



iMENTORS

SEVENTH FRAMEWORK PROGRAMME

Call (part) identifier: FP7-INFRASTRUCTURES-2012-1

Project full title: E-INFRASTRUCTURE MONITORING, EVALUATION AND TRACKING SUPPORT SYSTEM

Project Acronym: iMENTORS

Grant agreement for: Coordination and support action

Project Number: 313203

D2.3 Second Report on Database of e-infrastructure development projects in Sub-Saharan Africa

Deliverable Id :	D2.3
Deliverable Name :	D2.3 Second Report on Database of e-infrastructure development projects in Sub-Saharan Africa
Status :	Final
Dissemination Level :	PU
Due date of deliverable :	M15
Actual submission date :	M16
Work Package :	WP 2
Organisation name of lead contractor for this deliverable :	Stockholm University
Author(s):	Nils Jensen
Partner(s) contributing :	Louis Papaemmanuel, Kristin Curpas, Nikolaos Tsiavos

Abstract: D2.3) Second Report on Virtual Observatory for e-infrastructure development projects in Sub-Saharan Africa reports on the progress made in maintaining and updating the database for e-infrastructure projects in Sub-Saharan Africa, on the efforts made in successfully connecting with other databases, outline the stakeholders present on the platform in relation to the critical success criteria, and all the editorial management performed to maintain the platform updated and running.

iMENTORS is a project co-funded by the European Commission's DG CONNECT under the 7th Framework Programme. www.iMENTORS.eu

©Copyright by the iMENTORS Consortium

History

Version	Date	Modification reason	Modified by
0.1	09.07.2013	Initial draft	Louis Papaemmanuel
0.2	12.07.2013	Version 2	Kristin Curpas
	12.07.2013	Comments	Louis Papaemmanuel
	05.08.2013	Comments	Nils Jensen
0.3	06.08.2013	Version 3	Kristin Curpas
0.4	13.08.2013	Comments	Louis Papaemmanuel
0.5	15.08.2013	Version 5	Kristin Curpas
	17.08.2013	Quality check	Nils Jensen
1.0	17.08.2013	Deliverable ready to be submitted to the European Commission	Louis Papaemmanuel

Pending EC Approval

Table of contents

HISTORY	2
TABLE OF CONTENTS	3
LIST OF TABLES	4
LIST OF ABBREVIATIONS	5
EXECUTIVE SUMMARY	6
1 INTRODUCTION: WHAT IS THIS DELIVERABLE?	7
1.1 THE PROJECT: IMENTORS	7
1.2 THE WORK PACKAGE (SUBSET OF THE PROJECT).....	8
1.3 THE DELIVERABLE (SCOPE, OBJECTIVES, METHODOLOGY)	8
1.4 RELATION TO OTHER WP2 DELIVERABLES AND INTENDED AUDIENCE	9
1.5 INTENDED AUDIENCE OF THIS DOCUMENT:.....	9
1.6 RELATION OF WP2 TO OTHER WPS.....	10
2 METHODOLOGY AND EDITORIAL MANAGEMENT:	11
3 IMPROVING THE ONLINE OBSERVATORY: AN UPDATE SINCE D2.1	12
4 DATA COLLECTED AND PROGRESS ACHIEVED IN RELATION TO THE CRITICAL SUCCESS FACTORS	15
4.1 ORGANISATIONS	15
4.1.1 UNITED NATIONS ENTITIES	16
4.1.2 INTER-GOVERNMENTAL BODIES/INSTITUTIONS	17
4.1.3 GOVERNMENT AGENCIES AND PUBLIC INSTITUTIONS.....	17
4.1.4 HIGHER EDUCATION AND RESEARCH INSTITUTES	17
4.1.5 CIVIL SOCIETY ORGANISATIONS (NON-PROFIT)	19
4.1.6 PRIVATE SECTOR COMPANIES.....	19
4.2 RESEARCH INFRASTRUCTURES AND E-INFRASTRUCTURES	20
4.2.1 NETWORKING INFRASTRUCTURES	20
4.3 PROJECTS (AS FRAMEWORKS FOR PROVIDING FUNDING)	22
5 INTEGRATING IMENTORS WITH THE WORLD BANK PROJECT DATABASE	24
6 CONCLUSION: WHERE WE ARE AND WHAT'S NEXT	26

List of tables

Table 1: Intended Audience of D2.3..... 9

Table 2 Milestones of WP2 10

Table 3 African NRENs..... 18

Pending EC Approval

List of abbreviations

CANARIE	Canada's Advanced Research and Innovation Network
D	Deliverable (of the project)
DG	Directorate-General
DG CONNECT	European Commission Directorate General for Communications Networks, Content and Technology
DoW	Description of Work (of the project)
DSV	Department of Computer and System Sciences of Stockholm University
EC	European Commission
EFTA	European Free Trade Association
e-IRG	e-infrastructures Reflection Group
ESFRI	European Strategy Forum on Research Infrastructures
EU	European Union
FP7	Seventh Framework Programme
Gov2u	Government To You
GEANT	Pan-European Data Network
HEIs	Higher Education Institutes
HIV	Human Immunodeficiency Virus
ICT	Information and Communication Technology
ICT4D	ICT for Development
i.e.	That is
ISP	Internet Service Provider
IT	Information Technology
M	Month (of the project)
MDG	Millennium Development Goal
NEPAD	New Partnership for Africa's Development
NGO	Non-Governmental organization
NOC	Network Operations Centre
NRENs	National academic and research networks
PDF	Portable Document Format
PEST analysis	Political, Economic, Social, and Technological analysis
PRD	Programme Director
PU	Public view
RedCLARA	Latin American Advanced Networks Cooperation
RENs	Research and Education Networks
RRENs	Regional Research and Education Networks
SAB	Stakeholders Advisory Board
Sida	Swedish International Development Cooperation
SMEs	Small and Medium Enterprises
S&T	Science and Technology
Spider	Swedish Program for ICT in Developing Regions
SU	Stockholm University
SWOT analysis	Strengths, weaknesses, opportunities and threats analysis
T	Task (of the project)
TEIN	Trans-Euroasia Information Network
UN	United Nations
UNECA	The United Economic Commission for Africa
USAID	The United States Agency for International Development
URL	Uniform Resource Locator
VO	Virtual Observatory
WP	Work Package
PRACE	Partnership for Advanced Computing in Europe

Executive Summary

The overall objective of iMENTORS is to enhance the coherence and effectiveness of international actors involved in e-infrastructures development projects and initiatives in Sub-Saharan Africa.

Following up on Deliverable 2.2, the purpose of D2.3, entitled 'Second Report on the Virtual Observatory for e-infrastructure development projects in Sub-Saharan Africa' is to report on the progress of the consortium in relation to the maintenance and updating of the database for e-infrastructure development projects in Sub-Saharan Africa. The document also outlines the efforts of the consortium in successfully connecting with external databases, and outlines the stakeholders already present on the platform. It further elaborates on the progress made in achieving the critical success criteria, and briefly discusses the editorial management that has been performed to maintain the platform updated and running.

D2.3 is split in five parts: 1) Introduction, 2) Methodology of the data collection process, 3) Improving the online observatory: an update since D2.1, 4) Data collected and progress achieved in relation to the Critical Success Factors 5) Integrating iMENTORS with the World Bank Database, 6) Conclusions: where we are and what is next.

This is a deliverable of WP2: Ecosystem Identification & Establishment. Among the main objectives of WP2 are to identify and attract strategically selected stakeholder organizations, their key groupings and subgroups, their interests and their level of engagement in our project; to gather and analyse relevance of all past and on-going e-infrastructure development and ICT projects in Sub-Saharan Africa; to connect with stakeholders and maintain a strong ecosystem which will add value to the project by using and updating the information and engage with others on the platform; to develop a community of practice: by creating an online community for support to policy development and programme implementation. The critical factors set to evaluate the success of the projects with regard to the above mentioned objectives are that 80% of e-infrastructures and ICT related projects over the past 5 years identified by month 20; that 20% of the identified community is online by month 12 of the project, and 60% by month 24.

The initial draft was created by Stockholm University, the leader of WP2, which was sent to the relevant participants and the project coordinator for comments and review. The final version was finalised by incorporating the comments and suggestions in the deliverable. The information included in this deliverable is interrelated with all WP2 deliverables. The intended audience of this deliverable are the iMENTORS partners, the stakeholders and the European Commission (EC).

1 Introduction: What is this deliverable?

After giving a small description of iMENTORS and of the work package, this section defines the scope and methodology of the deliverable. Finally, the relation of the deliverable to other deliverables is explained, with a brief paragraph to describe its intended audience.

1.1 The project: iMENTORS

The overall objective of the project is to enhance the coherence and effectiveness of international actors involved in e- infrastructures development projects and initiatives in Sub-Saharan Africa.

The project will:

- Provide policy support by identifying and monitoring all on-going e-infrastructure projects in Sub-Saharan Africa and enable users to perform impact assessments
- Enhance aid coordination and collaboration by providing insight on e-infrastructure development projects and through the platform's collaborative features (knowledge sharing) for development of new e-infrastructure development projects
- Promote of e-infrastructures of common interest to Europe and Sub-Saharan Africa through extensive dissemination activities (workshops, conferences, communication)

Specific objectives:

Build: Create a virtual observatory, acting as one-stop-shop data warehouse providing up-to-date information on all e-infrastructure related development programmes and initiatives of the past five years in Sub-Saharan Africa to enhance the effectiveness and coherence of national and EU research policies and international cooperation in the field of research infrastructures:

- Identify and collect informational assets (data and sources): Identify and connect with key stakeholders, Gather and analyze relevance of all past and ongoing e-infrastructure development and ICT projects in Sub-Saharan Africa, Populate the virtual observatory
- Create the virtual observatory (platform), User interface Decision support system.

Develop a community of practice for support to policy development and programme implementation by creating a social hub facilitating interaction and knowledge sharing, to improve collaboration among different stakeholder groups, and offer them opportunity to create synergies and plan future projects.

- Launch a space for Collaboration
- Position iMENTORS as the knowledge broker
- Decision support to policy development enabling users to produce queries across several online databases, and to evaluate e-infrastructure proposals from multiple perspectives in a structured manner.

Sustain: Build and maintain a strong stakeholder ecosystem around the iMENTORS project, which will ensure long-term viability of the project and will enhance the development of e-infrastructure in Sub-Saharan Africa through:

- Community awareness
- Standardization of practice
- Political and Institutional support.

iMENTORS is a project co-funded by the European Commission's DG CONNECT under the Seventh Framework Programme (FP7). For more information visit: www.iMENTORS.eu

1.2 The Work package (subset of the project)

The project will be implemented through four interrelated Work Packages (subset of the project):

- WP1: Project Management will ensure the correct and timely execution of the project
- WP2: Ecosystem identification and establishment activities that generate the ecosystem including stakeholders, data gathering, editorial, entry of data and validation
- WP3: Platform customization and integration includes integration of the geospatial semantically enabled platform with a decision support system and ontology enhancement
- WP4: Dissemination and Sustainability includes activities such as workshops, conferences, training, newsletters and overall stakeholder engagement for sustainability.

This is a deliverable of WP2: Ecosystem Identification & Establishment. The overall objective of WP2 is to collect information stakeholders and e-infrastructure projects as well as to categorise them in a standardised way to visualise them meaningfully in the online virtual inventory.

1.3 The Deliverable (scope, objectives, methodology)

The purpose of this document is to report on the data collected on all e-infrastructure stakeholder organisations, e-infrastructure development projects, as well as research and networking infrastructures. The document will briefly outline the methodology and the sources used to gather the data collected, and report on the progress made in successfully connecting the external databases with iMENTORS to facilitate a semi-automatic data harvesting process.

D2.3 is split in five parts: 1) Introduction, 2) Methodology of the data collection process, 3) Improving the online observatory: an update since D2.1, 4) Data collected and progress achieved in relation to the Critical Success Factors 5) Integrating iMENTORS with the World Bank Database, 6) Conclusions: where we are and what is next.

The initial draft was created by Stockholm University, the leader of WP2, which was sent to the project coordinator for comments. The final version was finalised by incorporating the comments and suggestions in the deliverable. The information included in this deliverable is interrelated with all WP2 deliverables. The intended audience of this deliverable are the iMENTORS partners, the stakeholders and the European Commission (EC).

A Quality Management table was also created to describe the process used to ensure the quality of the deliverable (see Appendix A).

1.4 Relation to other WP2 deliverables and intended audience

Information included in this deliverable is interrelated with all WP2 deliverables:

- D2.1 Report on criteria data indicators and specifications for the updating of iMENTORS
- D2.2 First Report on Virtual Observatory for e-infrastructure development projects in Sub-Saharan Africa (M6):
- D2.4: Final Report on Virtual Observatory for e-infrastructure development projects in Sub-Saharan Africa (M30)

1.5 Intended audience of this document:

Table 1: Intended Audience of D2.3

Group of readers	Reasons for reading
iMENTORS partners	All the activities of the project are split homogeneously within the partners of the consortium, which also applies to the consortium's efforts regarding the collection and classification of data. Similarly, the efforts undertaken in WP3 (integration and customisation) are also split equally. This document will inform all the activities that are carried out in WP2 and controls the tasks carried out to integrate iMENTORS with external data sources, a task undertaken in WP3. Consequently, the deliverable targets the majority of individuals working in both these work packages.
iMENTORS stakeholders	To create the deliverable, particularly regarding the validation of the criteria used to carry out the categorisation of data, the consortium undertook several rounds of consultation with the members of the Stakeholder's Advisory Board.
The European Commission (EC)	To inform the Project Officer, the reviewers and other interested stakeholders from the EC, about project's dissemination plan, progress so far and anticipated activities regarding dissemination for the following period of the project.

SPIDER (Stockholm University) is the leader of WP2.

The ultimate responsibility of each work package and of each activity constituting the various work packages is assigned by the various partners to the Work Package Leaders (responsible for coordinating contributions to their individual WP) and the Task Leaders (responsible for coordinating the tasks within their individual task within each WP), and will directly report to the PRD, for collating progress updates from the WP Leaders on project activities with respect to expected technical achievements, results, schedule.

The responsibilities for WP leaders consist of:

- Coordinating tasks and activities according to WP objectives
- Setting up a project plan specific to the WP, covering work package management, quality management and risk management
- Ensuring the smooth running of the individual WP as well as co-ordination with other WPs
- Monitoring progress with respect to goals, milestones, and adequacy of results
- Reporting to the Project Director of any possible deviations identified due to scheduling, unsuitability or risks affecting the quality of project results and/or objectives
- Coordinating task leaders.

More specifically, roles and responsibilities in WP2 as defined in the DOW are the following:

- Identify the main actors and the sources of data
- Collecting and Categorizing the Data
- Validating and Populating the Virtual Observatory (platform)
- Carrying out the preparatory work to identify and integrate external project databases
- Editorial management

The following table lists the milestones of WP2:

Table 2 Milestones of WP2

Milestone	Delivery Date	
MS1 Initial Platform Launched	M06	The database created (D2.2), and the User guide is published
MS2 Second Version launched (with decision-support system)	M15	D2.3, D3.2, D 3.3, D4.2, D4.4 and D1.2 have been delivered. 20% of all stakeholders are found on the platform.
MS3 Third Version Launched (with process for the elicitation of stakeholder preferences)	M18	D3.4 delivered on time.
MS4 Decision Support system is fully operational, data collection has been finalized	M24	D1.3, D3.5 delivered on time - 60% of stakeholders are on the platform
MS5 The project is complete	M30	Final review has been accepted

1.6 Relation of WP2 to other WPs

WP2 is highly related to the work carried out in WP3, especially as regards the development of the decision support system as its smooth functioning depends on the quality and amount of data stored on projects and organisations in the e-Infrastructure repository. Additionally, WP2 shares responsibility for the integration of external databases, as all preparation and research is carried in within the Work package.

WP2 is contributing to the dissemination and public relation objective of WP4, when communicating with organizations for the purpose of collecting data for the e-Infrastructure repository.

2 Methodology and Editorial Management:

The iMENTORS Work Package 2 (WP2) team follows a data collection methodology strategically calibrated to meet the following critical success factors: 20 % of all stakeholders identified on the platform by month 12 of the project and 60 % by month 24, as well as 80 % of the e-infrastructure and ICT related projects over the past 5 years identified and recorded by month 20.

Meeting these goals, required giving priority to the task of identifying and recording the key stakeholders involved with e-infrastructure development in Sub-Saharan Africa. The assumption that there are more stakeholder organizations than there are infrastructure projects was the guiding principle behind the methodology adopted by the research team, since given the structure of the database, it is more efficient to record the organizations first and then link them back to the projects. Establishing a profile for each of the key organizations on the platform renders the later task of mapping e-infrastructure and other ICT projects to the relevant organizations far more manageable.

To create the organizational profiles, the team first engages in an extensive online survey focused primarily on identifying and recording the following information about each organization: name, abbreviation, year established, website URL, organization type, country, contact information, headquarters or main campus, description, and physical address. This data is collected in a spread sheet, which is then handed over to a separate team member to upload to the database. Once a critical mass of data is uploaded, the next phase, data validation, can be completed most efficiently via a crowd sourcing process, where the iMENTORS team handles the role of editorial management and support.

In addition to identifying and recording the organizational stakeholders, another primary focus of the WP2 team has been to record major e-infrastructure projects, including submarine and terrestrial cables and internet exchange points. The methodology used to record the infrastructure is substantially similar to the methodology used to collect the data for the organizations: the results of an extensive online search are first recorded in an offline spread sheet which is then uploaded by another team member to the database for further validation online if necessary.

This general methodology has proven robust, and can be extended as the WP2 team moves forward to record and upload more information on research infrastructure and other ICT development projects.

3 Improving the online observatory: an update since D2.1

The often fuzzy boundaries existing between e-infrastructures and research infrastructure facilities prompted the consortium (in accordance with some selected members of the Stakeholders' Advisory Board) to create an additional category to classify research infrastructures and e-infrastructures. This should enable the consortium to categorise distributed GRIDs or single sited e-infrastructures as infrastructures, and the projects that have led to their establishment will be categorized as projects. The main rationale is that projects that potential lead or contribute to the establishment of e-infrastructures (i.e. Africa-Connect) may have very distinguishable objectives or additional components (and therefore evaluation criteria) than the e-infrastructures themselves.

This work resulted in devising a complex, yet robust, classification scheme that facilitates the re-use of data and that maintains the possibility of creating a plethora of visualisations on the map. The classification schema is far more comprehensive and advanced than all other schemata examined and other mapping projects. The difficulty of this task was linked to the fact that there was no readily available classification or distinction between the different components of e-infrastructure to draw from – which caused important delays in the development of the database, and consequently also impacted data collection. By distinguishing between the different entities and working our way up from the smallest organizational unit, we can create many connections between the different entities of the database. In addition, we have built our database in a way that gives us the freedom and flexibility to add new functionalities or make minor alterations to it, without disrupting the entire project. Such liberty gives the opportunity to deliver a platform that can truly respond to stakeholder's needs since we are not restricted to merely one or two feedback processes, but can adapt to many recommendations as we are progressing towards meeting our objectives.

Adding a new dimension: Virtual Research Communities

Following the period of stakeholder consultation, it was suggested that in order to serve the needs of the scientific community (as opposed to merely mapping investments), the system would need to help bridge the gap that exists between scientists that are located in various parts of the African continent, with the resources available for carrying out their work.

As such, while iMENTORS would proceed with mapping the different applications available to users, one of the most important contributions from the workshop at the IST Africa conference in May 2013, was the need to add the Virtual Research Community (VRC) dimension to our database.

The concept of VRCs is at the cutting edge of the use of technology to support research. As a result, there are still no commonly agreed definitions, nationally or internationally: in the United States, for example, VRCs are frequently referred to as 'Collaboratories' with 'Cyberinfrastructure' synonymous with 'e-Infrastructure'. The EGI ecosystem is in constant flux, comprising Service Providers, Technology Providers, Funding Bodies and most importantly, Users. The EGI User Community can be decomposed via two main interconnected types: Virtual Organisations (VOs) and Virtual Research Communities (VRCs). Being sometimes difficult to distinguish or define, the European Grid Initiative proposed the following definition:

"A VO is a group of people with common interests and requirements, which need to work collaboratively with other members of their collaboration and/or share resources (e.g. data, software, expertise, CPU, storage space) regardless of geographical location. They join a VO in

order to gain access to resources with a set of rules and policies that govern the access and security rights for the users, resources and data in question.”¹

On the other hand, a “VRC is a group of large scale scientific research collaborations, either covering multiple VOs or simply belong are part of a larger domain area. The VRC model allows a community to have bi-directional interactions through defined points of contact across broader domain areas. Within the community, researchers can collaborate, communicate, share resources, access remote equipment or computers and produce results as effectively as if they, and the resources they require, were physically co-located.

To record such Virtual Communities (both VOs and VRCs), iMENTORS will not create separate entities, since despite their reported difference; they bear significant resemblance in their attributes and characteristics, and shall be distinguished from each other through a tagging option.

The task of mapping VRCs will be carried out in parallel with the recording of projects and research infrastructures, in the second reporting period, and will allow creating several new interconnections between (i) users, (ii) the distributed Research Infrastructures to which they are affiliated to, (iii) the applications & (iv) the supported middleware, and (v) the projects that led to their development (if applicable).

VO/VRC profiles will contain the following information:

1. Name [free text]
2. Acronym [free text]
3. Discipline(s) [tags]
4. Description of Activities [free text]
5. Homepage URL [free text]
6. Members [interlinking to users or interlinking with other VRCs]
7. RIs with which they belong to [interlinking to RI]
8. Applications they use [interlinking to RI]
9. Supported middleware [interlinking to RI]
10. Projects that contributes to their development [interlinking to projects]

The different relationships in the iMENTORS database

To this point, the iMENTORS database has the potential to contain a very large number of interconnections between the different entities that are recorded. To illustrate our point, all potential interconnections are listed below.

Direct relationships with Users:

Users can be associated to:

- a. Projects: as (i) Project managers and (ii) Project officers;
- b. Research Infrastructures as (iii) contacts;
- c. Organisations as (iv) staff.

Direct relationships with Organisations:

¹ EGI/NGI Policy Session: Survey Analysis EGI User Forum: 13 April 2011: Vilnius, Lithuania.

Organisations can be associated to:

- a. Projects: as (i) donors, (ii) lead beneficiary/main applicant, (iii) as project partners;
- b. Research Infrastructures / e-infrastructures: as (iv) owners of RI/e-Infrastructures;
- c. Other organisations:
 - a. in the framework of Research and Education Networks, organisations can be marked as (v) belonging to a NREN, and NRENs can be marked as (vi) belonging to a RREN.
 - b. (vii) as consortia managing projects together, or (viii) in the framework of ventures, as well as (ix) in a mother-child relationship (i.e. European Commission → DG CONNECT, or United Nations → UNDP).

Direct relationships with Projects:

Projects can be associated to:

- a. RIs as (i) contributing to their development;
- b. Networking infrastructures: as (ii) contributing to their development;
- c. Other projects to create a relationship between (iv) a project and its follow up project, a project and its (v) predecessor, and (vi) a project and its sub-projects (i.e. a project that is performed in several phases with distinct objectives from one another).

Direct relationship with RIs:

RIs can be associated to:

- a. Organisations as (i) belonging to one or more organisations;
- b. Users as (ii) being administered by them.
- c. Each other (iii) to depict a situation in which a single sited RI forms part of a distributed infrastructure under a separate framework,

Direct relationships with Networking Infrastructures:

Networking infrastructures can be linked to:

- a. Organisations (i) as belonging to them,
- b. Projects (ii) as having being funded by them

Direct Relationship with VRCs:

VRCs can be linked to:

- a. Users (i) as being part of the VRCs
- b. Other VRCs (ii) to depict cases in which VRCs are composed of already established Virtual Organisations (which have been mapped in the same way in iMENTORS)
- c. RIs to associate it with (iii) applications they use, and (iv) with the supported middleware, and (v) the RI to which they are associated to (i.e. EGI).
- d. Projects (vi) that contributed to their establishment.

4 Data collected and progress achieved in relation to the Critical Success Factors

4.1 Organisations

Identifying and recording the key stakeholders involved in e-infrastructure projects in sub-Saharan Africa has been the primary focus of the WP2 team's efforts. At the time of deliverable D2.2, we reported 217 organizations identified and recorded. We are now able to report an additional 1953 organizations, an approximately eight-fold increase. At the time of writing, 1666 of these 1953 organizations (approximately 85%) have been uploaded to the database.

The different types of organizations that are related to e-Infrastructure project include both national and international organizations, governmental and non-governmental organizations, private and public organizations as well as national and multinational donor agencies. The present section will list all the data collected so far on the actors that compose or potentially compose the world of e-infrastructures. Each sub-section corresponds to a specific type of actor, organised according to the classification scheme outlined in Deliverable 2.1.

The breakdown of the different types of organizations recorded in the spread sheets can be seen in the chart below. Universities compose the largest group recorded, followed by the cable operators and ISPs. The latter two comprise the biggest part of the private sector organizations recorded.

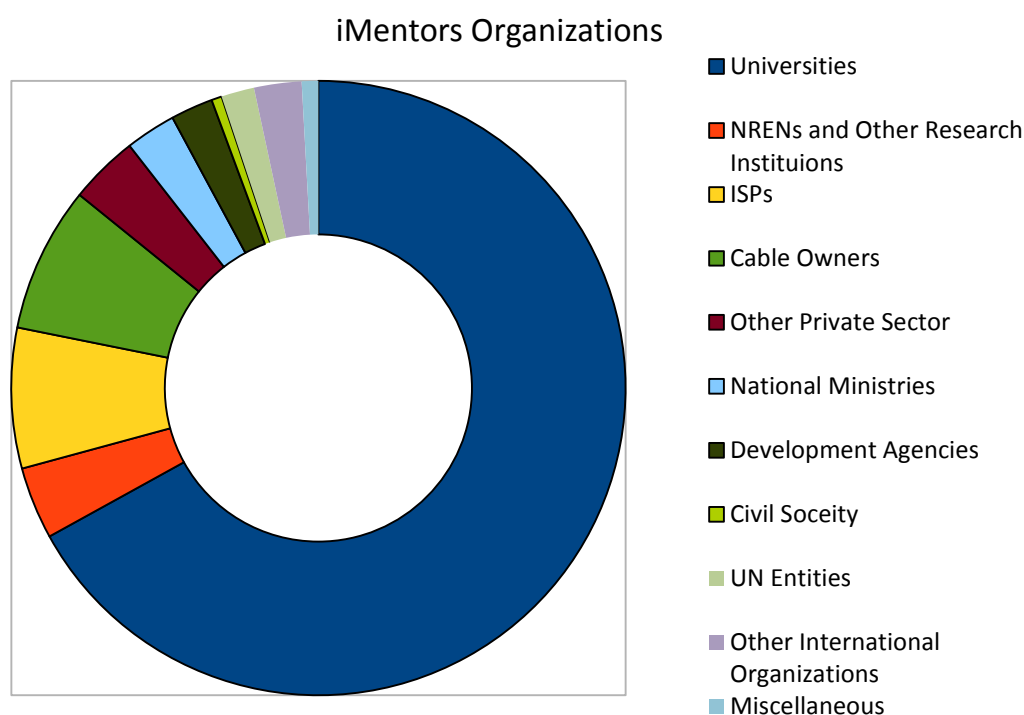


Figure 1 Data Collected on Organisations

Once each organization is uploaded, it is tagged in the database as belonging to one or more of the following six categories: United Nations Entities, Other Intergovernmental Bodies/Institutions, Government Agencies and Public Institutions, Higher Education and Research Institutes, Civil Society Organizations (non-profit), and Corporate & private Sector (for profit). Since one organization can be tagged as part of more than one category (for example, a public university that is both a Higher Education and Research Institute and a Public Institution), the chart above is based on the more narrow categorizations. In this way, we hope to avoid the problem of double-counting, although we cannot completely eliminate the risk that some categories may overlap (ISPs and Cable Owners, for example, may include some entities that fall into both categories).

Of the six categories, we have substantially completed recording the publically available online information of the organizations in the following four categories: Higher Education and Research Institutes, UN Entities, Other International Organizations, and Government Agencies and Public Institutions. The size of the remaining two categories - Corporate sector companies and civil society/non-profit organizations - is very difficult to estimate. However, we have made substantial inroads in recording two large groups of private organizations involved in networking infrastructure projects: cable owners and ISPs. Because many of the same cable operators are involved in both submarine and terrestrial cables, and we have finished recording the submarine cables, we do not anticipate adding a substantial number of organizations to this category. Finally, additional civil society/non-profit organizations will be added as necessary in the next phase when we shift our focus to recording projects.

In sum, despite the difficulties involved in estimation, we have made significant progress in identifying and recording the key stakeholders. Given the exhaustive online surveys we have performed in four of the six categories, including the extensive work involved in completing the Higher Education and Research Institutes category - the largest category by a substantial margin - we are confident that we have exceeded the critical success factor of 20% of the identified community online by month 12 of the project, and are well on our way to having 60% identified and uploaded by month 24.

More specific information about the data collection related to each of these subgroups follows in the subsections below.

4.1.1 United Nations Entities

A total of 32 United Nations Entities have been recorded and uploaded into the database. This is a relatively small category, and we have recorded most, if not all, relevant entities.

Recommendations for increasing the quality of data: Location and contact information for local and regional offices for these entities may be added, a task best completed in the process of recording projects.

4.1.2 Inter-Governmental Bodies/Institutions

A total of 46 inter-governmental bodies were recorded and uploaded (not including UN Entities, which are tagged separately). Additional Inter-Governmental Bodies/Institutions will be added when found related to a recorded project or infrastructure.

Recommendations for increasing the quality of the data: A separate profile for each relevant organizational unit should be added, along with information pertaining to local and regional offices.

4.1.3 Government Agencies and Public Institutions

Forty-nine (49) national ministries and 41 development agencies have been recorded and uploaded, for a total of 100 additional organizations tagged under the category of government agencies and public institutions. Note that public universities are also tagged under this category, but they will be treated separately under the category of higher education and research institutions for reporting purposes.

We have primarily focused on recording the national ministries most relevant to iMENTORS, including Science and Technology, Education, Infrastructure, Health, and Finance. National Assemblies and Parliaments may also be added, along with a "Government of ..." entry for each country for the purpose of linking all ministries and agencies together.

Recommendations for increasing the quality of data: Information on government agencies and ministries is available online in some countries, but for others, it is not easily accessible, and the address is not always accurate or recognized by Google maps, which makes it difficult to add to the iMENTORS map. More research is needed in those cases, and directly contacting the embassies may be the best way to proceed.

4.1.4 Higher Education and Research Institutes

A total of 1303 entities tagged as higher education and research institutes were recorded, and of those, approximately 1016 (approximately 78%) have been uploaded to the database. The number of universities alone reported was 1233, making this the largest single category of organization. Research and Education networks comprised a second type of organization tagged under this category, as detailed below.

4.1.4.1 Universities

A total of 1233 universities were recorded in all countries of Africa with the exception of Western Sahara, French Mayotte, and St. Helena, where no information was available. The universities subgroup accounts for the largest percentage of recorded organizations (see Figure 1, above).

Method used and source of information: Wikipedia and the website classbase.com, supplemented by general Google searches, were the primary methods used to identify the universities in each country. After identifying the universities, the information to be recorded in the database was found on the universities' individual websites.

The quality of the data varied widely depending on the country. In countries such as Kenya and Egypt, for example, most of the universities had comprehensive websites with easily accessible information. In smaller countries such as Guinea-Bissau or Sao Tome and Principe, the universities' websites were often under construction or non-existent, in which case information was very difficult to find; the data collection for the majority of countries fell somewhere between these two extremes, with information readily available for larger, state universities and more difficult to find for smaller technical colleges.

Overall, about 84 % of the universities had websites, and we were able to record email addresses for about 71% of the universities. Only about 8 % of the universities we recorded lacked both an email address and a telephone number.

Recommendations for increasing the quality of the data on universities: Calling or emailing the universities via the contact information provided in order to verify the information collection and request that they supply any missing information and/or corrections. In some cases, when both email address and telephone numbers are not available on the website, we will look for the facility of sending messages from the website itself.

4.1.4.2 Research and Education Networks

Method used & source of information: The individual sites of the NRENs, as well as Wikipedia pages and the information contained on the website of the UbuntuNet Alliance.

The NRENs were identified by a list of participants in the Africa Connect research infrastructure project, and information about these participating NRENs was easily accessible. However, information on the NRENs that are not yet consolidated or matured is not retrievable through online searches.

Recommendations for increasing the quality of the data on NRENs: the consortium has begun collaborating with the UbuntuNet Alliance (Dr. Margaret Ngwira) to begin consolidating the data collected so far. iMENTORS has also signed a Memorandum of Understanding with the eI4Africa project, which provides a valuable comparison. Additionally, in case further data is missing, the consortium will begin contacting the NRENs by phone or email. The table below lists all the RENs in Africa recorded on the platform.

Table 3 African NRENs

NREN	Country	RREN	Information Status
Eb@le (DRC)	D.R.Congo	UbuntuNet Alliance	Comprehensive
RITER	Cote d'Ivoire	UbuntuNet Alliance	Comprehensive
EthERNET	Ethiopia	UbuntuNet Alliance	Comprehensive
GabonREN	Gabon	UbuntuNet Alliance	Comprehensive
GARNET	Ghana	WACREN	Comprehensive
KENET	Kenya	UbuntuNet Alliance	Comprehensive
MAREN	Malawi	UbuntuNet Alliance	Comprehensive
MaliREN	Mali	WACREN	Comprehensive
MoRENnet	Mozambique	UbuntuNet Alliance	Comprehensive
NAMREN	Namibia	UbuntuNet Alliance	Missing
RwEdNet	Rwanda	UbuntuNet Alliance	Missing
SomaliREN	Somalia	UbuntuNet Alliance	Comprehensive
SUIN	Sudan	UbuntuNet Alliance	Comprehensive
TENET	South Africa	UbuntuNet Alliance	Comprehensive
TERNET	Tanzania	UbuntuNet Alliance	Comprehensive
RENU	Uganda	UbuntuNet Alliance	Comprehensive
TogoREN	Togo	WACREN	Missing
ZAMREN	Zambia	UbuntuNet Alliance	Comprehensive
SnRER	Senegal	WACREN	Comprehensive
ngREN	Nigeria	WACREN	Comprehensive
RIC	Cameroon	WACREN	Limited
iRENALA	Madagascar	UbuntuNet Alliance	Comprehensive
EUN	Egypt	EUMEDCONNECT	Missing

MARWAN	Morocco	EUMEDCONNECT	Comprehensive
RNU	Tunisia	EUMEDCONNECT	Comprehensive
CERIST	Algeria	EUMEDCONNECT	Comprehensive

4.1.5 Civil Society Organisations (non-profit)

A total of 10 civil society (non-profit) organizations were recorded and uploaded. This is a potentially very large category that is difficult to estimate and which cannot be recorded following the same methodology than for recording donors, higher education institutes or cable operators, for the simple reason that repositories containing information on every Civil Society organisation that is involved in such initiatives do not exist. As such, further entries tagged under this category will be added in connection with the recording of specific projects in phase two of our data collection project, on a case by case basis.

4.1.6 Private Sector Companies

A total of 344 private sector companies have been recorded, and all of those have been uploaded to the database. Additional Private Sector companies will be added when found related to a recorded project or e-infrastructure.

Much of the research carried out in identifying Private Sector Companies since deliverable 2.2 was submitted focused on identifying Cable and Telecommunication Operators, and Internet Service Providers (ISPs), as detailed further below. This data quality for this set was very high: only 6 private sector companies (approximately 2%) were missing a website, and only 8 (approximately 3 %) were missing both an email address and a telephone number.

4.1.6.1 Cable Operators

A total of 141 cable operators and investors in telecommunications from across the globe (North and South America, Europe, Asia, Africa and the Middle East) have been recorded so far.

Method used & source of information: The information was recorded from the operators' commercial websites and/or Wikipedia pages, which were found by performing a Google search for the name of each operator associated with submarine or terrestrial cable projects.

Most of the operators had easily accessible information on commercial websites. The only difficulties arose with websites that lacked an English version. In these cases, most of the information was retrieved by using Google translate.

Recommendations for increasing the quality of the data on cable operators: Contacting the cable operators via phone or email.

4.1.6.2 Internet Service Providers

A total of 136 Internet Service Providers (ISPs) in the following countries; Ghana, Gabon, Gambia, Kenya, Zimbabwe, Zambia, Uganda, Western Sahara, Kingdom of Lesotho, Liberia, Nigeria, Tanzania (partly).

Method used & source of information: Using internet search engines and membership lists of Internet Exchange Points, main internet service providers were mapped.

Many private entities regularly change ownership, so it is often difficult to track down the main ISP. These commercial entities have good quality websites but could be engaged to give more details of past and present projects.

Recommendations for increasing the quality of the data on ISPs: There is a clear interest for Internet Service Providers to want to be listed in the iMENTORS Database. Many of them have Corporate Social Responsibility policies and research investments or projects that they would like to showcase, and attract partners for.

4.2 Research Infrastructures and e-infrastructures

The type of entities that comprise this category can be associated with the following attributes: High Performance Computing, Instruments (sensors, telescopes, etc.), Data Infrastructure, Distributed Grids, Applications, Middleware.

At this stage, the consortium possesses limited information in this category, although the processes required to proceed with the mapping have already been established:

- Distributed Grids, Middleware & Applications will be retrieved with relative ease by re-using the results of the Chain-REDS project. A partial integration with the Chain-REDS database will ensure that the data in our own database is up to date at all times.
- Single-sited High Performance Computing: it is expected that the data will be retrieved through the information acquired on the distributed grids. Additionally, there are a few flagship HPC initiatives, such as the Supercomputing Centre in Tanzania (donated by India), and the High Performance Computing Centre in South Africa.

The consortium will proceed with identifying all research infrastructures in the first half of the second reporting period.

4.2.1 Networking Infrastructures

A total of 112 network infrastructures have been recorded and uploaded online. The categories of e-infrastructure recorded include submarine cables, terrestrial cables, and internet exchange points (IXPs), of which the submarine cables constitute the largest group.

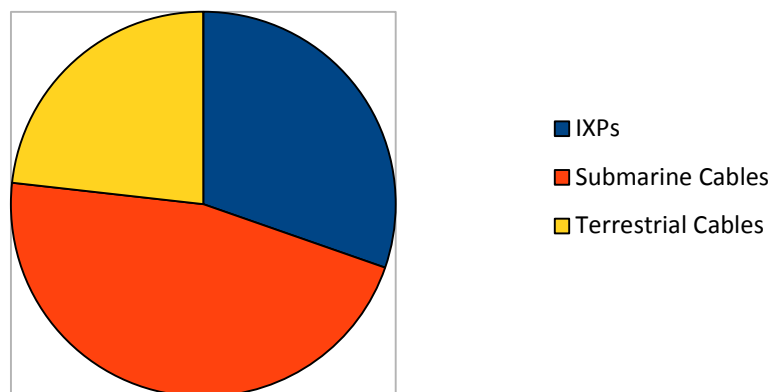


Figure 2 Types of Network Infrastructure

We have exhausted the publicly available information with respect to both IXPs and submarine cables (that is, we have recorded 100% of the information accessible online with respect to these two categories), and estimate that we have recorded approximately one-third of the terrestrial cables for which online documentation can be retrieved. More specific details about the data collection process for each category of network infrastructure follows in the subsections below.

In the next phase of our data collection process, we will focus on finishing recording the terrestrial cables and then turn our attention to research infrastructure, including high performance computing, distributed grids for e-Science, instruments such as telescopes and sensors, and software and middleware.

4.2.1.1 Terrestrial Links

A total of 26 terrestrial cables within and between the countries of Algeria, Angola, Benin, Botswana, Burkina Faso, Chad, Cameroon, Central African Republic, Cote d'Ivoire, Republic of Congo, Equatorial Guinea, the Democratic Republic of Congo, Gabon, Ghana, Kenya, Lesotho, Madagascar, Mauritius, Namibia, Niger, Nigeria, Malawi, Rwanda, São Tomé & Príncipe, Senegal, South Africa, Somalia, Sudan, Tanzania, Togo, Uganda, Zambia, and Zimbabwe

Method used & source of information: The Hamilton Research map and database were used to identify the main terrestrial lines by selecting posts tagged with the keywords "long haul fiber optic." Once identified, further information was retrieved by searching for the cables using Google, and also via Wikipedia and the AfTerFibre project.

Very few of the cables have their own website information or specific contact information apart from the general contacts found on the website of the cable's operator. In addition, it is often very difficult to determine the progress made on different lines as many projects are rolled out in phases or announced as extensions to existing lines.

Recommendations for increasing the quality of the data on terrestrial cables: Contacting the cable operators by phone or email. In addition, a more thorough search will be conducted on a country-by-country basis using the collection of maps in the AfTerFibre project along with the Hamilton Research maps.

4.2.1.2 Submarine Cables

A total of 52 submarine cables were recorded, the largest group in the category of network infrastructure (see Figure 2, above). The cables had landing points in Algeria, Angola, Benin, Cameroon, Cape Verde, Comoros, Cote d'Ivoire, Democratic Republic of Congo, Djibouti, Egypt, Equatorial Guinea, Gabon, Ghana, Guinea-Bissau, Kenya, Liberia, Libya, Madagascar, Mauritania, Mauritius, Mayotte, Morocco, Mozambique, Namibia, Nigeria, Reunion, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, Sudan, Tanzania, Togo, and Tunisia.

Method used & source of information: The submarine cables were identified via the Hamilton Research map and database and website submarinecablenetworks.com. Once identified, a Google search returned websites for individual cables, a Wikipedia page for the cable, and/or a press release or other information about the cable on the cable operator's website. These different sources of information were combined to yield the data needed for the iMENTORS database.

Many of the submarine cable projects lacked their own website and contact details. In these cases, the majority of the information came from the websites and press releases of the

cables operators, and it was also often difficult to find full information about the capacity, length and value of the project.

Recommendations for increasing the quality of the data on submarine cables: Contacting the operators via phone or email to verify the information collected and request corrections.

4.2.1.3 Internet Exchange Points

A total of 34 Internet Exchange Points (IXPs) were recorded across the African continent and the data analysed for inclusion in the iMENTORS project.

Method used & source of information: There is currently no comprehensive and update list of IXPs available for Africa as a region. We used internet search engines and paired available information with contact details provided by one of SPIDER's contact to have as comprehensive list as possible.

The type of information available on IXPs varies greatly between countries and exchange points. Whereas a number of IXPs use a very transparent and detailed way of reporting activity, others merely report themselves to exist, with no physical location, administrator or separate website. A number of IXPs lack a homepage but there does exist a traffic graph URL that can be used to monitor activity at the exchange point. Much of the information readily available about various IXPs is outdated by a few of years. It is therefore possible that some IXPs no longer exist or are to be replaced. We did remove the IXPs that are no longer in use.

It was nearly impossible to find out anything about the ownership status of IXPs (where the IXP is not a project of one specific national internet service provider). Details of funding/investment and figures for IXPs to operate illuminate in their absence.

Recommendations for increasing the quality of the data on IXPs: Contacting all IXP administrators by email presenting the information we have for each IXP and requesting their corrections and updates. Information about funders is often difficult to obtain. Specific numbers is difficult to access, especially where the owner is a private company. One possibility may be to use the initial list to cross reference projects as they become known and/or reported to the iMENTORS database. Funding structures would need to be completed and confirmed by the IXPs administrators and/or funders.

4.3 Projects (as frameworks for providing funding)

Data collection carried out following a pre-defined methodology, accurately designed according to the need to balance priorities in relation to our critical success factors: we have achieved our target of gathering 20% of data on actors, and are confidently reaching our 60% target. As noted above, up until this time, the iMENTORS WP2 team has directed its efforts primarily toward identifying and recording key stakeholders to populate the virtual observatory. The team has met or exceeded the critical success factors for this first phase of data collection, which this report now brings to a close.

During the next phase of data collection, the iMENTORS team will direct its efforts toward identifying and recording specific e-infrastructure and ICT related development projects of the past five years. These projects will be linked back to the data for the sponsoring and participating organizations already recorded and uploaded to the database during the first phase of data collection, and additional organizations will be added only as needed.

At this stage, the iMENTORS platform contains 139 e-infrastructure projects. On the basis of the research carried out, and the projects that we have currently collected, the international

actors that are funding the largest share of global and regional e-infrastructure projects are the European Union (DG Development and Cooperation in association with DG CONNECT), Canada's International Development and Research Council (IDRC), and the World Bank. Data on projects can be retrieved with ease for all three actors:

- a. For EU projects the CORDIS database offers an advanced method to search for projects with a wide array of filtering options to look for initiatives according to their funding scheme, framework programmes, geographical location and many other attributes.
- b. Projects that are funded by the IDRC are all listed in its online database (IDRIS+) and albeit not exportable, the projects are all retrievable with relative ease thanks to its search options. In any case, the iMENTORS team working on data collection has engaged in discussions directly with the IDRIS+ support staff to determine whether it would be possible to gain access to the raw data through a dump file, without the need to deploy a crawler.
- c. Projects that are funded by the World Bank are also all listed in the World Bank database. The iMENTORS team developed a functionality to integrate the database, and import all projects through its public API. All projects are currently on the platform, and await validation by the iMENTORS team in order to be publically displayed as individual entries.

Due to the fact that these databases (with the exception of the World Bank) offer no means to export data, and offer no API to extract it, we are currently in the process of contacting the support teams for each database to discuss the opportunity to retrieve the raw files directly from them. Alternatively, we will investigate how to deploy a crawler that will run through a pre-defined list of results (by finding the correct query to generate the target results). We expect to generate a large number of results in this way.

In addition, the iMENTORS database has been built in such a way that it dissociates physical networks from the funding or investment required to build them. The physical infrastructure is recorded first, and then must be linked with its funding or investment source. Therefore, in addition to mapping the physical network as described in the "networking infrastructure" subsection above in this report or any other research infrastructure, the research team will need to record the associated investments as projects which are managed by consortia or legal entities created for this specific reason. For networking infrastructures, iMENTORS has acquired the data needed by purchasing the Hamilton Maps. Each network is associated with different news sources which typically also provide financial data. For research infrastructures, the team will proceed by a web-survey and by contacting the administrators of each infrastructure.

For all other bilateral or multilateral funding agencies which do not have such project databases or publish no information regarding the projects which they are funding, there is no alternative than contacting each of these institutions independently and requesting their cooperation and assistance in retrieving these projects. One important factor which will determine the success of this method relates to the internal records of the relevant institutions. In many circumstances, National Development and Cooperation Agencies only partially align their practices with the international standards of the International Aid Transparency Initiative (IATI) and fail to report on individual 'project-level' initiatives. Thus, many reports include project identifiers which are not unique, a practice which renders such projects very difficult to trace.

5 Integrating iMENTORS with the WORLD BANK project database

The iMENTORS initiative aims to create a comprehensive database of ICT projects that have taken, are currently taking and will take place in Africa. There are some databases on the internet that contain data on this area, but they are usually focusing only on a specific country and/or organization. In order to make the most of the new opportunities that ICTs provide for interoperability and flow of information between organizations and platforms, a system for the synchronization of iMENTORS with other databases was developed. This system will also provide certain economies of scale, which are important for the efficient long-term run of the project.

In the first phase, a review of publicly available databases took place, in order to select an appropriate one for a test run. Relevant databases were evaluated in terms of the amount, the quantity and the quality of data they provided. The developer team (WP3) along with the research team (WP2) examined them carefully to determine which one was the best candidate for synchronization. From a data compatibility perspective, the data contained in any candidate database needed to closely match the data requirements of the iMENTORS typology; from a data validation perspective, the database further needs to provide unique donor project identifiers and offer project level data; whilst from a technical standpoint it was highly desirable to connect through an existing API. So far, the only suitable candidate proved to be the World Bank Database, as it fulfilled all three criteria. ²

A framework was created to allow the various categories and data from API to be “translated” to the corresponding entries in iMENTORS. The developing team not only created the synchronization tool, it also incorporated it directly in the iMENTORS platform. The only synchronization option at the moment is with the API, but the system can be expanded to include more databases, which in turn will become directly available as an option on iMENTORS.

It must be noted that the synchronization process is not a fully automated one, for two reasons. First, databases and websites often undergo changes and updates. A project category that is similar in both databases now might not have the same form next year. Second, iMENTORS has the responsibility to review the data it contains for accuracy and relevance. The initial filtering of information takes place by having the user select the appropriate project filters on the API website. The API produces a web link that leads to an XML page with the requested information. The user feeds this link to the synchronization tool, which then proceeds to translate and enter the data in iMENTORS. All the data entered this way are placed under the new “under review” status, created specifically for this process. As the final step, the user can check the newly-entered data and, if they are correct, move them to “reviewed” status, making them publically available on the iMENTORS website.

The process combines clever data selection, instead of unproductive complete data duplication between databases, time-saving on the data uploading process and quality checks. The tool’s ability to handle large amounts of information, the synchronization process only needs to take place once to twice per year, after API has received its update of yearly reports. The speed of the process and its regularity result in significant savings in work hours and a steady inflow of information.

²<http://search.worldbank.org/api/v2/projects>

On a second stage, the research team will focus on extending the number of external databases that are connected with iMENTORS. We are currently in the process of speaking with both the IDRC and the IATI database administrators and support staff to meet our objectives. In a last phase, iMENTORS will evaluate how to draw data from the CORDIS database. The aim is to connect all three databases because it has proven to be an extremely efficient and functional way to collect data, with comprehensive productivity gains.

Pending EC Approval

6 Conclusion: Where we are and what's next

During the time period covered by this report, the iMENTORS WP2 team recorded 1841 organizations and 139 e-infrastructure projects for a total of 1953 entities, representing an approximately eight-fold increase over the number of reported entities in the last deliverable. Of these entities, 1666 (approximately 85%) have been uploaded to the iMENTORS database.

In addition, existing data has been enriched and uploaded, and a synchronization tool was incorporated directly into the iMENTORS database in collaboration with the WP3 development and customisation team. The synchronization tool will allow for integration with the database at the World Bank, in turn generating many new project entries. More work remains to extend the tool to allow integration with additional databases, but the tool will be a great aid to the team in meeting the key target for the next reporting period: having 80% of all e-infrastructures and ICT related projects recorded.

The exhaustive online survey completed to enable the recording of all entities in the category of higher education and research institutes and other key organization categories allowed us to meet the critical success factor of having 20% of the online community recorded by month 12 and puts us well on our way to meeting the 60% goal by month 24. However, due to the sheer number of organizations necessary to record, we were unable to focus on mapping the SIDA and EU-funded (FP7) e-infrastructure projects we identified as a goal for ourselves in the previous deliverable. We will thus return to this goal in the next phase of the data collection process with renewed focus.

The WP2 research team also met with success in devising a method for recording and mapping networking infrastructure to, from, and within Sub-Saharan Africa. For the next reporting period, we are planning to have the data now recorded mapped online. Further, we will seek to collect data on hospitals, particularly advanced and university hospitals in need of connections to a fibre optic cable backbone for the purpose of telemedicine and research.

Finally, the WP2 team will continue to work on uploading the existing data recorded on the spread sheets to the iMENTORS database. We have met with great success overall in our efforts during this phase of data collection to identify and record the key stakeholders involved with e-infrastructure development in Sub-Saharan Africa. We expect to meet with similar success in identifying and recording specific projects during the next phase of data collection.

Pending EC Approval



iMENTORS is a project co-funded by the European Commission's DG CONNECT under the 7th Framework Programme.

www.iMENTORS.eu

©Copyright by the iMENTORS Consortium