents (with their consent) opens new horizons for research in SHS and for the whole scientific community (epidemiology, ergonomics, information technology etc). Unlike other survey methods (face-to-face, telephone, self-administered), there is no limit to the number of persons that can be surveyed simultaneously, given the size of the sample, which allows surveys to be conducted and data to be accessed very quickly. The lack of intermediaries also means reducing the cost of collecting the data and better controlling the rules and conditions of questionnaire procedure (follow-ups, aids, interruptions, retrieval of questionnaires).

These new tools establish a whole new playing field for research in SHS. Because they seized this opportunity, the Netherlands (formerly lacking the old-generation SRC, like France) is now at the forefront of quantitative research in SHS, having set up the LISS panel at the University of Tilburg, an internet panel that is representative of the Dutch population (inspired by Knowledge Networks, US). A groundbreaking project, the LISS panel is also at the most advanced level methodologically, particularly in the composition of the sample which is constructed to represent the Dutch population and based on providing internet connections to respondents who otherwise wouldn't have access. This procedure avoids the significant number of biases in traditional internet surveys, which are conducted solely with people who have an internet connection (Couper 2000).

Archival and diffusion – for reuse – of qualitative studies: an imperative for development and innovation in SHS

Although the French SHS field is characterised by its overwhelming use of qualitative methods (in political science for example see Billordo 2005), compared to what happens overseas (particularly in the United-States), the French contribution in terms of methodological innovation and expertise remains limited. This is notably explained by the complete lack of archiving for qualitative studies. Yet the drawbacks associated with not capitalising on these studies, already stressed in the Cribier report some years ago (Cribier 2003), remain relevant today. Worse, most of the qualitative data (interviews, observation notes etc.) collected since the 1960s have disappeared. What remains has been conserved in poor conditions and is at high risk of disappearing soon too, if nothing is done quickly. In addition to the fact that the research community shows a lack of understanding and interest in these questions, this situation can be further explained by the fact that public archives have little interest in this kind of material (Braibant, 1996; Duclert, 2002), and higher education and research institutions have little means to archive data themselves. Apart from a few isolated initiatives, notably in oral history (Duclert, 2002), there is at the current time no real policy on the scientific archiving of data from qualitative studies. Moreover, in sociology or political science the rare pioneering experiments in terms of archiving and sharing qualitative data in fact come from applied research[1]. France's tardiness on this issue is even more paradoxical given that the creation of a databank centralising all qualitative studies in SHS at the national level is a goal frequently expressed by public institutions that have financed several reports to this effect (Braibant, 1996; Silberman, 1999; Demazière *et al.*, 2001; Cribier, 2003).

We can see however, all that the archiving and availability of this data could bring to the field of SHS: preserving its history but also establishing it as a scientific field and allowing it to progress. Publications based on qualitative studies generally provide no access to the data upon which the analysis is based, and they only describe the research protocols in a limited fashion (Moravcsik 2010). This does not allow readers to directly compare the methodological strategies implemented and their impact on the results that are obtained. More generally, "making do" seems to be the rule, as is conformity to established principles; methodological innovation is notably curbed because it is impossible to make use of a detailed knowledge of previous experiences. The

[1] This is the case with EDF R& D, with the database "Verbatim", which was developed in 1998 on the model of Qualidata, and which contained 72 studies and 778 interviews from 1994 to 2007. It is only accessible to researchers of EDF R&D. For a description of this initiative see Le Roux, Vidal, 2000a, Le Roux, Vidal, 2000b Dargentas, Le Roux, 2005.

amelioration of research protocols is therefore hindered and the accumulation of research, in methodological terms remains limited.

Finally, beyond questions of methodology, the absence of archiving makes it impossible to reanalyze studies. Yet publishing, in qualitative research as elsewhere, supposes first and foremost making choices, retaining for example only those aspects of observations that are meaningful in the theoretical framework that has been elaborated. A study is always susceptible to multiple analyses however, above all if it is put into perspective or even directly compared to other studies and other data. The absence of data archiving deprives the SHS field of any possibility of qualitative comparison.

These different arguments[2] have been presented by the creators of the pioneering qualitative data bank ESDS Qualidata, a service of the UK data archive[3]. The initiators and managers of Qualidata launched a major debate[4] on the justification of archiving and the methodological conditions of secondary analysis of qualitative data (see the list of bibliographical references complied by Qualidata which includes more than fifty texts: http://www.esds.ac.uk/qualidata/support/reusearticles.asp), in which French researchers have – unfortunately – not taken part until now[5]. Their arguments, which are positivist in perspective, are treated sceptically by those working from a critical constructivist epistemological perspective. For these researchers, any information gathered or data constructed are interpretable and analyzable only by the person who directly participated in the study, any other usage would be considered fallacious (Duchesne, forthcoming). We position ourselves between these two poles, in the sense that we believe that it is possible to archive qualitative data so that it may be used by others, on the condition that the study be considered as a whole and that the context and the metadata be as fully informed as possible. In spite of resistance from a non negligible portion of the scientific community, the example of Qualidata has ended up establishing itself – most European countries are putting in place databanks of this type, combining qualitative and quantitative studies in the same structures, and trying to centralise archiving for reasons of both accessibility, and the quality of the staff responsible for this work[6].

In response to this, scientific events dealing with questions of qualitative data analysis have abounded in France in recent years, demonstrating a particular sensitivity to these issues on the part of the French research community[7]. Above all, a new initiative was launched in 2009 as a result of a partnership between the CNRS and Sciences Po (via the TGE-Adonis), who financed an exploratory project to create qualitative data bank in SHS, based on the development of the internet (with the expertise of Medialab at Sciences Po). The aim of this project is to be able to concretely address the arguments of researchers who remain sceptical as to the benefits of archiving and secondary analysis[8].

The web as a corpus

The uses of the web have become increasingly varied since the 1990s when the dominant form was the website. With the spread of high-speed connections, mobile internet access and rich interaction technology (which come together in the term Web 2.0), the web has become an essential form of media in which social interaction is woven (Presner & Schnapp, 2009). However, it is also a digital medium in which every element of data is easily traceable. An increasing number of new data sources from the web are now available, but their reputation is not one of quality. The

[2] Data archiving is also a powerful way of encouraging collaboration between researchers (McDermott 2010).

[3] Located at the University of Essex, the UK data archive is a public institution established in 1967, combining qualitative and quantitative data in social sciences. The two archives were integrated into the Economic and Social Data Service (ESDS) in 2001. http://www.esds.ac.uk/qualidata/about/introduction.asp

[4] Cf. for example the Qualitative Social Research forum (http:/ / www.qualitative-research.net)

[5] With the exception of contributions from the Verbatim team and notably: Dargentas M. et al.

[6] For example in Finland, Ireland, and Switzerland.

[7] For example: study days organised by the CAPAS group in Grenoble on research ethics; a study day organised by the DIM (data, infrastructure, methods) research priority of Sciences Po on the formatting and overview of interviews in social sciences (December 2007) and computer assisted textual analysis (June 2009); the study day organised by the GRHISPO in Amiens on the usage of interviews in social and historical sciences (June 2009); or the recent conference organised at Limoges University on the theme of interviewers' and interviewees' rights ("Droit d'enquêter / Droits des enquêtés" September 2009), etc. See also the issue n°63 of the journal *Genèses* dedicated to research archives (*Genèses* 2006)

[8] Information can be found on: **http://blogs.sciences-po.fr/recherche-donnees-qualitatives/**

digital environment introduces a new paradigm in which the versatility and the ease of information production imply a loss of control over the authenticity and the quality of the available data. Yet the quantity, and above all the diversity, of the traces left by these social interactions represent new opportunities for social sciences (Lazer et al., 2009).

The acknowledgment of the fact that digital media are playing a stronger and stronger role in collective debates is radically transforming the way social science look at them. Until few years ago, social scientists conceived electronic media as new terrains for old methods. Notions such as "cyberculture" (Negroponte, 1996), "virtual communities" (Rheingold, 2000), "online identities'" (Turkle, 1995) have been introduced to harness the novelty of new media. Such honourable enterprise, however, has been defeated by the speed at which digital technologies have infiltrated modernity. Electronic interactions have become so pervasive that they can no longer be conceived as a separate social space. No longer limited to a specific sector, digital interactions are now ubiquitously weaved into the fabric of collective existence (Roger, 2004). Such interactions have been described by the emerging domain of Web studies observing the use of the web as a social media (Cardon 2007).

Yet, the web also stores histories of intentions, matters of concern as well as cultural preferences (Lynch 2002; Latour, 2005; Venturini, 2009a). The web can also be studied as a corpus offering various types of data (hyperlinks (Park 2003), meme tracking (Leskovek 2009), etc.), in which the most promising ones (and less used) are the verbatims left on blogs, forums and other web platforms. This material has yet to be exploited by social scientists as a new field of empirical research where data are not prompted by researchers but are more or less spontaneous expressions or comments left on the web. So far, only private companies specialised in marketing have been able to develop tools to investigate this field (influence and sentiment analysis on the web), leaving social scientists dependent on their expensive and proprietary tools.

2. TECHNICAL AND SCIENTIFIC DESCRIPTION OF THE ACTIVITIES

2.1. ORIGINALITY AND INNOVATIVE FEATURES OF THE EQUIPMENT PROJECT

The *Données, Infrastructures, Méthodes d'Enquêtes en Sciences humaines et sociales* (DIME-SHS) project aims to provide France with an innovative intermediate structure for the collection, enrichment and diffusion of data that is intended not only to allow France to catch up in this area, but also to position itself as an actor meeting the highest standards of international competition. Its originality lies in the fact that it considers the collection, documentation and diffusion of data (from three major sources in SHS) as a whole, and takes full advantage of the possibilities internet and new technologies provide, especially in terms of costs but also in the exploration and diffusion of classic data and the creation of new data. Considering the different data in conjunction with the various modes of their collection will allow researchers to conduct innovative studies (for example studying the evolution of eating behaviour over time, along with life styles and Body Mass Index, whilst also analyzing web forums especially those dealing with eating disorders) and to develop multi-method research. This will allow France to take its place at the highest international level and to offer researchers using the services of DIME-SHS the highest degree of scientific rigor. Bringing together these three activities will reduce their running costs, through the numerous synergies made possible by the growing place of internet and information technology in the collection, enrichment and diffusion of data in SHS.

Collecting data with questionnaires

The SRC component of the DIME-SHS aims to put in place a panel that is representative of the French population, to be interrogated over the internet via a mobile terminal, complemented by an online and telephone survey programme. Other than the reduction of the costs that result from collecting data directly from the web and the respect of the most stringent methodological requirement in terms of sampling, the use of mobile terminals has a number of advantages: innovative studies that make the most of smartphones, ease of use compared to a computer (mobile phone penetration rate at 95%, 2009 ARCEP), increased possibility for participation in proposed surveys than via a computer (mobility), object more personal than a computer. Furthermore, offering a smartphone is a powerful incentive for limiting the attrition.

Archiving and diffusion of qualitative studies: innovating for greater respect of the nature of these data

The module BEQuali of the DIME-SHS has the objective of extending the French initiative by finalising the integration of the databank into a larger resource centre at the national level. The project we are developing is innovative on two levels. Firstly, it is based on the whole study rather than just the data, in other words it aims to limit the risks of decontextualised treatment of data. A "study of the study" (in-depth interviews with the researcher submitting the data, or a member of the research team who worked on the study) will be conducted in order to give meaning to the archived documents and to reconstitute *ex post* the research dynamic. Secondly, it makes a more complete use of the potentials of the web in order to promote this data through tools for navigation and exploration that (re)establish all the possible links between the documents offered.

Web corpora

The web data module of DIME-SHS will be the first structure in France to offer researchers a service providing collection of spontaneous expression data from the web, a service which is today solely available in the private sector. This service will provide new opportunities in the diversity of data, in the frequency of data collection and the size of corpora available. Moreover, the inclusion of web corpora within an equipment structure that also incorporates both survey based on quantitative and qualitative studies, allows for a true methodological complementarity between research projects. These web methods would allow researchers to get a better glimpse of the problem question before setting up a questionnaire based study, or inversely, they would complement a classical study with a qual-quant analysis of web verbatim.

Partners

The DIME-SHS project is based on the complementarity of the different partners that have come together around this project and who each bring their internationally recognised experience in data, infrastructures, and survey methods. The success of such an innovative equipment project

indeed depends on the exchange and the complementarity of these various forms of methodological expertise, which the equipment will in return help develop and reinforce. The methodological research that DIME-SHS partners will be able to conduct will help maintain the competitiveness of the project and the quality of the research it enables.

The collection and diffusion of DIME-SHS's quantitative data will benefit from the expertise in conducting studies on very different topics of the laboratories linked to the PRES Sorbonne Paris Cité (CDSP, CEE, CERLIS, CEVIPOF et OSC)[9], and of those of GENES (the research component of which is the Centre for Research in Economics and Statistics, CREST). INED[10] has solid experience in conducting studies on sensitive subjects (sexuality, reproductive health etc.) and in using innovative protocols (mixed mode surveys, internet, paper, telephone etc.). Moreover, this activity will make the most of the competences of Telecom ParisTech (research methods and telecommunication engineering).

The acquisition, enrichment and diffusion of qualitative studies will profit from the experience of GRETS (belonging to EDF R&D) in archiving qualitative surveys and in secondary analysis (with the PACTE laboratory, Grenoble). The different laboratories at Sciences Po that are involved in DIME-SHS (CDSP, CEE, Cevipof but also the OSC, CSO, the Centre for History, and the medialab) have contributed to the exploratory programme (funded by Sciences Po and TGE-Adonis at CNRS) constructing a qualitative studies database – of which the results will become the foundations of the project we intend to create. It will also benefit from the experience of the Social and Economic sciences department of Telecom ParisTech.

The collection, analysis and visualisation of the web data will be based on the expertise of Sciences Po's médialab which has been designed to develop methodological expertise for the use of digital technologies within the framework of research in SHS. This aspect of the project will also have the expertise of Telecom ParisTech which has been involved in the conception of web corpora collection and analysis programmes from representative samples and tools measuring internet usage. EDF R&D will also contribute their recognised skills in the area of the analysis of web data, particularly the analysis of heterogeneous data and audio mining analysis (analysis of emotions in the voice). The diffusion of the data collected and acquired within DIME-SHS will benefit from the experience of the GIS Réseau Quetelet, which will also ensure that the highest standard in matters of data confidentiality are met.

Together, the skills and expertise of these different partners will allow us to offer a programme of collection, enrichment and diffusion of data in SHS that responds to the three major issues in SHS research, whilst making the most (in terms of cost and innovation) of the social and technological revolutions of the web, the mobile web and smartphones. It will also contribute to improving the methodological training of French students, hence their employability. The possibility to conduct barometric studies will foster partnerships between the public and the private sector.

[9] Centre for socio-political data (Centre de données socio politiques), UMS Sciences Po - CNRS, evaluated at A+ level by AERES. Centre for European Studies (Centre d'études européennes), Sciences Po, research team under contract with the CNRS (EA 4459). Centre for research on social connections (Centre de recherche sur les liens sociaux), UMR CNRS, Paris Descartes and Paris III, Sorbonne Nouvelle, A+. Sciences Po Centre for Political Research (Centre de recherche politique de Sciences Po), A+. Observatory of sociological change (Observatoire sociologique du changement), UMR CNRS et Sciences Po, A+; Centre for the Sociology of Organisations (Centre de sociologie des organisations), UMR CNRS et Sciences Po, evaluated at A+ level by AERES.

[10] INED was evaluated at A+ level by AERES.

2.2. DESCRIPTION OF THE PROJECT

2.2.1 SCIENTIFIC PROGRAMME

DIME-SHS is an "Equipment for Excellence" which aims to collect, document, and diffuse quantitative, qualitative, and web data making the most of the new possibilities for reducing costs and conducting innovative surveys that information and communication technologies present.

Quantitative data

The creation of Survey Research Centres (SRC) in North America and in Europe aimed to overcome the limits and reduce the costs of academic studies outsourced to private survey companies. If the ongoing success of SRCs has demonstrated that it is possible for universities to collect data at a lower cost and higher quality, it is nevertheless impossible to apply the solutions of the 1960s to the situation in 2010. Two other parameters must be taken into account: 1) the declining response rates to surveys (academic or commercial); 2) the technological revolution of the internet and mobile connection possibilities.

The SRC component of the DIME-SHS equipment aims to provide France with the internationally competitive equipment for quantitative data collection that researchers in SHS are lacking. It is based on two complementary protocols: a web panel and a telephone call centre. The web panel is inspired by American and Dutch experiments and is based on the principle of combining the highest degrees of statistical quality and rigor with the advantages offered by recent technological evolutions. Like the LISS panel at the University of Tilburg, the web panel will be based on a representative sample of the French population selected according to a random sampling plan conducted by INSEE. Each person selected to participate in the programme will receive a smartphone and a corresponding plan, in exchange for their participation. Using a smartphone will allow for the combination of internet surveys and mobile protocols.

Internet surveys substantially reduce the costs of survey-based research because by definition they take place without using interviewers. This type of approach also allows for surveys to be conducted much more rapidly because there is no waiting list effect (limitation of the number of interviewers compared to the number of interviewees). Moreover, the people interviewed can respond to the online survey when they wish to, and not when they are solicited by interviewers by telephone or at their home. Internet surveys provide richer studies (multimedia) and different ways of posing questions (sound, text, images).

To these advantages that are specific to internet surveys, can be added those that result directly from the choice of a mobile terminal, the originality of the DIME-SHS compared to other equivalent projects in the US or in Europe. Mobile terminals provide three types of advantages over traditional computers. Firstly, the mobile telephone is totally integrated into daily life, it is even one of the pieces of technology that has been the most widely and rapidly adopted by people in France (penetration rate of mobile telephones was 95% in 2009 according to ARCEP). As an object of daily life it is much easier to use than a computer, especially the latest generation which is equipped with operating systems that are much more intuitive to use than PDAs or the first generation of smartphones, of which the interfaces were very close to traditional computers. These mobile terminals make internet IT accessible to the wider population and appear to be ideal for collecting data. Because they are mobile, they are most often carried constantly by the participants, who can then respond to surveys when and where they wish, not only at their homes. But such terminals also make it possible to envisage more original studies that make the most of the smartphone technology: movement studies (geolocalisation of telephones), time use studies (simplified time use diaries or alternative methods such as the day reconstruction method), following biomarkers (self-administered blood and saliva samples see Finch, Vaupel et Kinsella 2001). If this kind of survey were to be conducted, the highest level of confidentiality standard would be applied.

Finally, providing mobile terminals and paid plans ought to reduce the non-response and attrition rates (the loss of respondents who were intended to participate in a panel). This is primarily because the telephone number is known to the researchers, but also because of the incentive associated with the provision of a smartphone and the payment of a telephone plan. Such incentives reduce the non-response rate all the more given that these are unconditional gifts (i.e. they are given at the beginning and not at the end as a reward) and reasonably valuable (Mack et al. 1998, Singer and Kulka 2002).

To these advantages can be added those of a panel study that allows the researcher to follow individuals over time. This feature would make possible to monitor changes in voting intentions and

to relate them to both individual characteristics and campaign events. Whereas this kind of survey is common in the US (1st electoral panel in 1948, University of Columbia) and in the UK (since the 1990s), there have been only 4 such surveys in France (Le Hay 2010), the last two covering only three months. From a methodological perspective, the longitudinal study of a given population enables the quality of the information collected to be ameliorated (repeating certain questions, checking for coherence, possibilities of modification) and the specificities of these individuals to be more adequately taken into account during the analysis. But methodological researches need to be conducted in order to use these facilities efficiently and reliably (Laflamme and Mohl, 2007).

Although innovative, the web panel must be complemented by a telephone survey centre. This centre will have the objective of treating studies for which a different population than that of the panel is necessary, those for which the questions appear less easily understandable on a computer, or those that do not fit into the panel schedule. It will also be able to conduct studies on the panel itself (mixed mode). Tools for conducting online studies will also be available on the DIME-SHS website (with methodological assistance).

Like any important project, the creation of a call centre and a mobile web panel raise several legal, technical and economic problems. DIME-SHS will make sure it works with the best legal experts on questions related to the confidentiality of personal data. It will also have to establish a partnership with a telephone services provider in order to avoid technical problems and be able to offer the services proposed to the panel. The application development platforms on mobile telephones are now accessible through application programming interfaces (APIs) but it is not always possible to control how they are used, it depends on the supplier or even the service provider. An agreement in the interests of both parties must be found in order to protect the independence of the scientific objectives and the deontology of the collection of personal data.

Qualitative data

The difficulties in constructing a sustainable qualitative study database are multiple and have been well identified by comparable projects in Europe (Dargentas et alii, forthcoming). We will thus refer to the solutions they have adopted, adapting them to the French context – institutional, legal and scientific – and using them as the basis for further innovation, as we did in the exploratory project conducted since 2009. This experiment, which is at the test stage of a prototype composed of three studies chosen for the range of difficulties they present in terms of archiving, has shown that it is possible to sufficiently document data *ex post*, in order to conserve their validity.

In keeping with experiments conducted elsewhere, mostly in Europe, we choose to integrate the BEQuali module into a previously established resource centre, the CDSP, which allows us to benefit from the previously accumulated experience in the area of quantitative methodologies and to create economies of scale once the programme is up and running. The archiving, diffusion and reuse of these data do pose specific difficulties however: the particularity of the materials, and the particularity of the documentation and enrichment procedures for this data – which can only be overcome by bringing in a specialised team. Moreover, the realisation of a "new generation" database built on study-websites with their own navigation tools allowing for the online exploration and exploitation of the data, means specific IT developments. In order to do this, we will begin with the prototype developed by the medialab, and benefit from the shared IT resources of the three components of DIME-SHS. We will also take into account the concrete uses of this programme in the future, which means adapting to the expectations and real-life practices of users.

Bringing together studies at the national level is a challenge in France, because unlike in Switzerland or in the United Kingdom, there is no directory of researchers and no inventory of research projects currently underway. Researchers hold the intellectual property rights to their studies (Cornu, 2003; Mallet-Poujol, 2004; Descamps, 2007, etc.) and only the case of researchers in the public service requires the agreement of the CNRS (or other bodies such as INRA etc…). To begin with, we already have significant potential in the (numerous) researchers in the laboratories of Pres Sorbonne Cité. Moreover, the REANALYSE project (financed as part of the Programme *Blanc 2010* and partly born by Sciences Po) which will begin in autumn 2010, was especially designed to convince our colleagues[11] of the feasibility – in the epistemological as well as methodological sense

[11] The same reluctance has impeded the acceptability of these archive centres (and with them secondary analysis), comes back in the acceptability studies carried out in several countries (Opitz, Mauer, 2005 ; Medjedović, forthcoming)

– of secondary analysis, as well as the potential benefits of such a programme for research, in particular the amelioration of the protocols and methods of qualitative research, the possibility of being able to access a larger, more diverse range of data than is actually the case, and to access them in a holistic and thus more respectful way. Capitalising on qualitative data will also enable researchers to challenge some of the inherent limits of qualitative studies. It will help them (especially young researchers) better prepare new studies, but also to produce new findings by "doing new with old" – and all that at a lower cost.

Building the archive, we are paying close attention to the ways of involving researchers who are submitting work for archival and diffusion, through the guarantees of data anonymity, the completion of the "study on the study", and the possibility for those submitting work to follow the reuse of their data and benefit from it. Having one's study archived should be seen as a form of recognition of its quality as is becoming the case in Britain. Similarly the reanalysis it gives rise to will be a means of publicity for the original work and a way of further exploring its results. Eventually, it will enable an improvement in scientific work, along with a reduction in the costs of public funding to research. The implementation of the programme essentially requires the means of bringing together the necessary skills: the difficulties have been identified and the solutions are on hand.

The Collection of Spontaneous Expression Online

This element of the DIME-SHS facility will aim to develop the techniques and methods related to the collection of spontaneous expression data from the web: tracking and mapping service for locating pertinent online discussion according to the subject of the research and semi-automatic extraction and data cleaning methods.

The only initiative that resembles this in France is in the private sector within the company Linkinfluence, which offers a service analysing the presence of brands or institutional advertising on the web. The research partnership which was signed by the medialab and Linkinfluence in 2010 has enabled the recent development of research using Linkinfluence's data. The first of the difficulties encountered by the researchers is related to the inability to master the available web corpus. The methodological approaches of Linkfluence and DIME-SHS differ on this issue. Linkfluence set up a corpus covering the ensemble of blog communities for a number of countries, based on a detailed topographical analysis. This corpus is collected and indexed regularly and provides the basis for the different analyses of each client. The DIME-SHS equipment on the other hand will construct one corpus per research project although web resources may overlap from one project to another. The methods of identification and data extraction will of course be brought together and reused in each project. Thus the DIME-SHS equipment project would give researchers the ability to develop their own data collection methods, to give them full control of the construction of the corpus to be analysed.

By addressing the web data collection issue, the ambition of DIME-SHS aims at bridging the gap between traditional social sciences data and web ones. Web corpora are to be thought in a perspective that is a midway between qualitative and quantitative. On one hand, we are dealing with qualitative verbatim (of a specific kind because they are spontaneous), on the other hand, the sheer number of verbatim texts leads us towards a more quantitative perspective. The enormity of the web data brings a reversal of perspectives when approaching the different corpora.

In particular the representativeness of the data, a key point in classical survey method has to be adapted to the web context. It is clear that the traceability of digital brought new possibilities of profiling but two directions have to be opposed: on one hand, one can by investigating on a limited amount of profiles (applying classical qualitative methods in the digital space), discover a large amount of personal data (see Le Tigre); on the other, a large majority of spontaneous expression on the web does not come with any socio-demographic identification of authors, hidden behind pseudonyms. DIME-SHS will focus on the second way, building large corpus of verbatim using semi-automated data collection methods; and will investigate new ways to think representativeness as by adding the qualification of web expression platforms (with both topologic and qualitative analysis) from which the verbatim has been taken. According to the vast possibilities of places of expression in the web, the place one chose to express a point of view isn't trivial.

This movement towards a qualitative-quantitative scale as a result of the web corpora also changes the analysis methods of the spontaneous expression content taken from the web. The issue here is to assist the researcher by giving him/her the means to research and explore content, which means incorporating the capabilities of specialized software for the automatic analysis of language. Regarding the issue of analysis, the equipment will draw on another "Equipment for

Excellence", MUSCA, which develops analysis techniques for large bodies of text (Sciences Po's médialab is a partner of the MUSCA project).

From a technical point of view, the collection of spontaneous expression data from the web is a superficially simple problem. Digital technology ensures easy access to and duplication of information, but the versatility and diversity of content considerably complicates the task. The initial step will be to overcome the difficulty of cleaning the web content so as to extract verbatim text from the web pages. The problem presents itself differently depending on the means and site of expression one is interested in. Although initiatives exist in the search for automatic data cleaning methods, a great variety of web content seems to resist this approach. Adapting extraction methods to the particular platform or site ensures a much better quality of extraction. This method particularly allows the adding of various types of information to the verbatim such as the date of publication and details related to its author. The DIME-SHS equipment will have to invest as much in the development of automatic cleaning methods as in the development of heuristic data cleaning tools specific to the key platforms in the collection campaign. It's by using both methods in combination that one can obtain the better results. One of the important advantages of this equipment for the scientific community lies in its expertise and capital of extraction heuristics, ensuring that the data collected meets sufficiently high standards.

The great versatility of the web also creates challenges for the validity of corpora and extraction methods through time. This is an obstacle for the one-off initiatives of web data exploitation used in research today. Only a large-scale project dedicated to this task can reach the critical mass required to ensure the maintenance of corpora and of extraction methods. The technical team, which is dedicated to the DIME-SHS equipment, will be able to ensure "software" maintenance in the same way that any substantial scientific facility necessitates technical maintenance.

The legal aspects of the management of the data collected by this programme will be key to its success. The legal framework for the development of a database of verbatim spontaneous expression extracted from the web remains unclear, but is similar to issues of web archiving that are currently under discussion in France as we approach the publication of the report on the decree regarding the legal deposit responsibilities for the web of the BNF and the INA. Furthermore, as with any such qualitative database, the management of this data implies a rigorous privacy policy. The background and experience of the partners involved will encourage the development of policies that respect the laws currently in force.

Conditions of Access

The DIME-SHS equipment will be accessible to the whole national and international (comparative studies) scientific community as well as public bodies (administrative services, ministries) and the private sphere (associations, businesses). DIME-SHS will be available to the academic community in the SHS: demographers, economists, legal specialists, geographers, historians, linguists, political scientists and sociologists. The SHS are also in partnership with a host of other scientific disciplines and as a result, DIME-SHS will be able to contribute to the development of research in other areas (health, risk analysis, or environment for example) and work with other non-academic bodies: state services, local governments, NGOs and businesses. Data collection projects will be selected according to their scientific qualities and their potential in terms of exploitation. They will be dependent on a financial contribution to the running costs and the cost of depreciating equipment. This contribution will be determined in relation to the project (length) and the situation of the user. A limited number of data collection opportunities, for which no financial contributions will be required, will be kept for doctoral students, in line with its vocation as a teaching support and as a tool for the development of the methodological abilities of future generations. Conditional on their quality, data collection projects of the partners of DIME-SHS will have priority. The financial contribution required for using data collection services could be reduced by the creation of a system of institutional, academic or other subscriptions (as is the case of the European Consortium of Political Science). Data collected will be made available free of charge for scientific use on the portal of the GIS Réseau Quetelet.

Position on an international level

On a national level, DIME-SHS will be a major element in the national policy for data and infrastructure in the social sciences. Indeed, the "Very Large Research Infrastructure" (Très Grande Infrastructure de Recherche) GIS Réseau Quetelet is a partner of DIME-SHS. Moreover, three partners of DIME-SHS (Sciences Po, GENES, INED) are also partners in this frame. Centred on innovative and responsive studies, data collection procedures by and on the web, and the

development of archiving and of the diffusion of qualitative data, DIME-SHS will put forward projects which are complementary to the GIS Réseau Quetelet which focused on the archiving of data and support for large international studies (ESS/European Social Survey, SHARE/ Survey of health, ageing and retirement in Europe). The National Data Committee in Humanities and Social Sciences (*comité de concertation pour les données en sciences humaines et socials* - CCDSHS) will also be part of the DIME-SHS governance, ensuring a tight coordination with the national directions for a official data policy for research in the humanities and social sciences it proposes. DIME-SHS will also complement the CNRS's "Very Large Facility" (*Très Grand Équipement*) for the social and human sciences, Adonis. Moreover, the means of diffusion of the qualitative, quantitative and qual-quant data produced by the DIME-SHS will be set up for inter-operational use with other the frameworks put forward by Adonis.

DIME-SHS also appears to be complementary to other cutting edge projects. Firstly, it is closely linked with the other projects at PRES Sorbonne Paris Cité. DIME-SHS will underpin two advanced laboratory-led projects and more particularly the "Santé Globale et Politiques de Santé" (Global Health and Health Policy) project which aims to develop a multi-disciplinary centre of excellence for the analysis of health and health policy, specifically in relation to its social factors. The internet panel will enable researchers at the laboratory to develop innovative research which links everyday practices to long term health consequences (for example the analysis of the links between food consumption, lifestyle and obesity). DIME-SHS will be able to offer its various methodological frameworks, archiving methodologies, and documentation and diffusion of web studies to the "Transforming Worlds" laboratory for excellence project. DIME-SHS also complements the high-tech facility GeoMedia Mapper. In particular, the geo-localising data that will be collected in the framework of the panel will be able to be analysed with the tools that will be developed by GeoMedia Mapper. They will also be enriched by linking up with other geographical data thanks to the tools and standards which will be developed by the "Equipment for Excellence" project at the Centre Belgrand, supported by the PRES Paris Est. Finally, the DIME-SHS also complements the INED "laboratory for excellence" which develops research in demography and related disciplines through the articulation and integration of study research methods with the production of studies, to improve their quality.

The DIME-SHS also strongly complements other "Equipment for Excellence" projects. Medialab is a project partner of the MUSCA EquipEx (Multi-scale collaborative mapping of large digital corporat, lead by the Ecole Polytechnique) that aims to endow France with a facility for the macro scale collection of data and intensive calculation. DIME-SHS also complements the EquipEx project "Centre for the Secure Access of Data" of the PRES Paris Sud which aims at making highly secure public statistics data available to the research community.

Expected benefits and evaluation criteria

DIME-SHS will provide innovative and cost effective collection data capabilities at the highest methodological standards for scientific projects. It will also give the opportunity to develop methodological expertise, among others through the development of quality indicators of the studies (bias measurement, total error, quality of the statements etc.), which will serve to document future studies and will aim to set the standard for research in the SHS. Following a promising research line, paradata (information on the data collection process) will be used for enhancing data collection efficiency (Laflamme and Mohl, 2007) and thorough comparisons that will be conducted between data collection modes on one side and between results issued from the panel and those obtained by more classical survey means on the other. It will also make these data as well as the qualitative studies acquired available for secondary analysis. As such, the success of DIME-SHS can be assessed on the basis of the number of research projects conducted with the help of the services it proposes as well as the number of publications based on the data collected and acquired with its help. DIME-SHS will also contribute to better methodological teaching and as such will improve students' training and employability. The standards and methodological excellence promoted by DIME-SHS will also benefit to the private sector in two main ways. Firstly, results of methodological experiments will be made public and as such will stimulate innovation. Secondly, it will provide the private and public sector with students trained according the highest methodological standards, which will improve in the longer run their competitiveness.

.

2.2.2 STRUCTURE AND BUILDING OF THE EQUIPMENT

DIME-SHS is an "Equipment for Excellence" which aims to *collect, document, and diffuse* data making the most of the new possibilities for reducing costs and conducting innovative surveys that information and communication technologies present. Gathering these activities for three types of data makes it possible to bring about important synergies. To highlight them, we present the equipment DIME-SHS in terms of collection, documentation, and dissemination of data for the humanities and the social sciences and distinguish between the different types of data only when relevant. The equipment will be distributed in three sites: Sciences Po, INED, and GENES. This multisite distribution is a necessary condition to reach the critical mass of expertise needed to build and manage the equipment at the required level of methodological excellence. In addition, this will strengthen the pre-existing links between these three members of the GIS Réseau Quetelet.

- *Structure*

Element 1 – Data collection

Quantitative data

At the exception of the call centre of the IEP of Grenoble (PACTE), which is not opened to other researchers, there is no academic structure to collect quantitative data currently in France. In Europe, only the CentERdata at the University of Tilburg in the Netherlands that manages the LISS panel is close to the mobile web panel of DIME-SHS. The collection of quantitative data will be based on two complementary protocols: a panel and a telephone call centre. The panel will be selected according to a probability sample representative of the French population. Each respondent will be offered a free smartphone and a corresponding plan in return for replying to a certain number of questions every year, number that will be determined after an experiment stage. The targeted size of the mobile web panel is 10,000, each respondent staying for 24 months in the panel. Every year, a quarter of the sample will be renewed. We plan to have a subsample (representative of the French population) of respondents who will stay longer. The telephone call centre will have a capacity of 20 simultaneous interviews and will be equipped with call recording and monitoring systems that will be used to conduct methodological research and to teach. It will be located in the future building that GENES will occupy in 2014 on the Saclay campus. DIME-SHS's telephone platform will be equipped with hardware and software compatible with PACTE's telephone call centre in order to be able to mobilize them jointly to carry out phone-based studies on a greater scale or more rapidly than would be otherwise possible. In addition to these two protocols, we will make available on the website of the equipment tools to conduct web surveys and methodological guidelines to use them.

Qualitative data

For qualitative data, collection must be understood as the acquisition of all the materials related to qualitative surveys already conducted. At first, most of the acquisitions will be based on approaching researchers individually but at a later stage, we expect researchers to contact us directly for submission. An initial report on qualitative studies considered to be archived will be prepared by a member of the qualitative archive team, emphasising not only the quality of the data work, but also the intrinsic quality of the study and its potential for reuse. The archiving decision will be made by a scientific committee based on this report.

Web data

The web data collection service will assure two missions: identify and qualify the online expression platform building a corpus of sources by project using web mapping methods, crawl, filter and clean the sources to collect spontaneous expression verbatim.

The web data collection will be accompanied by reflection as to the uses of the web, the specificity of social interactions that occur in different areas (blogs, comments, discussion forums, social networks, messaging). The whole chain (data identification, constitution of the corpus and then interpretation) must be thought out in line with web media.

The software developed in the equipment (web mapping, crawling, filtering…) will be based on and released as open source projects allowing other researchers to reuse the tools but above all to access the very details of the algorithms verifying methodological validity.

Element 2 – Data documentation and preservation

Documentation of data is essential to foster secondary analysis. We are committed to respect the documentation international standards when they exist and will contribute to elaborate them when they do not. Questionnaire survey data will be documented following the DDI (Data Documentation Initiative) standard using NESSTAR and possibly Questasy. NESSTAR is a platform to document quantitative studies based on DDI developed within the Council of European Social Science Data Archives (CESSDA) and used by the CDSP since its creation. Questasy is a web application developed by CentERdata of the University of Tilburg. Based on DDI 3, this platform makes it possible to make the most of the latest version of the DDI standard to document longitudinal data. There are no established documentation standards for qualitative and web data and we will join in current discussions to help building them. Researchers involved in DIME-SHS will carry on with the discussions within CESSDA (the UK Qualidata and the Swiss FORS archives) to develop different ways of promoting studies (diffusion of study sites model) and to establish documentation and referencing standards on the basis of the Text Encoding Initiative (TEI) standard and on Qualidata. Expertise in the documentation of quantitative and qualitative data will benefit to the development of standard for the quali-quanti data collected on the web.

Element 3 – Data dissemination

The data collected (panel, telephone, qualitative, web) will be referenced on the portal of the GIS Réseau Quetelet and made available, free of charge, with innovative web tools and interfaces. The diffusion model for data is part of the open science research movement, a new model that is becoming increasingly common in the social sciences and humanities, either quantitative (Freese 2007 and King 2007) or qualitative (Moravscik 2010). Data will be available after a possible exclusivity period defined with researchers. Data disseminated will respect the highest level of confidentiality, if necessary by restricting access to researchers who signed an extended confidentiality agreement.

- *Building*

We will start with a period of testing (2 years) during which we will conduct experiments to fine tune technically (mobile web panel) and methodologically the project. At the end of this stage, we will have a fully functional small scale version of the equipment. During this test stage, only the partners of DIME-SHS will be able to conduct methodological and substantial surveys. The telephone call centre will be constructed at the end of this test stage for practical reasons (availability of office space) but also because it is a mature equipment that does not require such an extensive test stage. The next stage of the construction will consist in strengthening the equipment (increase of the size of the panel in particular). We will also start to open gradually the data collection services to the scientific community. We expect this second construction stage to last 6 years during which the services provided will be revised, fine-tuned, and expanded. This second stage will also be necessary in order to be able to produce to the scientific community examples of original, cost-effective, and high quality surveys and publications based on them so as to arouse the interest for the equipment and give the necessary time to researchers to raise funds to conduct them with DIME-SHS. These 6 years will also enable us to build up a backlog of funded surveys to be conducted with the equipment and to refine the economic model so as to make it viable.

The construction stage is estimated to end after 8 years after which the financial contributions required to collect data, together with the resources provided by the partners, will cover the running and replacement cost of DIME-SHS.

2.2.3 TECHNICAL ENVIRONMENT

Two resources will be mutualised. First, IT resources as not only do NICT provide innovative ways of collecting and disseminating data, but they also can be mutualised between the different activities and data of the project, bringing about significant economies of scale. The core IT team will be composed of experts of the different fields of computer science that are necessary to deliver the different services offered by DIME-SHS. To collect data (web panel) and to disseminate all the data collected and documented (quantitative and qualitative surveys, web data) we need one web interface designer and two web developers. To preserve and manage data, one data manager is required. Management resources will also be mutualised. The management team, described in more details in section 4.1.1, will be composed of a Scientific Director (scientific management), an Executive Director (day-to-day running), a lawyer (for the first two years of the project), a marketing manager and an administrative and financial assistant.

Element 1 – Data collection

Quantitative data

The collection of quantitative data will be based on smartphones and a telephone call centre. Marcel Das, the Director of the CentERdata, will contribute to the practical setting up of the DIME-SHS facility as a Consultant. The innovative and cost effective questionnaire data collection methods rely on continuous monitoring of the representativeness of the panel and of the different sub-panels, on limiting survey non-response and attrition, on controlling possible biases introduced by survey modes, etc. The survey team will comprise one statistician (refreshment of the panel, sub-sample design, statistical quality of the different samples, computation of probability weights for each survey, etc.), three panel managers (relations with panel members, follow-up, etc.), and one call centre manager. The survey team will work in close collaboration with the method team which will be composed of six survey methods experts (bias measurement, questionnaire design, multi-mode experimentations, etc.). In order to make the most of the expertise provided by GENES, INED, and Sciences Po, and to strengthen the methodological ties between these partners, the data collection team will be based on these three sites. The panel team (one statistician and three panel managers) will be located at INED where it will benefit from the expertise of the Surveys Department. The call centre manager will be hosted by GENES where the facility will be located. Half of the survey team will be based at Sciences Po, half at INED. Interface will be designed by the mutualised core IT team. We will coordinate with PACTE to ensure the highest level of compatibility between the equipment and software of the two telephone call centres so as to make it possible to use them jointly on research project that require higher samples on faster data collection.

Qualitative data

The acquisition of qualitative surveys will be ensured by the team in charge of documenting them. This requires the necessary skills, both to interest and convince them to pool their studies (highlighting the need for the person(s) in charge of this task to have an established scientific reputation) and then of course to evaluate their quality.

Web data

The collection of spontaneous expression data from the web requires the gathering of various expertise such as: Sociology of the web communities and of the use of social media; Extraction technology and the cleaning of web data; Web crawling technology and strategies; Linguistics IT; Databases and data-mining; Software architecture and distributed systems; System administration. The collection of web data requires extraction algorithm designed jointly with IT and researchers but also the use of computing resources (a server farm and a data centre, resources that will be used by the other services and activities of DIME-SHS). The software architecture will be composed of elements for the collection and cleaning of web data which can be applied to a server farm performing parallel calculations. This calculation node will host the data treatment algorithms necessary for the maintenance of the corpus and the functioning of the equipment. Further analysis of the corpus will be able to take advantage of collaborations with other facilities such as those existing at the Sciences Po medialab and more particularly the "Equipment for Excellence" MUSCA. The storing and the conservation of the data will be handled by a specialised data centre.

Web data will be collected by four computer scientists: one web API developer, one crawling developer, and two web data structure experts. In practical terms they will complete the core IT team.

Element 2 – Data documentation and preservation

Documentation is not only a technical question but involves also theoretical and practical expertise in survey methods and humanities and social sciences. As a result, a certain specialisation in certain types of data is difficult to avoid but we will strive towards developing synergies. The documentation team will be composed of two quantitative archive managers and four qualitative ones. Combining quantitative and qualitative documentation expertise will give rise to unprecedented synergies to develop the new standards that are required by the qual-quant materials collected on the web.

Data collected will be stored and preserved using a SAN and a data warehouse with the highest level of security and replication policies. This storage equipment will be centralised in Sciences Po where similar hardware are in use to keep the running cost as low as possible.

Element 3 - Dissemination

Data will be disseminated through the GIS Réseau Quetelet portal, which is the reference data access point for the social sciences. The portal will be redesigned by the GIS Réseau Quetelet to disseminate the data collected and documented within DIME-SHS. Qualitative survey data will be disseminated through tailor made web tools referred to in a new section of the RQ portal. We will develop new means of dissemination for web data based on the quantitative and qualitative expertise put together in this project.

All the partners involved in DIME-SHS confirm that the part of equipment depending on the competencies of their institution can be implemented and put into service as soon as it is delivered and that the means necessary to accompany its functioning, personnel included, will be implemented. The commitment letter of all the partners can be found in appendix 6.4.

3. DISSEMINATION AND EXPLOITATION OF RESULTS

- *Dissemination*

New information and communication technologies are at the heart of the DIME-SHS project, for the collection of data, but also for making this data available to the scientific community. The data collected (panel, telephone, qualitative, web) will be documented according to international standards and made available on the web with innovative tools and interfaces. The diffusion model for the data is part of the open science research movement, a model which is still relatively new within the social and human sciences but which is becoming common in economics, political science and sociology (see the special edition of *Sociological Methods & Research*, 2007 regarding the replication and access to data, in particular Freese 2007 and King 2007).

According to this perspective, the data and the procedures (statistical programmes for example) are an integral part of the results published in the articles or in books and for this reason must be documented and accessible, possibly on particular conditions (after a period of exclusivity, or for replication only etc). The same questions are raised by qualitative methods, where the scarce use of quotations and the inability to consult the analysed data is increasingly called into question (Moravscik 2010). A web portal presenting the "Equipment for Excellence" will provide access to the totality of the data which has been collected and archived by the DIME-SHS. Online tools will enhance the different types of data managed by DIME-SHS.

After their documentation according to the XML DDI standard (Data Documentation Initiative) the questionnaire study data (web panel and telephone) will be diffused by the CDSP on a NESSTAR platform. This will enable the user to access information about the studies (meta-data), to access the questions posed, to conduct initial analyses (recoding, frequency tables, two-way tables, graphics) and to download the data in one of the formats accepted by the main statistical software (Stata, SAS, SPSS, etc.). Micro data from one study may be downloaded for free once the exclusivity period – decided in consultation with the authors – has passed and once CDSP recording has been carried out.

The external coordination of the BEQuali module, as the development and the maintenance of the links between the team partners and the scientific community as a whole, will be the object of our full attention, given that the success of the project also depends on our capacity to spread the word and gain acceptance from our colleagues. Convincing the scientific community of the advantages to be gained from the pooling of studies will only happen with time, and will need to be supported by efforts to promote the programme, particularly through the participation in various scientific events and publications, as well as the development of modules to accompany the teaching of qualitative methods.

We are currently engaged in collaborations on a national level, notably with the REANALYSE project which will serve to demonstrate the interest of secondary qualitative analysis. We have also begun the creation of an international collaboration through the CESSDA of which CDSP is a member, and through the development of contacts with Qualidata or FORS so as to develop different ways of promoting studies (diffusion of study sites model) or the establishment of documentation and referencing standards. In the case of the archiving of qualitative studies, no specification which is comparable to the DDI norm for quantitative data has truly been developed, which enables it to respect the particular characteristics of this data. Collaboration with other partner resource centres of the CESSDA (in particular Qualidata) should be envisaged to fine tune such a specification, or at least to test its feasibility on the basis of a Text Encoding Initiative (TEI) and on the existing Qualidata initiative (DeXT project, for Data Exchange tools http:/ / www.data-archive.ac.uk/ dext/ about/ introduction.asp)

- *Exploitation plan*

If DIME-SHS offers a structure oriented to research, the inquiries completed with it must thus meet certain specific criteria: 1) Have scientific interest; 2) To be in line with the project diffusion policy (see above); 3) Respect the rules concerning data confidentiality.

The proposals will be examined by the Steering Committee (see 4.1.2). The key target groups who will have access to this equipment are identified according to the principle objective of DIME-SHS:

1. Academics from universities, research organizations research, department of private company etc;
2. Entities responsible for conducting public policy (administration etc.);
3. Civil society organizations (private or non-profit).

Restricting the perimeter of activity of the DIME-SHS equipment is the only way of ensuring its final vocation remains as a service for researchers in social and human sciences. This raises the question of how the establishment of a structure like this is to be financed by such a restricted community. In order for DIME-SHS to be quickly operational and to confirm its legitimacy it would be necessary to give evidence of a quick and complete launch, and thus a substantial financial effort to construct the equipment. Not launching the programme in these conditions would run the risk of not being able to launch it at all, due to an insufficient amount of data collected to get the programme going.

However, and in order to ensure the sustainability of the equipment beyond the 10-year period covered by the call for projects, the economic structure of DIME-SHS will be built in such a way as to ensure the self-financing of the project and its endogenous development from the end of the investment period. The economic structure will be based on the sale of the contracts. It must satisfy the following criteria: 1) Ensure the financial independence of the equipment; 2) Provide a cost per study that is reasonable and compatible with the budgets of academic research; 3) Respect competition laws in being careful to ensure profitable service provisions (no dumping due to obtaining of public funding).

Concerning the second point, DIME-SHS will permit to use public funds dedicated to research (ANR, FP7…) in a more rational way. As an example, the electoral panel project of the European Studies Centre (Centre d'Etudes Européennes) at Sciences Po (initial 20 minute interview followed by 12 repeated interviews of 10 minutes) could be conducted over a longer period, be produced in better methodological conditions and for a lesser cost than the 150,000 euro the private survey company is asking. Similarly, the archiving, documentation and diffusion of qualitative studies will enable the development of re-analysis. In addition to the savings made from the more extensive use of original data, it will allow the possibility of comparative studies – over time and between different social and cultural contexts. A very recent example is the book by the British sociologist Mike Savage (2010) who re-analysed qualitative studies led by English sociologists between the 1930s and 1960s, to understand the transformations in British society after the Second World War.

On the last point, to the extent that the activity domain of the programme will be limited to a niche group, there is no risk of unfair competition with the existing private sector: investment will encourage the emergence of new "clients" who, without DIME-SHS, would not have otherwise turned to surveys in social and human sciences.

Research and methodological advances that will be conducted thanks to the DIME-SHS equipment will be of potential interest for all stakeholders involved in the collection of data, especially private survey companies. The Dutch branch of TNS has thus developed some of its procedures following experiments conducted within the framework of the LISS panel. In fact, the research of the LISS panel, particularly the methodological aspects, is freely available on the web. DIME-SHS will also diffuse the results of the methodological experiments on the internet and will therefore be able to contribute to the debate among both public and private actors about the collection and analysis of data. The DIME-SHS Equipment for Excellence will also contribute to the competitiveness of private companies through professional training in cutting edge research methods, which it will conduct.

In addition to the adult education, the development of university training is also one of the natural offshoots of this project. Indeed, accumulated methodological expertise will also enable students to become better trained in SHS research methods and in their regeneration through recent technological innovations. The students and doctoral students of the DIME-SHS partners will be able to be trained in questionnaire methods on a practical level, as one panel per year will be reserved for teaching. Training in questionnaire-based studies will also benefit the call centre, which will be equipped with recording facilities. Archiving, documentation and availability of qualitative studies will serve to encourage further reflection on the validity of scientific work using qualitative methods. It will also promote the development of secondary analysis, a practice which is still marginal in France and which will subsequently provide a significant boost to the teaching of interview methods, particularly in the phase of analysis and the handling of collected data.

With improved methodological training and armed with the experience acquired through the use of this equipment, these students will be more competitive in the job market and will be able to complete their theses in better conditions. Their diverse experiences at the cutting edge of innovation will enable them to be easily integrated into private survey companies, research companies etc. but also to establish their own start-ups. Students who go on to doctoral study will be able to put their experiences into practice and develop innovative research methods at the highest methodological level. A certain number of web panel studies will be reserved for doctoral

projects. In both cases the DIME-SHS facility will enable doctoral students to be better trained and to increase their international competitiveness, both in terms of publications and employability.

The first significant studies which will make use of the various components of study development (panel and web corpus) will take place in the second year of the overall project. These studies could focus on two themes ensuring a synergy between classic data collection methods with the web and the panel. The first theme is the study of food behaviour problems, centred on factors associated with anorexia, bulimia, obesity and thinness. The following of the subjects over time, repeated questioning related to their dietary habits, their food spending, their perception of their problems, their professional or academic difficulties etc., will be accompanied by a study of the internet forums dealing with these issues, thus offering original complementary data. This synergy of collection methods and the longitudinal nature of the study will supply entirely new data on this question. A similar study about the consumption of psychoactive substances (alcohol, tobacco, cannabis and other illicit drugs) will lead to new information about consumption patterns (retrospective and prospective) and their links to life events (studies, love life and family, professional issues etc.). One aspect of both of these studies will be the detailed exploration of the social gradient often observed in such problems (i.e. social or economically disadvantaged people have greater problems of this sort) and its anomalies. The other aspect of this research will be updating of the vulnerability factors (the severity of the damage brought on by these problems) as well as risk factors (the presence of difficulties at home or in their entourage).

- *Development Strategy*

The inevitable development of studies through and on the internet, as well as the archiving and diffusion of qualitative data via the web, renders the creation of a platform such as DIME-SHS essential. This new type of infrastructure will have to exist within the framework of the European Strategy Forum on Research Infrastructures (ESFRI) which structures, coordinates and partly finances the development of European research infrastructure. Currently the two principal SHS structures are the Council of European Social Science Data Archives (CESSDA) and the European Social Survey (ESS). Founded in the 1970s, the CESSDA coordinates centres which document, archive and diffuse survey data. More recent (since 2001) is the European Social Survey, which is conducted every two years according to a very strict methodology – an achievement which earned them the 2005 Descartes prize which rewards Outstanding Scientific and Technological Achievements Resulting from European Collaborative Research.

With the DIME-SHS web panel, France will be in a strategic position for the second generation of European research infrastructure which is currently emerging. Based around the LISS panel at the University of Tilburg, a nascent European web panel network is being developed (University of Mannheim and GSIS in Germany, the Catholic University of Louvain in Belgium and the University of Lausanne and FORS in Switzerland). This network could allow the development of comparative studies in Europe at a much lower cost than those of the ESS and with an equally rigorous approach. The example of the ESS is interesting because it demonstrates that a project of innovative data collection, at the highest and most contemporary methodological level, can bid for European funding within the framework of ESFRI as a key research infrastructure. Although the modalities of co-financing of the European Research Infrastructure Consortium has not yet been finalised, it is likely to take the shape of a participation in the funding of national facilities and coordination structures. Funding obtained in this way could enable DIME-SHS to ensure its development and to maintain its activity at the highest scientific standards, but also to be competitive enough to become the key structure in European coordination. The 10-year long funding programme would thus enable France to take up a strategic position at the European level by endowing it with innovative equipment, which would enable it to conduct research (particularly methodological research), thus providing all the scientific legitimacy necessary to cement France's position as a European leader.

Finally, the methodological expertise that will be developed through the DIME-SHS project will be operated in the context of collaborative or individual European projects of the 7th Framework Programmes for Research and Technological Development. Other than hosting doctoral students and post-docs from the Marie Curie Institute, DIME-SHS could be linked to comparative research projects in Europe or respond to European tenders (DG, ESF, etc). Financed by an ERC Starting Independent Researcher Grant (2010-2014, 1.5 million euro) the "Trust" project by Yann Algan, on the cooperation and the mechanisms of trust, mobilises and combines studies by questionnaire, qualitative interview and the analysis of web discourses. In both cases (financing as part of a larger structure or as a partner in European projects) European funding could be a considerable growth engine for the DIME-SHS programme in the medium and long term.

4. PROJECT MANAGEMENT

4.1. MANAGEMENT

4.1.1 RELEVANT EXPERIENCE OF THE PROJECT COORDINATOR

Given the size of the consortium, the multisite organisation and the ambition of the project, DIME-SHS will have to implement an efficient management structure that enables decision making and flexibility in order to ensure the best collaboration possible between the different activities and teams. The management structure is centred on the experience of an institution and the expertise of a scientific director.

Sciences Po will be in charge of the administrative, legal and financial management. Involved in more than 35 European projects and 18 ANR projects, this top-ranking University, specialising in research in the social sciences and humanities in France, has developed a solid and productive experience in project management.

The **Scientific Director** will be Laurent Lesnard, Director of the *Centre de données socio-politiques* (CDSP), Senior Research Fellow at the National Centre for Scientific Research (CNRS) at the *Observatoire sociologique du changement* (OSC) at Sciences Po.

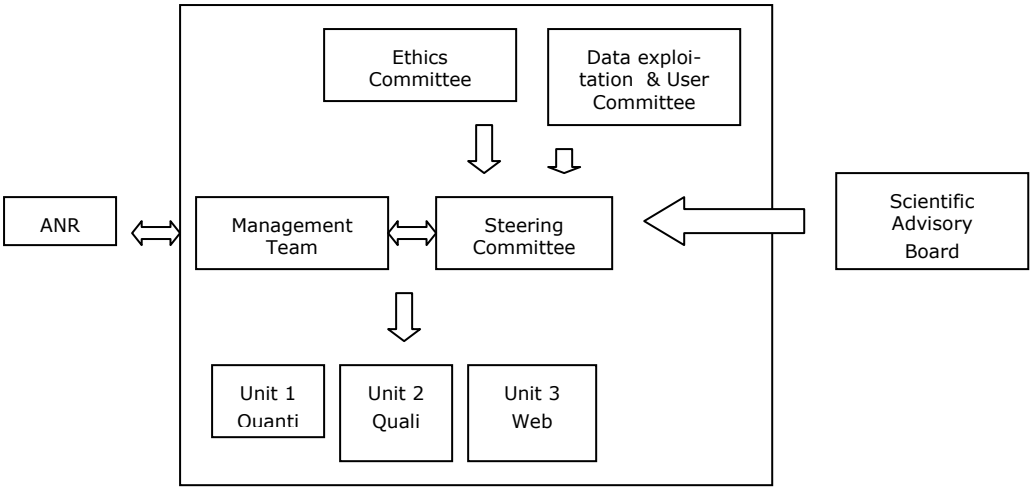
At the CDSP, Laurent Lesnard organises the DIM (Data, Infrastructures, Methods) seminar which aims to contribute to the development of methods for the social sciences in Sciences Po. He has also initiated and is supervising a series of innovative projects (question database and advanced research tools, interactive mapping of election results, etc). He is a member of the *Comité du label* (Quality Label Committee) of the French National Council for Statistical Information (CNIS) where the statistical quality of questionnaire surveys projects from government statistical departments and Insee is assessed.

Laurent Lesnard was involved in two of the social sciences infrastructures project pre-selected by the European Strategy Forum on Research Infrastructures (ESFRI). He represented Sciences Po during the European Social Survey (ESS) Infrastructure Preparatory phase which aimed at laying the groundwork for the ESS to become a European Research Infrastructure. He was also involved in the CESSDA preparatory phase project of which the objective was to develop the CESSDA as a European Research Infrastructure.

Laurent Lesnard will manage the scientific concerns of the DIME-SHS and will chair the Steering Committee. He will be assisted by an **Executive Director** to be recruited in charge of the everyday running of the project, in particular the administrative, legal and financial issues.

In addition to the Scientific Director and the Executive Director, the **Management Team** will be strengthened by: 1) a lawyer responsible for the elaboration and the implementation of a consortium agreement, the negotiation of grant agreements and contracts, the monitoring of all legal aspects related to the collection, dissemination and exploitation of data (first two years of the project); 2) a marketing manager, responsible for business development of DIME-SHS in France and abroad, 3) an administrative and financial manager, who will enhance the team once the equipment is set up (after the 2-year test stage). The Management Team will ensure the overall management of the project: ensuring the administrative, legal and financial management; administration and control of the budget (one contact per partner will be identified); overall responsibility for the relationships between the project partners and the ANR; project supervision and reinforcement of progress according to the time schedule; updating of the working plan and related budget with the agreement of the Steering Committee; preparation of the annual reports and final report; preparation of the meetings (decisions, agenda), drafting the minutes and implementation of decisions taken; ensuring exchange of relevant information received in relation to the project, between the partners, as well as information regarding the results achieved within it; development of the activity of DIME-SHS.

4.1.2 COORDINATION MODALITIES



Management Team

See above.

Steering Committee

The Steering Committee is the decision-making body of DIME-SHS. It supervises the methodological orientations as well as the functioning of the equipment. It assesses the collection requests and the research projects, modifies the work plan and the budget on the basis of the UEC and SAB's opinions. The Steering Committee is composed of: one representative of each partner and one representative from Insee and the leaders of the units. The three units, one for each component of the equipment, are in charge of the proper implementation of the project.

Scientific Advisory Board

The Scientific Advisory Board (SAB) advises the Steering Committee and the Management Team about the methodological orientations, the functioning of DIME-SHS, the projects of future development and its scientific results. The SAB is composed of internationally recognised experts.

Ethics Committee

For the setting up of the Ethics Committee, the *Comité de concertation pour les données en SHS* (CCDSHS) will be consulted. The Ethics Committee ensures that the ethical rules and the anonymity of the data collected and distributed by DIME-SHS are respected. Maintaining the anonymity of the participants in the web panel is a very important issue and will be closely monitored by the Ethics Committee in relation with the CNIL. The fact that the qualitative studies will be rendered anonymous raises many questions, of which the Ethics Committee will have to validate the answers.

Data exploitation and User Committee

The Data exploitation and User Committee will be in charge of estimating the effects of the inquiries related to data collection and of the projects proposed to DIME-SHS. The analysis will rely on a study of their potential in terms of exploitation of the results (see §3 below) and feedback from DIME-SHS users. It will incorporate, among others, representatives of firms interested in these questions.

Rules for use of equipment

The services proposed by the structure are available to any person, team, administration, local government body, firm, association etc, whatever its geographical location, provided that the request has scientific merit, as well as merit for the exploitation of the results and that it is adapted to the equipment. Requests to use the DIME-SHS structure will be selected by the Steering Committee, following the recommendations of the SAB and the UEC, regardless of the affiliation of the holder (public or private). The fee for these services will be determined in the test phase of the equipment. It should cover the research expenses, the functioning of the equipment and depreciation expenses. The activities, services and selection criteria of the projects will be presented on the website of DIME-SHS.

A consortium agreement governing the shared rights and obligations of the partners will be implemented during the first stage of the project. The issues of governance, management of the budget or intellectual propriety will be laid out in this agreement.

4.2. COLLABORATION ORGANIZATION

4.2.1 PARTNERS DESCRIPTION, RELEVANCE AND COMPLEMENTARITY

Partner 1 : **Fondation Nationale des Sciences Politiques (Sciences Po)**

The *Fondation Nationale des Sciences Politiques* (Sciences Po) is a fully-fledged, self-governing research university specialised in the social-economic sciences and the humanities which enrols some 9,600 students per year, including 40% of foreign students from more than 40 countries. Sciences Po is the **leading research university in the social sciences in France** with 50 full-time professors, 190 researchers, 80 foreign professors invited each year and 300 academic partnerships with universities around the world**.** Sciences Po is home to a doctoral school offering 7 graduate programmes. Based on a **multi-disciplinary** approach, it associates and combines skills and know-how from the different **social sciences** – in particular economy, history, political science and sociology and the humanities.

With a million-volume collection the library is the richest in humanities and social sciences in continental Europe. The active publishing house with more than 1,000 titles in its catalogue and six academic journals, *Les Presses de Sciences Po*, focuses on disseminating research results of the nine research centres, five of which are closely linked to the *Centre National de la Recherche Scientifique* (CNRS). Sciences Po facilitates laboratories' participation in the European Research Area. Today, 15 of the research teams of Sciences Po have been involved in more than 35 European projects and 18 ANR projects.

Research centres involved

Centre de Données Socio-Politiques (CDSP), Médialab, Centre d'Etudes Européennes de Sciences Po (CEE), Centre de Sociologie des Organisations (CSO), Observatoire Sociologique du Changement (OSC)

Partner 2: **Groupe des Ecoles Nationales d'Economie et de Statistique (GENES)**

The Groupe des Ecoles Nationales d'Economie et de Statistique (GENES) includes :

- Two professional engineers schools: the Ecole Nationale de la Statistique et de l'Administration Economique (ENSAE), which is going to be relocated on the Saclay campus, and the Ecole Nationale de la Statistique et de l'Analyse de l'Information(ENSAI), located on the European Brittany Campus.
- A professional training centre, the CEPE.
- A research centre, the CREST including nine laboratories:
  - a public policies evaluation research department;
  - an industrial economy laboratory;
  - a finance laboratory;
  - a macroeconomy laboratory;
  - a micro-econometrics laboratory;
  - a quantitative sociology laboratory (LSQ);
  - a statistics laboratory;
  - a survey methodology laboratory;
  - a statistical engineering laboratory.

About 100 scholars are permanent or associate members of the CREST.

The GENES teaches about 600 students every year. These students mainly become professionals and executives in banks, insurance companies, international organizations or the French public sector.

Strong links have been developed with the London School of Economics, Oxford, Cambridge, Imperial College, Pompeu Fabra, University of Mannheim, University of Humboldt, Columbia, NYU, Chicago, Harvard, Princeton, Cornell, Stanford.

Research centres involved

The quantitative sociology laboratory (LSQ) of CREST

Partner 3: **Institut national des études démographiques (Ined)**

The Ined is a public scientific and technological establishment (EPST) under the joint supervision of the Ministry of Education and Research and the Ministry for Employment, Labour and Social Cohesion. Founded in 1945, it was the first institute for demographic studies in Europe. The role of the Ined is to study demographic problems in all their aspects, by collecting the relevant documentation, conducting surveys and carrying out experiments or following experiments conducted abroad. The Ined adopts a multidisciplinary approach, and its areas of expertise cover not only the study of purely demographic phenomena such as nuptiality, fertility, mortality and migration, but also demography applied to society, economics, public health, human geography, history and more. It comprises 200 persons, including 60 tenured researchers, 110 technicians and engineers, plus around 30 PhD students and associate researchers. INED comprises 11 research units and several departments (surveys, statistical methods, library and documentation, publications, international comparisons, etc.). It was rated A+ by the AERES in 2009. INED is a member of CESSDA (Council of European Social Science Data Archives), and maintains close relations with the ICSPR (Inter-University Consortium for Political and Social Research), an international consortium of about 700 academic institutions and research organizations, which provides leadership and training in data access, curation, and methods of analysis for the social science research community.

Research centres involved:
The department of sampling design and surveys.

Partner 4 : **Université Paris Descartes**

With its 9 Units of Formation and Research (UFR) and its IUT (Institut Universitaire de Technologie), the University Paris Descartes cover the whole of knowledge in social sciences and health. The only university from the Île-de-France joining together medicine, pharmacy and dental, its pole health is recognized in Europe and in the whole world for the quality of its formations and the excellence of its research. The Paris Descartes University enrolls more than 33000 students each year.

The Faculty of Human and Social Sciences - Sorbonne is one of the 9 Units of formation and Research. With more than 3000 students, the Faculty is made up of three departments of teaching (sociology and ethnology, sciences of education, sciences of languages) and 10 research laboratories. It is one of the main Faculties in human and social sciences in Paris.

The PhD Program in Humanities and Social Sciences enrolls 300 PhD candidate each year and delivers about 70 Phd in Sociology, Linguistic, Demography and Educational Sciences each year.

The international dimension is one of the priorities of the CERLIS. Around 60% of the activities of the CERLIS are concentrated on Europe, especially in Italy with numerous conferences and books. It has also developed strong links with Brazil, Canada and Japan and Israel.

Research centres involved
Centre de Recherche sur les Liens Sociaux (CERLIS)

Partner 5 : **Telecom ParisTech**

Telecom ParisTech is one of France's leading graduate engineering schools and is considered the leading school in the field of Information Technologies. Its disciplines include all the sciences and techniques that fall within the term "Information and Communications": Computer Science, Networks, Communications, Economics and Business, Electronics, Signal and Image Processing, Social and Human Sciences as well as the study of economic and social aspects associated with modern technology.TELECOM ParisTech scientific policy is in phase with the rapid evolution of the sector of Information Technologies and the central role played by networks and information systems today, internet and the web and many of their areas of application (distance learning, health, environment, etc.)

Over the last few decades, TELECOM ParisTech has gained a recognized position internationally in its technical domains. In addition to working on the development in the basic sciences, research at TELECOM ParisTech also aims at broadening its field of study, moving in particular towards system integration, innovative services on the internet and in other media, and the analysis of users communication practices and their social impact, studies on TELECOM ParisTech thus covers all aspects of Information Technologies. Research projects are organized principally in collaboration with universities and major research groups, notably the CNRS (Centre National pour la Recherche Scientifique). Applied research is pursued through contractual agreements between TELECOM ParisTech and its corporate partners, such as France Télécom R&D, STM, Alcatel –Lucent, SFR, SIEMENS, with which TELECOM ParisTech has long established ties, as well as with other telecom operators and manufacturers of telecommunications equipment. TELECOM ParisTech is active in technological licensing and patents for work done in its laboratories.

TELECOM ParisTech has agreements for academic cooperation with partners all over the world. In Europe, these are closely connected in exchange networks and various kinds of programs (ERASMUS, SOCRATES, ATHENS). TELECOM ParisTech's European partners are the universities of : Berlin -TU, Aachen, Stuttgart (Germany), Bristol, London Imperial College (Great Britain), Delft (Netherlands), Karlskrona (Sweden), Leuven - Louvain-la-Nouvelle (Belgium), Lausanne, Zurich (Switzerland), Linz (Austria), Madrid, Barcelona (Spain), Lisbon (Portugal), Turin (Italy), Trondheim (Norway).

Research centres involved
LTCI

Partner 6 : **EDF R&D**

The vocation of EDF R&D is to contribute to improving performance among EDF Group operating units and to identify and prepare new growth drivers for the medium and long terms.

EDF R&D has a committed policy of working with partners in France and Europe, especially the countries where the Group is active, as well as in other parts of the world. The ICAME department is one of the 16 R&D departments of the EDF group. Its vocation is to accompany the operations departments by proposing new services and material – and this for all clients, whether they are individuals, businesses, industrials or local government. New offers are developed in light of the observation of and insight into the clients in their different facets, whether they are consumers or citizens. To complete this task, the ICAME department has developed the following skills:

- skills in social and human sciences (sociology, anthropology and semiologiy). This is essentially where the knowledge of the client is based, around social trends, links between consumers and businesses, observation of environmental issues.

- technical skills for the conception and elaboration of service provisions in the objective of appropriating different technologies which will provide support in getting the service to the consumer.

- skills in integration that draw on the knowledge of the client and on the technology available to put together proposals offering both service and price.

- innovation is stimulated by skills devoted to increasing the capacity for innovation through creative methods and through the knowledge of different innovating sources.

- Finally, skills in statistics and information systems with the objective of studying the innovative tools and methods that tomorrow will facilitate access to and knowledge of client data in information systems.

The policy of partnerships is part of a co-development strategy, both with industrial partners such as Electric Power Research Institute (EPRI) (United States), The Energy Technologies Institute (Great Britain), KEMA (Netherlands), and with research bodies such as Manchester University (Great Britain), Karlsruhe Institute of Technology (Germany), Ecole Polytechnique Fédérale de Lausanne (Switzerland), Imperial College (London)

Research centres involved: GRETS Research Group on Energy, Technology and Society

Partner 7 : **GIS Réseau Quetelet**

The Réseau Quetelet, in cooperation with its partners, provides researchers from France and abroad interested in data treatment with the requisite access to databases in the following domains:

- major data, censuses and other databases of French National Statistics;
- major French research data;
- privileged access to international data

The Réseau Quetelet's partners are the CMH-ADISP (Centre Maurice Halbwachs Data Archives of National Statistics), the CDSP (Centre for Socio-political Data), and the INED data service (National Institute of Demographic Studies). It also includes 3 Universitaries Platforms at Lille, Lyon and Caen.

The Réseau Quetelet, moreover, is a member of CESSDA, the European network of databanks for research which is now in the process to become a European Research Infrastructure Consortium (ERIC) as a result of the ESFRI process.

The Réseau Quetelet is under the responsibility of the National Data Committee in Humanities and Social Sciences (CCDSHS), which itself is in charge of national data policy at the Ministry of Research, in partnership with other ministries. The Committee coordinates access to microdata and provides assistance for training in the use of data as well as in data-gathering and documenting research.

The Réseau Quetelet will change soon its status into a "Groupement d'Intérêt Scientifique" held by major universities and research institutions, EHESS (Ecole des Hautes Etudes en Sciences Sociales), CNRS (Centre national de la Recherche scientifique), INED, (Institut national d'études démographiques), Sciences Po (Fondation nationale des sciences politiques de l'Institut d'études politiques de Paris), PSE (Paris School of Economics), Université de Lille 1, Université de Caen-Basse Normandie, IRDES.

The Réseau Quetelet is a member of CESSDA (Council of European Social Science Data Archives). It also maintains close relations with the ICSPR (Inter-University Consortium for Political and Social Research)

Actors' complementarity

The DIME-SHS project gathers key actors of the "knowledge triangle", referring to research, education, and innovation, in the social sciences in France : universities (PRES Sorbonne Paris Cité, and particularly Sciences Po and Université Paris Descartes) and « grandes écoles » (GENES' ENSAE and ENSAI, and Telecom ParisTech), research institute (INED) and private research (EDF R&D).

Beside, the success of this project rests on the gathering of these various expertise in methodological research and survey methods described below, but also in the different fields of SHS, many of which are represented by the different partners : economics, political science, history, and sociology.

A unique combination of expertise

As shown in the table below, each partner brings to the project a unique combination of expertise in quantitative, qualitative, and web data, and in collecting, documenting, or disseminationg them. Gathering these partners gives rise to synergies to conduct each of these activities more effectively across the traditional research boundaries (university vs grande école, public vs private, professors vs researchers). Complementarities between them will also enable them to contribute to the international scientific debate on documenting and disseminating web data. Their very nature makes web data close to qualitative materials, yet their number requires quantitative methods. Because DIME-SHS gathers expertise on both qualitative and quantitative data, as well as web experts, it will be able to contribute to the emerging standards for documenting and disseminating web data.

The consortium built around DIME-SHS will allow everyone to learn from the expertise of the other, and to develop common synergies.

	Collection	Documentation	Dissemination
Quantitative	Sciences Po GENES INED	GIS Réseau Quetelet Sciences Po INED	GIS Réseau Quetelet
Qualitative	Sciences Po Paris Descartes INED	EDF R&D Sciences Po	EDF R&D Sciences Po
Web	Sciences Po Telecom ParisTech		

Pooling this technical and scientific knowledge will make it possible to build an innovative and cost-effective equipment providing services at the highest level of methodological excellence. As DIME-SHS services will be available for teaching, it will also improve students' training in survey methods and methodology as they will be able to learn them by practicing them under the supervision of the highly skilled academics gathered in this project.

This will stimulate innovation in the private sector. Indeed, collecting, analysing, and interpreting heterogeneous data is crucial for companies to understand customers behaviour, detect trend reversals, assess the quality of the services they provide, etc. French companies will have access to the latest methodological tools and methods, and as such will be able to innovate by adapting them so as to be more competitive.

DIME-SHS will also be used to conduct methodological research in order to improve the services offered but also to expand knowledge on survey research. Methodological research is crucial to maintain the equipment at the highest level of expertise and to provide services that will enable researchers to conduct efficient and well-conceived surveys. It should also increase academics' productivity as access to high quality data is of the utmost importance for publishing in the best journals. Another outcome is the stepping up of methods teaching

4.2.2 QUALIFICATION, ROLE AND INVOLVEMENT OF INDIVIDUAL PARTNERS

Partner	Surname	First name	Position	Domain	Organization or company	Contribution in the project (4 lines max.)
Sciences Po	Lesnard	Laurent	Senior research fellow	Sociology	CNRS	Coordinator of the project
	Cousteaux	Anne-Sophie	Research coordinator	Documentation Statistics Sociology	Education Nationale	Method team manager
	Cornilleau	Anne	Research coordinator	Documentation Statistics Sociology	FNSP	Panel team manager
	Duchesne	Sophie	Research director	Political science	CNRS	Scientific coordinator of BEQuali
	Tiberj	Vincent	Senior research fellow	Political, science and sociology	FNSP	Scientific collaborator of the research survey centre
	Boullier	Dominique	Professor	sociology	Education Nationale	Scientific coordination of web data collection
	Girard	Paul	Research engineer	engineering	FNSP	Technological coordination of web data collection
	Jacomy	Mathieu	Research engineer	Web studies	FNSP	Web data collection software development leader
GENES	Caveng	Rémy	Associate researcher	Sociology	Education Nationale	Operations director of the telephone survey centre
	Gollac	Michel	Director of the LSQ	Sociology	INSEE	Scientific coordinator of the call centre survey centre
	Goux	Dominique	Senior research fellow	Sociology	INSEE	Methodology advice for users of the call centre (sample, survey design)
	Vallet	Louis-André	Research professor	Sociology	CNRS	Methodology advice for users of the call centre (national and international comparisons)

SES-Ined	Legleye	Stéphane	Head	Methodology, panel	INED	Web panel: conception, logistics, statistics
	Razafindratsima	Nicolas	Engineer	Survey methodology	INED	Senior statistical adviser
Telecom ParisTech	Valérie	Beaudouin	Research Professor	InfoCom	CNRS	Research on web panel and web data collection
	Dana	Diminescu	Associate professor	Sociology	Education Nationale	Research on web data collection and analysis
	Ludovic	Lebart	Research Professor	Data Analyse	CNRS	Web panel methodology
EDF R&D	Mathieu	Brugidou	Senior Researcher	sociology	EDF	Research on qualitative archiving
	Benoit	Habert	Senior Researcher	Linguistic	EDF	Research on web data methods
	Chloé	Clavel	Engineer	Web studies	EDF	Research on web data methods

Partner: Fondation Nationale des Sciences Politiques (Sciences Po)

- **Research centre involved:** Centre de données socio-politiques (CDSP – Socio-Political Data Center)

The CDSP is a joint unit of Sciences Po and CNRS (UMS 828) created in 2005. The main activities of the centre are to distribute well-documented data to researchers in social sciences, to contribute to the carrying out of international research projects and to support educational and training activities in the area of data collection and analysis.

The CDSP, as part of the French data archive network, the Réseau Quetelet, participates to the construction of a European research infrastructure for the circulation of data and metadata in social sciences.

Since the beginning, the CDSP takes as much as possible into account the development of Internet as daily tool of work and research. Thus, exploration and distribution of data are offered online and open tools are developed by the centre in order to facilitate the work of researchers.

The expertise of the CDSP is more and more solicited by students and researchers for constitution and design of original databases. The centre is also involved in several research projects in which it is asked to develop innovative information systems.

Individual partners:

Laurent Lesnard

Trained in sociology (Université Paris Descartes and École Normale Supérieure de Cachan) and in statistics and economics (University of Cergy and École Nationale de la Statistique et de l'Administration Économique), Laurent Lesnard is the Director of the *Centre de données socio-politiques* (CDSP) and a Senior Research Fellow at the National Centre for Scientific Research (CNRS) at the *Observatoire sociologique du changement* (OSC) at Sciences Po. He is also a research associate at the Centre for Time Use Research of the University of Oxford. Before joining the CNRS and Sciences Po, Laurent Lesnard was a Senior Research Officer at the Institute for Social and Economic Research of the University of Essex (UK) where he conducted research on the transformations of everyday life brought about the new ICTs His research is both substantial (time use and life course) and methodological (optimal matching) and is published in major international peer-reviewed journals (American Journal of Sociology, Sociological Methods & Research, Journal of the Royal Statistical Society Series A).

Anne-Sophie Cousteaux

Anne-Sophie Cousteaux was trained in sociology, statistics and computing at the University Paris 4 and held a master's degree in sociology from Sciences Po where she is finishing her PhD thesis on the social construction of men and women's health (Cousteaux, Pan Ké Shon, 2008, 2010; Cousteaux 2010). She is research coordinator at the Centre de Données Socio-Politiques since 2008. Her main activities concern the documentation of survey data in accordance with the international DDI standard and their distribution to researchers and students. She is also involved in the conception of different database (a survey question database for example) and participates in the realisation and the promotion of the European Social Survey in France (Chanvril, Cousteaux, Le Hay et al., 2009). In her research work, she shows a great interest for methodological questions in sociology such as, for example, the validity of the French socio-occupational classification (Cousteaux, 2004), the definition of women's social position (Cousteaux, 2006) or the application of multivariate descriptive methods to open-ended survey questions (Cousteaux, 2010).

Anne Cornilleau

Anne Cornilleau was trained in sociology, statistics and computing at the University Paris5 and held a master's degree in sociology from Sciences Po. She is research coordinator at the Centre de Données Socio-Politiques since 2005. Her main activities concern the documentation of survey data in accordance with the international DDI standard and their distribution to researchers and students. By her expertise in quantitative data documentation, she supports the constitution of the qualitative survey bank BEQuali.

Helping researchers with the constitution of original databases is also part of her activities. In addition, she is involved in the CESSDA European infrastructure which consists in developing and promoting data access in social sciences. She participates also in the different stages of the realisation of the European Social Survey in France since the round 3 (2006) and contributes to its

promotion. She is also involved in Sciences Po master program as lecturer in introduction to statistics and methodology for social sciences.

The CDSP team will supervise scientifically DIME-SHS.

- **Research centre involved**: European Studies Centre (CEE, Centre d'études européennes de Sciences Po)

Founded in 2005, the Centre d'études européennes is one of the nine research centres of Sciences Po. Its scientific project rests on a multidisciplinary and transversal approach to policy and politics. Its principal fields of research unfold into three transversal axes: to understand the European construction and its impact, to grasp the policy/politics articulation, and to invest in methodological questions. It assembles some sixty researchers, both senior (FNSP, CNRS) and junior (doctoral, post-doctoral and research assistants), as well as about twenty associate researchers from France and abroad.

Individual partners:

Sophie Duchesne

Sophie Duchesne is Research Professor at the National Centre for Scientific research (CNRS) and works at Sciences Po. She works on political identities, more particularly national and European ones in a context of globalization. She recently coordinated a comparative research on attitudes towards European integration using collective interview, in collaboration with Oxford University members and colleagues from Louvain-la-Neuve. She has always been specifically interested in research methods, especially interview techniques and analyses. She teaches at Sciences Po and at Oxford.

Vincent Tiberj

Vincent Tiberj, is associated research professor at Sciences Po in the Centre for European Studies (CEE). He is specialized in electoral sociology and particularly focused on the social and political values change (Michelat et Tiberj 2007, Tiberj 2008) in France and in the US. His works relies heavily both on sociology and psychology. He has also focused on quantitative methodology issues, using regularly experiments in survey research. In this brand he was co-principal investigator of a survey aiming at mapping the political representations of immigrant-origin French (Brouard et Tiberj 2005, Brouard et Tiberj 2010) and also work with James Stimson on the importation in the French context of the American public policy mood. He also co-ordinates the "Parcours Avancé de Méthodes en Sciences Sociales", a teaching track in quantitative and qualitative methods at the Sciences Po doctoral school .

The CEE team will advise DIME-SHS on qualitative and quantitative data collection questions.

- **Research centre involved**: médialab

The médialab of Sciences Po is positioned at the crossover of social sciences and digital technologies. The médialab was created in May 2009 with the explicit goal of exploring how data and resources provided by information and communications technologies can be harnessed for the benefit of social sciences. The médialab is a site dedicated to digital research, a fertile ground for collaborations between social scientists, engineers and designers. It is a high-tech facility, a hub for vanguard research, a scientific toolkit at the disposal of the Sciences Po academic community, though it extends beyond this community in establishing a platform for launching national and international collaborations.

Individual partners:

Dominique Boullier

Doctor of sociology of the School of the advanced studies in social sciences (EHESS), Dominique Boullier is one of the best French specialists in the sociology of communication and information technologies. He was Professor of the Universities of the University Rennes 2 and Professor of the Universities at the University of technology of Compiegne (UTC). He was also the creator and director of the mixed unit of services of CNRS LUTIN (Laboratory of the Uses in Numerical Technologies of Information) in the "Cité des Sciences et de l'Industrie" of La Villette and

director of the Laboratory of Anthropology and Sociology (LAS) at the University Rennes 2. He is also a research entrepreneur who is particularly active and effective in making a success of theoretically ambitious scientific projects led in direct link with non-academic partners (companies, administrations and communities).

Paul Girard

Paul Girard was trained in Université de Technologie de Compiègne where he choose the cultural industry engineering specialisation and graduated as an Information Technology engineer in 2004. He is interested in collaborations between non-technical domains and R&D laboratories as a innovative way of thinking about relationships between digital technologies and society. He joins the médialab laboratory in Sciences Po as technical manager in March 2009. Since then he participates in collaborative research projects between social sciences researchers, médialab team and industrial and academic partners.

Mathieu Jacomy

Mathieu Jacomy is an engineer trained in Université de Technologie de Compiègne where he isalso a graduate in Sciences and technology of cognition and cooperation master. He is a founding member of the association Webatlas (2007) gathering engineers and researchers investigating the web as a research object. He is the creator and still participates in two open source software projects Navicrawler and Gephi. The combination of the two software applications is one of the most used methodological chains for web mapping projects. Mathieu Jacomy is therefore one of the most skilled and promising web studies scientists. After being in charge of the R&D team of TIC-Migrations research program in la Maison des Sciences de l'Homme in collaboration with Telecom ParisTech (2006-2010), he joins the médialab team in Sciences Po to lead and develop the web mining research activities.

The médialab team will manage all the operations related to analysing web data (extraction, cleaning, documentation, analysis, dissemination).

Partner 2: Groupe des Ecoles Nationales d'Economie et de Statistique (GENES)

Individual partners:

Rémy Caveng

Rémy Caveng is a maître de conférences at the Jules Vernes University at Amiens. His main research topic is the making and use of quantitative data. Before becoming an academic, he worked during ten years in several private survey institutes, where he was in charge of organizing and managing fieldwork. Further, as a sociologist, he has studied the employment, work and careers of survey interviewers. More specifically, he has studied how their employment status and working conditions influence the quality of the data they collect. In 2008-2009, he worked as an expert for INSEE to study the data collection of the French Labor Force Survey in order to improve its quality.

Michel Gollac

Michel Gollac is the director of the LSQ at CREST. His personal current research interests include work organization, working conditions, work satisfaction and health at work; wage inequalities and the subjective perception of wages. He is the coordinator of the ANR granted SalSa project, which includes an original survey on the last topic. He is also the chairman of an international interdisciplinary expert panel in charge of designing a statistical system (including a survey) on psychosocial risks at work. Before becoming a sociologist, Michel Gollac has been an official statistician and he conceived or directed several large surveys. He teaches sociological surveys design in the Sociology master of the Ecole des Hautes Etudes en Sciences Sociales. He is an advisor of the Demographic and Social Statistics Director at INSEE.

Dominique Goux

Dominique Goux is a senior research fellow at the National Institute for Statistics and Economic Surveys (Insee). She currently studies leisure spillovers within households, and the influence of election results on local social policies. She was previously the head of the employment statistics division at Insee, in charge of the French Labor Force Survey and other main national social surveys (education and qualification surveys, working conditions surveys, literacy surveys, etc.). She has been associate professor of applied econometrics for ten years at the Ecole Normale Supérieure and she is now teaching "survey methodology and design" in the Master Public Policy and Development in the Paris school of economics.

Louis-André Vallet

Louis-André Vallet is Research Professor in the French National Centre for Scientific Research (CNRS) and holds a PhD in sociology from the University of Paris-Sorbonne (1992). He works within the Quantitative Sociology Unit of the Centre for Research in Economics and Statistics (CREST). His research interests and publications mainly concern social stratification, intergenerational mobility and the sociology of education with corresponding knowledge of large surveys such as the Labour Force Surveys, the Education, Training and Occupations Surveys and the National Education Panel Studies. He also teaches a course on Sociology and Statistical Survey within the INSEE School (ENSAE). On the methodological side, he has expertise in statistical modelling of categorical variables. From 2003 onwards, he has been the French country representative in the Steering Committee of the "Quantitative Methods in the Social Sciences" programme, a large training programme of the European Science Foundation. He is currently a member of the editorial boards of *Revue Française de Sociologie, European Sociological Review* and *Social Forces*.

The GENES team will be in charge of the telephone call centre: management of the platform and methodological advices (survey design, sample, questions wording to ensure comparability, etc.)

Partner 3: **Institut national des études démographiques (Ined)**

- **Research centre involved:** Service des enquêtes et des sondages

The department of sampling design and surveys (17 members) provides help and assistance to the researchers for their research projects. All methodological aspects are concerned: sampling design, questionnaire building, data collection mode, logistics, weighting and quality measurement. The collaboration will provide many opportunities to develop our expertise in methodology and facilitate some research projects involving general population surveys. The department was a pioneer actor in the field of sensitive subjects and hard to reach populations in France (homelessness, drug abuse, sexuality, mixed data collection modes, etc.). The department has competences in statistics and logistics and data archiving

Individual partners:

Stéphane Legleye

Stephane Legleye is the head of the department since 2009. He is a statistician of the National Institute for Statistics and Economic Studies (INSEE) and epidemiologist. He has worked in the field of drug use and abuse monitoring for ten years at the French monitoring centre for drug and drug addiction. He is also an associated researcher at the National institute for medical research (Inserm). He built up and analysed many quantitative surveys in general population as well as among homeless people, dealing with mental health, lifestyle, tobacco, alcohol, drug use. His competences comprise survey methodology and statistical methods. He is the principal investigator of some major national or cross-national surveys on drug use in France and in Europe.

Nicolas Razafindratsima

Nicolas Razafindratsima is a former student of the National School for Statistics and Economic Studies (ENSAE) and has a PhD in demography. He is statistician in the department since ten years. His competences comprise survey sampling design, weighting and quality measurement as well as survey methodology, demographic and economic analyses. He worked on many large and innovative surveys such as Study on family history, Migrations between Africa and Europe (MAFE), Sexuality in France (CSF), and Contraception cohort. He is a member of the National Council for Statistical Information (CNIS). He published many articles on survey methodology and was recently the associate editor of a special issue on hard to reach population of the journal « Methodological innovations online ».

The INED team will focus on managing the panel and conducting methodological research to ensure the highest methodological quality for users.

Partner 4: **University Paris Descartes**

- **Research centre involved:** Centre de recherché sur les liens sociaux (CERLIS)

The CERLIS laboratory carries out quantitative and qualitative investigations recognized. Several sociologists, trained with the techniques of investigation by questionnaires and the statistics, lead research to it on the family, education, the modern design of the individuals, work and health. The laboratory ensures the training of students of level Master and PhD. The laboratory is closely associated with training in "Sociologie d'enquête" (Master in sociology): the students of this Master are trained with the quantitative and qualitative advanced methods in sociology.

Individual partners:

Olivier Martin

Head of PhD Program in Humanities and Social Sciences (Université Paris Descartes), Head of Master "Sociologie d'enquête", Olivier Martin is full professor of sociology and statistics at the Faculté of Social Sciences – Sorbonne (Université Paris Descartes). He graduated from ENSAE (1990), as Statistician and economist, and he received a PhD in Mathematics for Human Sciences from EHESS (1996). His PhD treated the roles of mathematics and statistics in the development of human sciences during the first half of the twentieth century. He is specialized in the conception and the analysis of quantitative studies. He has published several quantitative studies in sociology, several articles dealing with the epistemology of statistics and a textbook in statistics for sociology. He co-ordinates the teaching of statistical and probabilistic methods in Licence, Master and PhD Programs. He built up and analysed many quantitative surveys. His competences comprise survey methodology and statistical methods.

François de Singly

François de Singly is head of CERLIS (Centre de recherché sur les liens sociaux) and full professor of sociology at the Faculté of Social Sciences – Sorbonne (Université Paris Descartes). He works in the field of sociology of family and private relations. He has been building up and analysing many quantitative surveys for 30 years (INSEE, INED, CNAF…). He has published several quantitative and qualitative studies in sociology and a well-known textbook about the conception of surveys in sociology.

Christophe Giraud

He graduated from ENSAE (1993), as Statistician and economist, and received a PhD in Sociology from Paris Descartes University (2001). He is the head of the Statistical Observatory of students and "Maître de conferences" in sociology at the Faculté of Social Sciences – Sorbonne (Université Paris Descartes). His competences comprise survey methodology and statistical methods. He is teaching statistics and quantitative methods. He participates in the teaching program in statistical and probabilistic methods (Licence, Master and PhD Program).

Muriel Letrait

Muriel Letrait is a statistician. She works as Research Engineer (CERLIS). Her competences comprise survey methodology and statistical methods in social sciences.

The Université Paris Descartes team will bring its expertise in survey methodology. It will also develop new teaching methods based on the services provided by DIME-SHS.

Partner 5: Telecom ParisTech

- **Research centre involved:** SES

The Social Science Department, whose specialization is Information and Communication Technologies, is one of the Telecom ParisTech laboratories. Ethnography of everyday activities has become a major area of expertise of the Department, thanks to the renewal of its qualitative studies: activities are now recorded by different means and interviews are considered as a moment during which the interviewees are confronted with the recorded traces. Some of the Department researchers have also been involved in the conception of collection and analysis devices for Web data and in the conception of tools for measuring Internet uses. At last, the Department has the strong experience in quantitative research, with two statisticians who have worked for a long time on the methodology of the studies at the CREDOC (Research Center for the Study and the Observation of Living Conditions).

Individual partners:

Valérie Beaudoin

Valérie Beaudouin is professor at TP since 2008 in the Economics and Social Science department. She graduated from ENSAE (1991), as Statistician and economist, and she received a PHD in Linguistics from EHESS (2000). From 2003 to 2008, she headed at Orange R&D the Lab of economics and social sciences, setting up new, transdisciplinary techniques for large scale analysis of computer mediated communication and digital transactions. She specialised in the analysis of internet uses and digital sociability, in private as well as in professional contexts.

Dana Diminescu

Dana Diminescu is associate professor at TP since 2008 in the Economics and Social Science department, scientific director of the Research Program on the Use of ICT in Migrations (FMSH, Paris) and project coordinator of the "E-diasporas-atlas" project funded by the ANR. Her field of research: communication practices and uses in mobility situation, uses of Internet (tailing, archiving, mapping of the Web).

Ludovic Lebart

Ludovic Lebart is a former Directeur de Recherche (DR1) at the Centre National de la Recherche Scientifique (CNRS) and he is since 1989 Associate Professor at Télécom-ParisTech, Paris, France. Doctorat d'Etat in Mathematics (University Pierre et Marie Curie, Paris, 1976), Ingenieur civil des Mines (1964), Elected member of the International Statistical Institute (1979). From 1978 to 1988: Deputy Director of CREDOC. Dr Lebart works concern the analysis of qualitative and textual data, the methodology of survey design and processing. He has published several books and articles about Data Analysis and Survey Methodology and co-authored the software SPAD. For a list of publications (books, chapters of books, editorships, articles) see the website: www.lebart.org.

Partner 6: **EDF R&D**

Individual partners:

Mathieu Brugidou

Mathieu Brugidou is senior researcher at the EDF Research and Developpement and was Associate senior researcher at PACTE, a CNRS social science research unit. He received a Ph.D and habilitation (HDR) in political science from Sorbone-Paris-I University in 1992. His research interest include analysis of political discourse and public opinion and methods of analysis of textuel data applied in social science (non-directive interview and open-ended question). His books include *L'opinion et ses publics* (2008) and *L'élection présidentiel : enjeux et discours* (1995).

Benoît Habert

Benoît Habert is Professor of Computational Linguistics at the École Normale Supérieure in Lyons. He is currently working at EDF Research and Development, where he leads a project in digital archiving. He was deputy-head of the TGE Adonis, where he was responsible for a pilot project for long-term preservation of spoken data. Part of his research was devoted to building and using digital corpora and to the role of 'instruments' in/for SSH. His publications include: (1998) *De l'écrit au numérique: constituer, normaliser, exploiter les corpus électroniques*, with C. Fabre & F. Issac (Paris : InterEditions/Masson); (2005) *Instruments et ressources électroniques pour le français* (Paris: Ophrys); (2009) *Construire des bases de données pour le français* (Paris: Ophrys); see also: http://www.revue-texto.net/Corpus/Publications/Habert/Habert_Portrait.html.

Chloé Clavel

Chloé Clavel is a research engineer currently working at EDF Research and Development. She defended her thesis in 2007 on the acoustic analysis of emotional manifestations. She works now on audio and text mining for the analysis of various data (satisfaction survey, blog, forum, call-centre data, etc.).

5. FINANCIAL ASSESSMENT

Budget of the project

The budget distinguishes 6 different types of costs (mixing fixed and variable costs):

- Hardware & Software (including hardware, software, licences, smartphones and cell-phone plans for the collection of quantitative data)
- Information system running costs
- Office rentals (necessity of installing new offices because of the project)
- Service offers for communication and market studies
- Workforce
- And overheads (environment costs are overheads taken in charge by the partners)

This budget reflects the very specific nature of the project which aims to develop an equipment based on innovative technologies and massive workforce to collect, to create, to archive and to disseminate data. For the 10 years of the project, and according the three different steps of its work plan (testing, development and operation), the budget will evolve as presented below:

DIME-SHS - Budget

TEST	DEVELOPMENT	OPERATION

	2011-2012	2013-2018	2019-2020	TOTAL
Equipment	1 723 246,40 €	3 314 272,04 €	742 246,40 €	5 779 764,84 €
Information system running costs	152 100,00 €	465 533,95 €	172 367,89 €	790 001,84 €
Office rentals	80 000,00 €	340 000,00 €	120 000,00 €	540 000,00 €
Service offers	30 000,00 €	50 000,00 €	- €	80 000,00 €
Financed workforce	804 266,93 €	3 569 327,42 €	1 314 874,81 €	5 688 469,16 €
Environment cost	241 280,08 €	1 070 798,23 €	394 462,44 €	1 706 540,75 €
Workforce to finance	2 804 115,30 €	11 377 379,19 €	4 051 112,86 €	18 232 607,34 €
Overheads	6 084,00 €	15 162,44 €	- €	21 246,44 €
TOTAL	**5 841 092,71 €**	**20 202 473,26 €**	**6 795 064,40 €**	**32 838 630,37 €**

The target panel is set at 15,000 people for the quantitative data collection. This size is necessary for a useful and operating equipment. A large investment in hardware devices will be done to reach this goal. Detailed information on this topic is given in the § 6.3. (financial detailed document).

A new facility will be set up in the heart of Paris to install part of the equipment. Another facility will be hosted in Saclay for telephone surveys. This explains the amount presented for the Office Rentals.

4286,4 person/months will be dedicated to the project during the first 10 years (an average of more than 35 people involved per year), which 38% will be funded by the partners.

The balance of the budget between each partner is, by nature, not equal, as presented below. The main reason is that the core of the equipment will be based at Sciences Po. The others partners will benefit of it, will reinforce the part of their own equipment linked to the project and will contribute to the project by their expertise and knowledge of high value, with an also massive workforce dedicated to build the equipment.

Budget / partner

	Total
Part1: Sciences Po	23 306 010,92 €
Part2: GENES	1 541 231,77 €
Part 3: INED	6 178 503,48 €
Part4: CERES	740 816,00 €
Part5: Telecom Paris Tech	803 898,86 €
Part6: EDF R&D	- €
Part7; Quetelet	268 169,34 €
Total	32 838 630,37 €

Financing of the project

To finance this project, 3 sources will be used: self-financing progression (turn-over generated by the equipment), partner own-financing and the EQUIPEX subsidy.

The main objective of this project is to build an equipment that will serve the research community for a long period of time. That is why a turn-over progression will be generated by the services. The turn-over progression is based on the following hypothesis.

To build self-financing progression, 2 low hypotheses have been done:

- First, the average price of 55 k€ / study is low compared to the prices of the market. The reason is that it is not yet possible to know an average price of the services sold (to be determined in the future). Nevertheless, the average price for a private study is now by 150 k€. So there is a still reasonable error margin.
- Last, the self-financing progression ratio is lower than the turn-over/budget ratio to minimize risk.

In the first phase of the project (called TEST), there is no possible turn-over. So it is the self-financing progression.

In the second phase (DEVELOPMENT), the turn-over will start. The required self-financing progression hypothesis is presented below:

Self-financing progression (DEVELOPMENT)

	2013	2014	2015	2016	2017	2018
Average cost of a study	50 000,00 €	50 000,00 €	50 000,00 €	50 000,00 €	50 000,00 €	50 000,00 €
Number of study done (capacity)	*15*	*20*	*30*	*40*	*50*	*60*
Number of invoiced study (low hyp.)	*7*	*15*	*25*	*35*	*45*	*55*
Turn-over	*385 000,00 €*	*825 000,00 €*	*1 375 000,00 €*	*1 925 000,00 €*	*2 475 000,00 €*	*3 025 000,00 €*
Turn-over/Budget	13%	25%	41%	57%	73%	80%
Self-financing progression hypothesis	0%	0%	5%	10%	30%	60%

At last, in the ultimate phase (2 years OPERATION), it is assumed that the equipment will be self-financed.

As for the own-financing contributions of each partner, it is detailed below. It represents almost half the cost of the Budget.

TEST	DEVELOPMENT	OPERATION

	2011-2012	2013-2018	2019-2020	TOTAL	TOTAL (%)
Part1: Sciences Po	555 097,14€	2 080 420,02€	744 502,89€	3 380 020,05€	44,92%
Part2: GENES	0,00€	896 044,84€	375 385,62€	1 271 430,46€	16,90%
Part 3: INED	200 213,55€	637 686,53€	225 539,08€	1 063 439,16€	14,13%
Part4: CERES	139 073,45€	442 953,35€	156 665,21€	738 692,00€	9,82%
Part5: Telecom Paris Tech	151 162,87€	481 458,54€	170 283,85€	802 905,26€	10,67%
Part6: EDF R&D	0,00€	0,00€	0,00€	0,00€	0,00%
Part7: Quetelet	0,00€	198 103,46€	70 065,89€	268 169,34€	3,56%
Total	**1 045 547,01€**	**4 736 666,74€**	**1 742 442,54€**	**7 524 656,28€**	100,00%

As for the own-financing contributions of each partner, it is detailed below. It represents almost half the cost of the Budget.

Partners' financing

	Total
Part1: Sciences Po	9 426 142,43 €
Part2: GENES	1 391 442,04 €
Part 3: INED	2 847 226,02 €
Part4: CERES	739 435,40 €
Part5: Telecom Paris Tech	803 253,02 €
Part6: EDF R&D	- €
Part7; Quetelet	268 169,34 €
Total	**15 475 668,26 €**

The EQUIPEX subsidy is the last financing source of the project, necessary to launch the project. It is detailed bellow:

DIME-SHS - Subsidy base

TEST	DEVELOPMENT	OPERATION

	2011-2012	2013-2018	2019-2020	TOTAL
Workforce to finance	2 804 115,30 €	9 316 493,76 €		12 120 609,06 €
Software and licences	131 050,00 €	180 637,50 €		311 687,50 €
Service offers	30 000,00 €	39 000,00 €		69 000,00 €
Office rentals	80 000,00 €	198 000,00 €		278 000,00 €
Hardware	402 196,40 €	239 486,63 €		641 683,03 €
Mobile Internet equipment	1 190 000,00 €	2 199 575,00 €		3 389 575,00 €

TOTAL INVESTISSEMENT (A+B+C)	**4 637 361,70 €**	**12 173 192,89 €**		**16 810 554,59 €**

Information system running costs	152 100,00 €	379 061,08 €		531 161,08 €
Overheads	6 084,00 €	15 162,44 €		21 246,44 €
TOTAL FONCTIONNEMENT (D+E)	**158 184,00 €**	**394 223,52 €**	**- €**	**552 407,52 €**

TOTAL GENERAL (A+B+C+D+E)	**4 795 545,70 €**	**12 567 416,41 €**	**- €**	**17 362 962,11 €**

The distribution of costs is based on the financial rules of the Equipex call.
The total investment cost is 16,810,554.59 euros, and the operating costs are 552,407.52 euros. The subsidy required is 17,362,962.11 euros in total. A part is dedicated to the partners, as described in the § 6.3. The figures are indicated in the file's coordinator into the document A, according the rules of the call.

To ensure the validity of the business plan developed, a development strategy has been established at international level.

DIME-SHS - Distribution of the financing

	TEST	DEVELOPMENT	OPERATION	
	2011-2012	**2013-2018**	**2019-2020**	**TOTAL**
Self-financing	- €	2 898 390,11 €	5 052 621,86 €	7 951 011,98 €
Self-financing (%)	0,00%	14,35%	74,36%	24,21%
Partner own-financing	1 045 547,01 €	4 736 666,73 €	1 742 442,54 €	7 524 656,28 €
Partner own-financing (%)	17,90%	23,45%	25,64%	22,91%
EQUIPEX subsidy	4 795 545,70 €	12 567 416,41 €	- €	17 362 962,11 €
EQUIPEX subsidy (%)	82,10%	62,21%	0,00%	52,87%
TOTAL BUDGET	**5 841 092,71 €**	**20 202 473,26 €**	**6 795 064,40 €**	**32 838 630,37 €**
TOTAL BUDGET (%)	100,00%	100,00%	100,00%	100,00%

To ensure the validity of the business plan developed, a development strategy described in §3 has been established. The follow up, adjustments and distribution of budget among partners will be defined according our management structure (§4), and will be established in the consortium agreement.

6. APPENDICES

6.1. LITERATURE REVIEW REFERENCES

Billordo, L. 2005. 'Publishing in French Political Science Journals: an Inventory of Methods and Sub-fields', French Politics 3(2): 178–186.

Braibant, G. 1996. Les archives de France. Paris: La Documentation française.

Cardon D. et al. (2007). « The Strength of Weak Cooperation:an Attempt to Understand the Meaning of Web 2.0 ». MPRA Paper. Mars. http://mpra.ub.uni-muenchen.de/4581/.

Caveng, R. 2009. "Inversement des positions et ré-enchantement de l'interaction". Actes de la recherche en sciences sociales 178:88-97.

_____ . 2007. "La "liberté" des vacataires en instituts de sondages". Interrogations, 4:59-77.

Clavel, C. and G. Richard. 2010. Reconnaissance acoustique des émotions, in Systèmes d'Interaction Emotionnelle, chap. 5, Hermès.

Cornu, M. 2003. « Droit des biens culturels et des archives » (www.educnet.education.fr/chrgt/biensculturels.pdf).

Couper, M. P. 2000. "Web surveys: A review of issues and approaches." Public opinion quarterly. 64 (4):464.

Cribier, F. (avec Feller, E.). 2003. Projet de conservation des données qualitatives des sciences sociales recueillies en France auprès de la « société civile », Rapport présenté au ministère délégué à la Recherche et aux nouvelles technologies, Paris, CNRS-EHESS, LASMAS.

Dargentas, M., M. Brugidou, D. Le Roux D and A-C. Salomon (eds.). forthcoming. L'analyse secondaire en recherche qualitative : une nouvelle pratique en sciences humaines et sociales. Paris: Lavoisier.

Dargentas, M. and D. Le Roux. 2005. Potentials and Limits of Secondary Analysis in a Specific Applied Context: The Case of EDF-Verbatim. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 6 (1). Article en ligne : http://www.qualitative-research.net/fqs-texte/1-05/05-1-40-e.htm

Demazière, D., C. Dubar, F. Leimdorfer, M. Reinert, and P. Wald. 2001. Rapport sur la faisabilité d'une banque de données qualitatives en sciences humaines et sociales, Laboratoire Printemp, Université Versailles Saint-Quentin.

Descamps, F. 2007. « L'entretien de recherche en histoire : statut juridique, contraintes et règles d'utilisation ». *Histoire@poltiique. Politique, culture et société*. 3.

Duchesne, S. forthcoming, « Développement de l'analyse secondaire et des méthodes d'analyse qualitative : une chance à saisir ? », in *L'analyse secondaire en recherche qualitative : une nouvelle pratique en sciences humaines et sociales*, Dargentas et alii eds. Paris: Lavoisier.

Duclert, V. 2002. « Archives orales et recherche contemporaine. Une histoire en cours », *Sociétés et représentations*. 13: 69-86.

Duclos, D. and H. Y. Meynaud. 1996. Les sondages d'opinion. Paris: La Découverte.

Finch, C. E., J. Vaupel and K. Kinsella. 2001. Cells and Surveys: Should Biological Measures Be Included in Social Science Research? Washington: National Academies Press.

Freese, J. 2007. "Replication Standards for Quantitative Social Science: Why Not Sociology?" Sociological Methods & Research. 36 (2):153. « Sciences sociales : archives de la recherche ». 2006. Genèses. 63.

Kaufmann, J-C. 1996. L'entretien compréhensif. Paris, Nathan, coll.128.

King, G. 2007. "An introduction to the Dataverse Network as an infrastructure for data sharing." Sociological Methods & Research. 36 (2):173.

Laflamme, F. and C. Mohl. 2007. Research and Responsive Design Options for Survey Data Collection at Statistics Canada: American Statistical Association.

Latour, B. 2005. Reassembling the social: An introduction to actor-network-theory. New York: Oxford University Press.

Lazer, D. and al. 2009. "Computational Social Science," *Science*. 323(5915): 721-723.

Le Hay, V. 2009. « Le panel électoral française 2007 : enjeux de méthode », Bruno Cautrès et Anne Muxel (dir.), /Comment les électeurs font-ils leur choix ? : le panel électoral français 2007/, Paris, Presses de la Fondation nationale des sciences politiques, pp. 259-284

Leskovec J. et al. 2009. «Meme-tracking and the dynamics of the news cycle ». in Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, 497-506. Paris, France: ACM. doi:10.1145/1557019.1557077. http://portal.acm.org/citation.cfm?id=1557019.1557077.

Lynch, C. A. 2002. "Digital collections, digital libraries and the digitization of cultural heritage information. "*First Monday*, 7(5). Retrieved from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/949/870 on 12 January 2010.

Mack, S., V. Huggins, D. Keathley and M. Sundukchi. 1998. Do Monetary Incentives Improve Response Rates in the Survey of Income and Program Participation? Presentation. Proceedings of the Section on Survey Research Methods, American Statistical Association.

Mallet-Poujol, N. 2004. « Protection de la vie privée et des données personnelles ». Legamedia. (www.educnet.education.fr/chrgt/guideViePrivee.pdf).

Moravcsik, A. 2010. "Active Citation: A Precondition for Replicable Qualitative Research". Political science & politics. 43(1): 29-35

Negroponte, N. 1996. Being digital. New York: Vintage Books.

Pagès, F., M. Tribel and A. Bonis-Charancle. 2006. "Indélicatesse et manque de rigueur dans les sondages. Enquête auprès des professionnels du sondage." Bulletin de Méthodologie Sociologique. 89:31-48.

Park, HW. 2003. « Hyperlink network analysis: A new method for the study of social structure on the web » - Connections, 2003

Presner, T. and J. Schnapp. 2009. "Digital Humanities Manifesto 2.0." UCLA. Retrieved from http://manifesto.humanities.ucla.edu/2009/05/29/the-digital-humanities-manifesto- 20/ on 16 January 2010.

Raento, M., A. Oulasvirta and N. Eagle. 2009. "Smartphones: An Emerging Tool for Social Scientists." Sociological Methods & Research. 37 (3):426-454.

Rheingold, H. 2000. The Virtual Community. Cambridge, Mass.: MIT Press.

Rogers, R. 2004. Information Politics on the Web, Cambridge, Mass: MIT Press.

Savage, M. 2010. Identities and Social Change in Britain since 1940: The Politics of Method. Oxford: Oxford University Press.

Silberman, R. 1999. Les sciences sociales et leurs données, http://www.education.gouv.fr/cid1925/les-sciences-sociales-et-leurs-donnees.html.

Singer, E. and R. Kulka. 2002. "Paying respondents for survey participation." Pp. 105-127 in Studies of welfare population: data collection and research issues, Studies of welfare populations: data collection and research issues, edited by Michele Ver Ploeg, Robert A. Moffit, and Constance F. Citro. Washington: National Academies Press.

Toepoel, V., C. Vis, M. Das and A. Van Soest. 2009. "Design of Web Questionnaires: An Information-Processing Perspective for the Effect of Response Categories." Sociological Methods & Research. 37 (3):371-392.

Turkle, S. 1995. Life on the screen: Identity in the age of the Internet. New York: Simon & Schuster

Venturini, T. 2009a. "Diving in Magma: How to explore controversies with Actor-Network Theory," Public Understanding of Science, forthcoming.

6.2. PARTNERS' REFERENCES

- *Fondation Nationale des Sciences Politiques (Sciences Po)*

Centre de données socio-politiques

Lesnard, L. and Kan M-L. Forthcoming. "Investigating Scheduling of Work: A Two-Stage Optimal Matching Analysis of Workdays and Workweeks." Journal of the Royal Statistical Society Series A.

Lesnard, L. 2010. "Cost Setting in Optimal Matching to Uncover Contemporaneous Socio-Temporal Patterns." Sociological Methods and Research. 38 (3):389-419.

_____ . 2009. La famille désarticulée. Les nouvelles contraintes de l'emploi du temps. Paris: PUF.

Lesnard, L. and T. de Saint Pol. 2009. "Organisation du travail dans la semaine des individus et des couples actifs : le poids des déterminants économiques et sociaux." Économie et Statistique. (414):53-74.

Lesnard, L. and T. de Saint Pol. 2009. "Patterns of Workweek Schedules in France." Social Indicators Research. 93 (1):171-176.

Lesnard, L. 2008. "Off-Scheduling within Dual-Earner Couples: An Unequal and Negative Externality for Family Time." American Journal of Sociology. 114 (2):447-490.

Cousteaux, A.S. and J.L. Pan Ké Shon. 2010. « Is ill-being gendered? Suicide, risk for suicide, depression and alcohol dependence », Revue française de sociologie, vol. 51, supplément, p. 3-40.

Cousteaux, A.S. 2010. « Représentations de la santé et cycle de vie. De la recherche du bien-être au maintien des capacités», Notes & Documents, 2010-01, Paris, OSC, Sciences Po/CNRS.

Chanvril F., A.S. Cousteaux, V. Le Hay, L. Lesnard, C. Méchinaud and N. Sauger. 2009. « La parentalité en Europe. Analyse séquentielle des trajectoires d'entrée dans l'âge adulte à partir de l'Enquête sociale européenne », Dossiers d'études CNAF, n° 122 (novembre)

Cousteaux A.S. and J.L. Pan Ké Shon. 2008. « Le mal-être a-t-il un genre ? Suicide, risque suicidaire, dépression et dépendance alcoolique », Revue française de sociologie, vol. 49, n° 1, p. 53-92.

Cousteaux A.S. 2006. « Occupation, gender and social status. Questioning the gender neutrality of status scales in contemporary France », Document de travail CREST, n ° 2006-34.

_____ .2004. « Un essai d'évaluation de la nomenclature des catégories socioprofessionnelles : une approche temporelle », mémoire de DEA de sociologie, Institut d'études politiques de Paris, sous la direction de Louis-André Vallet

Centre d'études européennes:

Duchesne, S. "Développement de l'analyse secondaire et des méthodes d'analyse qualitative: une chance à saisir?" à paraître in Dargentas, M., Brugidou, M., Le-Roux, D., & Salomon, A.C. (Eds.) L'analyse secondaire en recherche qualitative : une nouvelle pratique en sciences humaines et sociales. Paris : Lavoisier. Collection : Tec & Doc

_____ .2009. « Les méthodes de la science politique » in Les Cahiers Français n°350, « La science politique », Mai-juin 2009, 14-18.

Donegani, J.M., S. Duchesne and F. Haegel. 2002. « Sur l'interprétation des entretiens », in Aux frontières des attitudes, Paris, L'Harmattan, p. 272-295

Duchesne S. 1996. « Entretien non préstructuré, stratégie de recherche et étude des représentations. Ou : Peut-on déjà faire l'économie de l'entretien `non-directif' en sociologie ? » in : Politix, numéro 35 (Entrées en politique), troisième trimestre 1996, p. 189-206

Brouard, S. and V. Tiberj. 2010. " Yes they can: an exprimental approach to the eligibility of ethnic minority candidate in France", dans Karen Bird, Thomas Saalfeld & Andreas Wüst (dir.) /The Political Representation of Immigrants and Minorities: Voters, Parties and Parliaments in Liberal Democracies/, Oxford, Routledge, pp164-180

Brouard, S. and V. Tiberj. 2005. /Français comme les autres ? Enquête sur les Français issus de l'immigration africaine et turque/, Paris, Presses de Sciences Po

Michelat G. and V. Tiberj. 2007. « gauche, centre, droite et vote : permanence et mutation d'un clivage », Revue Française de Science Politique/, vol. 57, n°3, automne 2007.

Tiberj, V. 2008. /La crispation hexagonale : France fermée contre France plurielle, 2001-2007, Paris, FJJ–Plon, février 2008

- *Groupe des Ecoles Nationales d'Economie et de Statistique (GENES)*

Caveng, R. 2010. « Salariat libéral et auto-exploitation. Les travailleurs des entreprises de sondages », Notes de l'IES, n°15, août-septembre 2010

_____. 2009. Renversement des positions et ré-enchantement de l'interaction. La relation d'enquête dans les sondages et les études de marchés », Actes de la recherche en sciences sociales, n° 178

_____. 2009. Gestion de la relation d'enquête et usages du questionnaire. Enquête qualitative sur la passation de l'enquête Emploi en continu, Rapport pour la Division de l'Emploi et des Revenus d'activité de l'INSEE

_____. 2007, « La "liberté" des vacataires en instituts de sondages », Interrogations, n°4

Gollac, M. and S. Volkoff. 2006. « La santé au travail et ses masques », Actes de la recherche en sciences sociales, n°163

Baudelot, C., M. Gollac et al. 2003. Travailler pour être heureux ?, Fayard

Gollac, M. 1997. " Des chiffres insensés ? Pourquoi et comment on donne un sens aux données", Revue française de sociologie, n° XXXVIII-1

_____. 1995, Donner un sens aux données : l'exemple des enquêtes statistiques sur les conditions de travail, Noisy-le-Grand, CEE

Gollac, M.et al. 1995. L'usage des méthodes statistiques dans l'étude du travail, Paris, La Documentation Française.

Goux D. and E. Maurin. 2007. « Close Neighbours Matter: Neighborhood Effects on Early Performance at School », Economic Journal, vol 117, n° 125

Goux D. and E. Maurin. 2005. "The Effect of Overcrowded Housing on Children's Performance at School", Journal of Public Economics, n° 89

Goux D. 2003. « Une histoire de l'enquête Emploi », Economie et statistique, n°362

Chardon, O. and D. Goux. 2003. « La nouvelle définition européenne du chômage BIT », Economie et statistique, n°362

Goux, D. 1994. « Emploi et chômage des 'non qualifiés' en France », Economie et Statistique, n° 273

Skinner, C. and L.A. Vallet. 2010. "Fitting Log-Linear Models to Contingency Tables From Surveys With Complex Sampling Designs: An Investigation of the Clogg-Eliason Approach", Sociological Methods & Research, 39(1)

Vallet, L.-A. 2006. "Modèle log-multiplicatif et explication de la croissance de la fluidité sociale" in Lavallée P. et Rivest L.-P. (dir.), Méthodes d'enquêtes et sondages : pratiques européenne et nord-américaine, Paris, Dunod

_____. 2005. "L'apport des échantillons longitudinaux d'individus : deux expériences étrangères", Journal de la Société Française de Statistique, 146(3)

_____. 2004. "Modéliser la variation d'une association statistique selon une troisième variable : avancées récentes" in Ardilly P. (dir.), Échantillonnage et méthodes d'enquêtes, Paris, Dunod

- *Institut national des études démographiques (Ined)*

Legleye, S., D. Piontek and L. Kraus. 2010 [to be published]. Psychometric properties of the Cannabis Abuse Screening Test (CAST) in a French sample of adolescents, *Drug Alcohol Depend*.

Razafindratsima, N., S. Legleye and C. Beauchemin. 2010. Biais de non-réponse et précision des estimateurs dans l'enquête Migrations entre l'Afrique et l'Europe (MAFE-Sénégal) *Colloque francophone sur les sondages* (Tanger).

Legleye, S., F. Beck, P. Peretii-Watel, N. Chau, and J.M. Firdion. 2010. Suicidal ideation among young French adults: association with occupation, family, sexual activity, personal background and drug use, *J Affect Disord*, 123, 108-15.

Beck, F., R. Guignard, C. Cavalin, F. Maillochon and S. Legleye. 2010. AUDIO-CASI vs CATI : L'influence du mode de collecte sur la mesure des pratiques addictives en population générale *Colloque francophone des sondages* (Tanger).

Legleye, S., I. Obradovic, E. Janssen et al. 2009 Influences of cannabis use trajectories, school failure and family background on school dropout at 17 in France, *European journal of public health*, [Epub ahead of print].

Legleye, S. and F. Beck. 2003. Les marges des enquêtes téléphoniques en population générale dans le domaine des drogues : liste rouge, personnes difficiles à joindre, téléphones mobiles : quels apports, quelles contraintes ? *Séminaire de méthodologie Inpes* (Besançon, Inpes).

Legleye, S. 2003. The Telephone in General Population Surveys on Drugs *European society for social sciences and drug* (Gent, ESSD).

Beck, F., S. Legleye ans P. Peretti-Watel. 2002 Le recours au téléphone dans les enquêtes en population générale sur les drogues *Journées de Méthodologie Statistique,* (Paris, Insee).

- *Telecom Paris Tech - Social Science Department*

Beaudouin, V. and J. Velkovska. 2010. *Dialogues vocaux entre clients et automates ou comment l'homme et la machine s'entendent dans la réalisation d'un service*. JADT 2010 - Statistical Analysis of Textual Data, Sapienza - Università di Roma, LED Edizioni Universitarie di Lettere Economia Diritto, 865-876.

Beauvisage, T., V. Beaudouin and H. Assadi. 2007. "Internet 1.0: early users, early uses", *Annals of Telecommunications,* Vol. 62, n°3-4.

Beaudouin, V. 2004. Mesurer les usages d'Internet. *In*: G. Hébrail, L. Lebart and J.-M. Petit, *Extraction et gestion des connaissances (EGC'2004)*. Clermont-Ferrand, RNTI. I, p. 25-34.

Beaudouin, V., S. Fleury, M. Pasquier, B. Habert and C. Licoppe. 2002. "Décrire la toile pour mieux comprendre les parcours. Sites personnels et sites marchands", *Réseaux,* Vol. 20, n°116, p. p. 19-51.

Diminescu, D., M. Renault, M. Jacomy and Ch. d'Iribarne. 2010. "Le web matrimonial des migrants L'économie du profilage au service d'une nouvelle forme de commerce ethnique" in Réseaux N°159 pp: 17-57, http://www.cairn.info/revue-reseaux-2010-1.htm,

Diminescu, D., D. Bounie and A. François A. 2010. Une analyse socio-économique des transferts d'argent des migrants par téléphone, in Réseaux N° 159 pp : 91-111, http://www.cairn.info/resume.php?ID_ARTICLE=RES_159_0091

Diminescu, D. and M. Renault. 2009. « TIC et parrainage dans les mouvements militants de défense des *sans-papiers* en France », TIC & Société », in TIC&Société Vol.3, n°1-2, http://ticetsociete.revues.org/613

- *EDF R&D*

Brugidou, M. and M. Moine. 2010. « Normes émergentes et stigmatisation- Une analyse comparative à partir des deux questions ouvertes sur les raisons de ne pas trier les déchets et de ne pas faire d'économie d'énergie » JADT 2010 Rome, 10 juin 2010

Brugidou, M. and M. Moine. 2008. « Le public en campagne. Analyse du discours des publics de la campagne électorale présidentielle à travers une séquence de question ouverte et fermée », JADT 2008- Lyon, 13 mars 2008.

Brugidou, M. 2008. L'opinion et ses publics, Presses de Sciences Po, collection référence,

Benvegnu, N. and M. Brugidou. 2008. « Prendre la parole sur Internet : des dispositifs socio-techniques aux grammaires de la discussion », Réseaux, septembre 2008, 51-82.

Brugidou, M. and C. Escoffier. 2007. « Questions ouvertes et opinion publiques discursives », in Marc X. et Tchernia J.-F. (dir.), Etudier l'opinion, PUG, Grenoble, pp. 91-111.

Le Roux, D., M. Brugidou and M. Dargentas. 2007. Developing Computer-Aided Secondary Analysis : consequences of such an Innovation for Sociologists in an Industrial Context, Impacts of software use methodology and analysis, CAQDAS 2007 Conference "Advanced in Qualitative Computing", Royal Holloway University of London, 19, 20 avril 2007.

Brugidou, M. 2005. « Vers une représentation discursive de l'opinion publique », Corpus, 4

_____ . 2003. « Argumentation and Values : an analysis of ordinary political competence via an open-ended question », Intenational Journal of Public Opinion Research,V15 n°4, winter 2003.

_____. 2001. « La combinaison des inférences statistiques, linguistiques et sociologiques dans l'analyse d'une question ouverte », Journal de la Société Française de Statistique, 4.

Brugidou, M.and B. Cautrès. 2001. « Compte-rendu de la journée d'études du cidsp, les nouvelles méthodes d'analyse des entretiens- Analyse assistée par ordinateur et capitalisation des

entretiens non-directifs de recherche 9 MARS 2001 - Maison des sciences de l'homme - Grenoble. BMS.

Brugidou, M., C. Escoffier, H. Folch, D. Le Roux, P. Morin-Andréani and G. Piat. 2000. «Les facteurs de choix et d'utilisation de logiciels d'analyse des données textuelles», JADT, Lausanne.

Brugidou, M. 1998. « Epitaphes, l'image de François Mitterrand à travers l'analyse d'une question ouverte posée à sa mort», Revue Française de Science Politique , vol. 48, n°1, p.97-120.

Habert, B. and C. Huc. 2010. Building together digital archives for research in social sciences and humanities, Social Science Information, 49-3, 415–-443, http://hal.archives-ouvertes.fr/hal-00466352/fr/

Clavel, C., I. Vasilescu and L. Devillers. 2011. *Fiction support for realistic portrayals of fear-type emotional manifestations,* Comput. Speech Lang. 25, 1 (Jan. 2011), 63-83.

Clavel C. and G. Richard. 2010. *Reconnaissance acoustique des émotions,* , in Systèmes d'Interaction Emotionnelle, chap. 5, Hermès;

Danesi, C. and C. Clavel. 2010. *Impact of Spontaneous Speech Features on Business Concept Detection,* a Study of Call-Centre Data, , ACM Workshop on Searching Spontaneous Conversational Speech, Oct. 2010

Kuznik, L., A.L. Guénet, A. Peradotto and C. Clavel. 2010. L'apport des concepts métiers pour la classification des questions ouvertes d'enquête, TALN

6.3. ESTIMATE FOR THE EQUIPEMENT

Total budget

DIME-SHS - Coût complet

EQUIPEMENT (TTC)	TEST 2011	DEVELOPPEMENT 2012	2013	2014	2015	2016	2017	2018	FONCTIONNEMENT 2019	2020	TOTAL
Matériel et infrastructures informatiques											
Partenaire 1 - SCIENCES PO	383 100,00 €	10 700,00 €	6 600,00 €	6 000,00 €	49 100,00 €	5 700,00 €	6 600,00 €	398 559,64 €	37 100,00 €	23 700,00 €	927 160 €
Partenaire 2 - GENES	- €	- €	- €	2 208,00 €	- €	- €	- €	2 208,00 €	- €	- €	4 416 €
Partenaire 3 - INED	7 357,20 €	- €	- €	- €	7 357,20 €	- €	- €	- €	7 357,20 €	- €	22 072 €
Partenaire 4 - PARIS V - CERES	708,00 €	- €	- €	- €	708,00 €	- €	- €	- €	708,00 €	- €	2 124 €
Partenaire 5 - TELECOM PARIS TECH	331,20 €	- €	- €	- €	331,20 €	- €	- €	- €	331,20 €	- €	994 €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Sous-total	391 496,40 €	10 700,00 €	6 600,00 €	8 208,00 €	57 496,40 €	5 700,00 €	6 600,00 €	400 767,64 €	45 496,40 €	23 700,00 €	956 765 €
Logiciels et licences											
Partenaire 1 - SCIENCES PO	90 900,00 €	40 150,00 €	11 550,00 €	24 900,00 €	42 150,00 €	57 900,00 €	74 250,00 €	150,00 €	900,00 €	150,00 €	343 000 €
Partenaire 2 - GENES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 3 - INED	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 4 - PARIS V - CERES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 5 - TELECOM PARIS TECH	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Sous-total	90 900,00 €	40 150,00 €	11 550,00 €	24 900,00 €	42 150,00 €	57 900,00 €	74 250,00 €	150,00 €	900,00 €	150,00 €	343 000 €
Equipements Internet Mobile											
Partenaire 1 - SCIENCES PO	595 000,00 €	595 000,00 €	479 500,00 €	479 500,00 €	437 500,00 €	437 500,00 €	392 000,00 €	392 000,00 €	357 000,00 €	315 000,00 €	4 480 000 €
Sous-total	595 000,00 €	595 000,00 €	479 500,00 €	479 500,00 €	437 500,00 €	437 500,00 €	392 000,00 €	392 000,00 €	357 000,00 €	315 000,00 €	4 480 000 €
TOTAL EQUIPEMENT (1)	1 077 396 €	645 850 €	497 650 €	512 608 €	537 146 €	501 100 €	472 850 €	792 918 €	403 396 €	338 850 €	5 779 765 €
FONCTIONNEMENT SYSTEME D'INFORMATION (TTC)	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Partenaire 1 - SCIENCES PO	76 050,00 €	76 050,00 €	76 050,00 €	74 250,00 €	76 050,00 €	76 050,00 €	76 050,00 €	87 083,95 €	85 283,95 €	87 083,95 €	790 002 €
Partenaire 2 - GENES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 3 - INED	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 4 - PARIS V - CERES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 5 - TELECOM PARIS TECH	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
TOTAL FONCTIONNEMENT SI (2)	76 050 €	76 050 €	76 050 €	74 250 €	76 050 €	76 050 €	76 050 €	87 084 €	85 284 €	87 084 €	790 002 €
LOYERS (CO-FINANCEMENT)	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Partenaire 1 - SCIENCES PO	40 000,00 €	40 000,00 €	40 000,00 €	40 000,00 €	40 000,00 €	40 000,00 €	40 000,00 €	40 000,00 €	40 000,00 €	40 000,00 €	400 000 €
Partenaire 2 - GENES	- €	- €	- €	20 000,00 €	20 000,00 €	20 000,00 €	20 000,00 €	20 000,00 €	20 000,00 €	20 000,00 €	140 000 €
Partenaire 3 - INED	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 4 - PARIS V - CERES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 5 - TELECOM PARIS TECH	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
TOTAL LOYERS (3)	40 000 €	40 000 €	40 000 €	60 000 €	60 000 €	60 000 €	60 000 €	60 000 €	60 000 €	60 000 €	540 000 €
PRESTATIONS DE SERVICE	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Partenaire 1 - SCIENCES PO	20 000,00 €	10 000,00 €	15 000,00 €	- €	10 000,00 €	- €	15 000,00 €	10 000,00 €	- €	- €	80 000 €
Partenaire 2 - GENES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 3 - INED	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 4 - PARIS V - CERES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 5 - TELECOM PARIS TECH	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
TOTAL PRESTATIONS DE SERVICE (4)	20 000 €	10 000 €	15 000 €	- €	10 000 €	- €	15 000 €	10 000 €	- €	- €	80 000 €

MASSE SALARIALE CO-FINANCEE (sans financement ANR)	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Partenaire 1 - SCIENCES PO	199 320,00 €	227 677,80 €	231 092,97 €	266 269,36 €	270 263,40 €	274 317,35 €	278 432,11 €	282 608,59 €	286 847,72 €	291 150,44 €	2 607 980 €
Partenaire 2 - GENES	- €	- €	- €	118 849,08 €	120 631,82 €	122 441,29 €	124 277,91 €	126 142,08 €	128 034,21 €	129 954,73 €	870 331 €
Partenaire 3 - INED	76 431,97 €	77 578,45 €	78 742,13 €	79 923,26 €	81 122,11 €	82 338,94 €	83 574,02 €	84 827,64 €	86 100,05 €	87 391,55 €	818 030 €
Partenaire 4 - PARIS V - CERES	53 091,60 €	53 887,97 €	54 696,29 €	55 516,74 €	56 349,49 €	57 194,73 €	58 052,65 €	58 923,44 €	59 807,29 €	60 704,40 €	568 225 €
Partenaire 5 - TELECOM PARIS TECH	57 706,76 €	58 572,37 €	59 450,95 €	60 342,72 €	61 247,86 €	62 166,57 €	63 099,07 €	64 045,56 €	65 006,24 €	65 981,34 €	617 619 €
Partenaire 6 - EDF R&D	3 997,27 €	4 057,23 €	4 118,09 €	4 179,86 €	4 242,56 €	4 306,20 €	4 370,79 €	4 436,35 €	4 502,90 €	4 570,44 €	42 782 €
Partenaire 7 - GIS QUETELET	- €	- €	24 462,00 €	24 828,93 €	25 201,36 €	25 579,38 €	25 963,08 €	26 352,52 €	26 747,81 €	27 149,03 €	206 284 €
TOTAL MASSE SALARIALE CO-FINANCEE (5)	**390 548 €**	**421 774 €**	**452 562 €**	**609 910 €**	**619 059 €**	**628 344 €**	**637 770 €**	**647 336 €**	**657 046 €**	**666 902 €**	**5 731 251 €**

FRAIS D'ENVIRONNEMENT CO-FINANCE (sans financement ANR)	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Partenaire 1 - SCIENCES PO	59 796,00 €	68 303,34 €	69 327,89 €	79 880,81 €	81 079,02 €	82 295,21 €	83 529,63 €	84 782,58 €	86 054,32 €	87 345,13 €	782 394 €
Partenaire 2 - GENES	- €	- €	- €	35 654,72 €	36 189,54 €	36 732,39 €	37 283,37 €	37 842,62 €	38 410,26 €	38 986,42 €	261 099 €
Partenaire 3 - INED	22 929,59 €	23 273,54 €	23 622,64 €	23 976,98 €	24 336,63 €	24 701,68 €	25 072,21 €	25 448,29 €	25 830,01 €	26 217,47 €	245 409 €
Partenaire 4 - PARIS V - CERES	15 927,48 €	16 166,39 €	16 408,89 €	16 655,02 €	16 904,85 €	17 158,42 €	17 415,80 €	17 677,03 €	17 942,19 €	18 211,32 €	170 467 €
Partenaire 5 - TELECOM PARIS TECH	17 312,03 €	17 571,71 €	17 835,29 €	18 102,81 €	18 374,36 €	18 649,97 €	18 929,72 €	19 213,67 €	19 501,87 €	19 794,40 €	185 286 €
Partenaire 6 - EDF R&D	1 199,18 €	1 217,17 €	1 235,43 €	1 253,96 €	1 272,77 €	1 291,86 €	1 311,24 €	1 330,91 €	1 350,87 €	1 371,13 €	12 835 €
Partenaire 7 - GIS QUETELET	- €	- €	7 338,60 €	7 448,68 €	7 560,41 €	7 673,82 €	7 788,92 €	7 905,76 €	8 024,34 €	8 144,71 €	61 885 €
TOTAL FRAIS D'ENVIRONNEMENT CO-FINANCE (6)	**117 164 €**	**126 532 €**	**135 769 €**	**182 973 €**	**185 718 €**	**188 503 €**	**191 331 €**	**194 201 €**	**197 114 €**	**200 071 €**	**1 719 375 €**

MASSE SALARIALE NOUVELLE (financement ANR)	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Partenaire 1 - SCIENCES PO	987 540,00 €	1 122 851,10 €	1 334 008,22 €	1 287 701,42 €	1 307 016,94 €	1 326 622,19 €	1 346 521,53 €	1 366 719,35 €	1 387 220,14 €	1 408 028,44 €	12 874 229 €
Partenaire 2 - GENES	- €	- €	- €	36 240,00 €	36 783,60 €	37 335,35 €	37 895,38 €	38 463,82 €	39 040,77 €	39 626,38 €	265 385 €
Partenaire 3 - INED	344 280,00 €	349 444,20 €	481 525,86 €	488 748,75 €	550 439,98 €	558 696,58 €	567 077,03 €	575 583,19 €	584 216,93 €	592 980,19 €	5 092 993 €
Partenaire 4 - PARIS V - CERES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 5 - TELECOM PARIS TECH	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
TOTAL MASSE SALARIALE A FINANCER (7)	**1 331 820 €**	**1 472 295 €**	**1 815 534 €**	**1 812 690 €**	**1 894 241 €**	**1 922 654 €**	**1 951 494 €**	**1 980 766 €**	**2 010 478 €**	**2 040 635 €**	**18 232 607 €**

FRAIS GENERAUX	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Partenaire 1 - SCIENCES PO	3 042,00 €	3 042,00 €	3 042,00 €	2 970,00 €	2 889,90 €	2 737,80 €	2 129,40 €	1 393,34 €	- €	- €	21 246 €
Partenaire 2 - GENES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 3 - INED	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 4 - PARIS V - CERES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 5 - TELECOM PARIS TECH	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
TOTAL FRAIS GENERAUX (8)	**3 042 €**	**3 042 €**	**3 042 €**	**2 970 €**	**2 890 €**	**2 738 €**	**2 129 €**	**1 393 €**	**- €**	**- €**	**21 246 €**

TOTAL GENERAL (coût complet) (I+II+III+IV+V+VI+VII+VIII+IX+X)	**3 056 020 €**	**2 795 543 €**	**3 035 607 €**	**3 255 401 €**	**3 385 103 €**	**3 379 390 €**	**3 406 624 €**	**3 773 698 €**	**3 413 318 €**	**3 393 541 €**	**32 894 247 €**

Subsidy

DIME-SHS - Assiette

	TEST	DEVELOPPEMENT						FONCTIONNEMENT		
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
CA généré (hypothèse basse avec prix unitaire de 55k€ volontairement faible)	0	0	385 000	825 000	1 375 000	1 925 000	2 475 000	3 025 000	3 575 000	3 850 000
CA généré / Coût complet	0%	0%	13%	25%	41%	57%	73%	80%	105%	113%
Hypothèse de montée en charge de l'autofinancement	0%	0%	0%	0%	5%	10%	30%	60%	100%	100%

INVESTISSEMENT	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Dépenses de personnel (financement ANR)											
Partenaire 1 - SCIENCES PO	987 540 €	1 122 851 €	1 334 008 €	1 287 701 €	1 241 666 €	1 193 960 €	942 565 €	546 688 €	0 €	0 €	8 656 980 €
Partenaire 2 - GENES	0 €	0 €	0 €	36 240 €	34 944 €	33 602 €	26 527 €	15 386 €	0 €	0 €	146 699 €
Partenaire 3 - INED	344 280 €	349 444 €	481 526 €	488 749 €	522 918 €	502 827 €	396 954 €	230 233 €	0 €	0 €	3 316 931 €
Partenaire 4 - PARIS V - CERES	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €
Partenaire 5 - TELECOM PARIS TECH	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €
Partenaire 6 - EDF R&D	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €
Partenaire 7 - GIS QUETELET	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €	0 €
Total Dépenses de personnel (A)	**1 331 820 €**	**1 472 295 €**	**1 815 534 €**	**1 812 690 €**	**1 799 528 €**	**1 730 389 €**	**1 366 046 €**	**792 307 €**	**0 €**	**0 €**	**12 120 609 €**
Dépenses de réalisation de l'équipement											
Partenaire 1 - SCIENCES PO	90 900,00 €	40 150,00 €	11 550,00 €	24 900,00 €	40 042,50 €	52 110,00 €	51 975,00 €	60,00 €	- €	- €	311 687,50 €
Partenaire 2 - GENES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 3 - INED	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 4 - PARIS V - CERES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 5 - TELECOM PARIS TECH	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Logiciels et licences	90 900 €	40 150 €	11 550 €	24 900 €	40 043 €	52 110 €	51 975 €	60 €	- €	- €	311 688 €
Partenaire 1 - SCIENCES PO	20 000 €	10 000 €	15 000 €	- €	9 500 €	- €	10 500 €	4 000 €	- €	- €	69 000 €
Partenaire 2 - GENES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 3 - INED	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 4 - PARIS V - CERES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 5 - TELECOM PARIS TECH	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Prestations de service	20 000 €	10 000 €	15 000 €	- €	9 500 €	- €	10 500 €	4 000 €	- €	- €	69 000 €
Partenaire 1 - SCIENCES PO	40 000 €	40 000 €	40 000 €	40 000 €	38 000 €	36 000 €	28 000 €	16 000 €	- €	- €	
Partenaire 2 - GENES											
Partenaire 3 - INED	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	
Partenaire 4 - PARIS V - CERES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	
Partenaire 5 - TELECOM PARIS TECH	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	
Loyers	40 000 €	40 000 €	40 000 €	40 000 €	38 000 €	36 000 €	28 000 €	16 000 €	- €	- €	278 000 €
Total Dépenses de fonctionnement induites par la réalisation de l'équipement (B)	**150 900 €**	**90 150 €**	**66 550 €**	**64 900 €**	**87 543 €**	**88 110 €**	**90 475 €**	**20 060 €**	**- €**	**- €**	**658 688 €**

Dépenses d'équipement											
Partenaire 1 - SCIENCES PO	383 100,00 €	10 700,00 €	6 600,00 €	6 000,00 €	46 645,00 €	5 130,00 €	4 620,00 €	159 423,85 €	- €	- €	622 218,85 €
Partenaire 2 - GENES	- €	- €	- €	2 208,00 €	- €	- €	- €	883,20 €	- €	- €	3 091,20 €
Partenaire 3 - INED	7 357,20 €	- €	- €	- €	6 989,34 €	- €	- €	- €	- €	- €	14 346,54 €
Partenaire 4 - PARIS V - CERES	708,00 €	- €	- €	- €	672,60 €	- €	- €	- €	- €	- €	1 380,60 €
Partenaire 5 - TELECOM PARIS TECH	331,20 €	- €	- €	- €	314,64 €	- €	- €	- €	- €	- €	645,84 €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Matériel et infrastructure informatique	391 496 €	10 700 €	6 600 €	8 208 €	54 622 €	5 130 €	4 620 €	160 307 €	- €	- €	641 683 €
Partenaire 1 - SCIENCES PO	595 000 €	595 000 €	479 500 €	479 500 €	415 625 €	393 750 €	274 400 €	156 800 €	- €	- €	3 389 575 €
Equipements Internet mobile	595 000 €	595 000 €	479 500 €	479 500 €	415 625 €	393 750 €	274 400 €	156 800 €	- €	- €	3 389 575 €
Total Dépenses d'équipement (C)	**986 496 €**	**605 700 €**	**486 100 €**	**487 708 €**	**470 247 €**	**398 880 €**	**279 020 €**	**317 107 €**	**- €**	**- €**	**4 031 258 €**

TOTAL INVESTISSEMENT (A+B+C)	**2 469 216 €**	**2 168 145 €**	**2 368 184 €**	**2 365 298 €**	**2 357 318 €**	**2 217 379 €**	**1 735 541 €**	**1 129 474 €**	**- €**	**- €**	**16 810 555 €**

FONCTIONNEMENT	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Dépenses de fonctionnement induites par la réalisation de l'équipement											
Partenaire 1 - SCIENCES PO	76 050 €	76 050 €	76 050 €	74 250 €	72 248 €	68 445 €	53 235 €	34 834 €	- €	- €	531 161 €
Partenaire 2 - GENES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 3 - INED	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 4 - PARIS V - CERES	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 5 - TELECOM PARIS TECH	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 6 - EDF R&D	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Partenaire 7 - GIS QUETELET	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €	- €
Charges récurrentes du système d'information (IV)	76 050 €	76 050 €	76 050 €	74 250 €	72 248 €	68 445 €	53 235 €	34 834 €	- €	- €	379 061 €
Total Dépenses de fonctionnement induites par la réalisation de l'équipement (D)	**76 050 €**	**76 050 €**	**76 050 €**	**74 250 €**	**72 248 €**	**68 445 €**	**53 235 €**	**34 834 €**	**- €**	**- €**	**531 161 €**
Frais généraux (4% des frais de fonctionnement) (E)	**3 042 €**	**3 042 €**	**3 042 €**	**2 970 €**	**2 890 €**	**2 738 €**	**2 129 €**	**1 393 €**	**- €**	**- €**	**21 246 €**
TOTAL FONCTIONNEMENT (D+E)	**79 092 €**	**79 092 €**	**79 092 €**	**77 220 €**	**75 137 €**	**71 183 €**	**55 364 €**	**36 227 €**	**- €**	**- €**	**552 408 €**
TOTAL GENERAL (assiette) (A+B+C+D+E)	**2 548 308 €**	**2 247 237 €**	**2 447 276 €**	**2 442 518 €**	**2 432 455 €**	**2 288 562 €**	**1 790 905 €**	**1 165 701 €**	**- €**	**- €**	**17 362 962 €**

Sciences Po - Coût complet

	TEST		DEVELOPPEMENT						FONCTIONNEMENT		
EQUIPEMENT (TTC)	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**TOTAL**
Hypothèses de base :											
Nombre d'enquêtes (toutes composantes de l'équipement confondues) :	*5*	*10*	*15*	*20*	*30*	*40*	*50*	*60*	*70*	*80*	
Panel (nombre de personnes) :	*3 100*	*6 000*	*9 000*	*12 000*	*15 000*	*15 000*	*15 000*	*15 000*	*15 000*	*15 000*	
Nombre de questions par enquêtes :	*Entre 50 et 100*	*Entre 50 et 100*	*Entre 50 et 100*	*Entre 50 et 100*	*Entre 50 et 100*	*Entre 50 et 100*	*Entre 50 et 100*	*Entre 50 et 100*	*Entre 50 et 100*	*Entre 50 et 100*	
Volumétrie des données en Téra :	*15*	*18*	*22*	*26*	*31*	*37*	*45*	*54*	*64*	*77*	
Renouvellement du panel	*3 500*	*3 500*	*3 500*	*3 500*	*3 500*	*3 500*	*3 500*	*3 500*	*3 500*	*3 500*	
Coût unitaire équipement Smartphone (les équipements ne sont pas réutilisables pour un autre	*150*	*150*	*120*	*120*	*110*	*110*	*100*	*100*	*90*	*80*	
Coût unitaire abonnement téléphonique (Forfait 2h mensuel avec internet illimité)	*20*	*20*	*17*	*17*	*15*	*15*	*12*	*12*	*12*	*10*	
Nombre d'études réalisées	*5*	*10*	*15*	*20*	*30*	*40*	*50*	*60*	*70*	*80*	
Nombre d'études facturées (hypothèse basse)	*0*	*0*	*7*	*15*	*25*	*35*	*45*	*55*	*65*	*70*	
CA généré (50k€ en moyenne)	*0*	*0*	*385 000*	*825 000*	*1 375 000*	*1 925 000*	*2 475 000*	*3 025 000*	*3 575 000*	*3 850 000*	
Matériels et infrastructures informatiques :											
Matériel réseau (distribution horizontale - dépend également de la structure des bâtiments) et remplacement du matériel tous les 7 ans (prise en compte aussi de l'indice syntex de 3% par ans) ==> Ramené à 20 000 euros si hébergé à Sciences Po, sinon 90 000 euros	20 000 €							24 597 €			**44 597 €**
Cœur de réseau (distribution verticale, desserte optique, gestion des Vlans - réseau niveau 3) et remplacement du matériel tous les 7 ans (prise en compte aussi de l'indice syntex de 3% par ans) : 0 euros si bâtiment de structure simple, sinon 80 keuro								- €			**- €**
Commutateurs niveau 1 - accélérateurs http et remplacement du matériel tous les 7 ans (prise en compte aussi de l'indice syntex de 3% par ans) ==> 50 keuro si non hébergement FNSP, sinon 0								- €			**- €**
Serveurs (SAN, serveurs de fichiers, serveurs de base de données, stockage multimédia, mails, stockage ftp pour chercheurs, doctorants, etc.) et remplacement du matériel tous les 7 ans (prise en compte aussi de l'indice syntex de 3% par ans). 300 keuro si solution avec redondance, 200	300 000 €							368 962 €			**668 962 €**
Bornes WIFI : 3 bornes en 1ère année	4 000 €										**4 000 €**
TOIP (téléphonie sous IP et plateforme téléphonique) :	5 000 €										**5 000 €**
Ordinateurs équipe et remplacement du matériel tous les 4 ans : ==> stations de travail type Mac Pro ou HP Work Station (même puissance) pour les spécialistes de méthode d'enquête et d'ingénierie ==> station de travail plus standard pour les postes plus administratifs (Imac ou stations de travail / portables plus standards)	26 100 €	5 700 €	6 600 €	- €	26 100 €	5 700 €	6 600 €	- €	26 100 €	5 700 €	**108 600 €**
Portables et postes volants (visiteurs, doctorants supplémentaires, etc.) :	5 000 €				5 000 €				5 000 €		**15 000 €**
Outils enquêtes mobiles (smartphones, tablettes, etc.)	5 000 €	5 000 €						5 000 €			**15 000 €**
Equipements audiovisuels des salles de travail et remplacement tous les 5 ans : sonorisation, mobilier technique sur la base d'une équipement de trois salles de travail	18 000 €			6 000 €	18 000 €				6 000 €	18 000 €	**66 000 €**
Nombre de salles équipées	3			1							
Total matériels et infrastructures informatiques (I)	**383 100 €**	**10 700 €**	**6 600 €**	**6 000 €**	**49 100 €**	**5 700 €**	**6 600 €**	**398 560 €**	**37 100 €**	**23 700 €**	**927 160 €**
Logiciels et licences :											
Logiciels bureautiques : 300 euros de logiciels par ordinateurs	900 €	150 €	- €	150 €	900 €	150 €	- €	150 €	900 €	150 €	**3 450 €**
Bases de données et outils de développements	70 000 €	30 000 €									**100 000 €**
Frais de licence plateforme MESS (Université de Tilburg) (Upfront de 30k€ puis 3% du CA sur 5	20 000 €	10 000 €	11 550 €	24 750 €	41 250 €	57 750 €	74 250 €				**219 550 €**
Total logiciels et licences (II)	**90 900 €**	**40 150 €**	**11 550 €**	**24 900 €**	**42 150 €**	**57 900 €**	**74 250 €**	**150 €**	**900 €**	**150 €**	**343 000 €**
Equipements Internet mobile											
Smartphones	525 000 €	525 000 €	420 000 €	420 000 €	385 000 €	385 000 €	350 000 €	350 000 €	315 000 €	280 000 €	**3 955 000 €**
Abonnements téléphoniques	70 000 €	70 000 €	59 500 €	59 500 €	52 500 €	52 500 €	42 000 €	42 000 €	42 000 €	35 000 €	**525 000 €**
Total Equipement Internet mobile (III)	**595 000 €**	**595 000 €**	**479 500 €**	**479 500 €**	**437 500 €**	**437 500 €**	**392 000 €**	**392 000 €**	**357 000 €**	**315 000 €**	**4 480 000 €**
TOTAL EQUIPEMENT (I + II + III)	**1 069 000 €**	**645 850 €**	**497 650 €**	**510 400 €**	**528 750 €**	**501 100 €**	**472 850 €**	**790 710 €**	**395 000 €**	**338 850 €**	**5 750 160 €**

FONCTIONNEMENT SYSTEME D'INFORMATION (TTC)	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Charges récurrentes du système d'information (IV)											
Connexion internet (2000 euros / mois) (dans l'hypothèse où ne connaissant pas l'emplacement, il faut rajouter une nouvelle connexion)	24 000 €	24 000 €	24 000 €	24 000 €	24 000 €	24 000 €	24 000 €	24 000 €	24 000 €	24 000 €	**240 000 €**
Matériel réseau (distribution horizontale - dépend également de la structure des bâtiments) et remplacement du matériel tous les 7 ans (prise en compte aussi de l'indice syntex de 3% par ans)	3 000 €	3 000 €	3 000 €	3 000 €	3 000 €	3 000 €	3 000 €	3 690 €	3 690 €	3 690 €	**32 069 €**
Serveurs (SAN, serveurs de fichiers, serveurs de base de données, stockage multimédia, mails, stockage ftp pour chercheurs, doctorants, etc.) :	45 000 €	45 000 €	45 000 €	45 000 €	45 000 €	45 000 €	45 000 €	55 344 €	55 344 €	55 344 €	**481 033 €**
Bornes WIFI : 3 bornes en 1ere année	600 €	600 €	600 €	600 €	600 €	600 €	600 €	600 €	600 €	600 €	**6 000 €**
TOIP (téléphonie sous IP) :	750 €	750 €	750 €	750 €	750 €	750 €	750 €	750 €	750 €	750 €	**7 500 €**
-PABX Cisco, téléphones (ligne livrée)											
Equipements audiovisuels des salles de travail et remplacement tous les 5 ans : tableaux blancs interactifs, sonorisation, modèle technique sur la base d'une équipement de trois salles de travail	2 700 €	2 700 €	2 700 €	900 €	2 700 €	2 700 €	2 700 €	2 700 €	900 €	2 700 €	**23 400 €**
Total Charges récurrentes du système d'information (base de 15% des coûts d'achat des matériels et logiciels et ceci sur la base d'années pleines) (IV) :	**76 050 €**	**76 050 €**	**76 050 €**	**74 250 €**	**76 050 €**	**76 050 €**	**76 050 €**	**87 084 €**	**85 284 €**	**87 084 €**	**790 002 €**
TOTAL FONCTIONNEMENT SI (IV)	**76 050 €**	**76 050 €**	**76 050 €**	**74 250 €**	**76 050 €**	**76 050 €**	**76 050 €**	**87 084 €**	**85 284 €**	**87 084 €**	**790 002 €**
LOYERS (CO-FINANCEMENT)	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**TOTAL**
LOYERS (V)											
Location des locaux à Sciences Po (Paris 7e)	40 000 €	40 000 €	40 000 €	40 000 €	40 000 €	40 000 €	40 000 €	40 000 €	40 000 €	40 000 €	**400 000 €**
Total Loyers (V) :	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**400 000 €**
TOTAL LOYERS (V)	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**40 000 €**	**400 000 €**
PRESTATIONS DE SERVICE	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**TOTAL**
PRESTATIONS DE SERVICE (VI)											
Communication (recrutement panel et partenaires)	20 000	10 000			10 000			10 000			**50 000 €**
Etudes de marchés			15 000				15 000				**30 000 €**
Total Prestations de service (VI) :	**20 000 €**	**10 000 €**	**15 000 €**	**- €**	**10 000 €**	**- €**	**15 000 €**	**10 000 €**	**- €**	**- €**	**80 000 €**
TOTAL PRESTATIONS DE SERVICE (VI)	**20 000 €**	**10 000 €**	**15 000 €**	**- €**	**10 000 €**	**- €**	**15 000 €**	**10 000 €**	**- €**	**- €**	**80 000 €**

MASSE SALARIALE CO-FINANCEE (sans financement ANR)	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Hypothèses de bases de calcul et d'évolution de la masse salariale du projet :											
L'équipe											
Directeur scientifique	*0,5*										*0,5*
Assistant de gestion		*0,5*									*0,5*
Responsable de la survey team	*1*										
Responsable de la method team	*1*										
Responsable Bequali	*0,5*										
Ingénieur informatique				*0,5*							
Nombre de personnes	**3**	**0,5**	**0**	**0,5**	**0**	**0**	**0**	**0**	**0**	**0**	**4**
Salaire mensuel brut en début de projet	**euros**										
Directeur scientifique	*4 800 €*										
Assistant de gestion	*2 800 €*										
Responsable de la survey team	*3 400 €*										
Responsable de la method team	*3 200 €*										
Responsable Bequali	*4 000 €*										
Ingénieur informatique	*3 500 €*										
Taux de charge patronales : 51%	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	
Taux d'augmentation ancienneté	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	
Taux d'augmentation fonction publique	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	
Augmentation individuelle : 1% de la masse salariale au service par an calculé sur la base de l'année précédente	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	
Masse salariale co-financée (y compris charges patronales et augmentation collective et individuelles Accords Sciences Po - cf. hypothèses ci-dessous en rouge) (VII)											
Directeur scientifique	43 488 €	44 140 €	44 802 €	45 474 €	46 157 €	46 849 €	47 552 €	48 265 €	48 989 €	49 724 €	**465 440 €**
Assistant de gestion	- €	25 368 €	25 749 €	26 135 €	26 527 €	26 925 €	27 329 €	27 738 €	28 155 €	28 577 €	**242 501 €**
Responsable de la survey team	61 608 €	62 532 €	63 470 €	64 422 €	65 388 €	66 369 €	67 365 €	68 375 €	69 401 €	70 442 €	**659 373 €**
Responsable de la method team	57 984 €	58 854 €	59 737 €	60 633 €	61 542 €	62 465 €	63 402 €	64 353 €	65 319 €	66 298 €	**620 587 €**
Responsable Bequali	36 240 €	36 784 €	37 335 €	37 895 €	38 464 €	39 041 €	39 626 €	40 221 €	40 824 €	41 436 €	**387 867 €**
Ingénieur informatique	- €	- €	- €	31 710 €	32 186 €	32 668 €	33 158 €	33 656 €	34 161 €	34 673 €	**232 212 €**
Total Masse salariale avec augmentations individuelles (VII) :	**199 320 €**	**227 678 €**	**231 093 €**	**266 269 €**	**270 263 €**	**274 317 €**	**278 432 €**	**282 609 €**	**286 848 €**	**291 150 €**	**2 607 980 €**
TOTAL MASSE SALARIALE CO-FINANCEE (VII)	**199 320 €**	**227 678 €**	**231 093 €**	**266 269 €**	**270 263 €**	**274 317 €**	**278 432 €**	**282 609 €**	**286 848 €**	**291 150 €**	**2 607 980 €**
FRAIS D'ENVIRONNEMENT CO-FINANCE (sans financement ANR)	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**	**TOTAL**
Taux d'environnement	30%	30%	30%	30%	30%	30%	30%	30%	30%	30%	
Frais d'environnement	59 796 €	68 303 €	69 328 €	79 881 €	81 079 €	82 295 €	83 530 €	84 783 €	86 054 €	87 345 €	782 394 €
TOTAL FRAIS D'ENVIRONNEMENT CO-FINANCE (VIII)	**59 796 €**	**68 303 €**	**69 328 €**	**79 881 €**	**81 079 €**	**82 295 €**	**83 530 €**	**84 783 €**	**86 054 €**	**87 345 €**	**782 394 €**

MASSE SALARIALE NOUVELLE (financement ANR)	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
Hypothèses de bases de calcul et d'évolution de la masse salariale du projet :											
L'équipe											
Directeur exécutif	*1*										***1***
responsable marketing		*0,5*	*1*								***1,5***
Juriste	*1*										***1***
Assistant de gestion		*0,5*									***0,5***
Ingénieurs en informatique	*7*	*1*									***8***
Ingénieur de recherche	*1*		*1*								***2***
Ingénieur d'études	*3*		*1*								***4***
Consultant	*0,5*										***0,5***
Survey method expert	*3*		*1*								
Nombre de personnes	**16,5**	**2**	**4**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**22,5**
Salaire mensuel brut en début de projet	***euros***										
Directeur exécutif	*4 000 €*										
Responsable marketing	*3 500 €*										
Juriste	*3 500 €*										
Assistant de gestion	*2 800 €*										
Ingénieurs en informatique	*3 500 €*										
Ingénieur de recherche	*3 000 €*										
Ingénieur d'études	*3 000 €*										
Consultant	*3 000 €*										
Survey method expert	*3 000 €*										
Taux de charge patronales : 51%	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	*51,0%*	
Taux d'augmentation ancienneté	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	*1,5%*	
Taux d'augmentation fonction publique	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	*0,0%*	
Augmentation individuelle : 1% de la masse salariale du service par an calculé sur la base de	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	*1,0%*	
Masse salariale à financer (y compris charges patronales et augmentation collective et individuelles Accords Sciences Po - cf. hypothèses ci-dessous en rouge) (VIII)											
Directeur exécutif	72 480 €	73 567 €	74 671 €	75 791 €	76 928 €	78 082 €	79 253 €	80 442 €	81 648 €	82 873 €	**775 733 €**
Responsable marketing	- €	31 710 €	63 420 €	64 371 €	65 337 €	66 317 €	67 312 €	68 321 €	69 346 €	70 386 €	**566 521 €**
Juriste	63 420 €	64 371 €	65 337 €	- €	- €	- €	- €	- €	- €	- €	**193 128 €**
Assistant de gestion	- €	25 368 €	25 749 €	26 135 €	26 527 €	26 925 €	27 329 €	27 738 €	28 155 €	28 577 €	**242 501 €**
Ingénieurs en informatique	443 940 €	514 019 €	521 729 €	529 555 €	537 499 €	545 561 €	553 745 €	562 051 €	570 481 €	579 039 €	**5 357 619 €**
Ingénieur de recherche	54 360 €	55 175 €	110 363 €	112 018 €	113 699 €	115 404 €	117 135 €	118 892 €	120 676 €	122 486 €	**1 040 209 €**
Ingénieur d'études	163 080 €	165 526 €	222 369 €	225 705 €	229 090 €	232 527 €	236 014 €	239 555 €	243 148 €	246 795 €	**2 203 809 €**
Consultant	27 180 €	27 588 €	28 002 €	28 422 €	28 848 €	29 281 €	29 720 €	30 166 €	30 618 €	31 077 €	**290 900 €**
Survey method expert	163 080 €	165 526 €	222 369 €	225 705 €	229 090 €	232 527 €	236 014 €	239 555 €	243 148 €	246 795 €	**2 203 809 €**
Total masse salariale sans augmentation individuelle :	824 460 €	957 325 €	1 111 639 €	1 061 997 €	1 077 927 €	1 094 096 €	1 110 507 €	1 127 165 €	1 144 072 €	1 161 233 €	**10 670 420 €**
Total Masse salariale avec augmentations individuelles (IX) :	**987 540 €**	**1 122 851 €**	**1 334 008 €**	**1 287 701 €**	**1 307 017 €**	**1 326 622 €**	**1 346 522 €**	**1 366 719 €**	**1 387 220 €**	**1 408 028 €**	**12 874 229 €**
TOTAL MASSE SALARIALE A FINANCER (IX)	**987 540 €**	**1 122 851 €**	**1 334 008 €**	**1 287 701 €**	**1 307 017 €**	**1 326 622 €**	**1 346 522 €**	**1 366 719 €**	**1 387 220 €**	**1 408 028 €**	**12 874 229 €**

FRAIS GENERAUX	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	TOTAL
FRAIS GENERAUX (X)											
Base forfaitaire de 4% des charges de fonctionnement de l'Assiette	3 042 €	3 042 €	3 042 €	2 970 €	2 890 €	2 738 €	2 129 €	1 393 €	- €	- €	**21 246 €**
Total FRAIS GENERAUX (X) :	**3 042 €**	**3 042 €**	**3 042 €**	**2 970 €**	**2 890 €**	**2 738 €**	**2 129 €**	**1 393 €**	**- €**	**- €**	**21 246 €**
TOTAL FRAIS GENERAUX (X)	**3 042 €**	**3 042 €**	**3 042 €**	**2 970 €**	**2 890 €**	**2 738 €**	**2 129 €**	**1 393 €**	**- €**	**- €**	**21 246 €**

TOTAL GENERAL (coût complet) (I+II+III+IV+V+VI+VII+VIII+IX+X)	**2 454 748 €**	**2 193 774 €**	**2 266 171 €**	**2 261 472 €**	**2 316 049 €**	**2 303 123 €**	**2 314 513 €**	**2 663 297 €**	**2 280 406 €**	**2 252 458 €**	**23 306 011 €**

6.4. COMMITMENT LETTERS

1. Fondation Nationale des Sciences Politiques (Sciences Po)

SciencesPo.
PARIS

I, Richard Descoings, Administrator of the *Fondation Nationale des Sciences Politiques*, confirm that, in the framework of the *Données, Infrastructures, Méthodes d'Enquêtes en Sciences humaines et sociales* (DIME-SHS) project, coordinated by the *Fondation Nationale des Sciences Politiques* (Sciences Po) and answering the *Equipements d'Excellence* (EQUIPEX) call, published by the *Agence Nationale de la Recherche* (ANR) in the framework of the *investissements d'avenir* programme of which the deadline is September 15th, 2010, the part of equipment depending on the competencies of the *Fondation Nationale des Sciences Politiques* can be implemented and put into service as soon as it is delivered, and that the means necessary to accompany its functioning, personnel included, will be implemented.

Paris, September 13th, 2010

For the institution,

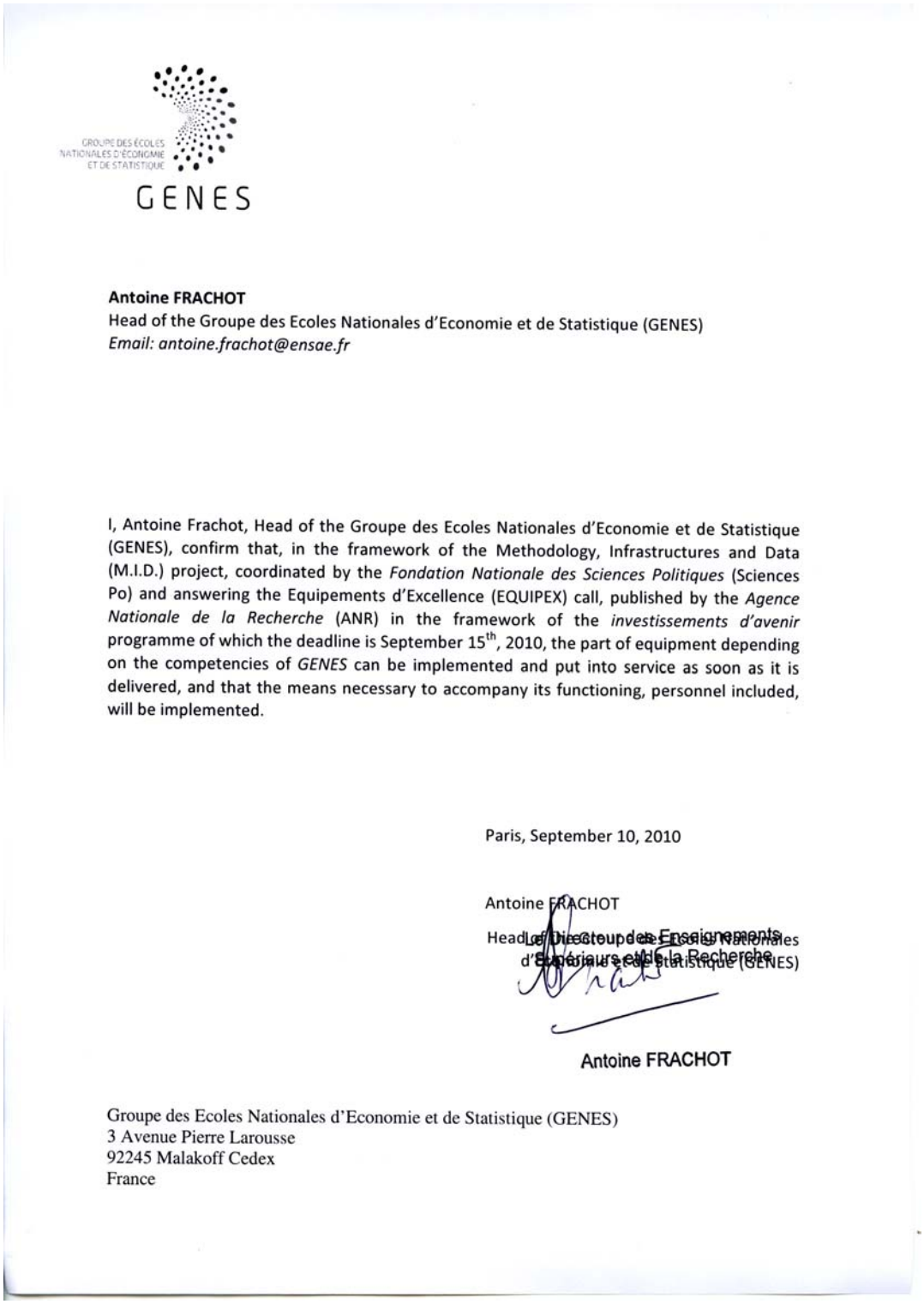
P/O

RÉPUBLIQUE FRANÇAISE
Fondation Nationale des Sciences Politiques
Ordonnance du 9 Octobre 1945

Richard Descoings
Administrator

27, rue Saint-Guillaume 75337 Paris cedex 07 France T/ +33 (0)1 45 49 50 50 www.sciences-po.fr

2. Groupe Nationales des Ecoles d'Economie et de Statistique (GENES)



GROUPE DES ÉCOLES NATIONALES D'ÉCONOMIE ET DE STATISTIQUE
GENES

Antoine FRACHOT
Head of the Groupe des Ecoles Nationales d'Economie et de Statistique (GENES)
Email: antoine.frachot@ensae.fr

I, Antoine Frachot, Head of the Groupe des Ecoles Nationales d'Economie et de Statistique (GENES), confirm that, in the framework of the Methodology, Infrastructures and Data (M.I.D.) project, coordinated by the *Fondation Nationale des Sciences Politiques* (Sciences Po) and answering the Equipements d'Excellence (EQUIPEX) call, published by the *Agence Nationale de la Recherche* (ANR) in the framework of the *investissements d'avenir* programme of which the deadline is September 15th, 2010, the part of equipment depending on the competencies of *GENES* can be implemented and put into service as soon as it is delivered, and that the means necessary to accompany its functioning, personnel included, will be implemented.

Paris, September 10, 2010

Antoine FRACHOT
Head of the Groupe des Ecoles Nationales d'Economie et de Statistique (GENES)

Le Directeur des Enseignements Supérieurs et de la Recherche

Antoine FRACHOT

Groupe des Ecoles Nationales d'Economie et de Statistique (GENES)
3 Avenue Pierre Larousse
92245 Malakoff Cedex
France

3. Institut nationale des études démographiques

INED
Institut national d'études démographiques

I, *Chantal Cases, director of INED*, confirm that, in the framework of the DIME – SHS : "Données, Infrastructures, Méthodes d'Enquêtes en Sciences humaines et sociales" project, coordinated by the *Fondation Nationale des Sciences Politiques* (Sciences Po) and answering the Equipements d'Excellence (EQUIPEX) call, published by the *Agence Nationale de la Recherche* (ANR) in the framework of the *investissements d'avenir* programme of which the deadline is September 15th, 2010, the part of equipment depending on the competencies of *INED* can be implemented and put into service as soon as it is delivered, and that the means necessary to accompany its functioning, personnel included, will be implemented.

Paris, 9/10/2010

Chantal Cases
Director of INED

INSTITUT NATIONAL
D'ETUDES DEMOGRAPHIQUES
133, boulevard Davout
75980 PARIS Cedex 20

INSTITUT NATIONAL D'ETUDES DEMOGRAPHIQUES
ETABLISSEMENT PUBLIC SCIENTIFIQUE ET TECHNOLOGIQUE • 133 BOULEVARD DAVOUT 75980 PARIS CEDEX 20 FRANCE
TEL. 33 (1) 56 06 20 00 • FAX 33 (1) 56 06 21 99

4. University of Paris Descartes

Letter under signature

5. Telecom – ParisTech

To be printed on the official letterhead of the institution
This statement shall be signed by the institution's legal representative and stating his/her name, function and stamp of the institution.

I, *Jean-Claude JEANNERET, General Administrator*, confirm that, in the framework of the Methodology, Infrastructures and Data (M.I.D.) project, coordinated by the *Fondation Nationale des Sciences Politiques* (Sciences Po) and answering the Equipements d'Excellence (EQUIPEX) call, published by the *Agence Nationale de la Recherche* (ANR) in the framework of the *investissements d'avenir* programme of which the deadline is September 15th, 2010, the part of equipment depending on the competencies of *Télécom ParisTech* can be implemented and put into service as soon as it is delivered, and that the means necessary to accompany its functioning, personnel included, will be implemented.

Paris, 10/09/10

Jean-Claude JEANNERET
General Administrator
INSTITUT TELECOM
46, rue Barrault
75634 PARIS cedex 13

Name, function of the legal representative

Jean-Claude JEANNERET
Administrateur Général
Institut TELECOM
46 rue Barrault
75634 PARIS Cedex 13

6. Electricité de France

Nos références HE/E70/10/0046/CM/DW
Interlocuteur C. MULLER – 01.47.65.37.38.
Objet : Lettre de soutien

Clamart, 14 septembre 2010

I, *Corinne Muller, manager of ICAME laboratory (EDF R&D)*, confirm that, in the framework of the Data, Infrastructures and Method Inquiry for Social Sciences, (D.I.M SHS) project, coordinated by the *Fondation Nationale des Sciences Politiques* (Sciences Po) and answering the Equipements d'Excellence (EQUIPEX) call, published by the *Agence Nationale de la Recherche* (ANR) in the framework of the *investissements d'avenir* programme of which the deadline is September 15th, 2010, the means necessary to accompany the functioning of the equipment, personnel included, depending on the competencies of *EDF R&D* will be implemented.

Director Department
Commercial Innovation and Market Analysis,

Corinne MULLER.

EDF R&D
1, avenue du général de Gaulle
92141 CLAMART Cedex
Téléphone +33147653032
Télécopie +33147655700
www.edf.fr
EDF-SA au capital de 8 129 000 000 euros
RCS PARIS B 552 081 317

7. GIS Réseau Quetelet represented by the EHESS

L'ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE
ÉCOLE DES HAUTES ÉTUDES EN SCIENCES SOCIALES
54, BOULEVARD RASPAIL, 75006 PARIS

PC/CC/37-2010

Paris, Tuesday 14th of September 2010

LE PRÉSIDENT

Letter of commitment

I, the undersigned, François Weil, President of the EHESS, the support institution of the Scientific Interest Group "GIS Quetelet", confirm, on behalf of the "GIS Quetelet" its interest for the "Data, Infrastructure, Methods of investigation in the social sciences and humanities (DIME-SHS) project, and welcome this initiative coordinated by the Fondation Nationale des Sciences politiques (Sciences Po). The DIME-SHS project will be submitted for financial support in the frame of the "Equipements d'excellence – Investissements d'avenir" call for proposals published by the Agence Nationale de la Recherche (ANR) – deadline 15th of september.

Concerning the means for the implementation of the DIME-SHS project, the GIS Quetelet is in favour of the allocation of a half part-time support personnel in 2013 for the dissemination activities of the outputs of the project.

François WEIL

TÉLÉPHONE : 01 49 54 25 01 / 02 - TÉLÉCOPIE : 01 49 54 24 96 - E-MAIL : preside@ehess.fr

8. PRES Sorbonne Paris Cité

SORBONNE
PARIS CITÉ

PÔLE DE RECHERCHE
ET D'ENSEIGNEMENT SUPÉRIEUR

www.sorbonne-paris-cite.fr

17 rue de la Sorbonne
75005 Paris

Labellisation d'un Équipex

Après consultation des membres de son Conseil d'orientation scientifique et pédagogique, le bureau du PRES Sorbonne Paris Cité a décidé de conférer son label au projet DIME-SHS, qu'il juge cohérent avec sa stratégie scientifique, et profitable à toute la communauté intéressée.

Le président de Sorbonne Paris Cité,

Jean-François Girard

9. INSEE

INSEE
DIRECTION GÉNÉRALE
INSTITUT NATIONAL
DE LA STATISTIQUE
ET DES ÉTUDES
ÉCONOMIQUES

Direction des Statistiques Démographiques et Sociales

Monsieur Richard Descoings,
directeur de l'IEP Paris,
27 rue Saint-Guillaume
75337 Paris Cedex 07

Dossier suivi par :
Stéfan Lollivier
Tél. : 01 41 17 53 64
Fax : 01 41 17 62 93
Mél : DG75-F001@insee.fr

Paris, le 10 septembre 2010
N° 1801/DG75-F001/

Monsieur,

J'ai lu avec un grand intérêt le dossier que votre institution présente – en tant que partenaire-coordinateur - au titre des Equipements d'Excellence. Je comprends que ce projet permettrait de doter les sciences humaines et sociales françaises de capacités nouvelles de réalisation d'enquêtes, de diffusion de données et de recherche méthodologique compétitives au niveau international.

L'institution que je représente a contribué de longue date à la mise à disposition de ses données d'enquête auprès du monde de la recherche, et s'est même impliquée directement dans la collecte de données en partenariat avec des opérateurs de recherche. Cette seconde modalité se heurtera dans les prochaines années à de fortes contraintes de moyens ; en revanche l'INSEE dispose maintenant du cadre juridique adapté pour fournir des échantillons destinés à des collectes de données réalisées à des fins de statistique publique.

Nous sommes convaincus, après examen du dossier, que le projet que vous présentez, en partenariat avec plusieurs organismes de recherche parmi lesquels l'INED avec qui nous avons une longue tradition de collaboration, pourrait permettre des avancées scientifiques importantes dans la connaissance et la compréhension des comportements des populations, en constituant une base longitudinale de données d'enquêtes indispensables à la recherche en sciences humaines et sociales.
Ainsi, nous apportons notre soutien à ce dossier. et vous confirmons notre accord sur le principe de fournir un échantillon destiné à la constitution de votre panel, dès lors que celui-ci aura été présenté au Cnis et reconnu comme nécessaire à la réalisation d'une enquête de la statistique publique.

Je vous prie d'agréer, monsieur le Directeur, l'expression de la considération distinguée

Cher ami, je me réjouis de cette collaboration !

Le Directeur général de l'Insee

Jean-Philippe COTIS

TIMBRE DG75-F001 - 18 bd Adolphe Pinard - 75675 PARIS CEDEX 14 - FRANCE - www.insee.fr
Tél. standard : 01.41.17.50.50 - N° SIRET : 120 027 016 00019 - Code APE : 84.11Z - Service Insee Contact : 09 72 72 4000 - (tarification "appel local")

RÉPUBLIQUE FRANÇAISE