Grant agreement no. 211714

neuGRID

A GRID-BASED e-INFRASTRUCTURE FOR DATA ARCHIVING/ COMMUNICATION AND COMPUTATIONALLY INTENSIVE APPLICATIONS IN THE MEDICAL SCIENCES

Combination of Collaborative Project and Coordination and Support Action

Objective INFRA-2007-1.2.2 - Deployment of e-Infrastructures for scientific communities

Deliverable reference number and title: D2.4 Report of implementation of the neuGRID protocol for data protection/safety

Due date of deliverable: month 36

Actual submission date:

Start date of project: February 1st 2008 Duration: 36 months

Organisation name of lead contractor for this deliverable: PROVINCIA LOMBARDO-VENETA - ORDINE OSPEDALIERO DI SAN GIOVANNI DI DIO FATEBENEFRATELLI

Revision: Version 1

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	PU
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Table of contents

EXECUTIVE SUMMARY

The present deliverable aims to report the WP2 "Privacy and ethical issues" monitoring of the implementation of the data protection protocol (D2.3) in the neuGRID e-infrastructure.
The proposal for data protection (D2.3) is briefly summarized pointing out the three key points (anonymous data, informed consent, secondary use) and the two main scenarios on which the protocol was developed. Furthermore, the relations with the Ethics Committees (the Independent CEIOC Ethics Committee set up for neuGRID Project, and the partners' Ethics Committees: UWE, VUmc, KI) regarding the submission and approval of the neuGRID data protection protocol are described.
The implementation of the data protection protocol was monitored through continuous contacts with technical partners via email, participation to in-person meetings and teleconferences, and revision of deliverables developed in other workpackages. The most relevant issues raised during the protocol implementation are reported and discussed. They relate to: ADNI dataset; DICOM header; single user case; technical difficulties in ensuring anonymization; and informed consent.

1THE PROPOSAL FOR DATA PROTECTION IN neuGRID

On month 6 (July 2008) the first deliverable of the WP2 was authored: D2.1 "Review document on data protection (legal and procedural issues)" based on the revision of the European legal framework on data protection with particular regard to Directive 95/46/EC and its implementation in Member States.
On month 12 (January 2009), starting from the review document on data protection (D2.1), a specific protocol for ensuring privacy in neuGRID was developed: D2.3 "Protocol for ensuring data protection/safety in neuGRID". This protocol was developed on the following key points: 1) data are considered as personal when the data subject may be identifiable; 2) specific informed consent of the person concerned is always required for the processing of sensitive data; 3) secondary use of data for scientific research purposes exempts from the duty to inform subjects only if the use of the data is consistent with the original purpose, the data are anonymized, and the provision of information is impossible or impracticable. In particular (Table 1):
1) Anonymous data. In the neuGRID protocol for data protection we have adopted a "relative" notion of anonymous data: data must be fully anonymous for the final users of neuGRIDgrid, and, as far as it

is technically possible, for the researchers of the coordinator laboratories (core labs) involved in the control of data quality. However, the centre that first collected the data needs to have the possibility to re-identify the subjects, if this is in the subjects' best interest, and the subject needs to maintain the possibility to withdraw from the projects. Any solution must consider all the interests of the research subjects: the interest to privacy, the interest to health, the interest to withdraw from the project. In fact, since full anonymity is the preferred solution to ensure the maximum degree of data subject privacy because no link with the subjects is maintained, the removal of any link between data and the subject concerned prevent both the possibility to inform the data subjects about research results that can be in the subjects' health interest, and the subjects' possibility to withdraw consent.

2) Informed consent. In the neuGRID protocol for data protection, informed consent for the processing of personal data should be given in a written form, in order to comply with the strictest legal and ethical requirements. As research subjects have the right to refuse the processing of their data, subjects need to be aware that their participation in the research project is possible only if they agree with the processing of their personal data. Research subjects have also the right to withdraw their consent at any time. If subjects revoke their consent any further processing of their data has to be considered unlawful, therefore their data and images must be removed from neuGRID data set as soon as possible. A temporary re-identification, that is possible only if data are only pseudonymized, will be put in place in order to allow the cancellation of the data. The right to refuse the data processing and the right to withdraw should be clear in the informed consent form.

3) Secondary use. In the neuGRID data protection protocol, in case of secondary use of clinical data and images, subjects need to be re-contacted and asked for their informed consent. A secondary use of data for scientific research purposes exempts from the duty to inform data subject only if all the following conditions are met:

- the use of data is compatible with the original purpose;
- the provision of information is impossible or impracticable; and
- the data are anonymized.

The compatibility of the secondary use and the impossibility or impracticability of providing information to the subjects must be evaluated case by case.

The aforementioned key points have been taken into consideration with reference to two major scenarios:

A) clinical data and images are collected from subjects specifically enrolled to be entered in neuGRID project;

B) clinical data and images were previously collected in different research projects and have been used, processed or communicated in or through the neuGRID e-infrastructure.

In the latter scenario two different possibilities have been foreseen:

1. the data are collected/ stored in neuGRID (scenario B1);
2. the data are used, but not collected/stored in neuGRID (scenario B2)

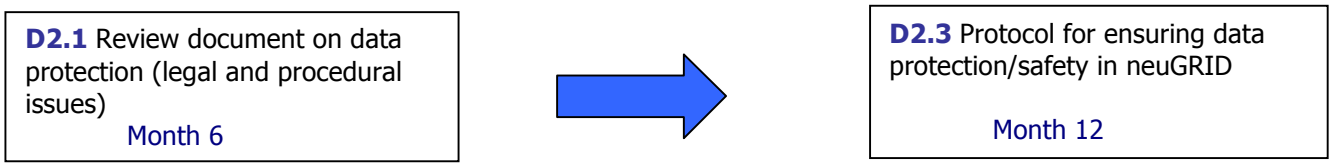
A specific procedure has been defined with regard to the different scenarios for both the informed consent and the anonymization of: clinical data, DICOM header and images.

For clinical data and DICOM header, in the absence of specific rules for the anonymization process in the European context, the USA HIPAA Privacy Rule (45 CFR Parts 160, 162 and 164) provisions have been taken into consideration. That Act indicates one of two different requirements must be met in order to ensure data protection: de-identification to a statistical standard; and de-identification by removal of 18 specific identifiers. For reason of feasibility in the neuGRID project we adopted the removal of the identifiers. In the selection of the identifiers that need to be removed we aimed to balance between guarding patient confidentiality and considering the needs of research which is performed to achieve results that will be beneficial to the medical community and the humankind. The suggested list resulted from the consideration of the HIPAA list of identifiers that must be removed, revised to guarantee the research quality.

For brain images, since from MR images of the brain it is possible through a rendering of images themselves discover the biometric features of the subject concerned and potentially determine his/her identity, the defacing of the data subject face has been suggested for ensuring a suitable protection of data subjects privacy.

As stated in the Description of Work of the neuGRID grant agreement a first draft of Deliverable 2.3 has been circulated at an early stage among partners for comments and suggestions that have been taken into account in writing the second draft. A late draft of the protocol was circulated for last comments among partners and was made available to the members of the Advisory Board before the first Advisory board meeting that took place in form of teleconference. The final version was submitted to the European Commission on 30th January2009 as scheduled.

On 31st March 2009 the protocol for data protection was presented to the European Commission for the first annual review.



PROPOSAL FOR DATA PROTECTION IN neuGRID (D2.3)

Three key points:

1. Anonymous data

Relative notion of anonymous data:
- For the final users and for the core labs researchers data are anonymous
- The collecting centres have the possibility to re-identify the subjects

2. Informed consent

Written specific informed consent of the person concerned is required for the processing of sensitive data

3. Secondary use of data

- Subjects need to be re-contacted
- Exception only if:
  - the use of data is compatible with the original purpose
  - the provision of information is impossible or impracticable
  - the data are anonymized

Two main scenarios for the neuGRID e-infrastructure:

A. Clinical data and images are collected from subjects specifically enrolled to be entered in the neuGRID project;

B. Clinical data and images were previously collected in different research projects and have been used, processed or communicated
in or through the neuGRID e-infrastructure
1. the data are collected/stored in neuGRID (B.1)
2. the data are used, but not collected/stored, in neuGRID (B.2)

Table 1

2 RELATIONS WITH THE ETHICS COMMITTEES OF neuGRID

According to the Technical Annex of the neuGRID grant agreement, the neuGRID protocol for data protection (as well as D2.2 Rules for Commercial Exploitation of Data) had to be submitted, for the approval, firstly to the Independent Ethics Committee set up for the neuGRID project (CEIOC) and in a second time to the partners' Ethics Committees: University of the West of England, Bristol - Faculty of Environment & Technology Research Sub-Committee – FETRESC; VU University Medical Center, Amsterdam – Medical Ethics Review Committee - METc; Karolinska Institute, Stockholm - Stockholm Ethics Board.

Because of the focus of the present deliverable is on the implementation of the data protection protocol in neuGRID, information that follow only refer to D.2.3 and not D2.2 (even if the submission to Ethics Committees also included D2.2).

2.1. CEIOC Ethics Committee

On 15th September2009 the D2.3 Protocol for Ensuring Data Protection/Safety in neuGRID, along with a summary in Italian of the protocol (as required by CEIOC), was submitted to the Independent Ethics Committee CEIOC. The protocol has been discussed in the CEIOC meeting on October 21st and CEIOC gave favourable opinion to the deliverables, with a request of some clarifications. With regard to the "Protocol for ensuring data protection/safety in neuGRID", paragraph 3.5 "Data Protection protocol. Open issue: subjects' capacity to give informed consent", the Ethics Committee asked for clarification of the "protective measures" that should be set up in the case of collection of data and images from subjects not fully competent to give a valid informed consent. In order to accomplish the request of CEIOC the original statement: "*In case of subjects not fully competent to give informed consent, rigorous protective measures should be set up conformant with local regulatory frameworks where applicable* " has been modified as follows: "*In case of subjects not fully competent to give informed consent, rigorous protective measures should be set up conformant with local regulatory frameworks where applicable. We regard as protective measures the participation of the caregiver family member (or the legally authorized representative in accordance with applicable law) in the informed consent process as well as the involvement of the local Ethics Committee in decisions about the enrolment of subjects not fully competent*".
The WP2 answer to the request has been considered satisfactory by the CEIOC Ethics Committee.
Moreover, CEIOC has been updated quarterly about the neuGRID project progress.

2.2 Ethics Committees of the neuGRID partners

Name and contact address of the persons in charge of the partners' local Ethics Committees have been collected through the neuGRID local principal investigators.
In order to submit the neuGRID protocol for data protection to the Ethics Committees of the neuGRID partners (UWE, VUmc and KI), the local submission procedures and possible local rules on data protection and on anonymization procedures have been asked by e-mail to the Ethics Committee

secretaries and collected.

On 7th January 2009 the deliverable D2.3 along with CEIOC favourable opinion to the beginning of the neuGRID project (opinion n.11/2008 – in Italian and English translation) and CEIOC favourable opinion on the protocol for data protection (opinion n. 44/2009) with request of clarification, WP2 answer to the request (10/12/2009) and letter of acceptance of the clarification by CEIOC (16/12/2009 - in Italian and English translation) have been submitted to the Ethics Committees of the neuGRID partners (UWE, VUmc and KI) for their opinion.

In order to facilitate the submission and approval procedure steady contact via email with Ethics Committees has been kept.

- Faculty of Environment & Technology Research Sub-Committee – FETRESC - (University of the West of England):

UWE's Ethics Committee asked Professor Richard McClatchey to complete a specific form in addition to the documents already submitted.

On July 19th 2010 UWE Ethics Committee (FETRESC) gave favourable opinion to the data protection protocol stating that: "*The Chair is content to **approve** the application subject to the following:· You notify the Faculty Research Ethics sub Committee in advance if you wish to make significant amendments to your original FETRESC application. · You notify the Faculty Research Ethics sub Committee if you terminate your research earlier than planned. · You liaise with the Deputy Chair of FETRESC, Tony Solomonides, to review the implications of any changes to the research protocol arising from changes instigated by partner organisations. Those changes judged significant should be reported to FETRESC for the ethical implications to be considered and appropriate actions instigated*.".

- Medical Ethics Review Committee – METc – (VU University Medical Center)

VUmc's Ethics Committee asked to fill in a privacy form in order to complete the submission procedure and WP2 proceeded to fill in the form that has been sent to the Ethics Committee (July 19th) in addition to the documents already submitted.

On 28th October 2010 VuMC Ethics Committee (METc) gave favourable opinion to the protocol submitted stating that: "*The METc VUmc hereby gives approval for the proposed research to be performed in the Netherlands at the VU University Medical Center. The study does not fall within the scope of the Medical Research Involving Human Subjects Act (WMO)*.".

- Stockholm Ethics Board (Karolinska Institute):

In order to complete the protocols submission to the local Ethics Committees WP2 filled in the application form for ethical approval as required by the Karolinska Ethics Committee and sent it to local researchers for the translation of the application into Swedish. Following translation the application form has been sent to the Karolinska Ethics Committee (May 12th) to be integrated with the documentation already submitted. On November 2010 Karolinska EC asked further integration and WP2 helped the local researchers to reply.
At present, we are waiting for Karolinska Ethics Committee opinion on data protection protocol.

3 IMPLEMENTATION OF THE PROTOCOL FOR DATA PROTECTION

In order to monitor the implementation of the protocol for data protection/safety into the neuGRID e-infrastructure and to deal with the rising technical questions, those responsible for WP2 of neuGRID

i) have kept continuous contacts with technical partners via email;
ii) have participated to in-person meetings and networking area teleconferences;
iii) have contributed to the revision of the deliverables developed in other workpackages with special regard to to D9.2 "User Requirements Specification (URS) document final release" and D6.2 "Implementation: service prototype report" where a specific section is related to the anonymization service (see D6.2 p.47-52).

In person meetings

Amsterdam 9th-12th February 2009. During the Amsterdam in-person meeting on February 2009 the protocol for data protection was fully presented and the discussion about the implementation of the protocol in neuGRID infrastructure started.
Issues related to ethics and privacy and to the implementation of the data protection protocol have been then discussed in the following in-person meetings:

- *Geneva 1st-4th December 2008*
- *Brescia 7th-10th September 2009*
- *Amsterdam 15th-18th March 2010*
- *Stockholm 7th-9th June 2010*
- *Brescia 6th- 8th September 2010*

The most relevant issues taken into consideration during the in-person meetings and via email refer to:

- ADNI dataset
- DICOM header
- Single user case
- Technical difficulties in ensuring anonymization
- Informed consent

3.1 ADNI dataset

With regard to the ADNI dataset three major issues deserve to be underlined:
a) neuGRID data challenge
b) the possibility to collect / store ADNI data in neuGRID
c) the possibility not to perform the defacing of the ADNI data.

a) neuGRID data challenge

During year 2 the ADNI data needed to be uploaded to the e-infrastructure in order to analyse them and to test, among other things, the technical feasibility to deface and the time needed to perform defacing (neuGRID data challenge). In order to proceed with the upload of the ADNI data to neuGRID, WP2 recommended to submit to the ADNI Data and Publications Committee a specific "data use agreement" request, indicating the purpose of the use of the data and the researchers involved in the task (i.e. all members/researchers of neuGRID consortium).

b) possibility to collect/store ADNI data in neuGRID

According to the data protection protocol the ADNI data case should fall into the scenario B2, i.e. clinical data and images are not collected nor stored in neuGRID but only temporary used in the grid. The provision is to give the final users the possibility to work on data already collected in other data sets using the neuGRID computational facilities.
However, from several discussions on ADNI data the idea to upload and store the data in neuGRID has been explored. This provision would give the user a facilitated use of the ADNI data in neuGRID. In this case, we would be in the scenario B1 of D2.3 that is data collected in previous research protocols are transferred into neuGRID. Following WP2 recommendation, the neuGRID consortium agreed that, in order to make the ADNI dataset available through neuGRID first of all an agreement

with the ADNI consortium is necessary. Moreover, as written in the neuGRID data protection protocol, it is necessary to verify if the original consent given by the subjects enrolled in the ADNI study is compatible with the collection of the data in neuGRID. This check should be performed before the ADNI data is made available by neuGRID and can only be performed by the ADNI consortium.
For now, the idea to give the final users the ability to access ADNI data directly through neuGRID is still a hypothesis that will be possibly taken into consideration in the future development of neuGRID.

c) possibility not to do defacing

The data protection protocol foresaw de-facing brain images to prevent recovering subjects' face and possibly retrace their identity. During the development of data protection protocol, the neuGRID consortium agreed to make the de-facing of all images collected or simply used in neuGRID i) in order to provide a common data protection treatment for all data and images used in neuGRID (including datasets eventually elsewhere available without de-facing); ii) on the assumption that the de-facing does not interfere with the possibilities of research.

However, in the first attempt to implement the de-facing it has become clear that the procedure to deface is very time consuming and costly, so that to implement the procedure is in practice very difficult and not really feasible on large datasets without limiting the concrete possibilities of research on the data.
For the above reasons, even if performing the defacing is clearly a better solution to protect subjects' privacy, the possibility of not performing the de-facing on data already available to the public without de-facing was taken into consideration and finally regarded as acceptable. In the balance between the risk of not having the concrete possibility of doing research and the risk of recovering of subjects' face, not defacing images already available to the users without scrambling does not seem to represent an inacceptable damage for subjects' privacy.
The possibility of not performing the de-facing applies in particular to the ADNI dataset, which is un-defaced and can easily be accessed by researchers under authorization by the ADNI data and publications committee. This provision could also apply to other datasets already available to final users without the de-facing of images.

It must be underlined that the decision of not performing the defacing of datasets already available without defacing can be implemented in the neuGRID data protection protocol, only under the condition that an amendment to the original protocol is developed and submitted to the pertinent Ethics Committees for approval. In fact, by now the protocol submitted to the Ethics Committees, and

already approved by three of them, provides for the de-facing of all images collected or simply processed in neuGRID.

Nevertheless it should be noted that currently (and till the end of neuGRID project) ADNI data are not available trough neuGRID for the final users. Only neuGRID researchers can perform technical tests falling under the first authorization for data challenge.

3.2 DICOM header

According to the data protection protocol the anonymization of DICOM header has to be made by removal of 18 identifiers. As stated, the suggested list results from the consideration of the HIPAA list of identifiers that must be removed, revised to guarantee the research quality, in particular at the point C) related to "elements of dates":

(A) Names;

(B) All geographic subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes;

(C) All elements of dates (except year) for dates directly related to an individual excluding: - birth date (month and year admitted); - exams/visits date (day, month and year admitted); - date of death (month and year admitted);

(D) Telephone numbers;

(E) Fax numbers;

(F) Electronic mail addresses;

(G) Social security numbers;

(H) Medical record numbers;

(I) Health plan beneficiary numbers;

(J) Account numbers;

(K) Certificate/license numbers;

(L) Vehicle identifiers and serial numbers, including license plate numbers;

(M) Device identifiers and serial numbers;

(N) Web Universal Resource Locators (URLs);

(O) Internet Protocol (IP) addresses;

(P) Biometrics identifiers, including finger and voice prints;

(Q) Full face photographic images and any comparable images; and
(R) Any other unique identifying number, characteristic, or code.

During the implementation phase of the data protection protocol a discussion took place regarding what is the better option between indicating the identifiers that need to be removed or the fields that should be kept. The solution to indicate what fields should be kept has been suggested, being the removal of identifiers foreseen in D2.3 not enough to guarantee privacy in the experience of some neuGRID researchers.
The pros and cons of the two possibilities, to remove or to keep, has been discussed many times but a decision on which is the better solution on the topic has not been reached.
The matter will be taken into consideration in further development of neuGRID. At the present the existing solution is the removal of the 18 identifiers as stated in the protocol.

3.3 Single user case

During the discussions on the future use of neuGRID infrastructure, the single user case has been presented as a possibility. The single user case is a single user that wishes to load his/her collected research images and use the computational power of neuGRID without sharing the images with other researcher.
This case has not been considered in the original scenarios provided by the neuGRID data protection protocol.
For this case the de-facing of brain images is not regarded as necessary, because there seems to be no risk of breach of privacy.
This topic will be taken into account in further development of neuGRID.

3.4 Technical difficulties in ensuring anonymization

In the data protection protocol strict guidelines for anonymization have been suggested both at collecting centres level and the core Labs level. Some technical difficulties to ensure anonymization as described in D2.3 have been underlined by technical partners, in particular related to an understanding of the role of the core Labs different from the one at the beginning of the neuGRID project. The not originally foreseen possibility to upload data without interference of the core Labs

makes quite impossible from a technical point of view to control the anonymization procedure applied to the data used in neuGRID.

If these difficulties are confirmed, the anonymization guidelines could only be suggested to the users of neuGRID but it would be impossible to guarantee that the rules are in fact observed. In this case, a disclaimer regarding the different responsibility of neuGRID and the one of users should be developed. Moreover, the possibility of offering tools to the neuGRID users for performing anonymization will be evaluated.

This point needs further discussion once the neuGRID project will be in a further step of implementation.

In any case, each amendment of the data protection protocol should be submitted to the local Ethics Committees for approval.

At the present the guidelines for anonymization in neuGRID are still the ones provided in D2.3.

3.5 Informed Consent

As stated in the neuGRID data protection protocol, the requirement of informed consent is of capital importance in order to protect the fundamental rights of a subject in the context of medical treatment and research. From an ethical and legal point of view informed consent protects subjects and their fundamental rights to integrity and self-determination.

For this reason, the provision in D2.3 is that subjects' consent for the processing of sensitive data must be explicit and given in a written form.

With regard to the processing of personal data, subjects must be informed of:

- the purpose of the data processing, including the specificities of the GRID (this information corresponds to the clear and full information about the clinical trial and the neuGRID project);
- the identity of the controller and of his/her representative, if any;
- the procedures adopted in order to guarantee anonymity;
- the possibility to withdraw their consent at any time asking for the cancellation of their data.

In particular the subjects should be informed that: - the link between data and subject identity is maintained only in the collecting centre, with the aim to ensure the possibility to re-contact the subjects, if it is in their best interest, and the aim to give the subjects the possibility to withdraw; -the data are subjected to a double-coding process, the first in the collecting centre and the second in one

of the neuGRID core labs; - every technical possibility is put in place for ensuring the full anonymity of the data subjects for the final users.
The informed consent should make reference to the national data protection law and to specific local rules, if any. The information sheet and the informed consent form, as well as the research project protocol, must be approved by competent Independent Ethics Committees.
Also for the case of secondary use of data already collected in other datasets (scenario B) data subject must be informed at least about:
- the new purpose of the processing (i.e. inclusion in neuGRID);
- the identity of the controller or his/her representative.
Moreover before the storage or use of data in neuGRID the dataset owner has to investigate whether the original protocol and the subjects' informed consent give this possibility.

During the three years of neuGRID project, the significance and importance of data subjects informed consent has always been underlined, being informed consent the essential condition in order to storage or use the data in neuGRID.
At the present the procedure established in D2.3 with regard to informed consent has not been implemented in that there is neither storage nor use of data in neuGRID.

4 CONCLUSION

Since the submission of the neuGRID data protection protocol (D2.3) to the European Commission continuous contacts and discussions have been maintained with technical partners in order to implement the protocol. Moreover, the data protection protocol has been submitted to the Ethics committees of neuGRID partners for opinion and frequent contacts have been maintained to follow the process of approval.
In consideration of the neuGRID project state of art, the protocol has been only partly implemented.
This deliverable gives a report of the data protection protocol implementation and of the issues raised and discussed during the neuGRID project years.