THEME [INFRA-2011-2.3.1.]
[Implementation of common solutions for
a cluster of ESFRI infrastructures in the
field of "Social Sciences and Humanities"]

Grant agreement for: Combination of CP & CSA[*]

Annex I - "Description of Work"

Project acronym: DASISH

Project full title: " Data Service Infrastructure for the Social Sciences and Humanities "

Grant agreement no: 283646

Version date: 2011-08-15 (including WP2)

Table of Contents

Part A

Workplan Tables

Project Number [1]	283646	Project Acronym [2]	DASISH

One form per project

General information	
Project title [3]	Data Service Infrastructure for the Social Sciences and Humanities
Starting date [4]	01/01/2012
Duration in months [5]	36
Call (part) identifier [6]	FP7-INFRASTRUCTURES-2011-1
Activity code(s) most relevant to your topic [7]	INFRA-2011-2.3.1.: Implementation of common solutions for a cluster of ESFRI infrastructures in the field of "Social Sciences and Humanities"		
Free keywords [8]		Social science, humanities, data archiving, meta data, data quality, persistent identifiers, authentication, single sign on, cross discipline, research infrastructure, integration and interoperability

Abstract [9]

DASISH provides solutions to a number of common issues relevant for the five ESFRI projects in social science and humanities, being CESSDA, CLARIN, DARIAH, ESS and SHARE and therefore is supporting infrastructure construction. These five infrastructures are presently at different stages in the process of establishing ERICs. DASISH has identified four major areas namely data quality, data archiving, data access and legal and ethics. The activities in these four major areas are based on a thorough analysis of the underlying architectures. The outcome forms the basis for educational activities and for outreach to the communities that are to benefit from the work.

Through DASISH the participating infrastructures will not only obtain new solutions for specific problems and a consolidation of their infrastructure building, but will work out solutions facilitating interdisciplinary cross-walks of their researchers. This will be of mutual benefit for the five infrastructures and the communities they serve.

Project Number [1]		283646	Project Acronym [2]		DASISH	

<div align="center">**List of Beneficiaries**</div>

No	Name	Short name	Country	Project entry month[10]	Project exit month
1	GOETEBORGS UNIVERSITET	UGOT	Sweden	1	36
2	MAX PLANCK GESELLSCHAFT ZUR FOERDERUNG DER WISSENSCHAFTEN E.V.	MPG	Germany	1	36
3	UNIVERSITAT POMPEU FABRA	UPF	Spain	1	36
4	Københavns Universitet	UCPH	Denmark	1	36
5	KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN - KNAW	KNAW-DANS	Netherlands	1	36
6	KING'S COLLEGE LONDON	KCL	United Kingdom	1	36
7	GEORG-AUGUST-UNIVERSITAET GOETTINGEN STIFTUNG OEFFENTLICHEN RECHTS	UGOE	Germany	1	36
8	OESTERREICHISCHE AKADEMIE DER WISSENSCHAFTEN	OEAW	Austria	1	36
9	UNIVERSITY OF ESSEX	UEssex	United Kingdom	1	36
10	TAMPEREEN YLIOPISTO	UTA	Finland	1	36
11	NORSK SAMFUNNSVITENSKAPELIG DATATJENESTE AS	NSD	Norway	1	36
12	THE CITY UNIVERSITY	CITY	United Kingdom	1	36
13	GESIS - LEIBNIZ INSTITUT FUR SOZIALWISSENSCHAFTEN e.V.	GESIS	Germany	1	36
14	NATIONAL UNIVERSITY OF IRELAND MAYNOOTH	NUIM	Ireland	1	36
15	UNIVERSITETET I BERGEN	UiB	Norway	1	36
16	UNIVERSITA CA' FOSCARI VENEZIA	UNIVE	Italy	1	36
17	STICHTING CENTERDATA	CentERdata	Netherlands	1	36
18	TARTU ULIKOOL	UT	Estonia	1	36

A3:
Budget Breakdown

Project Number [1]	283646		Project Acronym [2]	DASISH

One Form per Project

Participant number in this project [11]	Participant short name	Fund. % [12]	Ind. costs [13]	Estimated eligible costs (whole duration of the project)						Total receipts	Requested EU contribution
				RTD (A)	Coordination (B)	Support (C)	Management (D)	Other (E)	Total A+B+C+D		
1	UGOT	75.0	T	226,195.20	276,460.80	0.00	364,492.80	0.00	867,148.80	0.00	719,021.00
2	MPG	75.0	S	1,010,509.00	682,677.00	0.00	135,879.00	0.00	1,829,065.00	0.00	1,211,405.00
3	UPF	75.0	T	219,715.20	33,801.60	0.00	0.00	0.00	253,516.80	0.00	187,390.00
4	UCPH	75.0	T	131,648.00	157,664.00	0.00	22,752.00	0.00	312,064.00	0.00	226,925.00
5	KNAW-DANS	75.0	T	176,544.00	215,776.00	0.00	17,712.00	0.00	410,032.00	0.00	294,420.00
6	KCL	75.0	T	70,368.00	258,016.00	0.00	0.00	0.00	328,384.00	0.00	225,324.00
7	UGOE	75.0	T	66,416.00	308,360.00	0.00	17,232.00	0.00	392,008.00	0.00	273,259.00
8	OEAW	75.0	A	173,811.00	122,690.00	0.00	0.00	0.00	296,501.00	0.00	228,327.00
9	UEssex	75.0	T	85,299.20	109,670.40	0.00	0.00	0.00	194,969.60	0.00	137,316.00
10	UTA	75.0	T	58,800.00	202,860.80	0.00	0.00	0.00	261,660.80	0.00	179,763.00
11	NSD	75.0	S	411,096.00	318,600.00	0.00	18,417.00	0.00	748,113.00	0.00	565,131.00
12	CITY	75.0	T	477,481.60	91,147.20	0.00	19,084.80	0.00	587,713.60	0.00	438,149.00
13	GESIS	75.0	T	346,104.00	328,624.00	0.00	3,000.00	0.00	677,728.00	0.00	482,345.00
14	NUIM	75.0	T	0.00	31,308.80	0.00	0.00	0.00	31,308.80	0.00	20,937.00
15	UiB	75.0	T	0.00	216,230.40	0.00	0.00	0.00	216,230.40	0.00	144,604.00
16	UNIVE	75.0	T	284,640.00	0.00	0.00	0.00	0.00	284,640.00	0.00	213,480.00
17	CentERdata	75.0	T	332,291.20	150,067.20	0.00	3,000.00	0.00	485,358.40	0.00	352,575.00
18	UT	75.0	T	65,856.00	61,740.80	0.00	0.00	0.00	127,596.80	0.00	90,681.00
Total				4,136,774.40	3,565,695.00	0.00	601,569.60	0.00	8,304,039.00	0.00	5,991,052.00

Note that the budget mentioned in this table is the total budget requested by the Beneficiary and associated Third Parties.

* The following funding schemes are distinguished

Collaborative Project (if a distinction is made in the call please state which type of Collaborative project is referred to: (i) Small of medium-scale focused research project, (ii) Large-scale integrating project, (iii) Project targeted to special groups such as SMEs and other smaller actors), Network of Excellence, Coordination Action, Support Action.

1. Project number

The project number has been assigned by the Commission as the unique identifier for your project, and it cannot be changed. The project number **should appear on each page of the grant agreement preparation documents** to prevent errors during its handling.

2. Project acronym

Use the project acronym as indicated in the submitted proposal. It cannot be changed, unless agreed during the negotiations. The same acronym **should appear on each page of the grant agreement preparation documents** to prevent errors during its handling.

3. Project title

Use the title (preferably no longer than 200 characters) as indicated in the submitted proposal. Minor corrections are possible if agreed during the preparation of the grant agreement.

4. Starting date

Unless a specific (fixed) starting date is duly justified and agreed upon during the preparation of the Grant Agreement, the project will start on the first day of the month following the entry info force of the Grant Agreement (NB : entry into force = signature by the Commission). Please note that if a fixed starting date is used, you will be required to provide a detailed justification on a separate note.

5. Duration

Insert the duration of the project in full months.

6. Call (part) identifier

The Call (part) identifier is the reference number given in the call or part of the call you were addressing, as indicated in the publication of the call in the Official Journal of the European Union. You have to use the identifier given by the Commission in the letter inviting to prepare the grant agreement.

7. Activity code

Select the activity code from the drop-down menu.

8. Free keywords

Use the free keywords from your original proposal; changes and additions are possible.

9. Abstract

10. The month at which the participant joined the consortium, month 1 marking the start date of the project, and all other start dates being relative to this start date.

11. The number allocated by the Consortium to the participant for this project.

12. Include the funding % for RTD/Innovation – either 50% or 75%

13. Indirect cost model
   A: Actual Costs
   S: Actual Costs Simplified Method
   T: Transitional Flat rate
   F :Flat Rate

Workplan Tables

Project Number [1]	283646	Project Acronym [2]	DASISH

WP Number [53]	WP Title	Type of activity [54]	Lead beneficiary number [55]	Person-months [56]	Start month [57]	End month [58]
WP 1	Management	MGT	1	44.00	1	36
WP 2	Architecture and Quality Assessment	COORD	5	83.00	1	36
WP 3	Data Quality	RTD	12	199.50	1	36
WP 4	Data Archiving	COORD	11	67.00	1	36
WP 5	Shared Data Access & Enrichment	RTD	2	171.00	1	36
WP 6	Legal and Ethical Issues	COORD	2	68.00	1	36
WP 7	Education and Training	COORD	7	56.00	1	36
WP 8	Dissemination	COORD	4	34.00	1	36
			Total	722.50		

WT2:
List of Deliverables

Project Number [1]	283646	Project Acronym [2]	DASISH

List of Deliverables - to be submitted for review to EC						
Delive-rable Number [61]	Deliverable Title	WP number [53]	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D1.1	Project forms signed by partners	1	1	1.00	O	PP	1
D1.2	Establish reporting procedures	1	1	2.00	O	PP	2
D1.3	Mid-term management report	1	1	2.00	R	RE	18
D1.4	Final project management report	1	1	2.00	R	RE	36
D2.1	State-of-the-Art Report	2	5	9.00	R	PU	6
D2.2	Reference Architecture	2	5	34.00	D	PU	36
D2.3	Tools & Service Knowledge Registry	2	5	38.00	D	PU	36
D2.4	Quality Assessment of Reference Architecture	2	5	2.00	R	PU	27
D3.1	CAPI tool for occupation coding and user guide	3	12	10.00	D	PU	36
D3.2	Web-based tool for occupation coding and user guide	3	12	10.00	D	PU	36
D3.3	Operational question database and plug in	3	12	10.00	D	PU	31
D3.4	Operational questionnaire design & translation tool / databank	3	12	10.00	D	PU	34
D3.5	Standardized language	3	12	10.00	D	PU	36

Delive-rable Number [61]	Deliverable Title	WP number [53]	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
	management utility plug-in						
D3.6	Design of standardized sample management system	3	12	10.00	O	PU	24
D3.7	Report on keystroke analysis and implications for field work	3	12	3.00	R	PU	36
D3.8	Fieldwork monitoring application for decentralised surveys	3	12	5.00	R	PU	36
D4.1	Roadmap for Preservation and Curation in SSH	4	11	3.00	R	PU	6
D4.2	Report about Preservation Service Offers	4	11	3.00	R	PU	9
D4.3	List of Recommended Deposit Services for SSH	4	11	3.00	R	PU	33
D4.4	Comprehensive Policy-Rules for Data Management in SSH	4	11	3.00	R	PU	33
D5.1	Trust and PID Federation Services Report	5	2	6.00	R	PU	24
D5.2	Metadata Quality and Portal Progress Report	5	2	6.00	R	PU	30
D5.3	Workflow Requirements and Application Report	5	2	3.00	R	PU	36
D5.4	Annotation Framework Report	5	2	3.00	R	PU	30

Delive-rable Number [61]	Deliverable Title	WP number [53]	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D6.1	Report about New IPR Challenges	6	2	3.00	R	PU	18
D6.2	Sample merged paradata sets	6	2	3.00	D	PU	18
D6.3	Exemplary analyses of confidential paradata	6	2	3.00	D	PU	36
D6.4	Midterm report on the establishment of virtual centres	6	2	3.00	R	PU	18
D6.5	Handbook on legal and ethical issues for SSH data in Europe	6	2	4.00	O	PU	36
D6.6	Report about Preservation Policy-Rules	6	2	3.00	R	PU	30
D7.1	course modules	7	7	4.00	O	PU	36
D7.2	a compendium of workshop reports	7	7	3.00	O	PU	36
D8.1	Dissemination Strategy	8	4	2.00	R	PP	3
D8.2	Web-site up and running	8	4	2.00	O	PU	3
D8.3	Newsletter, bi-annual	8	4	6.00	O	PU	36
D8.4	Fact Sheets, ongoing, based on need	8	4	3.00	R	PU	36
D8.5	Conference list, ongoing	8	4	3.00	R	PP	36
D8.6	Final Conference	8	1	2.00	O	PU	36
D8.7	Qualitative Data Workshop Report	8	14	1.00	O	PU	6
D8.8	Quantitative Data Workshop Report	8	13	1.00	R	PU	35

Delive-rable Number [61]	Deliverable Title	WP number [53]	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
			Total	234.00			

Project Number [1]	283646	Project Acronym [2]	DASISH

One form per Work Package			

Work package number [53]	WP1	Type of activity [54]	MGT
Work package title	Management		
Start month	1		
End month	36		
Lead beneficiary number [55]	1		

Objectives

WP1 globally aims at ensuring a strong and efficient day-to-day management so that the project meets its objectives on time and within budget constraints. Important objectives within WP1 are:
• Ensure administrative and financial management of the project;
• Develop a spirit of co-operation between the partners;
• Ensure consensus management among the partners;
• Ensure work progress
• Keep the activities within the objectives;
• Ensure project reporting.
• Provide a smooth interface with the European Commission.

Description of work and role of partners

Task 1.1- Project coordination
This task encompasses the essential aspects of leadership and project management required for the successful planning, undertaking and finalising of any project. These aspects are especially important given the project will deliver 8 work packages, many with synergies to each other. The project involves 5 infrastructures and a total of 19 participants, and across 12 countries. In fulfilling this objective, monitoring of progress against agreed deliverables and milestones will be an obvious requirement to ensure that individual work packages effectively contribute to the overall timely delivery of the project. This will require the establishment of effective project management and communication structures and tools.
UGOT will manage the project. It will establish and maintain the central project office and undertake the full range of managerial, administrative, supporting and co-ordinating activities necessary and appropriate to a project of this size.
The DASISH scientific coordinator Hans Jørgen Marker, HJM, is an experienced project leader furthermore UGOT (SND) will employ a half time project manager and have the support of all relevant UGOT offices. HJM is assisted by Peter Wittenburg, PW, who serves as RTD coordinator. The HJM is taking care of the work in WP1, WP6, WP7 and WP8 and PW that of WP2, WP3, WP4 and WP5.
The Executive board EB in which all WP leaders are represented including HJM and PW who also lead WP1 and WP5 takes the relevant decisions to achieve the intended results.
The Advisory Board, AB, will include five international experts from the SSH domains to give advice about the relevance of the work being done.
The Strategic board, SB, will cover delegates of the 5 communities (ERICs when they have been established) and also delegates from the EC.
The work to be carried out within this task consists of:
• Ensuring quality management and coordination of the project;
• Preparing and leading of project meetings and reviews (propose agendas, prepare minutes). The meetings of the advisory board and the strategic board will be directly managed and financed by the coordinator
• Managing the fund transfers towards the partners;
• Dealing with contractual matters: maintaining the technical description of the work, the Consortium Agreements and any Contract Amendments;
• Handling the Cost Claim procedures and maintaining the financial budget status of each partner;

• Providing a management guideline specifying: the detailed roles and responsibilities of different committees, the methods and approaches for deliverables' quality insurance, the guidelines for progress reports, the guidelines for document management, templates, etc.;
• Giving guidance on all administrative matters;
• Supervising WP/tasks management in collaboration with WP leaders;
• Controlling the work progress by requiring monthly reports to the WP leaders: work progress,, problems, etc. in order to be able to be reactive and find solutions
• Providing progress and activity reports;
• Supervising the project contractual deliverables, keeping within the timeframes and budgets and ensuring that reports are prepared and submitted to the Commission in accordance with agreed procedures;
• Enabling Quality Assurance: to assure that the development process follows the quality rules for the project, to control actions planned and the time schedules, responsibilities and authorities need to be clearly defined;

Task 1.2 - The junction between strategic and operational level

The Executive board, EB, will assist in making the junction between the strategic and operational level. The EB will assist in following-up the project progress according to the work plan as defined, of preparing decisions that may be necessary to be taken if difficulties arise. In this sense, the management team together with the EB is at the junction between the strategic and operational level.

The Strategic board will provide the very crucial link to the five participating infrastructures (and as they begin to exist the respective ERICs)

Meetings and events:

The consortium meeting includes all partners and controls the work carried out at strategic level. There will be 3 consortium meetings, involving 30 participants each (1 representative from each of the 19 partners + project manager + 5 Advisory board members + 5 Strategic board members): the kick-off meeting and an consortium at the start of the second and third year. The organisation costs for the consortium meetings will be directly managed and financed by the coordinator. There will be also 2 executive meetings each year with SB members (8 participants). One of them will take place jointly with the annual consortium meeting.

Joint working groups and workshops will bring experts together from the interested communities to discuss specific issues of common interest and to synchronize activities.

At the end of the project there will be a final conference which be arranged under WP8

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
1	UGOT	27.00
2	MPG	9.50
4	UCPH	1.50
5	KNAW-DANS	1.50
7	UGOE	1.50
11	NSD	1.50
12	CITY	1.50
	Total	44.00

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D1.1	Project forms signed by partners	1	1.00	O	PP	1
D1.2	Establish reporting procedures	1	2.00	O	PP	2
D1.3	Mid-term management report	1	2.00	R	RE	18
D1.4	Final project management report	1	2.00	R	RE	36
	Total		7.00			

Description of deliverables

D1.1) Project forms signed by partners: Project forms signed by partners [month 1]

D1.2) Establish reporting procedures: Establish reporting procedures [month 2]

D1.3) Mid-term management report: Mid-term management report [month 18]

D1.4) Final project management report: Final project management report [month 36]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
MS1	Kick-off meeting	1	1	
MS26	Second consortium meeting	1	13	
MS27	Third consortium meeting	1	25	

WT3:
Work package description

Project Number [1]	283646	Project Acronym [2]	DASISH

Work package number [53]	WP2	Type of activity [54]	COORD
Work package title	Architecture and Quality Assessment		
Start month	1		
End month	36		
Lead beneficiary number [55]	5		

Objectives

• To acquire and make available knowledge and expertise on both the current data architectures and tools and technologies in use within all five domains and liaise with ongoing international initiatives working on data preservation and access issues – in particular the emerging data e-Infrastructure.
• To define & design a reference architecture, and to adapt it based on further insights gained by work packages 3, 4 and 5.
• To acquire an extensive list of requirements underlying the reference architecture, focused on integration and technical interoperability between heterogeneous systems, and determine which of these requirements would benefit the architecture and the domains the most.
• To provide a blueprint of the architecture and an assessment/validation procedure for the realization of an integrated architecture across the five domains based on existing components and foreseen results of work packages 3, 4 and 5.

Description of work and role of partners

Task 2.1 State of the Architectures Report (Leader: UEssex; Partner: UGOT, KCL)
The architectural solutions for data management, curation and access in the different SSH infrastructures must meet the requirements of the communities that they serve and are therefore very heterogeneous in nature. To make progress in creating a more integrated and interoperable SSH data domain DASISH needs to understand the current solutions to the community needs in great detail. It is likely that aspects of the research & data lifecycles for all the SSH communities coincide at certain points and for certain requirements, however these points of correlation have yet to be identified. This task will identify points of commonality in the five SSH infrastructures' community research & data lifecycles needs. Furthermore, current state-of-the-art solutions and best practices, in e-infrastructure for each community, will be identified and described. The goal is to describe the current solutions in each SSH community in a consistent manner which shall in turn inform the development of the reference architecture

The output of this task will be an analysis report of common infrastructure requirements that is the baseline for the integration and interoperability work.

The procedure will be as follows: (1) Analyse the current state of technology in use in all 5 disciplines; (2) Identify architectural commonalities (if any) and state-of-the-art solutions to community needs; (3) Describe the solutions in a uniform form; (4) Suggest potential architectures that could be utilized in the Reference Architecture.

Task 2.2: Reference Architecture (Leader: DANS; Partner: UGOE, MPIPL, UGOT, UEssex)
To improve integration and interoperability between the SSH infrastructures, a reference architecture will be created based upon the analysis of architectures (Task 1). This architecture will take into consideration the current heterogeneous nature of the existing architectures as well provide a direction for coherence, integration and interoperability. The aim of the reference architecture is not to impose a single architecture that requires all infrastructures to follow, but rather to create an architectural framework that facilitates both inter-infrastructure access to data as well as maintaining the unique needs, requirements and essence of each community's infrastructure – an architecture that offers a roadmap towards harmonization.

Since DASISH wishes to encourage the SSH communities to move towards increased compatibility, it must also ensure that it consistently follows its own reference architecture. To improve interoperability and standardization

of the activities carried out by DASISH this task will additionally undertake an ongoing evaluation of the activities of the work packages against the reference architecture.

The output of this task will be a reference architecture, in an open standard architecture format, against which DASISH activities and the SSH infrastructures' activities can be evaluated. Assessment reports evaluating DASISH activities against the reference architecture will be made during the planning stage of the activity.

The procedure for the reference architecture will be as follows: (1) Review the report from Task 1; (2) Create a reference architecture that maps as closely as possible to architectures identified in the state-of-the-art report; (3) Elicit expert feedback to the reference architecture from the SSH infrastructures; (4) When a planned activity is ready for assessment it will be assessed against the reference architecture and recommendations made to the activity lead partner if required. The assessment will in no way halt the activities progress.

Task 2.3: Tools & Services Knowledge Registry (Leader: KCL; Partner: UGOT, UPF, FSD, CentERdata, UT)
WP2 will work to build a cross-community methodological, process & tools knowledge registry aimed at research-projects in the SSH. The knowledge registry can best be regarded as a place where researchers from all disciplines as well as IT personnel can share their experiences with tools and services they have used while working on e.g. data acquisition and analysis, annotation, data aggregation and cross-referencing, etc. As such, it will provide a focal point for interdisciplinary knowledge transfer.
The knowledge registry will build upon, and aggregate, the expertise and knowledge already gathered by the participating SSH infrastructures plus it will include information about additional tools and methodologies that are not specific to any community. It will contain information on the applicability of tools for particular disciplines, how to integrate tools into technical environments, the standards being used (e.g. with regard to metadata formats, APIs and dissemination protocols), previous experiences from researchers and IT experts, and an assessment of the tools overall quality according to predefined benchmarks. Where available it will be based on existing metadata descriptions of the communities and be closely linked with the joint metadata registry being developed in WP5.
Content delivery to the knowledge registry will be based on a wider community effort. In order to spark general interest and foster the community building process, the task participants will provide initial contributions. To ensure ease of contribution a suite of existing social media tools and methodologies will be employed as well as aggregators and a knowledgebase platform.
Close collaboration with Work Packages 7 & 8 is required, as the results of this task will feed directly into education and training and will be dependent on effective and coordinated dissemination.

The output of this task will be an attractive portal with community provided general and technical information, as well as moderated and living discussion about tools and services and their use within the SSH domain.

The procedure will be as follows: (1) Define the subject areas that the knowledge-base will include and assess user requirements; (2) Assessment of existing SSH infrastructure knowledge-bases and the opportunities to aggregate specific information; (3) Assess what tools researchers use to communicate and disseminate knowledge; (4) By making use of standard technologies the portal will be rolled out and linked with the joint metadata registry in WP5; (5) Start a grassroots community around the portal by providing initial contributions (based on experiences of partner organizations) and applying social media platforms for widespread dissemination. This initial community building effort will be based on a beta-test phase for the platform; (6) Official launch of the knowledge registry platform; (7) Well-described and proven tools/services will be promoted by DASISH and will be integrated in educational and PR campaigns to spread its usage and users will be motivated to provide usage descriptions and ratings.

Task 2.4: Quality Assessment of Reference Architecture (Leader: DANS; Partner UGOT)
This task will assess whether the reference architecture matches the broader quality requirements for research infrastructures within the EU, whereas task 2.2 assess whether the activities of DASISH match the requirements of the reference architecture. With these two assessments, all the activities of DASISH will be assessed for consistency, compatibility and quality.
The objective of this task is to ascertain the quality requirements for research infrastructures, ensure that this information is fed into the creation of the reference architecture, and evaluate the final reference architecture against the quality guidelines.

The output of this task will be a report on the quality of the reference architecture, including recommendations for SSH infrastructures to meet guidelines and quality requirements for European Research Infrastructures.

The procedure will be as follows: (1) Identify all research infrastructure quality assurance requirements; (2) Establish a set of criteria with which to assess the reference architecture; (3) Assess the reference architecture against the quality criteria; (4) Report on the quality assessment of the reference architecture, including recommendations for SSH

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
1	UGOT	14.00
2	MPIPL	4.00
3	UPF	4.00
5	KNAW-DANS	17.00
6	KCL	15.00
7	UGOE	4.00
9	UEssex	9.00
10	UTA	5.00
18	CentERdata	3.00
19	UT	8.00
	Total	83.00

List of deliverables

Deliverable Number [61]	Deliverable Title	Lead beneficiary number	Estimated indicative person-months	Nature [62]	Dissemination level [63]	Delivery date [64]
D2.1	State-of-the-Art Report	5	9.00	R	PU	6
D2.2	Reference Architecture	5	34.00	D	PU	36
D2.3	Tools & Service Knowledge Registry	5	38.00	D	PU	36
D2.4	Quality Assessment of Reference Architecture	5	2.00	R	PU	27
		Total	83.00			

Description of deliverables

D2.1) State-of-the-Art Report: State-of-the-Art Report [month 6]

D2.2) Reference Architecture: Reference Architecture [month 36]

D2.3) Tools & Service Knowledge Registry: Tools & Service Knowledge Registry [month 36]

D2.4) Quality Assessment of Reference Architecture: Quality Assessment of Reference Architecture [month 27]

Schedule of relevant Milestones				
Milestone number [59]	Milestone name	Lead beneficiary number	Delivery date from Annex I [60]	Comments
MS2	Integration of investigation activities into architecture	5	18	
MS3	Registry implemented and linked with metadata portal	5	24	
MS4	Content from the community & infrastructures added to the registry	5	36	

WT3:
Work package description

Project Number [1]	283646	Project Acronym [2]	DASISH

Work package number [53]	WP3	Type of activity [54]	RTD
Work package title	Data Quality		
Start month	1		
End month	36		
Lead beneficiary number [55]	12		

Objectives

1. To develop software for Europe-wide surveys to collect and code occupation data more accurately, consistently and cost-effectively.
2. To create, refine and make widely available a multi-language questionnaire development, translation, documentation database tool to aid the design and implementation of cross-cultural research and ensure more effective access to questionnaire metadata.
3. To build up and computerise a fieldwork monitoring system for cross-national surveys which can centrally manage their distributed fieldwork, helping to optimise response rates and data quality.

Description of work and role of partners

While improved data protocols and data enrichment are essential elements of better science in the 21st century, they would have only limited utility in the absence of improving data quality itself. Too many important cross-national datasets have yet to take advantage of the remarkable gains in efficiency, precision and accessibility that can be derived from better use of technology. This work package aims to develop a set of tools that will allow these advantages to be shared across European data infrastructures and other survey programmes. WP8 contains specific proposals under DASISH for facilitating consultation with other large survey infrastructures in Europe about the tasks in WP3. It is anticipated these meetings would play a crucial role in helping to shape priorities and encourage adoption of the tools developed across the different cross-national survey programmes in Europe.

Many of the most rigorous sources of cross-national data in Europe – for example Eurostat, CESSDA, Eurobarometer, ESS, SHARE, European Values Surveys, European Election Study, European Quality of Life Survey and the Gender and Generations Survey – are aware of common imperfections within their data and of a persistent lack of equivalence between certain kinds of data in different nations. Both SHARE and ESS have attempted to understand the effects of cross-national differences in methodology on data quality (Cheshire, Ofstedal and Schröder 2010, Saris 2007, Stoop et al 2010). Our conclusion is that, while great strides have been made over the years, various underlying problems are still in need of robust remedies. If these can be delivered, they would be of longstanding value for both research and governance. We propose to tackle three of these key problems, each with a realistic chance of formulating remedies that would bring considerable benefit to a range of multinational surveys. This work package will be directed by the ESS team, with the work jointly undertaken by SHARE, CESSDA and ESS institutions, together with some elements to which other institutions will contribute. It comprises three principal tasks, as follows:

Description of work
Task 3.1: Develop software for improved coding of occupation (Task leader: City)
We will work with the developers of existing computer-assisted coding tools to produce software solutions which enable the accurate coding of occupation at the time of data collection. This would involve two different types of module. The first tool would be designed to be embedded in CAPI programs and would allow interviewers to select the 'best match' for the textual information they get from respondents in their different languages. A second tool would be for use on web surveys, building on the 'occupation search tree' that was developed by the 'Euroccupations' project.

Much serious social science data analysis relies heavily on the analyst's ability to divide samples according to the respondents' position in the occupational division of labour. It is for this reason that sustained transnational

efforts have been made to compile a detailed International Standard Classification of Occupations (ISCO). ISCO allocates each of many thousands of job titles into 433 unit groups that are subsequently grouped into a much smaller number of categories (e.g. social class) for detailed socio-economic analysis (Rose and Harrison 2010). Nonetheless, system differences between countries and human error often still seriously compromise their cross-national validity (ibid).

The basic data needed to derive social class are occupation, establishment size and employment status. Although often described by academics as a 'convenient proxy', this information requires up to eight questionnaire items – and a substantial amount of probing by trained interviewers - to capture it. Open format responses subsequently require expert coding into occupational unit groups. The trend away from paper and pencil interviewing towards computer-assisted and web-based surveys means that solutions to problems affecting quality are now more within reach than ever before. Computer-aided interviews make complicated, conditional routing possible through a complex but programmable tree structure of occupations. Automation has become even more necessary than before, as in web surveys interviewers are not present to facilitate the process.

Ground-clearing work has started on some of the technical problems, e.g. the CASCOT package and the Wageindicator web-survey, but these are largely developmental and by no means fully implemented in the field cross-nationally. SHARE has made parallel efforts but the problems detailed above persist. ESS data show that as few as 128 occupation codes actually cover 80% of respondents. So even if only these 128 could be coded via interactive computer methods, it would greatly reduce the considerable labour costs involved in their collection and coding, and greatly increase the chances of an accurate and equivalent classification cross-nationally. This task will build two computer-assisted tools for this purpose – one within CAPI and one web-based – to guide the respondent through a funnelling process leading to the best occupational title and, in turn to ISCO categories. CITY and UNIVE will supplement and synthesise the work already done by ESS and SHARE, both in the development of computerised coding tools and through participation in the cross-national Euroccupations project. A subcontract will be entered into by CITY for this development work which will be carried out in cooperation with the three social science infrastructures. We will produce for cross-national surveys an interactive system that will provide more comparable cross-national tools than are currently available. The subcontract will include specific programming for a coding module that could be run within any CAPI language. SHARE does already operate a single centralised CAPI system, the code produced under the subcontract will be written for this as directed by SHARE. The code will be open source. We will produce for cross-national surveys an interactive system that will provide more comparable cross-national tools than are currently available. The quality of the output from the different procedures (human coding, CAPI-CASCOT, web-tree) will be tested by UPF. CentERdata will oversee the programming of the CAPI and web-based tools.

Task 3.2: Develop a multi-language questionnaire development tool (Task leader: City)
Our aim is to create an advanced operational system for developing, translating and documenting the concepts and questions used in cross-national survey programmes for improving both quality and transparency. A key aim will be to ensure optimal interoperability across cross-national surveys in Europe via CESSDA.

The last two decades have seen advances in cross-national central coordination which have led to higher data quality (Jowell et al 2007). A key dimension of these quality gains has been improved transparency throughout the survey life cycle. This in turn enhances the quality of subsequent analysis because data users then acquire access to the premises and compromises involved in questionnaire construction. SHARE has developed a 'Language Management Utility (LMU)' which partially automates the translation task and caters for their 27 (or so) languages. The ESS has introduced state of the art design and documentation procedures, quantitative assessments of question quality, and an innovative committee-based translation system. However these fledgling initiatives remain quite separate from one another. Improvements and better integration remain crucial to the future development of survey practice. The work task will produce three compatible databanks that could stand as realistically adoptable models model for other cross-national survey vehicles. The aim would be to produce searchable meta data for the whole questionnaire design process on cross--national surveys from the initial conceptualization, through the design and pre-pretesting phases, onto the translation phase and finally to the question wording that is adopted across different countries and languages. No such comprehensive tool currently exists and this task will make a unique and ground breaking contribution to the field.

a) Questionnaire Design Documentation Databank
While good survey questions are fundamental prerequisites for good data, their development is often poorly documented or unavailable which is counter to the principles of good science. The ESS has developed and used a template completed during the process of questionnaire development (Fitzgerald, 2009), which reveals

to data users the rationale behind the design decisions made en route. But the template is neither automated nor interactive and needs painstaking software development to remedy. The new tool will be designed to make its outputs usable during the design process and will feed directly into translation and question databanks (b and c) below). The tool will also document the results of pre-testing on items developed to measure the required concepts and will thus provide a much needed source of evidence for other scientists in the field who want to address similar topics in the future. It will provide an open source 'plug in' accessible to other surveys. The design tool will be specified and tested by City who will also consider the questionnaire design procedures and needs of other cross-national survey programs in Europe. FSD will also provide input into the specification. NSD will be responsible for ensuring its compatibility with DDI and for the programming.

b) Translation Tool and Databank

ESS and SHARE have led the development of improved translation procedures, many of which have now been adopted by other cross-cultural surveys. It is effectively a 'committee-based' system, aiming to achieve optimal translations via staged perusal and amendment by different actors, whose interaction plays a major role in ensuring cross-cultural equivalence. This Task will start with SHARE's Language Management Utility (LMU) and will develop into a comprehensive translation tool to facilitate and faithfully document the process (Börsch-Supan & Jürges 2005). Its outputs will be usable during the translation design process and will then feed into the Question Databank. It will also be an open source 'plug in' available to other cross-cultural surveys. GESIS will be responsible for specifying and testing the tool in collaboration with MPG (MEA) and FSD, CentERdata will be responsible for programming whilst NSD will advise on compatibility with DDI and the Questionnaire Design Documentation and Question Databanks.

c) Question Databank

Access to survey questions in their original languages is often severely limited. So this task will create a fully-searchable database of translated questions which would be available to data users. Developed initially for ESS it will be designed to serve as a model for other surveys. It would include questions in all languages, associated interviewer instructions and meta data. The database would be constructed according to the long-term strategies for data sharing in the European Research Area which are being established through CESSDA (Gregory et al 2009). Once adopted by other research projects, users would be able, via the CESSDA network, to search across surveys for possible common solutions. NSD will be responsible for specifying the tool requirements in collaboration with City, UPF and FSD. The question database will be programmed with NSD which will ensure DDI compliance.

Task 3.3: Fieldwork monitoring system (Task leader: UMA)

Survey houses throughout Europe have different methods of real-time monitoring of fieldwork and potential response rate problems. This is a particular problem in cross-national surveys as early warning of differential response rate problems is difficult to achieve before the problems become entrenched and damage data quality. A centrally-administered control mechanism is thus necessary and, although within reach, it has not yet been constructed successfully. In addition, such a system needs to accommodate the different sorts of survey mechanisms that exist in different cross-national surveys. For instance whilst some cross-national surveys are able to use a single sample management and CAPI questionnaire system, others are still forced to rely on more distributed sample systems and paper and pencil questionnaires.

To address this problem, this task will draw on the work conducted by SHARE under previous framework programmes. SHARE developed its own "sample management system" (SMS) which all SHARE countries use. Now in its fourth revision, it works well even in those countries in which there was no precedent of computer-controlled processing. But at present the system is tailor-made for SHARE alone. So this Task will produce a very detailed design for a transportable, standardised system of monitoring employing harmonised metadata files which can aid central field supervision, control and monitoring. It will be of help to data quality and equivalence in other pan-European surveys too. And since, for a number of technical and organizational reasons, neither the ESS nor some other cross-national surveys will be able to take immediate advantage of such a shared electronic sample management system we will thus develop a more truncated but still automated centralized system. This will draw on the existing SMS and enable crucial elements of fieldwork progress to be fed into a central database accessible both to the central coordinators and the participating countries. MPG (MEA) will direct this task with City and GESIS ensuring that it is suitable for other surveys too.

In addition, SHARE already makes use of interview process data collected via keystrokes made on interviewers' PCs during fieldwork. This monitors how faithfully questions are administered, the duration of the questionnaire and, as a by-product, interviewer cheating. These key stroke files are augmented with interviewer characteristics such as age and education, and by neighbourhood characteristics to create a powerful tool for quality

assessment. This large and helpful source of data that were gathered in waves 1-3 of SHARE have, however, not been analysed beyond immediate SHARE applications owing to a lack of resources. This Task will therefore include such analyses and will deliver generally applicable "lessons learned" to inform better survey management for all pan-European surveys. GESIS will seek to analyse the more limited elements of these data that are applicable to the ESS and other cross-national surveys

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
2	MPG	26.50
3	UPF	12.00
10	UTA	5.00
11	NSD	40.00
12	CITY	38.00
13	GESIS	17.00
16	UNIVE	30.00
17	CentERdata	31.00
	Total	199.50

List of deliverables

Deliverable Number [61]	Deliverable Title	Lead beneficiary number	Estimated indicative person-months	Nature [62]	Dissemination level [63]	Delivery date [64]
D3.1	CAPI tool for occupation coding and user guide	12	10.00	D	PU	36
D3.2	Web-based tool for occupation coding and user guide	12	10.00	D	PU	36
D3.3	Operational question database and plug in	12	10.00	D	PU	31
D3.4	Operational questionnaire design & translation tool / databank	12	10.00	D	PU	34
D3.5	Standardized language management utility plug-in	12	10.00	D	PU	36
D3.6	Design of standardized sample management system	12	10.00	O	PU	24
D3.7	Report on keystroke analysis and implications for field work	12	3.00	R	PU	36
D3.8	Fieldwork monitoring application for decentralised surveys	12	5.00	R	PU	36
	Total		68.00			

Description of deliverables

D3.1) CAPI tool for occupation coding and user guide: Task 1: CAPI tool for occupation coding and user guide [month 36]

D3.2) Web-based tool for occupation coding and user guide: Task 1: Web-based tool for occupation coding and user guide [month 36]

D3.3) Operational question database and plug in: Task 2: Operational question database and plug in [month 31]

D3.4) Operational questionnaire design & translation tool / databank: Task 2: Operational questionnaire design & translation tool / databank [month 34]

D3.5) Standardized language management utility plug-in: Task 2: Standardized language management utility plug-in [month 36]

D3.6) Design of standardized sample management system: Task 3: Design of standardized sample management system [month 24]

D3.7) Report on keystroke analysis and implications for field work: Task 3: Report on keystroke analysis and implications for field work [month 36]

D3.8) Fieldwork monitoring application for decentralised surveys: Task 3: Fieldwork monitoring application for decentralised surveys [month 36]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead beneficiary number	Delivery date from Annex I [60]	Comments
MS10	Interim report on development of Web and CAPI tools for occupation coding	12	19	
MS11	Prototype databanks on questions, design process, translation	12	24	
MS12	Review and dissemination of standardized tools	12	33	

Project Number [1]	283646	Project Acronym [2]	DASISH

Work package number [53]	WP4	Type of activity [54]	COORD
Work package title	Data Archiving		
Start month	1		
End month	36		
Lead beneficiary number [55]	11		

Objectives

• to discuss the state of long term data preservation in the SSH domain and work out suggestions to overcome current bottlenecks and to establish trust in procedures
• to develop models for common deposit services which could be provided to all SSH researchers
• to work out policy rules, business and access models that can be taken up by a common data service layer providing long term preservation services

Description of work and role of partners

Task 4.1 State-of-the-art of Data Preservation and Curation (Leader: NSD; Partner: MPG (TLA), CenterData, OEAW, CITY, GESIS)
The situation of data in the SSH domain is very diverse. Some data is already taken care of very well such as some survey data in SS, some data about endangered languages etc. However, the general situation for all material collected by SSH researchers is not satisfying. This task will define the baseline for data curation and long term data preservation by analysing the current state of data management by interviewing SSH researchers or making use of existing investigation results, by talking to relevant initiatives and by studying relevant reports. The goal is to establish an appropriate state-of-the-art description and guidelines and a common roadmap for SSH institutions and researchers with respect to proper rules for data management and to determine the required services and measures.
The output of this task is a combination of a state-of-the-art report, guidelines for proper data management and a requirements specification for data preservation and curation.
Procedure: (1) Assess the state of data preservation and curation in the SSH domain by using all possible information; (2) analyse the results of the survey and work out guidelines and a roadmap for improving the state.

Task 4.2: Assessment of deposit services (Leader: NSD; Partner: UIB, MPG (TLA), GESIS, SND, OEAW,DANS,CITY,CenterData)
This task is very much related to task 1 in so far as it will analyse different deposit services that are offered by institutions and by the emerging data e-Infrastructures in all respects, i.e. technology, business models, long term preservation guarantees, accessibility, license conditions, etc. It will be influenced by architectural concerns in WP2 and with respect to legal and ethical aspects it will closely collaborate with WP6. The goal is to work out an overview of the offered services and their characteristics, to develop guidelines and recommendations for such services that will create the necessary trust relationship between the researchers and the data service centres. Particular concern will be put on interacting with the emerging common data service e-Infrastructure (EUDAT and similar) and initiate work on suitable service level agreements that meet the requirements.
Output: A report about existing data services that can act as a guideline for establishing competitive services as currently in the discussion.
Procedure: (1) Existing institutional and academic services will be analysed and described with respect to the relevant dimensions (IPR, interface, functionality, technology, legal and economic questions). (2) Based on this and the needs of the participating communities on the one hand and feasibility aspects on the other hand recommendations will be made for appropriate services for SSH researchers. (3) Negotiations will be initiated with possible service providers – in particular with the emerging common data infrastructure – to find feasible solutions.

Task 4.3: Deposit service convergence (Leader: DANS; Partner: NSD, GESIS, SND, CenterData, UGOE, MPG (TLA), UIB)

The first services that need to be offered to the SSH researchers are deposit services that guaranty easy deposit and access options that are coherent with the guidelines that have been worked out in tasks 1 and 2. With some service providers concrete interactions will take place which may include adaptations of their portals to make it attractive for SSH researchers. We do not expect that one service should be sufficient, both in practical and legal terms data are highly diversified and data depositing and copying will only take place when a functioning setup and a trust relationship has been established. Thus DASISH will assess the quality of given services, determine their specializations, carry out tests and promote their use. Partners are institutions that already have deposit services in place or who has showed an interest to offer one. The deposited data needs to be part of the preservation service described in 4.4., i.e. they need to be linked to preservation services.

Output: Selection and promotion of a number of high-quality deposit services for SSH researchers and concrete suggestions for service improvements.

Procedure: (1) Based on the outcome of tasks 1 and 2, we will investigate existing deposit offers. (2) With promising candidates negotiations will take place to adapt and/or extend their service offer so that it will meet the requirements. (3) Recommendations for adaptations and tests of candidate services will take place resulting in Service Level Agreements with the selected service providers. These adaptations will be done in iterative steps to ensure fast offers. (4) PR and training material will be developed and offered to WP7/8.

Task 4.4: Recommendation of a set of policy rules (Leader: NSD; Partner: UGOE, FSD, MPG (TLA), UIB, GESIS)

DASISH will not itself establish a long-term preservation service for all the data that is collected by its researchers, but it will rely on common data services offered by a network of strong data centres with national backing as it was requested by the HLEG on Scientific Data. The requirements and practical issues such as Service Level Agreements etc. for these services will need to be discussed with preservation service providers. Since long term curation and preservation in a distributed scenario will increasingly be based on "explicit and declarative policy rules " to allow assessing procedures and to maintain manageability, DASISH needs to work on such policy-rules based on the SSH requirements, the traditions and the different architectures underlying the SSH repositories. DASISH will carry out this work in close collaboration with experts who are already using such strategies and from the emerging collaborative data services infrastructure. Only in case that there will be no satisfying progress of the common data e-Infrastructure DASISH would look for other alternatives.

Output of this task will be recommendation of a set of policy-rules covering all aspects that are essential for the SSH domain and that can be used to negotiate with preservation and curation service providers.

Procedure will be as follows: (1) The scope of policy-rules and their requirements for the SSH domain will be determined and analysed in collaboration with other initiatives in particular in Europe and the US. (2) Concrete policy-rules will be established in the interaction with the emerging common data infrastructure or with other capable service providers who have experience with such a framework. (3) Only in case that there is no capable service provider DASISH will invest in setting up a framework based on iRODS to implement and test the policy framework with selected partners.

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
1	UGOT	5.00
2	MPG	6.00
5	KNAW-DANS	5.00
7	UGOE	6.00
8	OEAW	5.00
10	UTA	3.00
11	NSD	15.00
12	CITY	2.00
13	GESIS	6.00

WT3:
Work package description

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
15	UiB	7.00
17	CentERdata	7.00
	Total	67.00

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D4.1	Roadmap for Preservation and Curation in SSH	11	3.00	R	PU	6
D4.2	Report about Preservation Service Offers	11	3.00	R	PU	9
D4.3	List of Recommended Deposit Services for SSH	11	3.00	R	PU	33
D4.4	Comprehensive Policy-Rules for Data Management in SSH	11	3.00	R	PU	33
	Total		12.00			

Description of deliverables

D4.1) Roadmap for Preservation and Curation in SSH: Roadmap for Preservation and Curation in SSH [month 6]

D4.2) Report about Preservation Service Offers: Report about Preservation Service Offers [month 9]

D4.3) List of Recommended Deposit Services for SSH: List of Recommended Deposit Services for SSH [month 33]

D4.4) Comprehensive Policy-Rules for Data Management in SSH: Comprehensive Policy-Rules for Data Management in SSH [month 33]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
MS5	A publishable set of guidelines for data management, data preservation and data cu	11	6	
MS6	An overview of available models and solutions, described against	11	9	
MS7	A completed description of deposit service models and implementations with a ge	11	18	
MS8	Final set of recommendations complete with service level agreements	11	30	
MS9	A complete final set of policy rules	11	33	

Project Number [1]	283646	Project Acronym [2]	DASISH

One form per Work Package			

Work package number [53]	WP5	Type of activity [54]	RTD
Work package title	Shared Data Access & Enrichment		
Start month	1		
End month	36		
Lead beneficiary number [55]	2		

Objectives

• to establish a trust environment and a joint data metadata domain based on well-established AAI principles
• to establish a widely used tool/services portal and setup a framework for testing, evaluating and commenting on tools/services to spread information about useful components
• to define typical processing chains, test them based on existing technology and spread the information about how to create and use them • to create an interoperability framework based on typical use cases
• to create a data enrichment framework suitable for SSH researchers

Description of work and role of partners

Task 5.1: Establishment of Trust Federation (Leader: MPG (TLA)(5); Partner: GESIS (3), UGOE (3)) A trust federation between SSH service providers and national identity federations based on SAML-based AAI principles will be established that has the potential to include a wide range of European researchers. In close collaboration with eduGain/TERENA the existing trust federation established in CLARIN will be extended to other SSH centres that are DASISH partners. The existing agreements (SPF, IDF) will be analysed and adapted to cover the requirements of the interested SSH centres with the goal to extend the domain of SSH services and researchers that are supported by a single identity and therefore have a potential to also support single sign-on. As long as eduGain/TERENA do not have created a joint domain of European identity federations DASISH will continue with pragmatic solutions, nevertheless being ready to adopt eduGain/TERENA results. The output of this task will be a trust federation covering additional centres participating in DASISH based on agreements with the corresponding national identity federations. In case eduGain does not yet have a European harmonized solution, DASISH will go along the way taken by CLARIN in close collaboration with eduGain. The procedure will be as follows: (1) Assess which countries are involved and start interacting with the national identity federation representatives and the eduGain experts about an AAI roadmap and the agreements to be made based on the CLARIN blueprints. (2) Extend the AAI solution stepwise to the interested centres and countries by signing the appropriate agreements, by giving training courses that will help the local experts to integrate AAI components and by giving restricted help from DASISH experts.

Task 5.2: Robust PID Service (Leader: UGOT(6); Partner: GESIS (4), DANS (3), MPG (TLA)(1), UGOE(4), UCPH(2)) The requirements for PID services need to be gathered from the community centers and an analysis will be carried out for the different services currently being offered summarizing the state and type of three relevant services (DataCite, EPIC, PersID) in the light of the requirements. Particular concern is on business models, governance structure, functionality, robustness, availability, scalability and their persistency. All results will be handed over to the three service providers. Finally a report will be created (first part of D5.1) with recommendations to the DASISH communities to promote the usage of PIDs and to make the centers ready to participate in other collaborations where this is required.

The output of this task will result in reports that may lead to upgraded PID services fulfilling the DASISH requirements and that will foster the use of PIDs. The procedure will be as follows: (1) DASISH will make a comparison of three relevant PID services (DataCite, EPIC, PersID). (2) DASISH will work out requirements for such PID services based on the current usage and plans in the participating communities. (3) A report will summarize the results of this work and DASISH will create PR documents and education material that help promoting the use of PID services and that may lead to adaptations of the existing services. (4) Training courses will be offered that describe how to register and use PIDs.

Task 5.3: Metadata Quality Improvement (Leader: DANS (5); Partner: OEAW(3), UCPH (2), UEssex(3)) As stated earlier high quality metadata (and data) is the basis for smart eScience based computation. Often, however, the metadata descriptions are of poor quality. DASISH wants to start a campaign to improve the quality. DASISH will also analyze and compare the different metadata strategies of CLARIN (CMDI), DARIAH (TEI) and CESSDA (DDI) to take profit from each other's solutions where possible.

The output of this task is a report that describes in detail the various metadata solutions and the problems that hamper the creation of high quality metadata descriptions. In addition measures will be taken to raise awareness (flyer, training courses). The procedure will be as follows: (1) The metadata strategies offered by CLARIN, DARIAH and CESSDA will be analyzed and possibilities of cross-fertilization will be identified. (2) A special flyer will be created that addresses the quality issue and that is distributed in the SSH communities. (3) Training courses on proper metadata creations and use will be offered.

Task 5.4: Joint Metadata Domain (Leader: MPG (TLA)(24); Partner: GESIS(9), OEAW(8), DANS(6)) The analysis of the current metadata sets and frameworks used in CLARIN, DARIAH and CESSDA carried out in 5.3 will be taken as basis for the work on creating a joint metadata index. Standard harvesting (OAI PMH) and state-of-the-art gateway practices will be applied and an SSH portal offering filters, search and access methods will be created. All metadata elements being used will be registered in ISO 12620 based category registries if they are not yet registered in other open accessible and stable concept registries. The possibility of using a framework allowing users to easily create, store, share and manipulate relation sets based on RDF assertions will be tested which could be used by the mapping gateways and the search engine. Such flexible metadata framework would give users the possibility to map semantics according to their research needs and without losing information that is essential for them.

The output of this task will be an integrated metadata domain for the SSH disciplines based on standards, semantic mapping technologies and based on the existing experience. The procedure will be as follows: (1) The report from 5.3 (D5.2) will be studied and where necessary extended. (2) Un-registered concepts will be registered in open registries based on ISO 12620 and a proper and flexible semantic mapping scheme will be defined. (3) All centers will provide an OAI-PMH based metadata harvesting channel so that a joint index can be created. (4) Based on existing technology a suitable portal making use of smart search/browse/filtering techniques such as faceted browsing will be generated to give access to the joint metadata.

Task 5.5: Workflow Implementations (Leader: UPF(14); Partner: DANS(8), OEAW(6), UCPH(6), UT(6)) Web applications and web services that will carry out some useful analytical or visualization function and that can be orchestrated to powerful virtual workflows will play an increasingly important role in the emerging e-Research scenario. DASISH will not create new workflow specification or execution frameworks, but apply existing frameworks to typical workflows as they appear in the SSH domain. The responsible partners will make interviews with community members to identity a number of typical workflows that can be candidates for being dealt with by automatic processing chains and to derive their requirements. In particular workflows that exhibit a cross-disciplinary character should be selected as demonstration cases. Those that can be solved with existing solutions such as WebLicht, GATE or CQP will be tested and evaluated and the required adaptations to standard formats will be done.
The output of this task will be a report containing descriptions of successful applications of existing automatic workflow chains to problems in the SSH. The procedure will be as follows: (1) Interviews will be carried out with interested researchers from the participating communities to identify interesting use cases for processing chains. (2) The requirements will be analyzed and compared with the existing functionality and a selection will be made. (3) The selected use cases will be realized and tested. (4) The use cases including the complete process will be documented by making use of the described framework and made available for educational activities.

Task 5.6: Annotation Framework (Leader: MPG (TLA)(15); Partner: UT(8), KCL(8), UGOT(8), UCPH(1)) For all SSH researchers the manual annotation of web accessible material is of great importance since it often stores the results of intellectual work be it an annotation on a single fragment or the creation of typed relations between a number of fragments. Existing annotation frameworks (Annotea, OpenAnnotationArchitecture, etc) will be analyzed and compared with the requirements of SSH researchers. An annotation framework will be provided that can handle essential SSH requirements Typical annotation examples will be carried out to demonstrate the potential of the annotation paradigm to SSH researchers, the experiences will be entered at the usability portal and appropriate education/training modules will be created and made available.

The output of this task will be an annotation framework for web-based information that can be used by the SSH communities. The procedure will be as follows: (1) The requirements for an annotation service for SSH will be

specified. (2) Existing annotation frameworks will be analyzed and compared with the specified requirements. (3) An annotation framework covering essential functionality will be provided. (4) Tests with interested pilot users from the SSH domain will be carried out. (5) The annotation framework and the tests will be documented by using the described framework. (6) PR and education material will be worked out.

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
1	UGOT	18.00
2	MPG	45.00
3	UPF	14.00
4	UCPH	11.00
5	KNAW-DANS	18.00
6	KCL	6.00
7	UGOE	7.00
8	OEAW	17.00
9	UEssex	7.00
13	GESIS	16.00
18	UT	12.00
	Total	171.00

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D5.1	Trust and PID Federation Services Report	2	6.00	R	PU	24
D5.2	Metadata Quality and Portal Progress Report	2	6.00	R	PU	30
D5.3	Workflow Requirements and Application Report	2	3.00	R	PU	36
D5.4	Annotation Framework Report	2	3.00	R	PU	30
		Total	18.00			

Description of deliverables

D5.1) Trust and PID Federation Services Report: Trust and PID Federation Services Report [month 24]

D5.2) Metadata Quality and Portal Progress Report: Metadata Quality and Portal Progress Report [month 30]

D5.3) Workflow Requirements and Application Report: The application of workflow mechanisms to SSH problems will be reported [month 36]

D5.4) Annotation Framework Report: The state of the common annotation service will be reported [month 30]

Schedule of relevant Milestones				
Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
MS13	more DASISH centres are part of the SSH service provider federation	2	24	
MS14	PID services that fulfil essential SSH requirements are available	2	12	
MS15	Major metadata quality gaps have been indentified	2	24	
MS16	a first version of a joint metadata domain is ready	2	24	
MS17	a first set of workflow chains have been built and tested	2	24	
MS18	a first version of an annotation framework is available	2	24	

WT3:
Work package description

Project Number [1]	283646	Project Acronym [2]	DASISH

Work package number [53]	WP6	Type of activity [54]	COORD
Work package title	Legal and Ethical Issues		
Start month	1		
End month	36		
Lead beneficiary number [55]	2		

Objectives

• to identify legal and ethical issues, constraints and requirements for all data types occurring in the SSH domain as a result of the intended integration and linking
• to cope with the legal and ethical challenges imposed by new data types emerging in the social sciences,
• to look for professional long-run preservation strategies based on e-Infrastructures for data in the social sciences and humanities

Description of work and role of partners

Task 6.1: New Ethical and Legal Challenges (Leader: MPG (MEA); Partner: CITY)
Modern research in social sciences require to record much more innovative variables such as health information, financial status, political opinions and brain images. Also biomarkers need to be included to better understand health and well-being. Process generated data that cover large parts or even the entire life course are generated, e.g., by linking survey data to pension systems and health insurances. Finally, much sensitve paradata is generated while interviewing, such as keystroke data and contact protocols.
On the one hand, these data are very sensitive and thus pose new ethical and legal challenges with implied new and special protection requirements; they are of high scientific value on the other hand e.g. as life-course analyses play an ever more important role in the social sciences and non-response analyses require sophisticated paradata.
All these data extend the existing more traditional data to collections for which high security guidelines that go beyond protecting personality rights are required. For some level data, scientific-user databases can be released to the scientific community. Due to data privacy concerns, these data can only contain broad summary statistics of, e.g., employment biographies, but not the complete administrative records. This WP will apply the rules of "factually anonymised data sets" developed so far to typical SSH data. More detailed data will need to be handled in "data enclaves". The WP will develop rules on the precise procedures of such access in collaboration with the responsible data privacy officers.
The output of this task will be a report describing the IPR requirements for all new data types being recorded in the SSH and being shared in a cross-country usage, a report that will identify the challenges that need to be solved when offering sensitive data to a wider user group, and analyses of paradata requiring special legal and ethical considerations.
The procedure will be as follows: (1) Identify the various new data types including sensitive data; (2) Determining the IPR requirements for these new data types in a multi-country usage scenario. (3) Guidelines for appropriate measures will be defined and tools that help to take appropriate measures will be identified.

Task 6.2: Virtual L&E Competence Center (Leader: MPG (MEA); Partner: CITY, NSD, UIB, GESIS)
As a consequence of the sketched developments we foresee an increasing amount of routine legal and ethical procedures in the day-to-day operation of SSH data collection. This task will professionalise these efforts and create a virtual centre that collects and systemises all respective legal information from the member countries, create templates for letters and communication, model contracts for all involved partners, and standard procedures for the handling of legal issues.
This task will be tailored specifically to each member country x and will: (a) Map actual and where relevant future national legal restrictions and scopes for data collection and data dissemination (user access) that prevail in country x; (b) Ascertain the procedures for ethic review board approval for SSH data collection, storage and retrieval in country x; (c) Understand all legal ramifications of partnership and agency contracts

in country x; (d) Collect all legal procedures necessary to handle potential future restricted data (biomarkers, linked administrative records) in country x; (e) Explore the development and legal ramifications of low intrusion techniques for bio-markers (such as derived from dried blood spots and saliva) in country x; (f) Explore and report on the availability and quality of income, tax, social security and health insurance process data in country x; (e) Ascertain the possibilities of various anonymised linking mechanisms to social security and health insurance process data in country x; (f) Collect country-specific information on day-to-day legal issues, such as the formulation of survey agency contracts, ownership of addresses, tax issues (especially VAT), and partnership contracts. It is the task of this virtual centre to give legal advice and countercheck all contracts and legal arrangements.

The output of this task is a virtual competence center that is maintaining a comprehensive database about legal and ethical issues and that refers to relevant documents and experts.

The procedure will be as follows: (1) The requirements for a virtual competence center and its database organization will be specified. (2) The framework for the database and its user interface will be built. (3) The database will be filled stepwise with useful information.

Task 6.3: Preservation Challenges (Leader: FSD; Partner: KCL, NSD, MPG (MEA), UIB)
In WP4 (Data Archiving) we will collaborate with the emerging data e-Infrastructure on taking care of the long-term preservation of data sets. Given the various nature of the SSH data sets and its varying IPR characteristics (see above) we need to work out guidelines how they can be treated in such a multi-level architecture where for example data replication can be done across country borders etc. Maintaining access rights regulations for all replications is a major concern and needs to be guaranteed by appropriate mechanisms such as for example extending the information associated with persistent identifiers. IPR considerations will also influence policy rules for long-term archiving such as automatically executing anonymisation or summarization processes. This task therefore will work out guidelines for appropriate IPR management in the emerging data preservation landscape.

The output of this task is a report that describes the policy-rules that need to be considered when sensitive data will be preserved in a distributed data preservation and curation infrastructure.

The procedure will be as follows: (1) The problems that may occur when preserving SSH data in the emerging European data preservation eInfrastructure will be identified. (2) Policy-rules that guide the preservation and access rights while maintaining trust will be defined

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
2	MPG	33.00
6	KCL	3.00
10	UTA	6.00
11	NSD	9.00
12	CITY	3.00
13	GESIS	3.00
15	UiB	7.00
17	CentERdata	4.00
	Total	68.00

List of deliverables

Deliverable Number [61]	Deliverable Title	Lead beneficiary number	Estimated indicative person-months	Nature [62]	Dissemination level [63]	Delivery date [64]
D6.1	Report about New IPR Challenges	2	3.00	R	PU	18

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D6.2	Sample merged paradata sets	2	3.00	D	PU	18
D6.3	Exemplary analyses of confidential paradata	2	3.00	D	PU	36
D6.4	Midterm report on the establishment of virtual centres	2	3.00	R	PU	18
D6.5	Handbook on legal and ethical issues for SSH data in Europe	2	4.00	O	PU	36
D6.6	Report about Preservation Policy-Rules	2	3.00	R	PU	30
		Total	19.00			

Description of deliverables

D6.1) Report about New IPR Challenges: Task 1: Report about New IPR Challenges [month 18]

D6.2) Sample merged paradata sets: Task 1: Sample merged paradata sets [month 18]

D6.3) Exemplary analyses of confidential paradata: Task 1: Exemplary analyses of confidential paradata [month 36]

D6.4) Midterm report on the establishment of virtual centres: Task 2: Midterm report on the establishment of virtual centres [month 18]

D6.5) Handbook on legal and ethical issues for SSH data in Europe: Task 2: Handbook on legal and ethical issues for SSH data in Europe [month 36]

D6.6) Report about Preservation Policy-Rules: Task 3: Report about Preservation Policy-Rules [month 30]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
MS19	Establishment of virtual LE competence centres	2	24	

WT3:
Work package description

Project Number [1]	283646	Project Acronym [2]	DASISH

Work package number [53]	WP7	Type of activity [54]	COORD
Work package title	Education and Training		
Start month	1		
End month	36		
Lead beneficiary number [55]	7		

Objectives

☐ To establish a joint domain for training and education for the SSH infrastructures.
☐ To inspire researchers and developers to come up with new research methodologies and approaches using research infrastructures.
☐ To discuss and feedback with researchers in all SSH domains about the role of data infrastructures in research methodologies.

Description of work and role of partners

Task 7.1: Training modules (Leader: UGOE; Partners: NSD, MPG (MEA), KCL, OEAW, UIB)
This task develops a matrix of training modules for e.g. data quality, data archiving and data enrichment. Topics and materials for the modules will be taken from (a) experiences within the domains of the participating SSH ESFRIs, as well as (b) the findings of the work packages WP2-WP6. Target groups for the training modules are infrastructure-related initiatives within the participating SSH domains as well as to new communities in the SSH.
Output: Each module will be compiled and published in the shared system EduNET provided by ESS. ESS EduNet was created through EU funding from 2004 to 2010, aiming to promote access to the ESS for students and new academics. This interactive learning environment is applicable for general usage within the SSH, and will be ready to cater for content authored by DASISH across infrastructures. Furthermore, existing course material will be integrated or referenced.
Procedure: While the training modules will be published to be reused and extended over time, there will be two main iterations: one covering key experiences from the communities, and another one covering key findings from the other WPs. The second iteration will start as soon as the first deliverables of the other work packages are available. Each module will have an assigned editor, and test persons from another SSH community and/or WP will inform the work to ensure general applicability.

Task 7.2: Workshop Programme (Leader: GESIS; Partners: MPG (MEA), UGOE, OEAW, KCL, NSD)
This task gives a series of training workshops. To ensure the effectiveness of the training, there will be domain-specific as well as cross-domain events. This could include workshops for the following topics (as an overall output of the task):
1. Data management for data creating researchers, specific for social sciences
2. Data management for data creating researchers, specific for humanities
3. Metadata development, specific for social sciences
4. Metadata development, specific for humanities
5. Data registration for the whole DASISH community
6. Data quality and certification workshop for the whole DASISH community
Maximum interaction between task 7.1 and task 7.2 will be sought so that material from task 7.1 can be reused in training workshops and extended based on the feedback of participants.
Procedure: The workshop dates and locations will be determined as early as possible in the project to support an effective organisation and planning for the organizing institution(s) as well as for the target audience. If possible and feasible the workshops can be co-located with other conferences in the SSH context. The outcomes of the other work packages, as well as the appropriate training material bundled in WP7.1 are the basis for the workshops. Therefore, workshops will take place in the second and third year of the project.

Links to other Work Packages

WP2 (Architecture) - Link the Knowledge Base into the various communities.
WP4 (Data Archiving) and WP5 (Shared Access) - DASISH-tools for collaborative data processing
WP8 (Dissemination) - publicise training and education efforts through the dissemination communication channels (newsletter, website, etc.)

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
2	MPG	3.00
6	KCL	4.00
7	UGOE	19.00
8	OEAW	7.00
11	NSD	7.00
13	GESIS	12.00
15	UiB	4.00
	Total	56.00

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D7.1	course modules	7	4.00	O	PU	36
D7.2	a compendium of workshop reports	7	3.00	O	PU	36
		Total	7.00			

Description of deliverables

D7.1) course modules: course modules [month 36]

D7.2) a compendium of workshop reports: a compendium of workshop reports [month 36]

Schedule of relevant Milestones

Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
MS20	experience based course modules made available via EduNET	7	12	
MS21	course modules based on the outcome of the other WPS made available	7	36	
MS22	first workshop plan and schedule	7	15	
MS23	all workshops finished and reports published	7	36	

WT3:
Work package description

Project Number [1]	283646	Project Acronym [2]	DASISH

Work package number [53]	WP8	Type of activity [54]	COORD
Work package title	Dissemination		
Start month	1		
End month	36		
Lead beneficiary number [55]	4		

Objectives

☐ To develop a dissemination strategy which comprises activities designed for various target groups and various channels

☐ To carry out the components of the strategy successfully according to the goals (quantitative as well as qualitative) that will be defined in the strategy.

Description of work and role of partners

Task 8.1: Dissemination strategy (M1-3; Leader: UCPH; Partners: FSD, CITY, MPG (MEA), UGOE)
The work will consist in developing a dissemination strategy for the entire project, comprising the relevant target groups and channels.
Output: A Dissemination Strategy Report which will provide a framework for dissemination activities which will be flexible enough to accommodate lessons learned during the DASISH project period.
Procedure: In order to develop the most adequate and efficient strategy for DASISH we will examine the 5 current dissemination plans and assess the outcome of the activities carried out according to those plans in order to learn from them. Based on our findings we will create a plan which comprises the types and number of activities that will lead to the best and broadest impact within the resources spent on the effort.

Task 8.2: Web-site, newsletter and fact sheets (M1-36; Leader: UCPH; Partners: UGOT, CITY, MPG (MEA), UGOE, UT)
The work in Task 8.2 is to establish the project's general dissemination channels very early in the project and to make a plan for providing content for them and publishing them at regular intervals, according to the dissemination plan, throughout the project.
Output: General dissemination channels which are:
☐ A project web-site for external dissemination and a password protected site for internal project communication and dissemination.
☐ A design for and the bi-annual publication of a newsletter. The newsletter task also includes the appointment of an editorial committee, the appointment of correspondents and contributors from each infrastructure and the creation of mailing lists of subscribers.
☐ Fact sheets according to the dissemination plan. These fact sheets will be published on the project web-site for download by users and others, and they may also serve as hand-outs at events such as workshops and conferences.
Procedure: 1) the experience of the infrastructures w.r.t. this type of tasks will be put together 2) the best approach to each sub task will be chosen, and 3) the work will be executed in close collaboration with the project management and the other work packages.

Task 8.3: Conference Participation (M1-36; Leader: UCPH; Partners: UGOT, CITY, MPG (MEA), UGOE)
A good way of reaching far out is by participating in scientific conferences with paper presentations, poster sessions, workshops, panel discussions, and as invited speakers, etc. All infrastructures will need to contribute to this task by maintaining a list of relevant conferences in their specific disciplines, and seek to participate in the conferences.
Output: A conference list and participation strategy.
Procedure: 1) Contributions to list of relevant conferences, 2) Schedule conference presentations/papers

Task 8.4: Final conference (M1-36; Leader: UGOT; Partners: FSD, CITY, MPG (MEA), UGOE, UCPH)

The DASISH project aims at a final conference directed to attract SSH scholars and in particular young researchers and technologists. To make the program interesting enough, given the variety of wishes and requirements, we are looking for a multi-day (2-3 days) event where discipline-specific as well as common topics can be addressed fairly much in the same direction as NEERI 2010 was organized. Per community one member is in the preparation board in order to come to a balanced program.
Output: A conference that has attracted many researchers from the SSH domain.
Procedure: 1) The preparation group will design a program early enough to be decided by the EB; 2) the final conference will be organized.

Task 8.5: Qualitative Data Workshop (M6; Leader: NUIM)
In task 8.5 we will organise a workshop that will ensure the involvement and co-operation of qualitative social science researchers. The newly formed EQUALAN network brings together European researchers and archivists committed to preserving and organising qualitative data resources for sharing and re-use. Qualitative data – typically, texts and audio files created to address specific research questions - have distinct features that require different treatment from quantitative data. As a consequence, progress in making qualitative data available for re-use is at an earlier stage than quantitative data. DASISH provides an important opportunity to address the distinctive features of qualitative data in ways that will ensure efficient harmonization of access and promote trust.
Output: A workshop. We do not envisage a very large or a very long event – perhaps up to two representatives from each of up to sixteen organisations in EQUALAN, who would meet for a maximum of two days in, say, London or Brussels.
Procedure: The workshop will inform members of the network about DASISH and develop a plan for how they can be involved. It will examine ways in which cooperation with DASISH can promote enhanced integration and linking with SSH infrastructures, and promote user engagement and trust on the part of qualitative researchers

Task 8.6: Quantitative Data Workshops (M4-M36 Leader: GESIS; Partner City)

In task 8.6 we will organise two workshops that will ensure the involvement and cooperation of around nine of the main sources of rigorous European data (viz. Eurostat, CESSDA, Eurobarometer, European Social Survey, SHARE, European Values Surveys, European Election Surveys, European Quality of Life Survey and the Gender and Generations Survey) and if possible some of the most prominent commercial pan-European survey organisations, such as Gallup, IPSOS and SOFRES, will also be drawn in. This endeavour would link to the emerging US-European working group of "panel study leaders" which specializes in the problems created by large panels (attrition, consistency of instruments over time vis-à-vis state-of-the-art questions, etc.) but on this occasion focusing on cross-national issues.

Output: Two workshops will be convened. We do not envisage a very large or a very long event in each case – up to two representatives from each of up to twelve organisations who would meet for a maximum of two days. Procedure: The first workshop will inform all these key actors about DASISH and help shape the detailed work programme for how they can be involved and cooperate. The second workshop will focus on presenting and discussing the outcomes of the involvement and cooperation, thus enabling the widest possible buy-in and continuing relevance.

The workshops will be organised by GESIS. The agenda will be agreed with City following consultation with the participants in WP3. Up to 25 participants will have their travel and accommodation costs covered and hosting costs are also included. A report (D8.8) will summarise the events and outline how the DASISH deliverables were received by other survey programs.

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
1	UGOT	3.00
2	MPG	3.50
4	UCPH	10.50
7	UGOE	3.50

WT3:
Work package description

Person-Months per Participant

Participant number [10]	Participant short name [11]	Person-months per participant
10	UTA	3.25
12	CITY	3.50
13	GESIS	2.00
14	NUIM	1.50
18	UT	3.25
	Total	34.00

List of deliverables

Delive-rable Number [61]	Deliverable Title	Lead benefi-ciary number	Estimated indicative person-months	Nature [62]	Dissemi-nation level [63]	Delivery date [64]
D8.1	Dissemination Strategy	4	2.00	R	PP	3
D8.2	Web-site up and running	4	2.00	O	PU	3
D8.3	Newsletter, bi-annual	4	6.00	O	PU	36
D8.4	Fact Sheets, ongoing, based on need	4	3.00	R	PU	36
D8.5	Conference list, ongoing	4	3.00	R	PP	36
D8.6	Final Conference	1	2.00	O	PU	36
D8.7	Qualitative Data Workshop Report	14	1.00	O	PU	6
D8.8	Quantitative Data Workshop Report	13	1.00	R	PU	35
	Total		20.00			

Description of deliverables

D8.1) Dissemination Strategy: Dissemination Strategy [month 3]

D8.2) Web-site up and running: Web-site up and running [month 3]

D8.3) Newsletter, bi-annual: Newsletter, bi-annual [month 36]

D8.4) Fact Sheets, ongoing, based on need: Fact Sheets, ongoing, based on need [month 36]

D8.5) Conference list, ongoing: Conference list, ongoing [month 36]

D8.6) Final Conference: Final Conference [month 36]

D8.7) Qualitative Data Workshop Report: Qualitative Data Workshop Report [month 6]

D8.8) Quantitative Data Workshop Report: Quantitative Data Workshop Report [month 35]

Schedule of relevant Milestones				
Milestone number [59]	Milestone name	Lead benefi-ciary number	Delivery date from Annex I [60]	Comments
MS24	Web-site up and running	4	3	
MS25	Final conference	1	36	

WT4:
List of Milestones

Project Number [1]	283646	Project Acronym [2]	DASISH

List and Schedule of Milestones					
Milestone number [59]	Milestone name	WP number [53]	Lead benefi- ciary number	Delivery date from Annex I [60]	Comments
MS1	Kick-off meeting	WP1	1	1	
MS2	Integration of investigation activities into architecture	WP2	5	18	
MS3	Registry implemented and linked with metadata portal	WP2	5	24	
MS4	Content from the community & infrastructures added to the registry	WP2	5	36	
MS5	A publishable set of guidelines for data management, data preservation and data cu	WP4	11	6	
MS6	An overview of available models and solutions, described against	WP4	11	9	
MS7	A completed description of deposit service models and implementations with a ge	WP4	11	18	
MS8	Final set of recommendations complete with service level agreements	WP4	11	30	
MS9	A complete final set of policy rules	WP4	11	33	
MS10	Interim report on development of Web and CAPI tools for occupation coding	WP3	12	19	
MS11	Prototype databanks on questions, design process, translation	WP3	12	24	
MS12	Review and dissemination of standardized tools	WP3	12	33	

Milestone number [59]	Milestone name	WP number [53]	Lead benefi- ciary number	Delivery date from Annex I [60]	Comments
MS13	more DASISH centres are part of the SSH service provider federation	WP5	2	24	
MS14	PID services that fulfil essential SSH requirements are available	WP5	2	12	
MS15	Major metadata quality gaps have been indentified	WP5	2	24	
MS16	a first version of a joint metadata domain is ready	WP5	2	24	
MS17	a first set of workflow chains have been built and tested	WP5	2	24	
MS18	a first version of an annotation framework is available	WP5	2	24	
MS19	Establishment of virtual LE competence centres	WP6	2	24	
MS20	experience based course modules made available via EduNET	WP7	7	12	
MS21	course modules based on the outcome of the other WPS made available	WP7	7	36	
MS22	first workshop plan and schedule	WP7	7	15	
MS23	all workshops finished and reports published	WP7	7	36	
MS24	Web-site up and running	WP8	4	3	
MS25	Final conference	WP8	1	36	
MS26	Second consortium meeting	WP1	1	13	
MS27	Third consortium meeting	WP1	1	25	

WT5:
Tentative schedule of Project Reviews

Project Number [1]	283646	Project Acronym [2]	DASISH

Tentative schedule of Project Reviews			
Review number [65]	Tentative timing	Planned venue of review	Comments, if any

WT6:
Project Effort by Beneficiary and Work Package

Project Number [1]	283646	Project Acronym [2]	DASISH

Indicative efforts (man-months) per Beneficiary per Work Package

Beneficiary number and short-name	WP 1	WP 2	WP 3	WP 4	WP 5	WP 6	WP 7	WP 8	Total per Beneficiary
1 - UGOT	27.00	14.00	0.00	5.00	18.00	0.00	0.00	3.00	67.00
2 - MPG	9.50	4.00	26.50	6.00	45.00	33.00	3.00	3.50	130.50
3 - UPF	0.00	4.00	12.00	0.00	14.00	0.00	0.00	0.00	30.00
4 - UCPH	1.50	0.00	0.00	0.00	11.00	0.00	0.00	10.50	23.00
5 - KNAW-DANS	1.50	17.00	0.00	5.00	18.00	0.00	0.00	0.00	41.50
6 - KCL	0.00	15.00	0.00	0.00	6.00	3.00	4.00	0.00	28.00
7 - UGOE	1.50	4.00	0.00	6.00	7.00	0.00	19.00	3.50	41.00
8 - OEAW	0.00	0.00	0.00	5.00	17.00	0.00	7.00	0.00	29.00
9 - UEssex	0.00	9.00	0.00	0.00	7.00	0.00	0.00	0.00	16.00
10 - UTA	0.00	5.00	5.00	3.00	0.00	6.00	0.00	3.25	22.25
11 - NSD	1.50	0.00	40.00	15.00	0.00	9.00	7.00	0.00	72.50
12 - CITY	1.50	0.00	38.00	2.00	0.00	3.00	0.00	3.50	48.00
13 - GESIS	0.00	0.00	17.00	6.00	16.00	3.00	12.00	2.00	56.00
14 - NUIM	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.50	1.50
15 - UiB	0.00	0.00	0.00	7.00	0.00	7.00	4.00	0.00	18.00
16 - UNIVE	0.00	0.00	30.00	0.00	0.00	0.00	0.00	0.00	30.00
17 - CentERdata	0.00	3.00	31.00	7.00	0.00	4.00	0.00	0.00	45.00
18 - UT	0.00	8.00	0.00	0.00	12.00	0.00	0.00	3.25	23.25
Total	44.00	83.00	199.50	67.00	171.00	68.00	56.00	34.00	722.50

WT7:
Project Effort by Activity type per Beneficiary

Project Number [1]	283646	Project Acronym [2]	DASISH

Indicative efforts per Activity Type per Beneficiary

Activity type	Part. 1 UGOT	Part. 2 MPG	Part. 3 UPF	Part. 4 UCPH	Part. 5 KNAW-DA	Part. 6 KCL	Part. 7 UGOE	Part. 8 OEAW	Part. 9 UEssex	Part. 10 UTA	Part. 11 NSD	Part. 12 CITY	Part. 13 GESIS	Part. 14 NUIM
1. RTD/Innovation activities														
WP 3	0.00	26.50	12.00	0.00	0.00	0.00	0.00	0.00	0.00	5.00	40.00	38.00	17.00	0.00
WP 5	18.00	45.00	14.00	11.00	18.00	6.00	7.00	17.00	7.00	0.00	0.00	0.00	16.00	0.00
Total Research	18.00	71.50	26.00	11.00	18.00	6.00	7.00	17.00	7.00	5.00	40.00	38.00	33.00	0.00
2. Demonstration activities														
Total Demo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3. Consortium Management activities														
WP 1	27.00	9.50	0.00	1.50	1.50	0.00	1.50	0.00	0.00	0.00	1.50	1.50	0.00	0.00
Total Management	27.00	9.50	0.00	1.50	1.50	0.00	1.50	0.00	0.00	0.00	1.50	1.50	0.00	0.00
Work Packages for Coordination activities														
WP 2	14.00	4.00	4.00	0.00	17.00	15.00	4.00	0.00	9.00	5.00	0.00	0.00	0.00	0.00
WP 4	5.00	6.00	0.00	0.00	5.00	0.00	6.00	5.00	0.00	3.00	15.00	2.00	6.00	0.00
WP 6	0.00	33.00	0.00	0.00	0.00	3.00	0.00	0.00	0.00	6.00	9.00	3.00	3.00	0.00
WP 7	0.00	3.00	0.00	0.00	0.00	4.00	19.00	7.00	0.00	0.00	7.00	0.00	12.00	0.00
WP 8	3.00	3.50	0.00	10.50	0.00	0.00	3.50	0.00	0.00	3.25	0.00	3.50	2.00	1.50
Total Coordination	22.00	49.50	4.00	10.50	22.00	22.00	32.50	12.00	9.00	17.25	31.00	8.50	23.00	1.50
4. Other activities														
Total other	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Work Packages for Support activities														
Total Support	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	67.00	130.50	30.00	23.00	41.50	28.00	41.00	29.00	16.00	22.25	72.50	48.00	56.00	1.50

Activity type	Part. 15 UiB	Part. 16 UNIVE	Part. 17 CentERd	Part. 18 UT	Total
1. RTD/Innovation activities					
WP 3	0.00	30.00	31.00	0.00	199.50
WP 5	0.00	0.00	0.00	12.00	171.00
Total Research	0.00	30.00	31.00	12.00	370.50
2. Demonstration activities					
Total Demo	0.00	0.00	0.00	0.00	0.00
3. Consortium Management activities					
WP 1	0.00	0.00	0.00	0.00	44.00
Total Management	0.00	0.00	0.00	0.00	44.00
Work Packages for Coordination activities					
WP 2	0.00	0.00	3.00	8.00	83.00
WP 4	7.00	0.00	7.00	0.00	67.00
WP 6	7.00	0.00	4.00	0.00	68.00
WP 7	4.00	0.00	0.00	0.00	56.00
WP 8	0.00	0.00	0.00	3.25	34.00
Total Coordination	18.00	0.00	14.00	11.25	308.00
4. Other activities					
Total other	0.00	0.00	0.00	0.00	0.00
Work Packages for Support activities					
Total Support	0.00	0.00	0.00	0.00	0.00

| Total | 18.00 | 30.00 | 45.00 | 23.25 | 722.50 |

WT8:
Project Effort and costs

Project Number [1]	283646	Project Acronym [2]	DASISH

Project efforts and costs

Benefi-ciary number	Beneficiary short name	Estimated eligible costs (whole duration of the project)						Total receipts (€)	Requested EU contribution (€)
		Effort (PM)	Personnel costs (€)	Subcontracting (€)	Other Direct costs (€)	Indirect costs OR lump sum, flat-rate or scale-of-unit (€)	Access costs (€)	Total costs		
1	UGOT	67.00	497,408.00	6,000.00	40,810.00	322,930.80	0.00	867,148.80	0.00	719,021.00
2	MPG	130.50	738,729.00	0.00	56,115.00	1,034,221.00	0.00	1,829,065.00	0.00	1,211,405.00
3	UPF	30.00	145,548.00	0.00	12,900.00	95,068.80	0.00	253,516.80	0.00	187,390.00
4	UCPH	23.00	162,150.00	0.00	32,890.00	117,024.00	0.00	312,064.00	0.00	226,925.00
5	KNAW-DANS	41.50	236,550.00	3,000.00	17,845.00	152,637.00	0.00	410,032.00	0.00	294,420.00
6	KCL	28.00	193,200.00	0.00	12,040.00	123,144.00	0.00	328,384.00	0.00	225,324.00
7	UGOE	41.00	225,500.00	3,000.00	17,630.00	145,878.00	0.00	392,008.00	0.00	273,259.00
8	OEAW	29.00	208,800.00	0.00	12,470.00	75,231.00	0.00	296,501.00	0.00	228,327.00
9	UEssex	16.00	114,976.00	0.00	6,880.00	73,113.60	0.00	194,969.60	0.00	137,316.00
10	UTA	22.25	153,970.00	0.00	9,568.00	98,122.80	0.00	261,660.80	0.00	179,763.00
11	NSD	72.50	489,883.00	3,000.00	31,175.00	224,055.00	0.00	748,113.00	0.00	565,131.00
12	CITY	48.00	301,056.00	73,000.00	20,640.00	193,017.60	0.00	587,713.60	0.00	438,149.00
13	GESIS	56.00	308,000.00	36,000.00	93,080.00	240,648.00	0.00	677,728.00	0.00	482,345.00
14	NUIM	1.50	13,073.00	0.00	6,495.00	11,740.80	0.00	31,308.80	0.00	20,937.00
15	UiB	18.00	127,404.00	0.00	7,740.00	81,086.40	0.00	216,230.40	0.00	144,604.00
16	UNIVE	30.00	165,000.00	0.00	12,900.00	106,740.00	0.00	284,640.00	0.00	213,480.00
17	CentERdata	45.00	282,124.00	3,000.00	19,350.00	180,884.40	0.00	485,358.40	0.00	352,575.00
18	UT	23.25	69,750.00	0.00	9,998.00	47,848.80	0.00	127,596.80	0.00	90,681.00
Total		722.50	4,433,121.00	127,000.00	420,526.00	3,323,392.00	0.00	8,304,039.00	0.00	5,991,052.00

WT9:
Summary of transnational access / service provision per installation

Project Number [1]	283646	Project Acronym [2]	DASISH

Summary of transnational access / service provision per installation

Part. num.	Org. short name	Short name of infrastructure	Installation								Est. num. of users	Est. num. of proj.
			Num.	Name	Operator country code	Unit of access	Total Estimated costs	Estimated unit cost	Min. quantity of access to be provided	Access costs charged to the GA		
						Grand Total	0.00			0.00		

1. Project number

The project number has been assigned by the Commission as the unique identifier for your project. It cannot be changed. The project number **should appear on each page of the grant agreement preparation documents (part A and part B)** to prevent errors during its handling.

2. Project acronym

Use the project acronym as given in the submitted proposal. It cannot be changed unless agreed so during the negotiations. The same acronym **should appear on each page of the grant agreement preparation documents (part A and part B)** to prevent errors during its handling.

53. Work Package number

Work package number: WP1, WP2, WP3, ..., WPn

54. Type of activity

For all FP7 projects each work package must relate to one (and only one) of the following possible types of activity (only if applicable for the chosen funding scheme – must correspond to the GPF Form Ax.v):

• **RTD/INNO =** Research and technological development including scientific coordination - applicable for Collaborative Projects and Networks of Excellence

• **DEM =** Demonstration - applicable for collaborative projects and Research for the Benefit of Specific Groups

• **MGT =** Management of the consortium - applicable for all funding schemes

• **OTHER =** Other specific activities, applicable for all funding schemes

• **COORD =** Coordination activities – applicable only for CAs

• **SUPP =** Support activities – applicable only for SAs

55. Lead beneficiary number

Number of the beneficiary leading the work in this work package.

56. Person-months per work package

The total number of person-months allocated to each work package.

57. Start month

Relative start date for the work in the specific work packages, month 1 marking the start date of the project, and all other start dates being relative to this start date.

58. End month

Relative end date, month 1 marking the start date of the project, and all end dates being relative to this start date.

59. Milestone number

Milestone number:MS1, MS2, …, MSn

60. Delivery date for Milestone

Month in which the milestone will be achieved. Month 1 marking the start date of the project, and all delivery dates being relative to this start date.

61. Deliverable number

Deliverable numbers in order of delivery dates: D1 – Dn

62. Nature

Please indicate the nature of the deliverable using one of the following codes

R = Report, **P** = Prototype, **D** = Demonstrator, **O** = Other

63. Dissemination level

Please indicate the dissemination level using one of the following codes:

• **PU =** Public

• **PP =** Restricted to other programme participants (including the Commission Services)

• **RE =** Restricted to a group specified by the consortium (including the Commission Services)

• **CO =** Confidential, only for members of the consortium (including the Commission Services)

• **Restreint UE =** Classified with the classification level "Restreint UE" according to Commission Decision 2001/844 and amendments

• **Confidentiel UE =** Classified with the mention of the classification level "Confidentiel UE" according to Commission Decision 2001/844 and amendments

• **Secret UE =** Classified with the mention of the classification level "Secret UE" according to Commission Decision 2001/844 and amendments

64. Delivery date for Deliverable

Month in which the deliverables will be available. Month 1 marking the start date of the project, and all delivery dates being relative to this start date

65. Review number

Review number: RV1, RV2, ..., RVn

66. Tentative timing of reviews

Month after which the review will take place. Month 1 marking the start date of the project, and all delivery dates being relative to this start date.

67. Person-months per Deliverable

The total number of person-month allocated to each deliverable.

Part B

Table of Contents

B1. Concept and objectives, progress beyond state-of-the-art, S/T methodology and work plan

This project brings together all five research infrastructure initiatives in the social sciences and humanities (SSH) domain and even goes beyond. All five initiatives have made big progress in the preparatory phase each of which addressing their own specific issues technologically as well as organisationally. We first describe the state of the initiatives to indicate that all are ready to start the implementation phase and that by focusing on some specific issues that will profit from cross-fertilization. Second, we describe the concept and the major ideas that drive the DASISH project. Based on this we specify the concrete objectives of the project. Finally we describe how these objectives advances the infrastructures beyond the state-of-the-art.

B1.0.1 State of Infrastructures

For each research infrastructure we indicate the state of the technological and organizational work.

SHARE

The Survey of Health, Ageing and Retirement in Europe (SHARE) is a *multi-disciplinary*, *cross-national*, and *longitudinal* survey measuring the combined health, economic and social living conditions of persons aged 50 and over. It serves as the foundation of empirical ageing research for many disciplines, including demography, economics, epidemiology, gerontology, biology, medicine, psychology, public health, health policy, sociology and statistics. SHARE combines two major strengths which make it a unique infrastructure in the world. First, it is *ex-ante harmonised across countries* which allows comparing the effects of different welfare systems (e.g. pension and health care systems) on well-being, health, economic status, retirement, and social participation of middle-aged and older citizens. Second, it is *multi-disciplinary and fills an important research vacuum*, namely the interaction between health and medical factors on the one hand, and socio-economic factors on the other hand, both together driving individual and population ageing.

In 2004/05 a *baseline wave* of data of about 32,000 respondents aged 50 and over was collected in 11 European countries. The participating countries covered all EU15 regions: Nordic countries (Sweden, Denmark), Western European countries (Netherlands, Belgium, France, Germany, Austria, Switzerland), and the Mediterranean (Spain, Italy, Greece). Modules include *health variables* (e.g. self-reported health, physical functioning, cognitive functioning, physical measures such as grip strength, walking speed, peak flow and body mass index, health behaviour, use of health care facilities), *psychological variables* (e.g. psychological health, well-being, life satisfaction, control beliefs), *economic variables* (e.g. current work activity, job characteristics, job flexibility, opportunities to work past retirement age, employment history, pension rights, sources and composition of current income, wealth and consumption, housing, education), and *social support variables* (e.g. assistance within families, transfers of income and assets, social networks, volunteer activities, time use).

A *longitudinal follow-up collected data* in 2006/07. In addition to the original 11 SHARE countries, four new countries were included in this second wave: the Czech Republic, Poland, Ireland and Israel. Data releases (April 2005, June 2007, November 2008, August 2009), first substantive results, and a detailed description of the survey methodology are available under www.share-project.org. A third wave collecting unique *retrospective data on life histories* was fielded in 2008/09. Data and first results will become available during 2010.

SHARE has become a *world class example* for the successful integration of disciplines as well as countries. Since its first release of the prototype data, SHARE has attracted nearly 2000 registrations with many more users from all over the world, resulting in a host of scientific publications. SHARE has sparked similar survey projects in China, India, Japan, Korea, and Thailand. Particulary due to its electronic language management and ex-ante harmonization system, SHARE is regarded state-of-the-art in multidisciplinary comparative survey technology.

ESS

The European Social Survey (ESS) is an academically-driven, biennial cross-national survey series which is designed to chart and explain the interaction between Europe's institutions and the changing attitudes, beliefs and behaviour patterns of its diverse populations. It is designed to produce easily accessible data on European social change which serve both academic and policy needs. Its twin aim is to improve the rigour of quantitative social measurement throughout and beyond Europe, which it achieves not only by example but also by a combination of extensive dissemination and training.

The ESS covers 34 countries. These include 26 of the 27 EU member states (the exception is Malta), plus Croatia, Iceland, Israel, Norway, Russia, Switzerland, Turkey and Ukraine. Owing to the ESS's extensive spread across countries and the coordinated rigour of its measurements, it generates a transparent moving picture of multi-cultural change in Europe's political, social and moral framework. Several ESS questions and scales have already been replicated in the US, Australia and South Africa, thus providing a measure of cross-continental differences too.

The ESS was initiated in the 1990s when it was seed-funded by the European Science Foundation in response to what they saw as a paucity of high quality time series covering public perceptions of, and attitudes towards, the social, economic and political fabric of different societies. It has been funded since 2001 both by the European Commission's 5th, 6th and 7th Framework Programmes and by each of its 34 participating countries. In these respects the ESS is an exemplar of the European Research Area at work. Having already been widely acclaimed for its exceptional rigour and accessibility, the ESS received the 2005 Descartes Prize for "excellence in collaborative scientific research" – the first social science collaboration to win this coveted award.

The ESS's diverse range of data is available to all, whether for scrutiny or detailed analysis. Almost 35,000 registered data users have already quarried its data, generating over 20 books and over 200 journal articles. Its data are also used extensively by governments and think tanks for policy purposes, and to the general public, thus enabling citizens to investigate and understand more about their own society's values and behaviour patterns and to compare them with those in other societies.

The ESS is directed by a Central Coordinating Team (CCT) headed by the Centre for Comparative Social Surveys at London's City University in the UK. The CCT also comprises six other partner institutions: Norwegian Social Science Data Services (Norway), GESIS (Germany), Netherlands Institute for Social Research, Catholic University of Leuven (Belgium), University Pompeu Fabra (Spain), and the University of Ljubljana (Slovenia).

CESSDA

CESSDA differs from most other ESFRI projects by being an established infrastructure already. The services of CESSDA have been provided by the individual member institutions for decades. The cross national CESSDA cooperation has been active for more than 30 years. Over the years numerous common infrastructure initiatives have been completed by CESSDA members. In the last 15 years some of these projects have been supported by EC. Among them are such projects as NESSTAR, FASTER, MADIERA, METADATER, LIMBER, CESSDA PPP and others. The purpose of joining the ESFRI process is for CESSDA to intensify and formalise the cooperation.

In the CESSDA PPP the focus was on controlled vocabularies, metadata standards development, federated user access, authentication and authorisation, persistent identifiers, training and exchange, data accessibility and technical platforms for cyber infrastructure for social science and the humanities.

The process of establishing a CESSDA ERIC is well advanced. The basic documents are in place and the formal agreement is under way. The CESSDA ERIC will be a distributed activity with a formal seat in Bergen, Norway, with major activities in Germany, with other common activities undertaken in other member countries and national activities carried out by service providers in each member country.

Some of these issues clearly calls for common solutions that will be pursued in the future CESSDA ERIC which presently under negotiation between a group of countries containing the majority of the present CESSDA members. It is however also clear that some of these issues actually have relevance outside CESSDA circles. And that is exactly the reason why it makes sense for CESSDA to participate in the DASISH cooperation.

DARIAH

DARIAH's overall goal is to design a virtual bridge between various humanities and arts resources across Europe. The mission of DARIAH is to enhance and support digitally-enabled research across the humanities and arts. DARIAH aims to develop and maintain an infrastructure in support of ICT-based research practices and is working with communities of practice to:

- Explore and apply ICT-based methods and tools to enable new research questions to be asked and old questions to be posed in new ways;
- Improve research opportunities and outcomes through linking distributed digital source materials of many kinds; and
- Exchange knowledge, expertise, methodologies and practices across domains and disciplines.

Thus, DARIAH seeks to support all disciplines across the humanities, encouraging inter disciplinarity and the exploration and sharing of content, tools and methods.

The DARIAH network will be designed to be a decentralised network of competency centres (VCC – Virtual Competency Centres), which will allow services to stay close to end-users (researchers) both geographically and thematically. Common technologies (e.g. for authentication or federation of archive contents) and good practices (standardised formats, digital assets management workflows) will ensure coherence across the support services offered by the competency centres. The DARIAH ERIC is expected to be established during 2011 with France and Germany being co-hosts for the Coordination Office.

CLARIN

At the technological level CLARIN followed a two-track approach – on the one hand concrete prototypical solutions with a long-term potential were developed and on the other hand standards and concepts were pushed ahead with a long-term goal. Most of the he following activities are operational and can be mentioned in particular:

- a multi-national authentication and authorization trust domain
- an offer for a domain wide robust persistent identifier solution
- steps towards centres with a proper repository system based on requirements few offering language resource deposit possibilities
- a component-based flexible metadata framework based on the ISO 12620 standard and a professional implementation of the ISOcat concept registry
- a Virtual Language Observatory with a wide range of language resources and tools based on harvested metadata
- the basis for a common framework for workflow solutions based on a few concrete prototypes
- a number of standardization initiatives together with ISO TC37/SC4 have been pushed forward by CLARIN in collaboration with colleagues from the US and Asia mainly and where possible CLARIN already based its work on emerging standards

In addition much work has been done in national initiatives to improve the visibility and accessibility of language resources and tools. In total CLARIN can claim to be at the cutting edge of the technological work in its field worldwide. Much awareness about state-of-the-art methods has been spread by running a large number of training courses within and beyond the community.

At the organizational level CLARIN is moving towards an ERIC, i.e. the statutes are being finalized and it seems obvious that at least about 10 countries will be ready to give clear commitments to participate in funding the ERIC and the corresponding national CLARIN chapters.

In total we can claim that CLARIN is ready to start the implementation work.

Cross-Discipline Work

All 5 initiatives focused on the specific issues of their disciplines to demonstrate their potential in advancing the state towards an e-Science enabling infrastructure. The consequence was that cross-disciplinary fertilization was addressed only occasionally. Technical meetings were organized between CLARIN, DARIAH and CESSDA and we also can refer to the NEERI conferences 2009 and 2010 that served as an interaction platform. ESS has helped SHARE to develop a European sampling scheme. ESS, SHARE and CESSDA have worked jointly in matters of translation and questionnaire development procedures for cross-national social surveys. Due to the different agendas defined by the transformation to ERICs and the enormous time pressure for the work we yet missed opportunities to take up real activities with a broad coverage. In DASISH is we start this cross-fertilization and stabilize our infrastructure efforts.

B 1.1 Concept and project objective(s)

B1.1.2 Concept and Idea

General View

In the Social Sciences and Humanities research is increasingly driven by the availability of a variety of digital resources, which exhibit an escalating internal complexity as well as multifarious external relationships. The data production, management and dissemination processes are widely organized in a distributed manner, which is leading to a situation of great fragmentation that we need to take into account when designing research infrastructures for the corresponding scientific disciplines. This not only includes primary data such as audio/video recordings or collections of digitized primary sources in the humanities area and the results of surveys as in the Social Science area, but also the secondary data that is derived from these sources as well as the continuous enrichments resulting from multiple annotations for example. All these data need to be made persistently available to the users mostly via computer environments that implement discipline specific workflows and traditions.

Research in social sciences and arts and humanities has created a large quantity of digital material that represents a significant investment, both in terms of public funding and of intellectual effort. Given the current lack of infrastructure for sustaining this material, these resources are typically hosted in their home institutions using a variety of approaches and technologies. This situation incurs a number of risks. At the most basic level, without on-going maintenance a resource will cease to be usable at all as the technologies in which it was implemented become obsolete and unsupported. Even if hosting institutions apply digital preservation techniques to ensure that resources continue to be accessible – something that at present is not universal – this does not enable the collections to make full use of technological advances that might greatly enhance their use to researchers and their impact on research. Access to legacy resources may be limited to a simple download or by browser access in a website; in neither case does this facilitate advanced research services that will be more commonly used in future digital research such as mash-ups or data/text mining.

It must be remembered that the impact of research in the social sciences and humanities may still be felt many years after the original research was undertaken – the information produced has a long lifespan in intellectual terms. Sustainability does not just mean keeping the data alive, but enabling the exploitation of advances both in technology – making the data accessible in new ways – and in humanities and social sciences research – forging connections between resources that lead to new discoveries and broader impact. It is the enhanced visibility and re-usability of digital resources and tools/services that is at the heart of DASISH. Only like this we can ensure long-term interest and sustainability of these resources. In contrast to natural sciences where the quality of the primary data is guaranteed by expensive sensor equipment, the quality of data and metadata in SSH often is dependent on human efforts. It can be observed that the enhanced use of SSH data in terms of eScience paradigm is restricted by the lack of quality – thus a campaign to increase quality will be essential to make quantum-leaps.

In CLARIN, DARIAH, CESSDA, SHARE and ESS it is widely agreed upon that such data need to be made available via networks of community based solutions, organized and sustained by centres of expertise

organized within the communities[1], which can provide the stability and reliability that is required by the researchers. These centres are specialized on community related access and enrichment tasks, and, in order to comply with eco-principles, they should avoid duplicating basic services such as persistent identifiers, AAI (Distributed Authorization and Authentication Infrastructure) or long-term preservation services, which obviously cross over disciplines and infrastructures. The latter is also true for more content oriented services such as data quality assurance and access licenses.

A digital eco-system of infrastructures where community centres collaborate with e-Infrastructures could probably be sketched as in the figure below, but more effort should be spent across on-going initiatives to account for the variety of user needs as well as the possible synergies across disciplines.

Each of the SSH initiatives has defined the kind of services that they need to implement to satisfy the communities and each established requirements for their centres to get the infrastructure and their services operational. On the other hand we expect collaboration in particular with

- GEANT/eduGain/TERENA to establish a European trust domain allowing single identity and single sign-on mechanisms
- a new data e-Infrastructure that can act as a layer offering common data services as explained in the figure introduced by the HLEG on scientific data[2] (figure below)
- an easy-to-use and integrated infrastructure offering moderated computational services as they are relevant for SSH[3]

DASISH will interact closely with e-Infrastructures to push solutions that are driven by the community needs, thus time needs to be reserved to work out requirements and closely discuss them with all e-Infrastructures that address the practical needs. This will help a) to prevent solutions that are unnecessary complex and cannot be integrated with the community approaches and b) to speed up processes to come

---

[1] Community centres can take various organizational forms. Some centres in SSH have their own computing and storage facilities, but some just provide the knowledge and outsource their computational and storage needs to data or specialized computer service centres.

[2] http://**cordis**.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf

[3] At the recent EGI User Forum in Amsterdam a whole session was devoted to the SSH cluster proposal. At the end it was not obvious what EGI could do for DASISH. But it was agreed to keep in touch.

to pan-European solutions. On the other hand DASISH is relying on proper and speedy decisions with respect to the mentioned eInfrastructures to not need to invest time ourselves in setting up solutions.
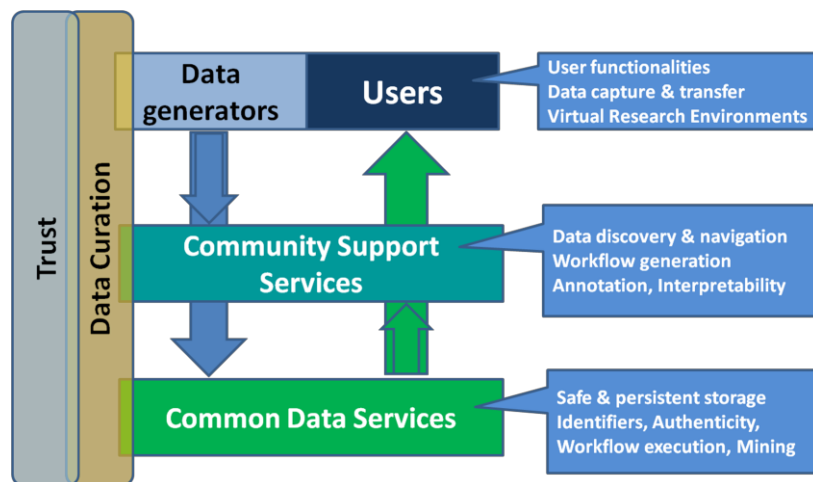
Specific Views

While all five initiatives share the topics mentioned in the general view we need to explain a difference in focus which is a consequence of the different natures of the SSH research infrastructures. CESSDA, CLARIN and DARIAH are focusing on networks of connected centres offering integrated and interoperable services, SHARE and ESS are much closer to the data creation and analysis process. Actually SHARE and ESS are situated more in the first layer of the previous diagram and act as data creators, while the other three research initiatives are located in the second layer.

Consequently the term "infrastructure" has a slightly different connotation which has been accepted in the ESFRI process. For SHARE and ESS the primary focus is on improving the quality of the content of the surveys to make them interpretable in a cross-cultural and cross-national context. "Infrastructure" here means primarily to set up a system that supports the needs in quality and efficiency improvement at data creation time and at monitoring the process. Needless to say that the results of the surveys will be deposited at CESSDA centres for example and that by improving the quality interoperability will be made much easier.

Challenges and Approach

All research disciplines are faced with great challenges with respect to data creation, management, curation and access. Four major challenges can be distinguished:

- How can we manage to preserve our cultural and scientific memory and keep the records of science accessible?
- How to come from a down-load first scenario to a truly web-based usage scenario to optimally access and enrich the stored data to tackle the many big and small research challenges?
- How can we improve the quality of our data to enable advanced and cross-disciplinary access and enrichment operations?
- How to establish trust of SSH researchers in the infrastructure services?

Finally the SSH community needs to improve the understanding how it will go ahead in future with respect to community-based infrastructures. Currently, we are building 5 different ERICs; this makes very much sense given the differences in research paradigms, community practices and the state of using technology. However, it needs to be investigated whether there will be possibilities of merging infrastructures to foster interdisciplinary work, to reduce the overhead and therefore costs. Since researchers' trust in infrastructures will be the key point for its acceptance, this can only be driven by a bottom-up learning process.

Given these major challenges DASISH organizes its work along a few major dimensions:

- understand the different usage scenarios and architectural solutions to identify ways to come to common SSH solutions for data and service integration and interoperability (WP2)
- identify major quality issues and take serious efforts and measures to improve quality with foci on survey quality enhancements (WP3) and the quality of metadata and data access (WP5)
- establish criteria for long term persistency and curation of data and interact preferably with the emerging data infrastructure to push the quick deployment of first generally available services (WP4)
- work on all aspects that will foster shared data access and enrichment starting with basic layers such as AAI based trust domain up to cross-disciplinary data enrichment frameworks (WP5)
- take care of legal & ethical issues that are of relevance for all SSH domains in a cross-disciplinary activity and work on simplified solutions (WP6)
- take a variety of measures in trust building and to engage the communities – in particular the young generation of researchers and perhaps even the public – in making use of advanced features of the infrastructures by education and training programs (WP7)
- disseminate the results according to proven channels (WP8)

B1.1.3 Objectives
The major objectives for DASISH can be extracted from what was described above:

WP1:
- to efficiently manage the project at administrative/organizational and content level

WP2:
- to get a deep understanding about the differences in the way data is managed and organized in the SSH with the goal of a stepwise harmonization towards a reference architecture, better integration and higher quality

WP3:
- to improve the quality of survey data to make it fit for the cross-country/cultural and lingual challenges

WP4:
- to offer robust deposit and long term preservation services for all SSH researchers preferably in collaboration with a European data infrastructure based on trust

WP5:

- to improve (meta) data quality and the infrastructure integration by sharing joint methods for basic services such as AAI, PID, metadata etc. as the basis for typical automated e-Research workflows
- to improve the visibility, accessibility and re-use of state-of-the-art tools and services that are useful in a cross-disciplinary context and provide a cross-disciplinary annotation platform

WP6:

- to bundle SSH forces to simplify access restrictions for researchers in particular due to the many IPR issues the SSH domain is faced with

WP7:

- to take serious measures to educate and train SSH researchers about how to efficiently apply state-of-the-art methods and technology by taking profit from each other's knowledge, and to involve the public and pupils in these efforts

WP8:

- to reach out to the participating communities and beyond

By defining these objectives DASISH addresses the following topics of the call:

- Many action points (infrastructure, quality, technology and knowledge exchange, exchange of tools/services, education, legal issues) address the common needs for the implementation phase and will use cross-disciplinary synergies.
- Some activities in the DASISH project will optimize the implementation plans and facilitate cross-walks due to the explicit and implicit harmonization and integration efforts (metadata, concept registries, persistent identifiers, distributed authentication based on single identity, etc.).
- Respecting the differences in community approaches and traditions DASISH will implement a few cross-disciplinary and ready-to-use services making use of common infrastructure solutions, but also respect the special needs of the survey-oriented work.
- DASISH sees itself as contributing to an overall eco-system of infrastructures and on purpose wants to link in particular with GEANT/eduGain to create a broad SSH based AAI trust domain and with the emerging data infrastructure[4]. Intense discussions with our EGI colleagues at NEERI 2010, at the EGI user forum and in bilateral meetings for example at CERN resulted in the agreement to continue the exchange – yet concrete collaboration could not be indicated.
- Needless to say that DASISH builds on all standardization efforts being undertaken by the communities and foster future standardization work and awareness.

All communities are very much interested in the DASISH collaboration, although the foci are slightly different due to the nature of the communities main work, thus, we assume sufficient commitment in particular since we also defined criteria for the centers who showed interest in participating in this consortium. This initiative was also used to motivate centers again to adapt their practices to better facilitate e-Research.Key participants from the communities are part of this consortium as well as centers

---

[4] Yet we do not know whether such a data infrastructure will emerge and whether the work will be oriented towards fast solutions, i.e. in case of problems DASISH will make separate agreements with a few data centers that are ready to deliver services as temporary and limited solutions.

who invested quite some adaptation effort and who already showed commitment in the preparatory phase.

B1.2 Progress beyond the state of the art

Integration and Interoperability

Worldwide we can observe a number of initiatives to network the humanities and the social sciences to improve the level of data integration and to enable interoperable data usage.[5] These coalitions share resources and use web-based collaborative environments based on standards to tackle new research questions. Europe can benefit from the high quality level of work and the density of interactions and awareness building in this respect already achieved now by the ESFRI process and its national counterparts. In contrast to the Bamboo project who first organized large conferences worldwide to find out what the humanities researchers' needs are and then turned completely to a technology project, all SSH projects followed a balanced approach by organizing the (sub) communities and immediately developing technology for useful services.

DASISH is a way to follow up naturally on this path knowing that infrastructure building as an enabler for e-Research is not a one-step-process, but a process of continuous adaptations to improve integration and interoperability[6]. Moving only one small discipline on this path would not be feasible, since the adaptations researchers and research organizations need to undertake are enormous - some speak about a change of culture in this respect. Going together, adopting the same type of "language", integrating into an eco-system of infrastructures and promoting synergies will speed up this process significantly.

Data Quality Benefits

Another enabler for smart eScience type of operations is data and metadata quality. With the current state of quality there will be severe limitations in all sorts of automatic treatment. Quality mainly is an issue to be tackled at data creation time, but also some improvements can be achieved during data life cycle management as is suggested in WP5. In this chapter we will focus on the special needs to be addressed by SHARE and ESS during data creation time.

Great strides have been made over the years (most notably by Eurostat), to create rigorous cross-nationally comparative surveys in Europe – such as the Eurobarometer, the European Election Study, the European Quality of Life Survey, the European Values Surveys, the Gender and Generations Survey, and of course ESS and SHARE. Yet, several key dimensions of data quality still show serious shortcoming even within these celebrated European data sets. Examples are lack of equivalence between certain items of data from different nations, incoherent definitions and sloppy treatment of the many different languages in the Union; the low and declining response rates, the difficulties of some surveys too maintain their panel stability. Most surveys are acutely aware of these common imperfections within their data but the key underlying problems are still in need of robust remedies.

---

[5] e.g. centerNet http://digitalhumanities.org/centernet/ ; Bamboo; http://www.projectbamboo.org/; Social Science Archives in Collaboration for Preservation as presented by J. Crabtree (http://www.asis.org/Conferences/RDAP10/RDAP10Program.html)

[6] As can be seen from the e-IRG Task Force Report on Data Management "interoperability" cannot be a goal as such due to the inherent complexity in particular at the semantic and cross-cultural level. What can be enhanced is the potential of creating interoperability more easily due to the adoption of standards and the explicit adherence to basic IT principles such as registering all concepts and schemas being used.

DASISH tacklea five of these key problems, each of which has a realistic chance of bringing considerable benefit to a range of multinational surveys within the social sciences: a stringent and operational definition of occupation codes, a key for understanding labour market attachment; a transparent IT-based methodology for developing questionnaires in many languages simultaneously; an IT-based improvement of fieldwork control to increase response and retention rates in Europe; a systematic approach to sampling taking account the vastly different statistical backgrounds in Europe; and a system of geocoded and other contextual data alleviating cross-national analyses.

These tasks of DASISH are jointly directed by ESS and SHARE and first applied to these two data sets. In addition to improving data protocols and data enrichment, they enhances the state-of-the-art in Europe by addressing those critical aspects of data quality which are common to Pan-European surveys. The project thus goes much beyond ESS and SHARE: the achievement of such improvements of the state-of-the-art in survey technology would be of long-lasting value both to better-informed European governance and to well-founded academic research.

General Benefits
The partners will contribute to and benefit from DASISH in a number of ways. Foremost, (a) the gained influence by 5 large-scale infrastructures, (b) enhanced sustainability through synergies between the ESFRI initiatives, (c) the increased availability of resources through interoperability, as well as (d) ensuring usability and implementation through combining education and training across all relevant disciplines:

- DASISH helps to stabilize and broaden the path towards establishing a European-wide trust federation driven by practical user community needs and thus not only facilitate access but also spread the message across communities.
- DASISH strengthens the existing strong solutions driven by scientific communities and allows cross-fertilization.
- DASISH helps to improve interoperability across SSH resources through standardization of repository federation protocols, an integrated PID environment, as well as mapping data descriptions.
- DASISH allows for coordinated approaches regarding data management, preservation and curation and therefore reduce appropriate efforts in the ESFRIs as well as establish a broad knowledgebase for new communities facing similar challenges.
- DASISH helps unify repository setups and thus facilitate access to data and its metadata and it will allow us to design abstract generic interfaces that will allow the SSH centres to more easily link up to e-Infrastructures and influence international standardization.
- DASISH helps spreading the messages about the relevance of high quality metadata and data and unifying the metadata infrastructure for research data in particular with respect to the registration of concepts and schemas in web-accessible registries.
- DASISH helps to significantly increase data quality in particular at survey creation time.
- DASISH detects commonalities in data management and curation with a focus on data quality based on its extensive domain experiences.
- DASISH creates a shared community across all SSH disciplines, which opens the possibility to incubate new research questions and methodologies, fosters exchange of experiences and best practices, and raises the visibility of the SSH domains.

- DASISH creates a network of tools and supplementary material (e.g. historic encyclopaedias), which is of great research value to all SSH fields.
- DASISH helps building an attractive portal making resources and tools/services much more visible to interested researchers and the public. It will also help to improve the skills to build smart filters oriented towards serving specialized sub-communities.
- DASISH helps to make many more researchers aware of the potential of digital infrastructure and virtual research environments, thus getting in particular more young people involved in creating new services that can be used by everyone across disciplines.
- DASISH allows us to offer some services that are common for all SSH researchers.
- DASISH helps integrating training and education, and thus provide users from all SSH domains with a single point of reference, which will increase the user focus and research capacities.
- DASISH allows us to tackle difficult IPR issues with much more power, since to a certain extent DASISH can claim that it represents the SSH communities at large in data related issues.
- DASISH combines the influence of some of the key researchers and institutions in the SSH community, which will help to advocate policy development and funding in the European Research Area.
- DASISH strengthens the world-leading position in SSH research by combining the capacities of all SSH infrastructure initiatives.

B1.3 S/T Methodology and associated work plan

B1.3.1 Overall strategy and general description

The work and discussions in WP2 guides the whole project, since all activities should be compliant with the master plan which is (a) determined by the work in the participating initiatives and (b) by the agreements to create a joint SSH domain and to do cross-discipline fertilization. WP2 therefore has a number of overarching activities such as assessing the state of technology and quality and develop a common roadmap, setting up a knowledge base and filling the knowledge base in particular with information about useful tools and services and by assessing the work done in the project.

The data and metadata quality campaigns have two strands: (1) One is specifically addressing the needs of the social surveys by setting up an appropriate "infrastructure" (WP3); (2) The other is addressing the general needs to improve metadata about all types of resources (WP5).

Integrated services are generated in work packages 4 and 5. WP4 is focusing on joint data storage and preservation services that allow all SSH researchers to deposit their valuable data and to achieve policy-based safe replication of data together with the emerging common data eInfrastructure (or if this does not exist with selected national centres). WP5 is focusing on creating joint service domains by (1) extending the existing distributed authentication domains and the persistent identifier services, (2) creating a joint metadata domain, (3) extending and testing out processing chains for SSH applications and (4) offering a generic annotation framework for web-based data.

All three work packages WP3/4/5 start with assessing the state and then start with concrete work to offer improved services as soon as possible. WP5 in addition will influence and help constructing the community based infrastructures.

All activities and services need to be paralleled by training and education programs and lead to dissemination activities. Part of this task needs to be carried out by the individual infrastructures themselves, but also in this respect DASISH should allow these infrastructures to benefit from each other and to save time. The training programs should allow the social scientists to add a block into their training program about multimedia annotation which is given by a CLARIN specialist for example. Also the common tools & services commentary (WP2) is seen as a start to benefit from each other's knowledge and it will be maintained by the individual infrastructures after the project end.

Very important is to join efforts in the area of IPR issues, since some of them such as a simplification of access rights and licensing issues can only be achieved by a large group. In this respect the building of a Virtual Competence Centre for IPR issues will be a major achievement and it will be built in a way that also the activities in the infrastructure initiatives can make use of this shared knowledge base.

B1.3.2 Timing of work packages and their components

| Task	Task Name	Y1 Q1	Y1 Q2	Y1 Q3	Y1 Q4	Y2 Q1	Y2 Q2	Y2 Q3	Y2 Q4	Y3 Q1	Y3 Q2	Y3 Q3	Y3 Q4
1.1	Consortium Agreement	█											
1.2	Project Management & Reporting	█	█	█	█	█	█	█	█	█	█	█	█
1.3	Administrative support for Workgroups and Workshops		█	█	█	█	█	█	█	█	█	█	█
1.4	Negotiations with third parties	█	█	█	█	█	█	█	█	█	█	█	█
1.5	Liaison with external partners (EC)	█	█	█	█	█	█	█	█	█	█	█	█
2.1	State of Architectures	█	█										
2.2	Reference Architecture		█	█	█			█	█			█	█
2.3	Tool and services knowledge registry			█	█	█	█	█	█	█	█	█	█
2.4	Quality Assessment of reference architecture					█	█	█	█	█	█	█	█
3.1	Coding of Occupation Software		█	█	█	█	█	█	█	█	█	█	█
3.2	Multi-language Questionnaire Software	█	█	█	█	█	█	█	█	█	█	█	█
3.3	Fieldwork Monitoring System	█	█	█	█	█	█	█	█	█	█	█	█
4.1	Data Management and Curation	█	█	█									
4.2	Assessment of Deposit Services		█	█									
4.3	Deposit Service convergence					█	█	█	█	█			
4.4	Recommendation of a set of policy rules							█	█	█	█	█	
5.1	Trust Federation	█	█	█	█	█	█	█					
5.2	Robust PID Services	█	█	█	█								
5.3	Metadata Quality Improvement			█	█	█	█	█	█	█	█		
5.4	Joint Metadata Domain			█	█	█	█	█	█	█	█		
5.5	Workflow Implementations						█	█	█	█	█	█	█
5.6	Annotation Framework			█	█	█	█	█	█	█	█		
6.1	New Ethical and Legal Challenges	█	█	█	█								
6.2	Virtual L & E Competence Centre					█	█	█	█	█	█	█	█
6.3	Preservation Challenges	█	█	█	█	█	█			█	█	█	
7.1	Training Modules	█	█	█	█	█	█	█	█	█	█	█	█
7.2	Workshop Programme	█	█	█	█	█	█	█	█	█	█	█	█
8.1	Dissemination Strategy	█											
8.2	Web-Site, Newsletter and fact sheets	█	█	█	█	█	█	█	█	█	█	█	█
8.3	Conference participations	█	█	█	█	█	█	█	█	█	█	█	█
8.4	Final Conference									█	█	█	█
8.5	Qualitative Data Workshop		█										
8.6	Quantitative Data Workshops		█	█	█	█	█	█	█	█	█	█	█

Legend: marked areas mark periods of activity; black boundaries indicate deliverables

B1.3.2.1 WP Dependencies

DASISH has 3 work packages that are vertical in the sense that they need to interact with all content work packages: Management, Dissemination and Legal & Ethical Issues. With respect to content aspects DASISH differentiates into three layers:

- the first layer is determined by getting a deep understanding about cross-disciplinary architecture and quality issues as an enabler of advanced services
- the second layer is determined to produce concrete services and quality enhancements
- the third layer is determined by using all results for education and training efforts and to make use of the services and harmonization effects in the five research infrastructures

The arrows indicate schematically the major influences of the results. Of course much more interaction will be necessary between the different activities. DASISH will react flexibly on new challenges by establishing working groups and task forces where necessary.
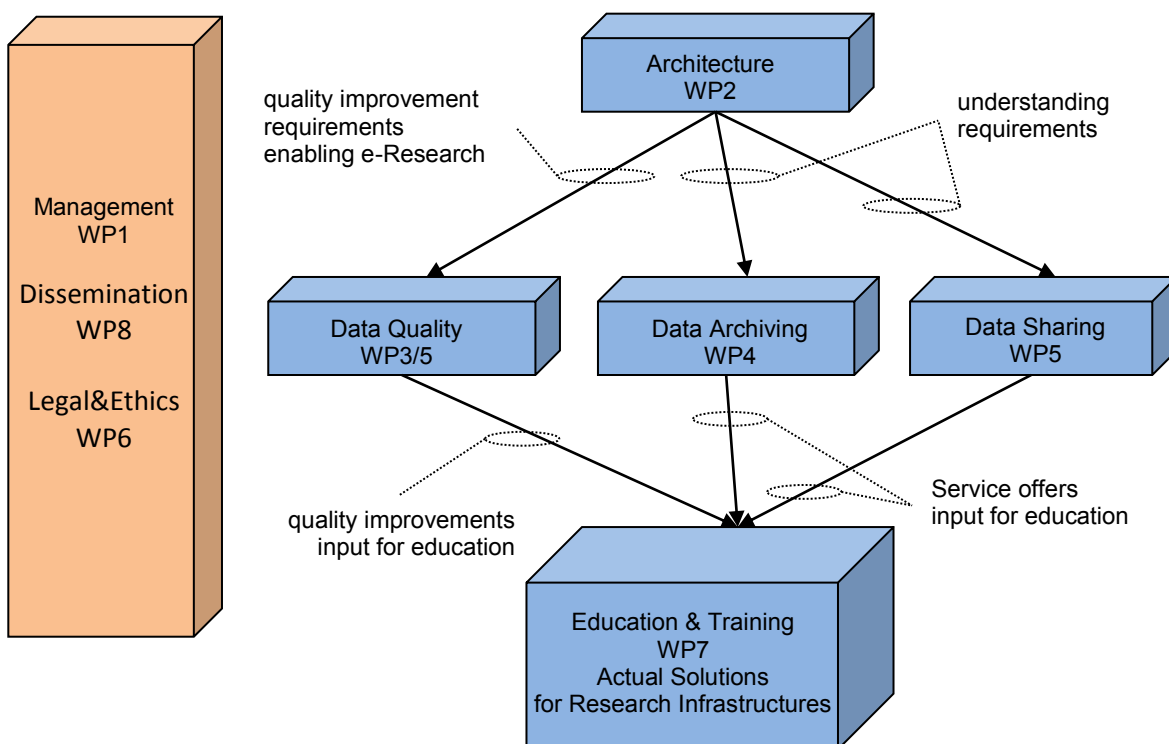
B2. Implementation

B2.1 Management structure and procedures

The governance structure ensures that

- there is an efficient management and leading structure
- the communities are represented as well as the work packages
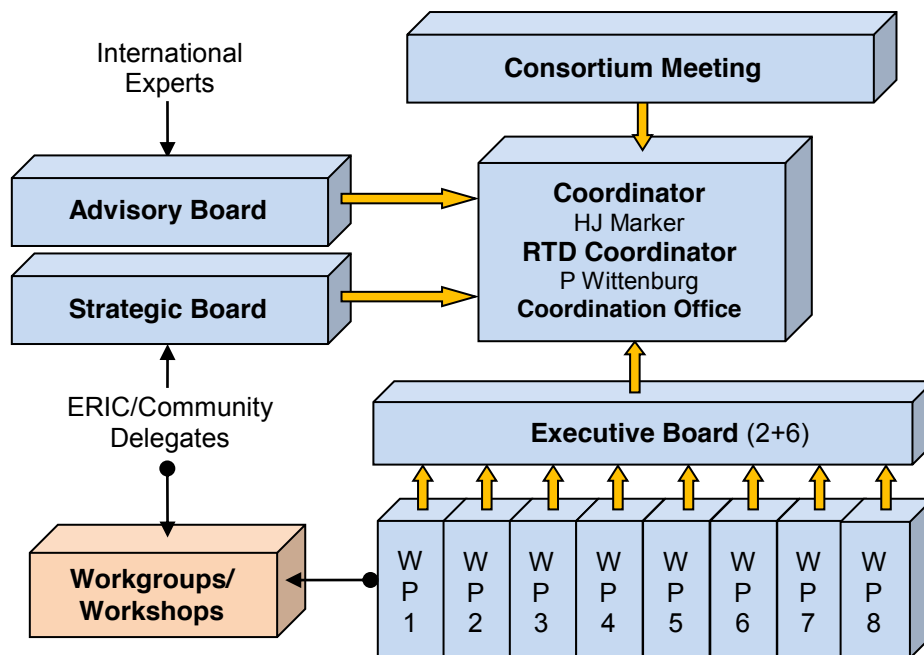
The EB in which all WP leaders are represented including the two coordinators who also lead WP1, WP5 takes the relevant decisions to achieve the intended results. The coordinators lead the DASISH work, take care that decisions are put in action, take care of the compliance with all advice and requests from the Advisory Board and the consortium meeting, take care of compliance with the grant agreement and interact closely with the EC. The coordinators are supported by a coordination team. The coordinator is taking care of the work in WP1, WP6, WP7 and WP8, the RTD coordinator that of WP2, WP3, WP4 and WP5.

The consortium meeting includes all partners and controls the work carried out at strategic level. The Advisory Board will cover experts international experts from the SSH domains to give advice about the relevance of the work being done. The Strategic board will cover delegates of the 5 communities (ERICs when they have been established) and also delegates from the EC.

Joint working groups and workshops will bring experts together from the interested communities to discuss specific issues of common interest and to synchronize activities.

B2.2 Beneficiaries

1. University of Gothenburg

University of Gothenburg is represented in DASISH by the Swedish National Data Service, SND. SND is the Swedish member of CESSDA and is responsible for collection, protection, documentation, availability and dissemination of digital resources of relevance for research in the humanities, social science and health science. SND takes part in standards development primarily through the DDI Alliance. SND develops tools for documentation of digital resources and for making them available to secondary research. SND provides guidance for researchers especially in connection with secondary research, legal and ethical aspects in connection with data base establishment and use of digital resources and documentation of digital

resources. SND is very active in international cooperation. As SND belongs to University of Gothenburg SND benefits from experience and resources of the university especially the Research and Innovation services.

Key personnel

Hans Jørgen Marker assumed the post of director of the Swedish National Data Service, SND, in 2009. Before that he worked at the Danish Data Archive, DDA, for 25 years, and was director from 1997 to 2003. He is president of CESSDA since 2010. From 2006 until 2009 he was chairman of the DDI Alliance expert committee, and he was Secretary of CESSDA from 1998 to 2010. He was President of the Association for History and Computing from 2001 to 2004. Hans Jørgen Marker has played a responsible role in connection with the EU projects NESSTAR, FASTER, MADIERA, METADATER and CESSDA PPP and he is connected to the present application Data without Boundaries as member of its steering committee.

Birger Jerlehag has worked for nearly 30 years in the "data archive world". He has mainly worked as a systems administrator but have also experience of documenting and dissemination of data. He is experienced in the methods and techniques used by the researchers of quantitative data. In his role as IT-coordinator he has contacts with and knowledge of national resources and institutions working with similar tasks. He plays a central role in supporting SND's users in the humanities.

2. Max Planck Society MPG

Max Planck Society (MPG) is represented in DASISH by two institutions The Language Archive (TLA), which is part of Max-Planck-Institute for Psycho Linguistics (MPIPL), and Munich Centre for the Economics of Aging (MEA), which is part of Max-Planck-Institute for Social Law and Social Policy (MPISLSP).

The Language Archive (TLA) is a new unit in the institute covering the archiving and tool/infrastructure developers, the team that participates in CLARIN activities. The unit was established to strengthen the sustainability of the investments made so far. The experts involved in TLA have been participating in a large number of national and international projects since 2000 under which are a number of tool, standardization and infrastructure oriented projects funded by the EC (MUMIS, ISLE, ECHO, INTERA, DAM-LR, LIRICS and CLARIN). In addition TLA is housing one of the largest language archives which currently offers about 74 Terabyte of language resources with all resources being metadata described and accessible via web-applications. The team also has been participating in a number of standardization efforts in the realm of ISO TC37/SC4 (component based metadata, persistent identifiers, 12620 data category registry and its ISOcat instantiation software, Lexical Markup Framework, structured annotation formats).

Key personnel

Wolfgang Klein, scientific director of the MPIPL and responsible for TLA, is a linguist and will give scientific advice for the DASISH work.

Peter Wittenburg, for 34 years technical director at MPIPL and now director of TLA, is technologist (electronics and IT) and has a long record in leading international projects and work packages. He will bring in 40% of his time to lead the RTD work in DASISH and linking DASISH work with CLARIN on the one hand and with e-Infrastructures on the other hand.

Daan Broeder, for many years senior developer for archive and infrastructure solutions and now deputy director of TLA, is technologist (electronics and IT) and has a long record in leading development tasks in

international projects. He will participate in DASISH development work and also link to the ongoing CLARIN work.

Menzo Windhouwer, now senior developer for knowledge based systems in TLA, is software engineering expert and was very active in designing and developing semantic interoperability solutions and in pushing ISO standards (12620, ISOcat). He will contribute to DASISH with his great expertise in standards and software solutions.

MEA was founded in 2001. It has rapidly developed into one of the world-renowned centres of excellence for the economics of ageing. Its general aim is to anticipate and monitor the micro and macroeconomic aspects of population aging.

The activities of the MEA are divided into four broad areas of research:

Social policy implications of aging, especially for pension systems and health care provision

Savings behaviour of households, especially in old age

Macroeconomic implications of an aging society, especially on growth and productivity

Public health and statistical analysis of data on health, ageing and retirement

MEA currently employs a staff of 28 scientists, some 40 research assistants, 5 part-time administrative staff. MEA has its own PC-based network with several servers, including facilities to handle very large data sets and confidential data. The institute has a library, a conference room, and all necessary technical support facilities. MEA publishes a quarterly newsletter and its own discussion paper series.

Key personnel
Axel Börsch-Supan is Director of the Munich Research Institute for the Economics of Aging (MEA) and Professor for Macroeconomics and Public Policy at the University of Mannheim, Germany. He received his Ph.D. in economics from the Massachusetts Institute of Technology in 1984. From 1984-89 he was assistant professor of public policy at Harvard's John F. Kennedy School of Government before he returned to Germany in 1989. Börsch-Supan is chairman of the Council of Advisors to the German Economics Ministry and chaired the pension reform unit of the German Social Security Reform Commission. He is speaker of the special research group on behavioural economics (SFB504) at the University of Mannheim, research associate at the National Bureau of Economic Research (NBER) in Cambridge, Mass., and member of the German National Academy of Sciences. He has written and edited more than a dozen books. His journal articles focus on saving and retirement behaviour, pension reform and population ageing and have been published e.g. in the American Economic Review, the Journal of Public Economics, the Journal of Econometrics, and in Economic Policy. Since 2002, he coordinates SHARE

3. Universitat Pompeu Fabra, UPF
The Institute for Applied Linguistics, IULA, created at the Universitat Pompeu Fabra in 1994, is a research and graduate training centre that gets together about 70 researchers. IULA members are UPF professors of the Department of Translation and Language Sciences, the Department of Information Technologies and Communication, and PhD students, research scholarship holders and researchers contracted for specific projects. IULA has been developing Language Resources and tools for their exploitation since its creation.

The IULA's research group "Tecnologies dels Recursos Linguistics" directed by Dr. Núria Bel, is currently involved in the following European projects: Common Language Resources and Technologies Infrastructure CLARIN, FLaReNet: Fostering Language Resources Network *(*ECP-2007-LANG-617001) and are coordinating the PANACEA project (7FP-ITC-248064) started January 2010 and devoted to the development of a web service-based platform to be used as a factory for the automatic production of language resources. The Research centre RECSM is an expertise centre for survey research. The head of the centre, Prof. Dr. Willem Saris, is member of the Central Coordinating Team of the ESS since the beginning. Besides him Dr Daniel Oberski is working on quality control and improvement of survey research.

Key personnel
Dr. Núria Bel (PhD in 1993 by Universitat de Barcelona) Associated Professor of the UPF joined the IULA in 2003 where she currently is Head of Research. Her current lines of research are related to tools for the automatic acquisition of language resources, area in which she has several international publications. In the past, she has actively participated in more than 10 EU and national funded research projects related to LRs such as: MULTEXT (LRE62050), LSGRAM (LRE61029), PAROLE (LE2-4017), as Technical Manager in the project SIMPLE (LE4-8346) and LIRICS, Linguistic Infrastructure for Interoperable Resources and Systems (e-content 22236).

Dr. Marta Villegas, Master in Computational Linguistics, Universitat de Barcelona, 1992 and PhD in *Informàtica* by Universitat Politècnica de Catalunya joined the IULA in 2007 as senior researcher. She has been involved in different European projects such as LE-PAROLE, LE-SIMPLE, IST-PEKING and EAGLES, ISLE related to the standardization and description of language resources. She has been actively involved in technical tasks of the project CLARIN.

4. Københavns Universitet (UCHP)
Centre for Language Technology, CST, is a department of the University of Copenhagen (UCPH). It employs a staff of around 20 scientists (linguists, computational linguists, engineers, computer scientists) working in many areas of language technology. In addition to basic research in lexicography, formal grammar, deep semantic analysis, MT, and multi-modality, UCPH has considerable experience in the development and evaluation/validation of a variety of HLT applications such as MT, and in collecting data resources such as large corpora, both in EU projects, in national research projects, and for commercial customers. UCPH has over the past 20 years been home to a portfolio of R&D projects among which the Centre co-ordinated the NEMLAR and MEDAR projects (support actions on Arabic language technology, resources, and tools), the EUROMAP and HOPE projects (support actions and dissemination on language technology in Europe). UCPH has been responsible for organizing and hosting many conferences, workshops and other events at national and international level. Currently UCPH is the co-ordinator of the Danish CLARIN RI-project whose goal is the construction of a Danish research infrastructure for the humanities, and UCPH is also the Danish partner in the European CLARIN project with the responsibility for leading the work package on governance, which prepares the future CLARIN ERIC.

Key personnel:
Bente Maegaard, Head of the Centre for Language Technology since 1991. Has coordinated a large number of EU and national projects (i.a. EUROMAP and HOPE projects, NEMLAR and MEDAR). Areas of expertise: HLT in general project management and leadership, dissemination.

Hanne Fersøe, Computational linguist, Deputy Head of the Centre for Language Technology. Areas of expertise: management and coordination, dissemination and user involvement.

Lene Offersgaard, Senior R&D Engineer, expert role in current projects: data integration in humanistic infrastructure (DK-CLARIN), user focused system design and SMT (EU LetsMT! project).

5. Koninklijke Nederlandse Akademie van Wetenschappen (KNAW-DANS)

DANS – Data Archiving and Networked Services – is an institute under the auspices of Royal Netherlands Academy of Arts and Sciences (KNAW), which is also supported by the Netherlands Organisation for Scientific Research (NWO). Since its establishment in 2005, DANS has been storing and making research data in the arts and humanities and social sciences permanently accessible. To this end DANS itself develops permanent archiving services, stimulates others to follow suit, works closely with data managers to ensure as much data as possible is made freely available for use in scientific research.

DANS manages the DARIAH initiative (see: http://www.dariah.eu), the European research infrastructure for the Arts and Humanities and is partner in the CESSDA project (RI for the Social Sciences) and CLARIN (RI for Language Resources). Digital preservation and data curation are important components of these initiatives.

Together with other stakeholders DANS implemented a system for the persistent identification of scientific data sets. For the persistent identifiers, use is made of Uniform Resource Name (URN) for which a namespace was obtained via the Dutch National Library. DANS took the initiative to establish the "Data Seal of Approval". The Data Seal of Approval and its quality guidelines may be of interest to research institutions, organizations that archive data and to users of that data (see://www.datasealofapproval.org).

Key personnel

Peter Doorn studied human geography in Utrecht and defended his PhD at that university. He taught computing for historians at Leiden university between 1985 and 1997. He was director of the Netherlands Historical Data Archive and head of department at the Netherlands Institute for Scientific Information Services (NIWI). He acquired and directed a considerable number of externally funded digitisation projects and other projects in the field of humanities computing.

Rutger Kramer studied Technical Informatics at the Technical University in Delft, where he did his Master's Thesis on the subject of photographic metadata. He has worked as a Scientific Programmer at the Netherlands Institute for Scientific Information Services and started as an Information Scientist at DANS. After gaining considerable experience with project management of software development projects, he became Head of the Software Development Group at KNAW-DANS.

Mike Priddy has a BSc in Microbiology and Microbial, MA in Visual Communication and a MSc in Computer Science. Before joining DANS as Project Manager in October 2010 Mike worked for 5 years designing, specifying and creating research infrastructures in an academic environment, latterly as part of the technical team developing the architecture for European ESFRI project DARIAH (Digital Research Infrastructure for the Arts and Humanities). Prior to working with research infrastructures he was instrumental in creating a number of notable Internet firsts in the UK in the 1990's before moving into technical & project management and R&D in Virtual Learning Environments (VLE) in higher education.

6. King's College London

King's College London is represented in DASISH by the Centre for e-Research. The Centre works at the intersection between research methods and practice, digital informatics, and e-infrastructure development and practice. The Centre has worked with a wide range of researchers on realizing their research ICT needs but has a particular tradition in building research infrastructure solutions for the arts and humanities, as it incorporates the staff and experience of the former Arts and Humanities Data Service Executive (http://www.ahds.ac.uk), and currently the Arts and Humanities e-Science Centre (http://www.ahessc.ac.uk), which is unique in the world. It is the main hub for research infrastructure projects in the arts and humanities in the UK and works closely with European and other international partners. Since 2010, the Centre together with the Centre for Computing in Humanities also runs an MA in Digital Asset Management.

Key personnel

Dr Tobias Blanke is co-leading the work at the AHeSSC and coordinating the Centre's teaching programme. He is secretary of the Humanities, Arts and Social Science Community Group (HASS-CG) of the Open Grid Forum (OGF) and leads the technical architecture work package for DARIAH. Before joining King's College in 2007, Tobias had worked at Credit Suisse First Boston in London, as a lead architect for a large financial data warehouse. From October 2010, he will lead the infrastructure work for the European Holocaust Research Infrastructure.

Dr Stuart Dunn co-manages the Arts and Humanities eScience Support Centre (AHeSSC). At AHeSSC, Stuart is responsible for supporting the projects funded under the AHRC-JISC-EPSRC e-Science Initiative. He works closely with a range of individuals and organisations involved in developing the Initiative's programme and disseminating its outcomes.

7. Goettingen State and University Library, University of Goettingen (UGOE)

The Goettingen State and University Library is one of the largest scientific university libraries and one of the leading e-Infrastructure innovators in Germany. It participates in numerous national and international activities in the context of metadata, semantic technologies, preservation, digitisation, education, and virtual research environments. As an example the UGOE is actively involved in building a national Grid-Infrastructure in D-Grid, coordinates the open access repository federation COAR and builds OpenAire, and establishes dedicated research environments for the humanities in its national, European and ESFRI activities. It further hosts one of the two largest national digitisation centres.

In DASISH UGOE will lead the training and education work package, and contribute to the technical reference architecture and the archiving work packages. UGOE has been working on projects related to data archiving and data infrastructure for almost 10 years, including most recently its participation in kopal, WissGrid and SHAMAN. Further it has established training courses in the field in the national network nestor and contributed to several other projects and conferences.

Key personnel

Dr. Norbert Lossau is Director of the Göttingen State and University Library, where he moved from his posts as Library Director at Bielefeld University and Head of the Oxford Digital Library, University of Oxford, UK. He is a member of several national and international steering committees and advisory boards.

Dr. Heike Neuroth is head of the Research and Development department of the Goettingen State and University Library. Her e-Infrastructure activities include a leading role in the DARIAH management board

as well as the board of the German Grid Initiative D-Grid. In addition to this, she participates in and leads various initiatives in the fields of digital libraries, long-term preservation, and metadata, and gives lectures in library and information sciences.

8. Österreichische Akademie der Wissenschaften (OEAW)

The Institute for Corpus Linguistics and Text Technology (ICLTT) is part of the Center for Linguistic, Text, and Audio-visual Research of the Austrian Academy of Sciences.

The purpose of the ICLTT is to develop new tools for text technological processing and use of linguistic corpora and new methods of corpus-based analysis of parallel corpora, special language texts and domain-specific terminologies. The ICLTT investigates the development of the German language in Austria by further developing, extending, and processing the Austrian Academy Corpus (AAC), one of the largest corpora for the German language worldwide.

The research priorities of the ICLTT include (1) interdisciplinary research on corpus-based terminology, translations and parallel texts; (2) international research infrastructures in the digital humanities (CLARIN und DARIAH); (3) new corpus-linguistic and lexicographical methods and technologies for corpus development and annotation, word-sense disambiguation, term extraction, etc.; and (4) comparative computational studies on literary corpora and historical editions.

OEAW contributes its long-standing experience in handling diverse data formats and ensuring data quality to DASISH, by participating in work packages WP3 on Data Quality as well as WP5 Shared Data Access + Enrichment. Furthermore, OEAW participates in WP7 Education/Training.

Key personnel

Prof. Dr. Gerhard Budin is director of the Institute for Corpus Linguistics and Text Technology (ICLTT) at the Austrian Academy of Sciences, as well as professor at and director of the Center of Translation Studies at the University of Vienna. He chairs Austria's participation in the ESFRI projects CLARIN and DARIAH with a team of Austrian institutions.

9. UK Data archive

The UK Data Archive (UKDA) is a centre of expertise in data acquisition, preservation, dissemination and promotion and is curator of the largest collection of digital data in the social sciences in the UK. Founded in 1967, it now houses several thousand datasets of interest to a wide range of researchers and provides resource discovery and support for secondary use of quantitative and qualitative data in research, learning and teaching. The UKDA is a member of CESSDA and has co-ordinated or participated in a number of EU funded projects including co-ordination of the CESSDA Preparatory Phase Project.

UKDA is a designated Place of Deposit by The National Archives allowing it to ingest and preserve public records. UKDA is based at the University of Essex in Colchester.

Key personnel:

Hilary Beedham is a member of the senior management group at the UK Data Archive. She leads the Resources and Management Services team which is responsible for UK Data Archive's routine managerial and administrative activities, for example: human resources; finance; contracts; and licences. She has

extensive experience of European developments in data services, gained through her management of a series of EC-funded projects, including the CESSDA Research Infrastructure Preparatory Phase Project.

Hervé L'Hours is the Preservation Metadata Implementation Manager for the UK Data Archive, researching and specifying a full metadata schema for the UK Data Archive and translating this into an ingest system. He has previously worked with the Online Historical Population Reports project (Histpop) and as a metadata consultant to the JISC Digitisation Programme. Hervé is on the board of the Data Seal of Approval.

John Shepherdson is a member of the senior management group at the UK Data Archive, and is head of Information Development and Programming. He provides strategic direction and senior management for UK Data Archive systems - supporting resource and data discovery, data access, data management and preservation of social scientific data. Prior to joining the UK Data Archive, John worked for the Research department of British Telecommunications plc for two decades. He is an acknowledged expert on multi-agent systems and wireless applications for virtual teams and has published extensively on this topic.

Matthew Woollard is Director of the UK Data Archive. His key responsibilities include, provide effective leadership and strategic direction of the UK Data Archive, developing and enhancing the organisation's national and international reputation, ensuring the effective management, operation and development of the Archive as an umbrella organisation hosting multiple national services relating to the acquisition, curation and dissemination of digital data; He is also responsible for the Archive's digital preservation policy and its implementation and maintenance framework.

10. Finnish Social Science Data Archive (FSD) - Tampereen yliopisto
The Finnish Social Science Data Archive (FSD) is a national resource centre for social science research and teaching. FSD archives, promotes and disseminates digital research data for research, teaching and learning purposes. The archive is funded by the Ministry of Education and is a separate unit of the University of Tampere.

FSD has been actively involved in the European data archive co-operation and has participated in projects which have aimed at providing data services and infrastructure for the European research community.

Key personnel
Taina Jääskeläinen, Master in both English and Library and Information Science, has worked at FSD since 2002. She has been involved in questionnaire translation of cross-national comparative surveys (ISSP, WVS, and CSES), has written on the challenges of multinational questionnaire translation and design, and has reviewed questionnaire drafts. She has been involved in several international projects promoting and developing tools for crossing language barriers in data documentation and search (multilingual social science thesaurus ELSST, controlled vocabularies for the international DDI metadata standard). She is able to read several languages.

Mari Kleemola is Information Services Manager and has worked at FSD since 1999. She has a Master's degree in Economics and a Bachelor's degree in Mathematics from the University of Tampere. She is the vice chair in the DDI Alliance Expert Committee. She is also the editor of the CESSDA website. Her expertise areas include the development and management of customer services and data documentation. She has recently been analyzing the OAIS model and costs models for digital data preservation.

Matti Heinonen is Senior Systems Analyst and the head of the technical services at FSD. He received Master's degree in Computer Science from the University of Jyväskylä 1999 and has worked at the FSD since then. His main responsibilities consist of development and maintenance of the archive's IT systems and software development of the web based applications. His tasks also include authentication and authorisation issues of data portals.

Arja Kuula, Master of Social Sciences (social policy, 1991) and Doctor of Social Sciences (sociology, 2000), has been working at FSD since 2000. Her present title is Development Manager. Her expertise areas include research ethics, data sharing, IPR and confidentiality issues and qualitative data, on all of which she has published and given presentations. She is or has been a member several national and international working groups dealing with research ethics, and qualitative data documentation, processing, archiving and reuse.

11. The Norwegian Social Science Data Services

NSD is a national multi-disciplinary research service facility and one of the largest archives for research data in the world. NSD provides a variety of services to researchers in Norway as well as internationally. NSD serves as a resource centre, assisting researchers with data collection, questionnaire design, social science data analysis, methodology, privacy issues and research ethics.

NSD has over the last 15 years been involved in several EU-financed projects to develop data infrastructures over the Internet, Nesstar, FASTER, LIMBER, Madiera and the CESSDA PPP project. Several important outcomes are of relevance, the software suite Nesstar, the Multi-lingual technology for search systems and portal software. For practical demonstration purposes, consult CESSDA portal

NSD has been the official data archive for the European Social Survey (ESS) since the first wave of the survey in 2002, and is one of seven scientific partners in the central coordination of the ESS.

NSD managed the data collection and documentation part of the FP5 RTD-project "Democratic Participation and Political Communication in Systems of Multi-Level Governance" (PartCom) and the follow-up FP6 specific targeted project "The determinants of active Civic Participation at European and national level" (CIVICACTIVE) NSD is presently participating as data manager in the FP6 project "EUROSPHERE: Diversity and the European Public Sphere. Towards a Citizens' Europe"

Key personnel
Bjørn Henrichsen   Director NSD

Atle Alvheim, Associate Director, Operation and Development

Vigdis Kvalheim,   Associate Director, Administration

Knut Kalgraff Skjåk,   Associate Director, Contract work

Kirstine Kolsrud,   Senior Adviser, ESS

Ørnulf Risnes, Senior Adviser, Nesstar Project Leader

Dag Kiberg, Senior Adviser,

Atle Jåstad, Senior Adviser, ESS EDUnet

Bjarne Øymyr, Senior Adviser Web development

12. City University London, UK (CITY)
City University houses the Centre for Comparative Social Surveys (CCSS) and has been the lead partner in the European Social Survey since its inception. The Centre specialises in cross-cultural and cross-national research, whether substantive or methodological.

Key personnel
Professor Sir Roger Jowell heads CCSS. He was previously the Founder Director of the National Centre for Social Research (NatCen) for nearly 30 years. He is also the Coordinator of the European Social Survey and the former Founder Director of the International Social Surveys Programme (ISSP).

Rory Fitzgerald is deputy director of CCSS. He is also a key figure in the ESS and a highly-respected survey designer and methodologist.

Dr Eric Harrison is Senior Research Fellow in CCSS who has most recently been responsible for the overall management of a series of networking and research activities under the infrastructure grant ESSi and before as a researcher at Essex University where he helped to develop the ESEC occupational classification.

13. GESIS – Leibniz Institute for the Social Sciences
The GESIS - Leibniz Institute for the Social Sciences, Germany, is an institution devoted to research and service, which supports and facilitates scientific work at every stage of the research process. Based upon extensive research programs GESIS provides services in data (archiving, provision, processing, research, social monitoring), methodology (consultation, developing complex methods), and specialized information (databases on social science literature research activities). In addition, GESIS delivers individual information and consulting services.

GESIS houses the oldest archive of social survey data in Europe and is among the founding members of CESSDA. It currently is one of the key partners in moving CESSDA to a new level of integration, and has led much of the work in the EU-funded "CESSDA preparatory phase project" (2008-2010). GESIS acts as host to integrate and disseminate some well-known international comparative surveys and is offering numerous added-value products derived from that.

GESIS is among the founding members of the European Social Survey consortium and has developed the conceptual and operational standards for sampling design, fieldwork monitoring and translation management. The ESS team supplies research-based knowledge, supports the national teams in implementing ESS standards, monitors the data collection process and evaluates compliance.

Key personnel
Markus Quandt (PhD) is Acting Director of GESIS Department "Data Archive for the Social Sciences".

Alexia Katsanidou (PhD) leads a team for International Data Infrastructures.

Brigitte Hausstein is head of the project "DOI registration agency for social science data da|ra".

Achim Wackerow is a DDI specialist and long-term member in the DDI Alliance.

Brita Dorer is specialised in questionnaire translation and heading the translation team of the ESS.

Dr. Matthias Ganninger is specialist for survey statistics and leads the ESS sampling team.

Verena Halbherr is involved in the ESS fieldwork monitoring.

Dr. Angelika Scheuer leads the ESS team at GESIS and is responsible for ESS Quality Enhancement

14. National University of Ireland Maynooth

The National University of Ireland Maynooth (NUIM) is represented in DASISH by the Irish Qualitative Data Archive (IQDA), which is hosted by the National Institute for Regional and Spatial Analysis. IQDA is the central access point for all qualitative social science data generated in Ireland. It was established as part of the Irish Social Science Platform with funding under the national Programme for Research in Third Level Institutions (PRTLI), Cycle 4. Under the auspices of NUIM, IQDA is part of a consortium that will develop a National Audio Visual Repository (NAVR) for Ireland with funding under PRTLI Cycle 5. The NAVR will provide a central access point for humanities and social science audio/visual data and research materials gathered and held in organisations across the island of Ireland. IQDA played a leading part in the establishment of the EQUALAN network, which brings together European researchers and archivists committed to preserving and organising qualitative data resources for sharing and re-use.

Key Personnel
Jane Gray is a Senior Lecturer and Head of the Department of Sociology at NUI Maynooth. She is a Research Associate of the National Institute for Regional and Spatial Analysis, where she is the programme leader for IQDA. In this capacity she is responsible for leading the development of the archive, and for establishing and promoting best practice in making qualitative data available for re-use. She is a core researcher in the recently funded project to develop a National Audio Visual Repository in Ireland. Jane Gray's research and teaching interests centre on questions relating to families, households and social change and she has published extensively in this area.

15. The University of Bergen
The University of Bergen (UiB) has a long-standing research record in computational linguistics and language technology. UiB combines competency in developing and exploiting language resources with theoretical, methodological and pedagogical knowledge and skills in the areas of general linguistics, Scandinavian studies, foreign languages, corpus and computational linguistics. UiB has participated in national and international co-operative projects aimed at language resources, tools and applications, including computational grammars and lexicons for Norwegian (NorGram), proofreading (SCARRIE), parsebanking (TREPIL), transfer-based quality translation (LOGON), research infrastructure for language (CLARIN), learner language corpora (ASK and ASKELADDEN) and others. Currently UiB is leading INESS, a Norwegian infrastructure project in syntax and semantics. UiB will recruit researchers for the execution of DASISH.

Key personnel:
Prof. Koenraad De Smedt has done research in computational linguistics and language processing applications. He has initiated the Telematics project SCARRIE (Scandinavian proofreading tools) and the TREPIL project on parsebanking funded by the RCN. He has coordinated two previous Marie Curie projects and currently coordinates CLARA, a Marie Curie ITN on Language Resources and their applications. He represents UiB as the Norwegian participant in CLARIN and will lead UiB participation in DASISH.

16. University of Venice "Ca' Foscari"

The **Department of Economics** at the **University of Venice "Ca' Foscari"** has a long standing tradition in applied microeconomic research. It is a large Department (about 40 academics) and the main research interests are equilibrium theory, environmental policy and public economics. The Department has been for several years a node in EU sponsored projects on saving, retirement and ageing.. It runs an international doctoral program (Doctorate in Economics and Organisation, VIU-Venice International University). The Venice working group has developed the SHARE questionnaire on employment and pensions. Assets and expectations have been developed by the Salerno team (CSEF) under the direction of Tullio Jappelli and Luigi Guiso.

Key personnel

Agar Brugiavini is Professor in Economics at the University "Ca' Foscari" of Venice. She obtained a Ph.D. in Economics at the London School of Economics, UK, and was a lecturer in Finance at the City University Business School (London, UK). She was a "Fulbright" Fellow at Northwestern University (USA) and she is currently responsible for the EU-sponsored RTN-Program "AGE" for the Venice node. She is also part of the NBER (National Bureau of Economic Research) international research group on "Social Security and Retirement Around the World" and a research associate of the Institute for Fiscal Studies (London, UK). Her major fields of research are in the economics of savings, pensions, pension reforms and labour supply. In her applied work she has acquired extensive knowledge of micro-data sets at household/individual level both for Italy and for the UK.

17. CentERdata

CentERdata, located on the campus of Tilburg University (The Netherlands), is a non-profit research institute specialized in (online) data collection and applied economic and methodological research. Staff members are experts in collecting and analysing (panel) data and putting these at the disposal of academic researchers. They manage and maintain several panels that regularly complete questionnaires through the Internet. A professional helpdesk and panel management staff support the fieldwork. Staff at CentERdata also has the expertise to analyse data using state-of-the art econometric models and techniques. The aim is to optimize panel research and model development and make optimum use of innovative methods.

CentERdata has years of experience in developing software to support panel research, both online and offline. Academically skilled IT-staff has experience with a great diversity of information systems. In the framework of the SHARE project, they have designed tailor-made software, such as the Blaise-based CAPI instrument, the sample management system (SMS) and the language management utility (LMU). This allows for language-independent questionnaire development in order to reach out to international target groups.

Key personnel

Marcel Das is Director of CentERdata and Professor of Econometrics and Data collection at Tilburg University. He holds a Ph.D. in Economics from Tilburg University (1998). As a director of CentERdata he has managed a large number of national and international research projects. Das has published a number of scientific publications in international peer reviewed journals in the field of statistical and empirical analysis of survey data and methodological issues in web-based (panel) surveys.

Alerk Amin works with CentERdata since 2007. He has a B.S. and an M.Eng. in Computer Science from Cornell University. He worked with various software companies in the US and has gained a lot of experience with a variety of programming languages and environments.

Maurice Martens works with CentERdata since 2004. He studied Applied Mathematics and Knowledge Engineering and has a Master's degree in Business Mathematics. Martens has extensive experience with various programming languages and environments. He focuses on web-based systems.

Arnaud Wijnant works with CentERdata since 2008. He studied Information Technology at Eindhoven University of Technology and did his final Master's project in the field of queueing theory in Switzerland. With his research he contributed to building the LHC particle accelerator at CERN in Geneva

18. University of Tartu

University of Tartu's Institute of Computer Science hosts more than 400 students across a range of Bachelors, Masters and doctoral curricula. The Institute's mission is to deliver world-class computer science education that is both relevant in practice and grounded on rigorous conceptual foundations.

On the research front, the institute is active in bioinformatics, natural language processing, cryptology, programming languages, distributed systems, and software engineering. Members of the institute publish in excess of 90 refereed papers per year. The Institute is a partner in the Estonian Center of Excellence in Computer Science <http://www.cs.ioc.ee/excs/> and the Software Technology and Applications Competence Centre — an R&D centre that conducts industry-driven research projects in the fields of software engineering and data mining. The Institute is a founder of the centre of Estonian Language Resources. The institute also participates in a number of EU-funded projects.

Key personnel

Tiit Roosmaa since 1993 University of Tartu, Institute of Computer Science, Assoc. Prof Since 1989 Vice Dean Faculty of Mathematics Since 2003 Head, Institute of Computer Science, University of Tartu

1987 PhD in Computer Science, Institute of Cybernetics, Tallinn

1978 University of Tartu, MSc in Mathematics, Faculty of Mathematics, University of Tartu

Kaili Müürisep 2007 - ... University of Tartu, Faculty of Mathematics and Computer Science, Institute of Computer Science, Chair of Language Technology; Senior Researcher Education

1996-2000 University of Tartu, Faculty of Mathematics, PhD (computer science)

1990-1996 University of Tartu, Faculty of Mathematics, BSc (computer science), MSc (computer science)

B2.3 Consortium as a whole

The DASISH consortium addresses the challenge of implementing harmonized solutions for the ESFRI infrastructures in the field of social sciences and the humanities. To be able to do so the consortium need both an intimate knowledge of relevant problems seeking solutions, a competence in the fields covered by the infrastructures to be able to prioritise between different problems and different solutions. Though such knowledge and competence might in theory be possessed by an outsider in the real world it is highly

unlikely. Similarly it is the case that though some institutions may have intimate understanding of a couple of the ESFRI infrastructures no one knows them all.

Building the consortium did then require relevant representation from each of the five infrastructures (CESSDA, CLARIN, DARIAH, ESS and SHARE) and also for practical reasons including as few partners as possible to provide the necessary diversity. Achieving this was naturally only possible by first discussing this matter on the infrastructure level.

Very early on in establishing the cooperation it became apparent that to establish the common solutions it was necessary to connect traditions the usually did not feel a great need to combine efforts. Among the five infra structures two are clearly humanities (CLARIN and DARIAH), two are clearly social science (ESS and SHARE) while CESSDA though predominantly out of a social science tradition to a higher extend than the two other social science infrastructures share the perspective of the two humanities infrastructures. Having realised this it also became clear that it would make sense to have CESSDA be the go-between among the other groups.

Having established this much of the structure the next necessary step was to identify the common problems to find solutions for. As resources would not be unlimited this step involved prioritising between issues that had different value to different communities and in this process to make the difficult choices between solutions of high value to a limited community and others with a lesser value to a larger community. The result of this process is the work plan presented in this application especially in item 1.1.2, 1.1.3 and 1.1.4.

Based on the work plan as represented in dimensions and work packages the next necessary step was collecting the teams that could actually fulfil the tasks outlined. Again there are issues of technical competence as well as understanding the community addressed. This explains why the work package teams are somewhat larger than what you would prefer when addressing a less complex community.

The result of this process has been a consortium which does not only cover the highest availably competence in a number of fields within social science, humanities, technical development and data archiving. The team also include a high degree of diversity in culture, nationality, gender etc.

So not only is the consortium clearly capable of fulfilling the tasks in front of it, it is also the only possible combination of abilities, competences and background that could do this.

Subcontracts

WP1

Subcontracts for audit will be entered into by UGOT €6000 and KNAW-DANS, UGOE, NSD, CITY, GESIS and CentERdata €3000 each.

WP3

A subcontract will be entered into by City. The purpose of the subcontract will be to play a key but specialist role in the development of the coding tools. Under the subcontract an occupational coding

module will be developed. This module can be embedded in CAPI interviews and can be adapted for different programs in use by the various fieldwork agencies. Versions will be produced for three European languages. These will be specified in the contract. The subcontractor needs to be able to show experience with working with this type of development tasks as well as practical issues with the programming. French, Dutch and Spanish are the most likely choices at this stage. Each of these three modules involves a fixed amount of programming hours. This subcontract will be for €70.000. The subcontractor will be selected on the basis of the best-price quality ratio under conditions of transparency and equal treatment.
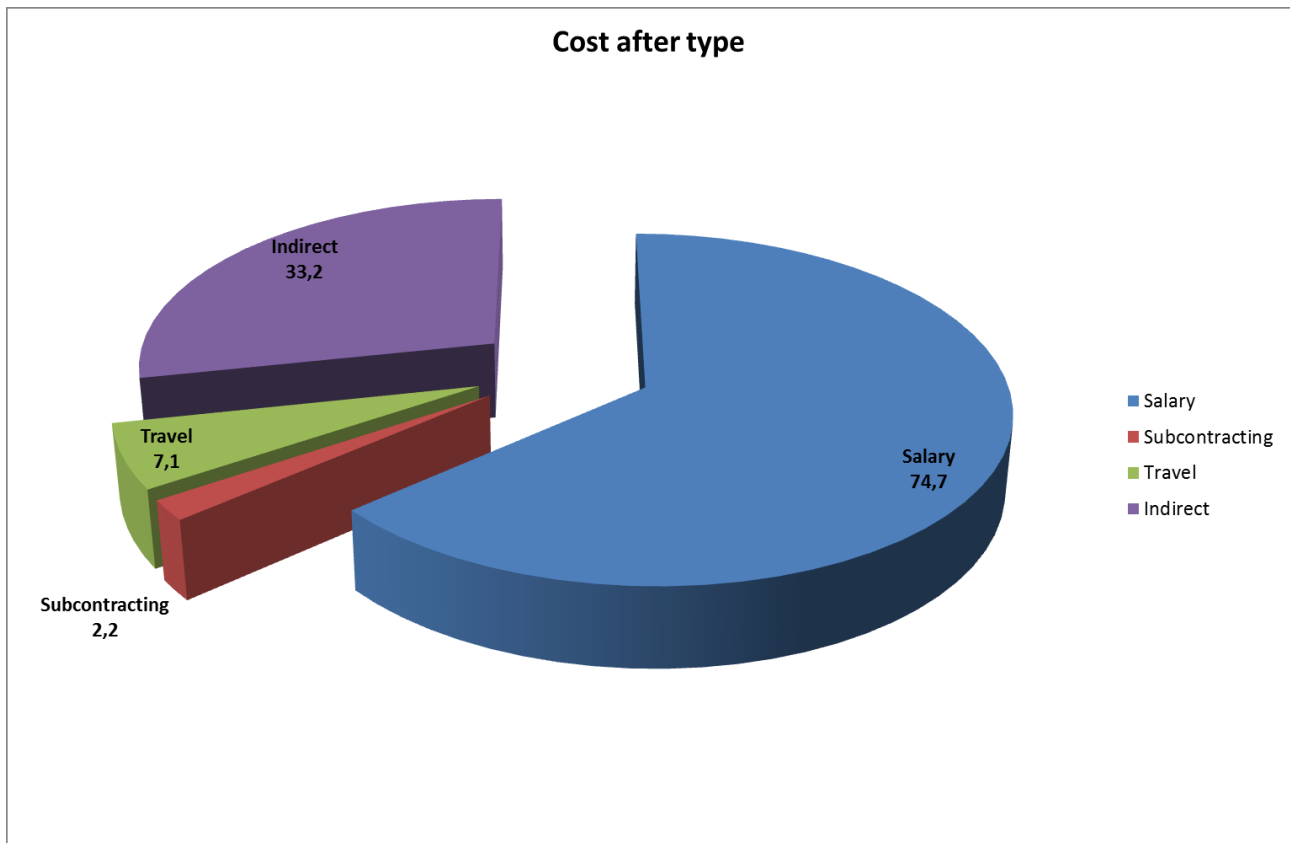
A subcontract in the size of €33.000 will be entered into by GESIS with translators who will be responsible for translating the occupational titles and tool text into the required languages.

B2.4 Resources to be committed

The total requested EU contribution for DASISH is 5.99 million Euros. This amount is very modest given the promised results. This is only possible because every deliverable includes a large amount of previous work carried out by the partners in other context such as e.g. the respective PPPs for the five ESFRI infrastructures.

Ressources pr WP
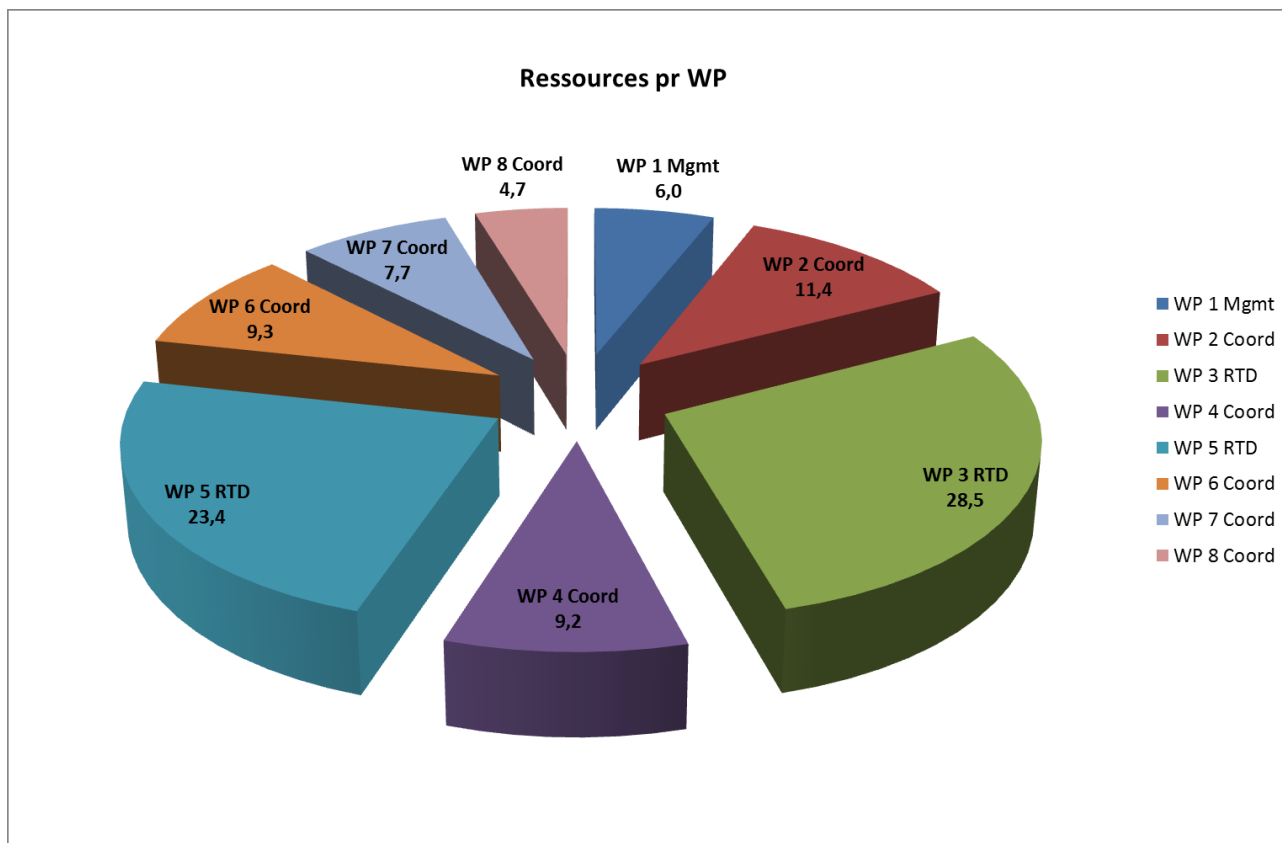
Just fewer than 4.1 million Euros of the total costs or 51 % is for personnel. DASISH is a staff intensive project. RTD and Coordination takes up almost all the effort there is relatively little travel and not much subcontracting involved.

Cost after type

Legend:
- Salary
- Subcontracting
- Travel
- Indirect

Chart labels:
- Indirect 33,2
- Travel 7,1
- Subcontracting 2,2
- Salary 74,7

Are also be complementing the community contribution by accepting the reduced indirect costs available for Coordination and Support actions under FP7 (which are considerably below the actual costs of hosting staff for DASISH activities and reflect a reduction compared to the indirect costs available for most i3 activities in FP6). For RTD activities institutions also contribute by generally receiving only 75% of the direct costs available

B3. Impact

The value of free exchange of research data cannot be overestimated.[7] The effect of free flowing data has already been identified by the participating ESFRI projects and is one of the paramount objectives when building infrastructures in the SSH. As the need for and positive effects of research infrastructures had already been made clear by ESS, SHARE, DARIAH, CESSDA and CLARIN individually, there is no need to sum up the advantages of increased availability of data. Important to stress, however, is the positive impact DASISH will have on the individual communities, cross-disciplinary research and European research.

B3.1 Strategic impact

B3.1.1 Expected Impact with regard to the Work Programme
DASISH combines all today's ESFRI projects in the SSH, and hence establishes a direct link to some of the most renowned researchers in SSH in Europe and beyond. The collaboration of these ESFRI projects in DASISH aims to enable the development and co-evolution of a consistent and pivotal eco-system of

---

[7] OECD Principles and Guidelines for Access to Research Data from Public Funding. OECD 2007.
http://www.oecd.org/dataoecd/9/61/38500813.pdf

research infrastructures - employing shared infrastructure systems where feasible and supporting each other in the creation of disciplinary virtual research environments. The collaboration across the SSH in DASISH facilitates the structuring of the scientific community and therefore plays a key role in the construction of an efficient research and innovation environment - as pointed out in the work programme[8]. A recent study underlined the potential of the SSH ESFRI projects in structuring the SSH communities, and the opportunities in pooling curation and access to research data in the SSH.[9]

Beyond this, DASISH contributes to the development and the implementation of the Digital Agenda for Europe[10], with actions including (1) fostering standards for the interoperability and sharing of research data; (2) coordination and pooling of resources in the SSH through interacting digital research infrastructures; as well as (3) supporting researchers in the adequate creation, management, curation and publication of research data. The particular benefits of data infrastructures in the efficiency and productivity of researchers - within and across working groups, institutions, and regions - has recently been pointed out by the High Level Expert Group on Scientific Data[11].

B3.1.2 Synergies within ESFRI and related initiatives

The DASISH goals will be pursued through the close cross-disciplinary collaboration of the ESFRI projects ESS, SHARE, DARIAH, CESSDA and CLARIN. DASISH has identified common areas in which the participating consortia can cooperate and make use of shared solutions. These areas include both, technologies such as registries, single-sign-on systems and persistent identifiers, and they also include coordinating efforts in the field of data quality, data archiving and on-line research. By developing new approaches and tools for the construction and execution of social science surveys, comparison and aggregation of survey results can be drastically enhanced greatly improving the usability, applicability and quality of social science data. By combining and exchanging insights in data archiving methods and technology, and offering generic procedures and applications to support this, both cost reduction and international alignment of requirements can be achieved.

At the same time, however, many research-related questions need to be dealt with within the communities and hence the individual ESFRI-projects themselves, as they are too specific to be taken out of the context of its discipline. Notwithstanding the variations in the form of virtual research environments to support discipline-specific research practices and resources, DASISH aims to tap into the synergies between infrastructure systems and frameworks across the SSH. DASISH will explore and exploit those technical, operational and advocacy synergies such as interoperability, persistent identification and publication of research data; convergence in best practice guidelines and e.g. certification of data management; as well as communication and training to support researchers in employing research infrastructures in and across the SSH. Integrating both, diverse disciplines in the SSH as well as local (e.g. national) efforts will add to the

---

[8] Capacities - Research Infrastructures, Work Programme 2011, Part 1. European Commission C(2010)4903 of 19 July 2010.

[9] European Commission, Evaluation of the impact of the Framework Programme on the formation of the ERA in Social Sciences and the Humanities (SSH), 2010, ftp://ftp.cordis.europa.eu/pub/fp7/ssh/docs/evaluation-fp-ssh_en.pdf

[10] Digital Agenda. http://ec.europa.eu/information_society/digital-agenda/index_en.htm
Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions A Digital Agenda for Europe /COM/2010/0245 final

[11] Riding the wave, How Europe can gain from the rising tide of scientific data, Final report of the High Level Expert Group on Scientific Data A submission to the European Commission, October 2010, http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf

visibility of European research, improve durability of research results and make it available to a larger group of potential consumers.

Impact Determining Factors and Assumptions

Close cooperation between the SSH ESFRI projects and within DASISH is paramount in creating the right conditions for maximizing impact for the SSH in Europe and beyond. DASISH needs to present itself as a coherent group that effectively caters to all SSH communities involved without losing sight of their specific context and requirements. Focus must be put on those areas where the most can be gained when combining forces or coordinating effort.

Another factor is the availability of actual data. Without a sufficient and representative amount of readily available datasets it cannot be ensured that common solutions will be built that can cater to the respective infrastructures. The influx of high quality research data has already been ensured by the individual research infrastructures, and DASISH will ensure that bridges will be built between infrastructures in order for this data to circulate. The more data will be available, both inside and across the domains, the higher the appeal will be for others to use the data and contribute their results.

Quality and trustworthiness of both the infrastructures as well as the data they serve will have an effect on the impact of this effort as well. Both the individual infrastructures as well as DASISH need to establish themselves as a haven for high quality research and durable, high availability services. Failure to address this issue will result in the infrastructures only playing a marginal role within the research life-cycles and can seriously stem their growth. The efforts in WP3 (Data Quality) and WP4 (Data Archiving) as well as the technical coordination in WP2 (Architecture & Quality Assessment) and development in WP5 (Shared Data Access and Enrichment) will address trust and quality throughout their activities.

Legal issues can form an obstacle for international cooperation and exchange of research. These issues have already been identified by the participating infrastructures, and it is foreseen that a lot can be gained by forming an inter-infrastructural consortium. Being able to act as a representative of the majority of European SSH research will give DASISH the mandate to negotiate with national governing bodies and other relevant players to break down existing barriers to data exchange and cooperation (cf. WP6 - Legal and Ethical Issues).

Perspectives on Impact: Roles and their Goals

DASISH combines the efforts of various researchers and infrastructure initiatives in the SSH. Beyond those contributing to DASISH and the individual ESFRI initiatives, there are various other stakeholders that DASISH connects to.

Researchers in all SSH domains will - through DASISH - have data and tools available that increase their efficiency and productivity. The mechanisms for interaction within their community and across communities may provide them with opportunities for collaboration and new insights. Eventually, their research will gain more visibility, and - vice versa - they will have the possibility to build on and re-use existing research activities and hence "stand on the shoulders of giants".

Funders and policy makers will - through DASISH - increase the impact of their programmes through better exchange between projects and by eliminating duplication of work. This, as well as the available data in DASISH will support evidence-based decisions and increase the return on investment of their work.

Being aware that community-building and take-up of technologies is an evolutionary process, DASISH will trigger and lubricate the creation of and interaction between data infrastructures in the SSH. In this effort, DASISH can build on existing programmes including the work of the European Grid Initiative, GEANT, PRACE/DEISA and in future the hopefully emerging data eInfrastructure. At the same time it will interact with e.g. Datanet[12] in the USA[13] and other related initiatives.

B3.2 Dissemination and/or exploitation of project results, and management of intellectual property

We see dissemination as an internal as well as an external outreach effort since the consortium in the DASISH project originates from two different research communities, the social sciences and the humanities.

The internal dissemination will serve the purpose of making the two different communities within the consortium aware of the similarities that we share and thus of the similar requirements we have to, and the common solutions we choose for our research infrastructures. The increased awareness will also help us better understand the differences so that it will be very clear in which contexts we will not be able to benefit from common solutions but should rather choose different and domain specific approaches.

More specifically the results of the DASISH work packages can and will be directly exploited by the future ERICs, and by the research communities, for example

- The work on architectures and quality issues (WP2) will be continued within the disciplines and will lead to a greater harmonization.
- The knowledge base and the tool commentary (WP2) may become a useful tool that will be maintained by the ERICs.
- All work on improving the quality of surveys (WP3) will directly feed into improving the survey work and will be continued by the SS disciplines.
- Concrete services for deposits (WP4) will be continued by the designated centers in the realm of their infrastructure projects.
- The work on preservation issues (WP4) will feed in to the emerging common data eInfrastructure and thus be maintained.
- The trust federation extension (WP5) will be maintained by the ERICs as long as GEANT/eduGain do not have a practical model for maintaining it.
- The enhanced PID services (WP5) will be maintained by the existing consortia such as EPIC.
- The joint metadata domain (WP5) will be maintained by CLARIN, who will particularly need to maintain this technology.
- The work on workflows (WP5) will improve the current approaches which are being set up within the communities.
- The annotation framework technology (WP5) will be maintained by CLARIN, since it is a basic technology that needs to be offered.

---

[12] NSF: Datanet. http://blog.jonudell.net/2007/10/02/the-nsfs-datanet-initiative/
http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=503141
[13] There have already been very productive interactions with strong US initiatives at the ASIS&T Summit on Research Data in Phoenix in 2010 and concrete collaborations have been agreed.

- Legal and ethical work devoted to the survey problematic (WP6) will be maintained by the social science communities, and in particular the Virtual Competence Center, which is of relevance to all SSH communities, will be maintained by CESSDA.
- The work on training modules (WP7) is of relevance to all communities, thus we expect that all will continue to create new modules and to share them.

The external dissemination will be for various and different target groups, but key in the effort will be the focus on future users of the research infrastructures who will be those to truly exploit the project results. These are the SSH researchers and scientists, including their students, who through the digital system of eco-infrastructures will be getting access to data, tools, computing power, implemented domain specific workflows etc. which will enable them to ask whole new types of research questions within their field, and find answers to them through the new possibilities offered to them by the results of the DASISH project. We foresee a new generation of researchers which will grow up with the research infrastructures. It is anticipated that they will be highly motivated to invest much time in the new eScience paradigm and will hopefully much better manage to tackle the huge challenges of our multilingual and multicultural societies.

Management of intellectual property rights will be handled through the consortium agreement where it will be written in proper legal terms. The basic principles, however, will be that the IPR stays with the part who developed it, and that software component will be made freely available as open source.

B4. Ethics Issues

The DASISH project is about research infrastructures and about designing and deploying joint services. Thus it is not about creating and manipulating sensitive data. I may include access to sensitive data by a broader group of researchers. All individual initiatives have been tackling their proper methods of dealing with sensitive data, however, when setting up trust domains etc. new approaches need to be found. Therefore DASISH has included legal and ethical aspects in WP6 and thus will devote much time on the corresponding aspects.

ETHICS ISSUES TABLE

(Note: Research involving activities marked with an asterisk * in the left column in the table below will be referred automatically to Ethics Review)

Research on Human Embryo/ Foetus		YES	Page
*	Does the proposed research involve human Embryos?		
*	Does the proposed research involve human Foetal Tissues/ Cells?		
*	Does the proposed research involve human Embryonic Stem Cells (hESCs)?		
*	Does the proposed research on human Embryonic Stem Cells involve cells in culture?		
*	Does the proposed research on Human Embryonic Stem Cells involve the derivation of cells from Embryos?		
	I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	yes	

Research on Humans		YES	Page
*	Does the proposed research involve children?		
*	Does the proposed research involve patients?		
*	Does the proposed research involve persons not able to give consent?		
*	Does the proposed research involve adult healthy volunteers?		
	Does the proposed research involve Human genetic material?		
	Does the proposed research involve Human biological samples?		
	Does the proposed research involve Human data collection?		
	I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	yes	

Privacy		YES	Page
	Does the proposed research involve processing of genetic information or personal data (e.g. health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)?		
	Does the proposed research involve tracking the location or observation of people?		
	I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	Yes	

Research on Animals [14]	YES	Page
Does the proposed research involve research on animals?		
Are those animals transgenic small laboratory animals?		
Are those animals transgenic farm animals?		
* Are those animals non-human primates?		
Are those animals cloned farm animals? I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	yes	

Research Involving ICP Countries [15]	YES	Page
Is the proposed research (or parts of it) going to take place in one or more of the ICP Countries? Is any material used in the research (e.g. personal data, animal and/or human tissue samples, genetic material, live animals, etc):		
a) Collected in any of the ICP countries?		
b) Exported to any other country (including ICPC and EU Member States)?		
I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	yes	

Dual Use	YES	Page
Research having direct military use		
Research having the potential for terrorist abuse		

---

14   The type of animals involved in the research that fall under the scope of the Commission's Ethical Scrutiny procedures are defined in the Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes Official Journal L 358 , 18/12/1986 p. 0001 - 0028

15   In accordance with Article 12(1) of the Rules for Participation in FP7, 'International Cooperation Partner Country (ICPC) means a third country which the Commission classifies as a low-income (L), lower-middle-income (LM) or upper-middle-income (UM) country. The list of countries is given in annex 1 of the work programme. Countries associated to the Seventh EC Framework Programme do not qualify as ICP Countries and therefore do not appear in this list.

I CONFIRM THAT NONE OF THE ABOVE ISSUES APPLY TO MY PROPOSAL	yes	

B5. Consideration of gender aspects

The participation of women in the production of knowledge in science is an important indicator for the extent to which the full potential of human resources is used in society. Despite a multitude of measures, laws and resolutions, women are still under-represented in science. There are an increasing number of women with university degrees, but the higher the career level, the fewer women are present.

All the five participating ESFRI infrastructures CESSDA, CLARIN, DARIAH, ESS and SHARE actively support the principle of gender equality in all aspects of their activities.

The partners are committed to gender equality in all areas of their work and recognize the need to maximize the opportunities for women to achieve higher representation in research and the different types of management (technical, financial and administrative). This will be promoted in the context of the 2002 Helsinki Group report on National Policies on Women and Science in Europe and the 2008 study: Benchmarking Policy Measures for Gender Equality in Science.

The project will promote opportunities for women by example. Women will play an important role in the management of the project. The work of four key partnering organizations (UPF, UCPF, NUIM and UNIVE) will be led by women and experienced women will be invited to join the project Advisory Board. Women will participate in all other planned activities. Among the centrally placed female participants in the DASISH work are:

- Heike Neuroth (UGOE)
- Nuria Bel, Marta Villegas (UPF)
- Taina Jääskeläinen, Mari Kleemola, Arja Kuula (FSD)
- Brigitte Mathiak, Alexia Katsanidou, Brigitte Hausstein, Brita Dorer, Angelika Scheuer (GESIS)
- Kirstine Kolsrud (NSD)
- Bente Maegaard, Lene Offersgaard, Lina Henriksen, Dorte Haltrup, Costanza Navaretta (UCPH)
- Agar Brugiavini (UNIVE)

Bente Maegaard will lead WP8 and a number of tasks will be led by the mentioned persons.