



# SHARE

## D6.1 INTEGRATED ROAD MAP I

---

Document ID:	<b>SHARE-D6.1-v1.2.doc</b>
Date:	<b>24/02/2007</b>
Activity:	<b>WP6: Integrated Road Map</b>
Document status:	<b>FINAL</b>
Document link:	<b><a href="http://eu-share.org/deliverables.html">http://eu-share.org/deliverables.html</a></b>
Confidentiality:	<b>Public</b>
Keywords:	<b>Roadmap, HealthGrid, Legal, Ethical, Social, Economic, Bottlenecks &amp; Challenges, Technology, Security, Applications, Epidemiology, Innovative Medicine, Community</b>

---

**Abstract:** This document describes the first integrated road map towards realisation of the HealthGrid vision. It brings together the technical road map D3.3 and the ethical, legal and socio-economic prospects of D4.2 and seeks to reconcile the two by resolving likely conflicts between technological developments and regulatory frameworks.

The proposed road map in this document will be tested against the use case scenarios for (a) Epidemiology, and (b) Innovative Medicine in D5.1. This paper is also presented to the HealthGrid community for critical analysis and debate towards definition of the final Integrated Road Map.



Document Log

Issue	Date	Comment	Author
0.0	15 02 2007	Contributions from D3.3, D4.2 and other earlier deliverables.	I. Andoulsi; I. Blanquer; V. Breton ; A. Dobrev; C. van Dooselare; J. Herveg; N. Jacq; Y. Legré; M. Olive ; H. Rahmouni; T. Solomonides; K. Stroetmann; V. Stroetmann; P.Wilson
1.0	21 02 2007	First complete draft	M. Olive; H. Rahmouni; T. Solomonides
1.2	26.02.2007	First release	Y. Legré

Document Change Record

Issue	Item	Reason for Change

**CONTENT**

<b>1. INTRODUCTION.....</b>	<b>4</b>
1.1. RATIONALE AND PURPOSE FOR <i>ROAD MAP I</i> .....	4
1.2. TASK 6.1 INTEGRATE ROADMAP COMPONENTS .....	4
1.3. APPLICATION AREA .....	5
1.4. REFERENCES .....	5
1.5. DOCUMENT EVOLUTION PROCEDURE .....	5
1.6. TERMINOLOGY.....	5
<b>2. EXECUTIVE SUMMARY.....</b>	<b>7</b>
<b>3. COMPONENTS OF THE ROAD MAP .....</b>	<b>9</b>
3.1. THE TECHNICAL COMPONENT .....	9
3.1.1. <i>Introduction</i> .....	9
3.1.2. <i>Milestones</i> .....	9
3.1.3. <i>Grid technology issues</i> .....	10
3.1.4. <i>MT1 – Testing grid middleware(s) with medical applications at a significant scale</i> .....	10
3.1.5. <i>MT2 – Production of a reference distribution of healthgrid services</i> .....	10
3.1.6. <i>Standardisation issues</i> .....	11
3.1.7. <i>Healthgrid deployment issues</i> .....	12
3.1.8. <i>MD1 – A computational grid production environment</i> .....	12
3.1.9. <i>MD2 – A data grid production environment</i> .....	12
3.1.10. <i>MD3 – A knowledge grid production environment</i> .....	13
3.1.11. <i>Communication issues</i> .....	13
3.1.12. <i>Conclusion</i> .....	14
3.2. THE ETHICAL, LEGAL AND SOCIO-ECONOMIC COMPONENT .....	15
3.2.1. <i>Legal and Ethical issues</i> .....	15
3.2.2. <i>The ELSE Roadmap</i> .....	17
3.3. THE INTEGRATED ROAD MAP I.....	20
3.3.1. <i>Phasing ELSE Requirements into the Technical Roadmap</i> .....	20
3.3.2. <i>Cross-Domain Impacts</i> .....	21
<b>4. HISTORICAL DEVELOPMENT OF THE ROAD MAP .....</b>	<b>23</b>
<b>5. APPENDIX.....</b>	<b>25</b>

## 1. INTRODUCTION

### 1.1. RATIONALE AND PURPOSE FOR ROAD MAP I

The principal aim of the SHARE project is to establish a research and technology development road map for healthgrids. This will be the next step in the concretisation of the vision first expressed on the establishment of the healthgrid community and the incorporation of the organisation HealthGrid. Through debate in the community and at the first two HealthGrid conferences, the vision evolved into the White Paper [1] which developed a number of themes to establish both the business and the scientific and technical cases for healthgrids. The road maps to emerge from the SHARE project will

- express certain measurable goals and objectives for the HealthGrid community;
- analyse the technical gaps to be bridged in order to achieve a number of staged technical objectives;
- explore the ethical, legal and socio-economic (ELSE) conditions of such developments and analyse the extent to which technology and its environment will need to be reconciled;
- articulate a strategy for the concurrent achievement of these goals and objectives subject to realistic contextual conditions.

Within the SHARE project, work packages 3, 4 and 5 deliver three essential elements of the desired road maps:

- WP3 provides a technical description of requirements, potential approaches to their fulfilment and likely technologies to be adopted in their delivery.
- WP4 provides a review of the ethical, legal, regulatory and socio-economic environment in which it is hoped to develop 'healthgrid' solutions; this context is largely expressed in terms of constraints on what may be possible. Thus the familiar tension between science and technology on one hand and regulation and control on the other, is reflected in the outputs of these two work packages.
- WP5 will explore these issues from the point of view of two critical applications to be used as test cases for the road map.

Thus, work package 6 has the responsibility of synthesizing the outputs of the other work packages and reconciling the issues that arise under WPs 3 and 4, absorbing the lessons of WP5, leading to a road map informed by these. Its approach thus involves use of methodologies that fit in with and complement those used in WPs 3, 4 and 5. The first task of Work Package 6 is to integrate the components from the other work packages into a preliminary road map for consultation. The description of this task below is taken from the SHARE Technical Annex.

### 1.2. TASK 6.1 INTEGRATE ROADMAP COMPONENTS

This task will produce the first draft of the complete HealthGrid RTD roadmap, the ultimate deliverable of the whole project. Its input from WP3 includes definition of RTD activities such as the following:

- definition of appropriate SLAs on high bandwidth connectivity between grid nodes to support a robust infrastructure;
- extension and finalisation of current grid architectures including distributed resources sharing and the grid common operating system;
- addressing open issues in respect of mandated levels of security;

- specification of high level services to comply with legal frameworks;
- required standards extensions;
- definition and specification of possible healthgrid tools for end users specific to medical research and healthcare.

The input from WP4 is expected to include a set of RTD activities, issues and flanking measures to address:

- concerns about the legality of using such technology to process sensitive personal data;
- the organisational and social aspects of integrating the technology into the health work space;
- the economic costs and benefits of investment in the technology and its deployment.

Other activities which may be defined by WP4 include an investigation of the potential for compliance with existing legal rules which regulate the collection, storage and manipulation of personal data, subject to governance and accountability standards and an investigation of the business case for grid computing in the health sector, including a cost-benefits analysis of healthgrid.

### 1.3. APPLICATION AREA

This document will be used towards the development of the final road map by the collaboration and for consultation within the wider community of healthgrid projects and individuals who have contributed to the HealthGrid vision, the White Paper and through their research.

### 1.4. REFERENCES

[1]	I. Blanquer, V. Breton, M. Cannataro, J. Herveg, T. Solomonides et al (September 2004). HealthGrid White Paper. Available at: <a href="http://whitepaper.healthgrid.org/">http://whitepaper.healthgrid.org/</a>
[2]	N. Jacq, F. Harris, V. Breton, J. Montagnat, R. Barbera et al (September 2005). First revision of EGEE application migration progress report (DNA4.3.2), EGEE Collaboration
[3]	J. Herveg and Y. Poulet. Directive 95/46 and the use of GRID technologies in the healthcare sector: selected legal issues

### 1.5. DOCUMENT EVOLUTION PROCEDURE

This is the first stable version of the proposed road map for research and technology development in healthgrid computing. It represents the fruit of the first year's consultation and debate on how best the community can realise the HealthGrid vision. The document itself will be updated incrementally through further consultation and interaction in WP6, and through testing against the use case scenarios of WP5..

Comments should be sent to Tony Solomonides at [tony.solomonides@uwe.ac.uk](mailto:tony.solomonides@uwe.ac.uk).

### 1.6. TERMINOLOGY

#### Glossary

SHARE	Supporting and structuring Healthgrid Activities and Research in Europe
WP1	SHARE project Management Activity
EU/EC	European Union/Commission

Roadmapping	An extended look at the future of a chosen field of inquiry, leading to an outline or map of how and by what means to achieve certain goals.
Ontology	The systematic description of a given phenomenon, which often includes a controlled vocabulary. Typically, ontology defines data entities, data attributes, relations and possible functions and operations.
Metadata	Metadata summarize data content, context, structure, inter-relationships, and provenance (information on history and origins). They add relevance and purpose to data.
Web Service	A software system designed to allow inter-computer interaction over a network to perform a task. Other computers interact with a web service, in a manner prescribed by the interface, using messages.
Middleware	A software stack composed of security, resource management, data access, accounting, and other services required for applications, users, and resource providers to operate effectively in a Grid environment.
Virtual Organisation	A group of grid users with similar interests and requirements working collaboratively and/or sharing resources regardless of geographical location.
Heterogeneous Data	Data potentially from multiple sources, and in different formats.
Data Mining	Automatically searching large volumes of data for patterns or associations.
Authentication	Verifying and confirming the identity of a grid user.
Authorisation	Restricting access to resources based on what a user has been granted access to.
Data Subject	The Person to whom the data relate.
Data Controller	The natural or legal person (i.e. a company) determines the purposes and means of processing the data.

## 2. EXECUTIVE SUMMARY

The principal aim of the SHARE project is to establish a research and technology development road map for healthgrids. This will be the next step in the concretisation of the vision first expressed on the establishment of the healthgrid community and the incorporation of the organisation HealthGrid. Through debate in the community and at the first two HealthGrid conferences, the vision evolved into the White Paper [1] which developed a number of themes to establish both the business and the scientific and technical cases for healthgrids. The road maps to emerge from the SHARE project will

- express certain measurable goals and objectives for the HealthGrid community;
- analyse the technical gaps to be bridged in order to achieve a number of staged technical objectives;
- explore the ethical, legal and socio-economic (ELSE) conditions of such developments and analyse the extent to which technology and its environment will need to be reconciled;
- articulate a strategy for the concurrent achievement of these goals and objectives subject to realistic contextual conditions.

The proposed road map has been developed from two major inputs:

- (a) a technical road map which sets out the key objectives for a viable 'knowledge healthgrid' to be achieved in a span of 10-15 years.
- (b) a conceptual map of ELSE conditions, constraints and requirements which must be addressed before a knowledge healthgrid can be deployed in a real healthcare setting.

The technical road map analyses its objectives into milestones according to a number of key criteria:

- is the proposed healthgrid aimed at supporting a research environment or a real healthcare delivery application?
- is the proposed healthgrid essentially a computational grid, a data grid or a knowledge grid?
- is the necessary development to achieve any given stage likely to be delivered by generic grid research or is it particular to healthgrid?
- is some prerequisite standard or other agreed framework necessary in the achievement of any particular milestone?

The conceptual map of ethical, legal and socio-economic issues considers the regulatory challenges that any real healthgrid must meet:

- legal challenges concerning rights to privacy and confidentiality, 'right to know' and duty of care;
- ethical challenges concerning primary and secondary use of data whether individual or aggregated;
- legal and ethical challenges concerning provenance and quality of information;
- legal, ethical and economic challenges to the use of healthcare data in commercial and public research, including questions of ownership of data;
- legal and ethical challenges in the communication of genetic information and the resultant 'lateral leakage' of information;



## D6.1 INTEGRATED ROAD MAP I

*Doc. Identifier:*  
**SHARE-D6.1-v1.2.doc**

*Date:* 24/02/2007

---

- legal and ethical challenges to the communication of medical data across borders;
- social and legal challenges concerning the formal professional competencies of different healthcare actors;
- legal, ethical and socio-economic challenges of 'exceptional cases', such as assisted reproduction, organ donation and transplantation.

The proposed integrated road map below represents the first attempt to bring all these concerns together into one strategic plan.

### 3. COMPONENTS OF THE ROAD MAP

#### 3.1. THE TECHNICAL COMPONENT

##### 3.1.1. Introduction

The first SHARE technical roadmap (D3.3), informed by the issues and successes of previous healthgrid projects, has defined a series of milestones for healthgrid research in order for the successful widespread deployment of grids for medical research. When combined with numerous ethical, legal and socio-economic (ELSE) issues identified by the first ELSE roadmap (D4.2), this will form an integrated roadmap for discussion with the rest of the healthgrid community to help coordinate, prioritise, and set targets for research efforts.

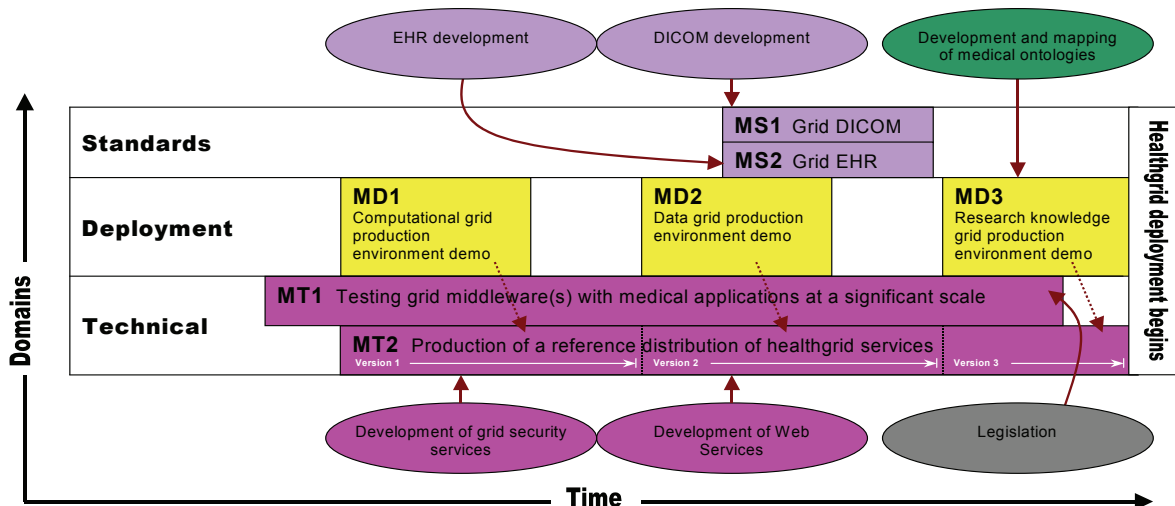
Before addressing the ELSE considerations, our initial step was to examine the technical roadmap in the context of previous research, incorporating findings from the use case scenarios defined by workpackage 5. This has led to some minor alterations to the technical roadmap regarding the concurrency of milestones in different domains, the distinction between research occurring within the healthgrid community and research from other fields, and communication with external research efforts and other influences required in order to reach particular milestones.

##### 3.1.2. Milestones

Five milestones were defined in D3.3; three deployment milestones for the creation of a demonstration production environment for a computational grid, a data grid and a knowledge grid for medical research, and two milestones dealing with specific standardisation efforts for medical imaging and electronic health records for grids. These remain in our version of the roadmap, although we have introduced some overlap to represent concurrency and a cyclical relationship between the standardisation milestones and the second deployment milestone of a data grid production environment.

Additionally, two ongoing research efforts were discussed by D3.3; the development of a tested, robust and scalable grid middleware based on Web Services that allows job and data management, and the availability of a robust, user friendly open source distribution of grid middleware with appropriate user support. When discussing grid technology issues, these two key requirements were again mentioned and it became clear to us that these should be represented as milestones in their own right.

The following diagram shows the milestones defined, their ordering and relationship with external research efforts.



The reasons for the existence and ordering of these various milestones can be explained by summarising the technology, standardisation, deployment and communication issues that will drive healthgrid research. An historical overview of the development of the technical road map has been set out in section 5 below.

### 3.1.3. Grid technology issues

There are a number of important development challenges remaining when evaluating grid technology for medical research. Of the many middleware distributions available, such as gLite, Unicore, GRIA, etc. none currently fulfil the requirements for a healthgrid. Several have not successfully demonstrated their scalability and robustness, and even those that have are limited by insufficient data management capabilities or job management services.

Perhaps one of the most important technology issues is that many are not entirely based on web services and therefore do not have standard interfaces. Progress is certainly being made in this area though, with GT4 and GRIA both providing services based on the Open Grid Services Architecture (OGSA).

Using existing web service standards, OGSA defines a grid service as a web service that conforms to a particular set of conventions. Grid services are defined in terms of standard Web Services Definition Language (WSDL) with minor extensions. This provides a common and open standards-based mechanism to access various grid services using existing standards, such as SOAP, XML, and WS-Security. This also allows additional services to be added and integrated easily, and OGSA will facilitate interoperability between grids that might have been built using different underlying tools.

However, scalability, particularly regarding medical applications, is still a concern. Middleware such as gLite and Unicore on the other hand have been deployed on large scale infrastructures in Europe and have demonstrated their scalability and robustness with medical applications, but are still awaiting migration to web services.

Particularly when it comes to the installation of grid nodes in medical research centres, a major barrier will be the complexity of the software, its management and installation. Reviews of previous large scale projects such as WISDOM have demonstrated that considerable knowledge about internal grid mechanisms are required in order to quickly resolve problems, and error messages are often insufficient<sup>[2]</sup>.

As a result of these issues, two technical milestones have been defined.

#### 3.1.4. MT1 – Testing grid middleware(s) with medical applications at a significant scale

Testing of grid middlewares for scalability and robustness should begin at an early stage, as any deficiencies in this area will only hamper deployment and the development of a reference distribution. It is anticipated that this will be an ongoing activity, with different generations of grid operating systems offering newer, faster and more stable capabilities. An important aspect will be compliance with EC country laws regarding reliability and the necessary mechanisms for recovery and backup, as shown on the diagram by the link with legislation.

#### 3.1.5. MT2 – Production of a reference distribution of healthgrid services

Shortly after testing has begun, work should commence on the production of a reference distribution. Standards for web services and particularly grid security services are still evolving, and therefore new versions of the distribution should be made available as standards are adopted and implemented. Regular updates will be required, but critically backwards compatibility must be retained allowing sites running previous versions to coexist in the same virtual organisation.

The reference distribution should be fast and simple to install on a grid node in a medical research centre, with management and configuration functions suitable for the target users. The users and administrators should not require in depth knowledge of grid mechanisms, and all functions should be made available through user-friendly interfaces.

From a stability and quality point of view, having minor and major versions would be a sensible approach. There will be considerable interaction between the work for MT2 and the work for the deployment milestones, as indicated by the dotted (internal) arrows shown. Although updates will be regular, critically version 1 should be in a stable, finalised state before the commencement of MD2, and likewise version 2 should be frozen before the start of MD3. This should ensure MD2 and MD3 are using fully tested and stable distributions from the outset.

### 3.1.6. Standardisation issues

The use of computer-based tools for clinical research has led to the definition of standards for the exchange of data in many areas. However, such standards are in many cases not universal, with different disciplines and countries adopting different standards.

The exchange of data between bioinformatics and medical informatics is an area where standards are particularly limited. Such standards will be required in order to build integrated data models and produce ontologies, both requirements of knowledge grids.

Medical imaging is an exemplary case, in which the adoption of Digital Imaging and Communications in Medicine (DICOM) for the acquisition, connection and storage of medical images has been accepted worldwide. However DICOM has no standardisation of storage, and DICOM servers are implementing their own data storage policies. To fill this gap the gLite middleware added a standard interface SRM to enable the storage element to interact with other grid services. Another standard interface SRM-DICOM was used to transform SRM requests to DICOM transactions. This has allowed medical data servers to be shared between clinicians (using the classic DICOM interface inside hospitals) and image analysis scientists (using the SRM-DICOM interface to access the same database) without interfering with clinical practice.

Medical records are another area where standardisation would have clear benefits, with HL7 being the favoured standard for the exchange of data. However, previous standards such as CEN/TC251 EN13606 focused more on the storage and structuring of clinical records and have prevented a wider uptake of HL7. The adoption of both DICOM and HL7 has increased due to initiatives such as IHE (Integrating the Healthcare Enterprise), which promote the coordinated use of DICOM and HL7 by publishing best practice guidelines.

A particularly important consideration for both of these standards is their compatibility with grid technologies, and how they could be implemented on a healthgrid. Both DICOM and HL7 are just starting to study the interface between their standards and web services technology.

Two standardisation milestones have therefore been created as examples: **MS1 – Grid DICOM** and **MS2 – Grid EHR**, although clearly standards in other areas of bioinformatics and medical informatics with respect to grid technologies will be required. It is important to recognise that EHR and DICOM standards are unlikely to be modified themselves in order to satisfy the requirements of grids, rather the healthgrid community must liaise with those developing these standards in order to define any extensions that may be needed.

As shown on the diagram, MS1 and MS2 overlap with MD2. The reason for this will be explained in the following section.

### 3.1.7. Healthgrid deployment issues

The deployment of grids in medical research centres has to date been very limited for a number of reasons. Cost, both in terms of infrastructure and manpower, is a significant barrier.

The legal and organisational barriers to accessing resources outside of a hospital must also be examined. Convincing healthcare management of the benefits of deploying a computational grid on a hospital or clinic IT infrastructure, which would not be composed of dedicated grid nodes and may already be working near capacity, may not be a simple task. Many medical centres may simply not have the necessary bandwidth or storage capabilities to make best use of grid technology, or may not have appropriate equipment to capture data in digital form.

Security will be an issue at all deployment phases. The grid infrastructure needs to provide secure mechanisms for access, authentication, and authorisation, as well as facilities for secure data storage and transfer. Critically, these must adhere to the legal requirements for the particular type of data concerned, with regulations differing from country to country in Europe. The grid operating system needs to provide access control to individual files stored on the grid, fine-grained access control for services and data encryption mechanisms. Mechanisms must also be in place to facilitate temporary, time-limited access to data and the revocation of credentials. The authentication of health professionals must be achieved in a practical manner, perhaps through the use of professional cards for authentication, but this should not prohibit single sign-on for grid services.

Security requirements are also likely to escalate, with remote queries, the use of patient records, and data mining of records occurring in MD2 and MD3. This is one of the reasons for the overlap between the technical and deployment milestones; as new security requirements are identified, mechanisms must be implemented, then tested and incorporated into the healthgrid distribution.

### 3.1.8. MD1 – A computational grid production environment

There have been successful deployments of computational grid medical applications such as WISDOM on general purpose grid infrastructures such as EGEE and DEISA, and therefore this initial deployment step would at first seem to be achievable with relatively little effort once stable and robust healthgrid middleware distributions are available. However, as mentioned previously there are a number of issues associated with the installation of grid nodes in hospitals and medical research centres, including economic factors, hiding the complexity of grid mechanisms from users, and security concerns. Participating institutions will require a significant amount of training and investment in manpower and new hardware to support grid computing. Adequate documentation and guidance must be available, and where grid infrastructures are distributed amongst geographically remote sites there must be sufficient communication mechanisms such as video conferencing to ensure problems and concerns raised during deployment are dealt with quickly and efficiently.

### 3.1.9. MD2 – A data grid production environment

Although several prototype data grids for medical research have been demonstrated by healthgrid projects, developing and maintaining a production quality data grid will require a number of issues relating to the distributed storage of medical data to be resolved. In European grid infrastructures, the distributed storage of medical images has been hampered by the limited data management services available, and so the continuation of improvements in this area will be important for the adoption of grids by the medical community. High speed links between data providers and consumers will be a prerequisite, particularly given the high volume of data predicted.

It's important to note the relationship between MD2 and the standards milestones. Data grids are likely to make use of both medical images and patient records, and as a result will not only make use of

standards for these but will also define requirements and make suggestions as research in this area progresses. The candidate standards may then be modified as a result, these are again tested in practice, and so on in a cyclical manner. As MD2 finishes, the final feedback on standards will be provided to MS1 and MS2 in order for these to be completed before the start of MD3. This kind of close collaboration between separate but related areas of research is an important feature of the roadmap, and will be essential for the successful completion of each milestone.

Another important concern for MD2 will be the integration of heterogeneous data from multiple sources. Whilst mechanisms for data integration have been demonstrated by previous projects, biomedical data can be exceptionally varied including images with associated metadata and free form text or hand written notes from patient records. There is also the issue of how to deal with missing, inaccurate or obsolete data.

### 3.1.10. MD3 – A knowledge grid production environment

After the issues with the distributed storage and querying of medical data have been resolved, the next task will be to deploy services that can build relationships between data items, and will provide appropriate representation to medical researchers. Particularly given that there have been no successful deployments of knowledge grids for medical research to date, MD3 will pose a significant challenge. The data concerned can be extremely varied in nature, structure, format and volume. Depending on the area of research, the synthesis of knowledge from data could require sophisticated data mining, integrated disease modelling and medical image processing applications, and may also involve the use of techniques from Artificial Intelligence (AI) to derive relationships between data from different sources and in different contexts.

The development of medical ontologies and the mapping between ontologies will be particularly important for the successful deployment of knowledge grids. These ontologies will allow relationships between concepts and nuances in meaning to be captured, greatly enhancing the opportunities for communication, knowledge sharing and reuse, and machine reasoning.

Most biomedical applications currently using ontologies deal primarily with decision support, such as assisting health professionals in diagnosis. For example, breast cancer diagnosis and treatment is served by a Breast Cancer Imaging Ontology. The ontology for the GALEN model, which aims at developing advanced terminology systems for clinical information systems, is designed to be re-usable and application independent. It is intended to serve not only for the classification of surgical procedures but also for a wide variety of other applications including decision support systems and natural language processing. In most cases, biomedical ontologies function as terminology vocabularies, containing the domain knowledge required to build the classes, rules and relationships according to which the several concepts interact with each other.

The development and mapping of medical ontologies has been represented as an external entity in the roadmap diagram for two reasons. As mentioned by D3.2, the development of medical ontologies is outside the scope of healthgrid research and will be a task for the medical research community, although communication with those involved with the development of these ontologies can only be beneficial. The other reason is simply that an ontology is only as valuable as the general support it has, as its primary purpose is to facilitate a common understanding of terms. As an ontology is adopted more widely, this increases the possibilities for a resource that supports it.

### 3.1.11. Communication issues

Communication is a very important issue for healthgrid research, both within the research community and between different interested groups. Communication with the biomedical community is vital, both to increase awareness of the potential of grid technology and to clarify requirements. From an



awareness point of view, the dissemination of success stories highlighting the usefulness of grids for medical research will be important for encouraging future adoption and investment.

Better communication and dissemination of ideas between healthgrid projects is essential. This is not only required in order to avoid the replication of work, but also to identify the best techniques to achieve specific aims, to identify key points requiring wider attention, and to propose standards. As the overlap between milestones on the diagram shows, better communication is required not only between projects working in the same area but also on areas that influence or rely on each other. Requirements and findings from deployment or standards for example must be known by those developing and testing middleware as soon as possible.

As mentioned before, liaising with groups developing biomedical and medical informatics standards will clearly be required, as will closer ties with grid technology providers and developers.

### **3.1.12. Conclusion**

We have proposed a number of minor alterations to the first technical roadmap defined in D3.3 prior to integration with ethical, legal and socio-economic considerations. These changes are primarily to do with concurrency and communication between milestones, communication with research and developments outside of the healthgrid community, and clarifying the technical tasks of testing middleware and producing a reference distribution of healthgrid services.

We believe these alterations better represent current and anticipated research into the use of grids for medical research, and highlight the importance of communication both within the community and with outside influences.

## 3.2. THE ETHICAL, LEGAL AND SOCIO-ECONOMIC COMPONENT

### 3.2.1. Legal and Ethical issues

The effort should not only be directed at technical and deployment issues related to healthgrid. Other challenges are also arising which are linked to the legitimacy and ethic of using such a technology and its social impact on the health work space.

Collection, storage and manipulation of personal data within the grid environment need to comply with current rules and legislation. The actual European rules and directives concentrate mainly on the reservation of rights to the data subject, i.e. the person to whom the data belongs. In the case of a healthgrid, the data subject is mainly the patient. The rights range from being informed about the data processing, the ability to express or withdraw his or her consent at any time of the process, the right to access and rectify the data, the right to be safe and not suffer any harm caused by the use of healthgrid services and products.

Addressing these issues should start at a very early stage of the healthgrid development process. Failure to reach a reasonable solution at the right time could result in freezing some of the deployment milestones suggested by the technical roadmap. Therefore a good structuring and planning to the actions that need to be done is necessary.

#### 3.2.1.1. Patient Data Processing

In the healthcare sector, the use of the grid technology implies automated processing of health personal data, especially for therapeutic purposes.

*“In this context, end-users will collect and dispatch the patient’s personal data to healthgrid services providers. These providers will process the patient’s personal data in order to complete the therapeutic [...] on behalf of the end-users, before sending them back or making them available to the end-users if needed”.*[3]

The key European principles relevant to the processing of personal data were first established by the Council of Europe and further developed in Directive 95/46/CE of the European Union<sup>1</sup>. The purpose of the Directive is to allow the free flow of personal data between the Member States of the European Union, to facilitate the establishment and functioning of the internal market and to protect the fundamental rights and freedoms of natural persons, and in particular their right to privacy with respect to the processing of their personal data.

The Directive thus banishes the processing of sensitive or medical data. However, the principle of this ban is not absolute; some texts in the Directives tend to grant permissions to the processing of personal data under exceptional conditions. For example, the explicit and valid consent of the data subject constitutes the most important source of legitimacy in the processing of medical data although it is subject to very strict conditions for its validity and the data subject may revoke this consent to the processing of his or her medical data at any time and without justification.

Processing is also allowed if

- it is necessary for the purposes of carrying out the obligations and specific rights of the controller in the field of employment law in so far as it is authorized by national law providing for adequate safeguards;

---

<sup>1</sup> For references to statutes, directives and other legal references, see SHARE Deliverable 4.2 at [6].

- it is necessary to protect the vital interests of the data subject or of another person when the data subject is physically or legally incapable of giving his or her consent;
- it is carried out in the course of its legitimate activities with appropriate guarantees by a foundation, association or any other non-profit-seeking body with a political, philosophical, religious or trade-union aim and on condition that the processing relates solely to the members of the body or to persons who have regular contact with it in connection with its purposes and that the data are not disclosed to a third party without the consent of the data subjects;
- it relates to data which are manifestly made public by the data subject or is necessary for the establishment, exercise or defence of legal claims;
- it is required for the purposes of preventive medicine, medical diagnosis, the provision of care or treatment or the management of health-care services, and when those data are processed by a health professional subject to the obligation of professional secrecy or by another person also subject to an equivalent obligation of secrecy.

Finally, the European Directive offers the opportunity to the Member States to add exemptions to those listed above, for reasons of substantial public interest. These exemptions should be subject to suitable safeguards. For instance, under Article 8, 4 of the Directive, national exemptions might be adopted for scientific research. For example exemptions could be added to the processing of medical data as long as the following conditions are verified by the data controller.

- The data processing must be legitimate. In other word, the data processing corresponds to one of the social justifications laid down by the European Directive in its Article 7. For example, the data was processed after an unambiguous consent was expressed by the data subject.
- The data controller must ensure the good quality of the data to be processed. For example, the data must be and remain accurate and if necessary up to date.
- The data controller must also respect the rights of the data subject and ensure that these are observed at all stages of the processing.
- The data controller must ensure that security and confidentiality requirements are met.
- The controller is required, prior to carrying out the processing, to provide the relevant national supervisory authority with certain items of information regarding the planned processing. The information recorded will then normally be accessible to data subjects or to third parties.

### **3.2.1.2. Personal Data transfer**

For EU Member States, the transfer of personal data between two or several controllers established on the territory of one member states or on the territories of several member states involves the problem of communication of personal data to third parties. Indeed, the transfer or disclosure of personal data to third parties is considered to be a processing operation and, as such, is subject to the processing legal requirements discussed in the previous section. However, the transfer of medical data is subject to more specific requirements. Medical data may not be communicated unless the conditions listed below are fulfilled:

- the medical data to be communicated are relevant for the communication purpose;

- the recipient of the communication is subject to confidentiality rules equivalent to those incumbent on healthcare professionals, and the communication is legally authorised and is realised for public health reasons or for another important public interest.

The national legislations of the different Member States of the European Union are by and large harmonised by now, and the transfers of personal data between these member states should not create any problem. However, the controller must refrain from transferring personal data to a recipient located in non-EEA countries, if the country involved does not ensure an adequate level of protection, unless the controller adduces adequate safeguards as regards the protection of the privacy, fundamental rights and freedoms of individuals and the exercise of the corresponding rights.

### **3.2.1.3. Liability Issues**

This section addresses questions of liability that may impact healthgrid participants. In particular, developers need to be aware of these so as to prevent any damage to patient health, rights and freedoms. Healthgrids are complex systems. They involve different actors such as doctors, specialists, hospitals, pharmaceutical companies, data controllers and processors, technicians, etc., located in different countries. Due to this complexity, the establishment of the person to be held responsible for a specific damage can be problematic.

The products and services offered by a healthgrid could turn out to be a cause of death of the patient or of delay in diagnosis and treatment. This may, for example, be due to the lack of regular testing and monitoring of products and services. The question here is who would be liable for such damage? With respect to healthgrid products (including medicinal products, pharmaceutical products or any other kind of products used in relation to a patient's health), there is currently no specific legislation applicable to the liability deriving from the delivery and the use of those products. However, one might apply the general principles relative to consumers' protection.

The basic principle has long been established that if a product does not conform to the offer made or causes damage, the consumer (or another person representing him or her) may claim for compensation. Any liability issue will thus normally depend on the general rules of law applicable in the different EU member states. Other special European regulations might as well be applicable in case of damages caused to consumers by defective or not corresponding products. Amongst these special legislations we can distinguish those which deal with prevention and those which allow to repair damages or to have an appeal when the delivered product is not in accordance with what was foreseen in the contract.

The situation is not also simple from the services perspective. Services supplied through healthgrids might cause damages to those who depend on it. These services could either be services that constitute the grid system or services provided by the internet within the grid domain, including health related websites.

Nowadays, there is for instance a lack of data and knowledge management services. A citizen might thus be seriously harmed or even die, if the information transmitted to the general practitioner treating him or her is not accurate or false, or if it is not supplied on time. As long as the related service is part of the grid infrastructure, Directive 2000/31 (on certain legal aspects of information society services, in particular electronic commerce, in the internal market, also called the 'Directive on Electronic Commerce' might apply).

### **3.2.2. The ELSE Roadmap**

Addressing the issues listed above is a challenge as the advice of medical and legal bodies is crucial. A well-planned and structured set of actions and steps is needed to make the processes easier. The ELSE roadmap could be the answer but it won't be enough unless it is harmonised with the technical roadmap milestones.

In this section we discuss the different milestones we suggested to constitute the ELSE roadmap.

### **3.2.2.1. Roadmap Milestones**

#### ***Ethico-Legal Milestones (MEL1, MEL2)***

MEL1 The primary concerns for MEL1 will be liability and determining what the responsibilities of the healthgrid actors are. Possible damage that could happen to the patient could be outlined along with some preventative measures. Logging and auditing must be addressed early to monitor whether enough testing was done to healthgrid services and products.

MEL2 Patient consent is crucial to the legitimacy of medical data processing and transfer, therefore verifying that the patient has unambiguously expressed his/her consent should take place prior to any data manipulation. A technical way of sorting out this problem could be recommended like for example adding a flag or metadata to the patient record to indicate whether he has any objection to the processing of his/her personal data.

Appropriate and user friendly ways of allowing patient access to data is also recommended. This will help patient not to feel totally dispossessed of data and information that concerns them and excluded from the data processing process. Thus more public trust will be added to research carried out within the healthgrid domain.

#### ***Data Protection Milestones***

MDP1 This milestone is concerned with patient privacy and how it could be best protected within the HealthGrid environment. Patient identification issues should be discussed and good analysis and evaluation of the current de-identification software and tools should be produced. We suggest a start with medical images de-identification as it might contain recognisable parts of the patient body. This action should start at an early stage so the deployment milestones could benefit from any recommendations.

MDP2 At this stage researchers need to make sure robust anonymisation, pseudo-anonymisation and other identity protection techniques are developed and deployed in the grid infrastructure. The eDiamond project suggests that a semantic understanding of the reasons why a person may be accessing particular pieces of data is crucial to the legitimacy of data processing in the HealthGrid environment.

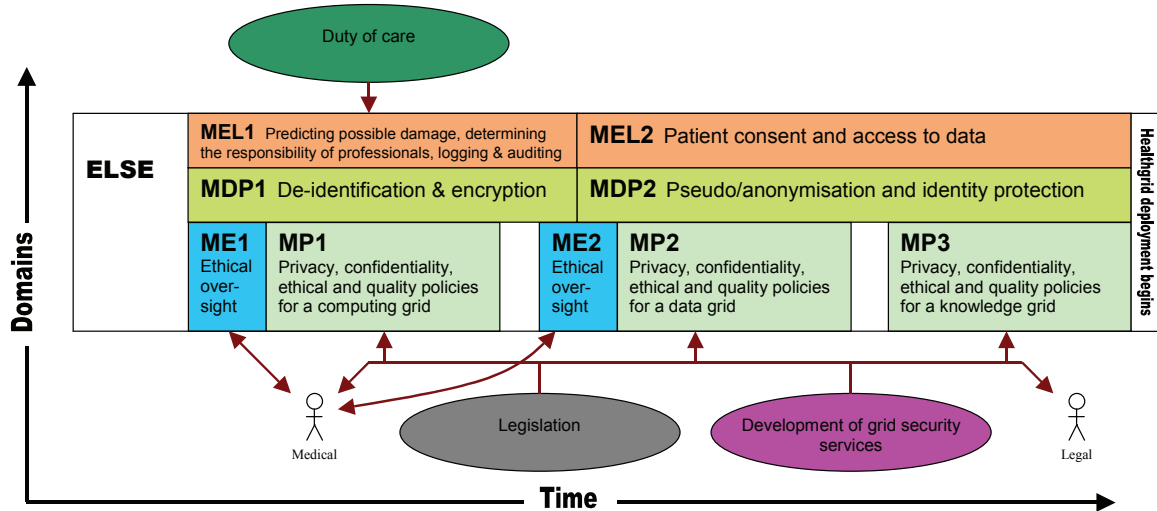
#### ***Ethical Control Milestones (ME1, ME2)***

Before MD1 and MD2 should start, the requirement for ethical oversights and monitoring should be determined. ME1 will focus on the requirements and tools to facilitate oversight, with automation being explored. ME2 will satisfy the arrangements for automated ethical control for a data grid which will be more complex with long-term data storage.

#### ***Policy Milestones (MP1, MP2, MP3)***

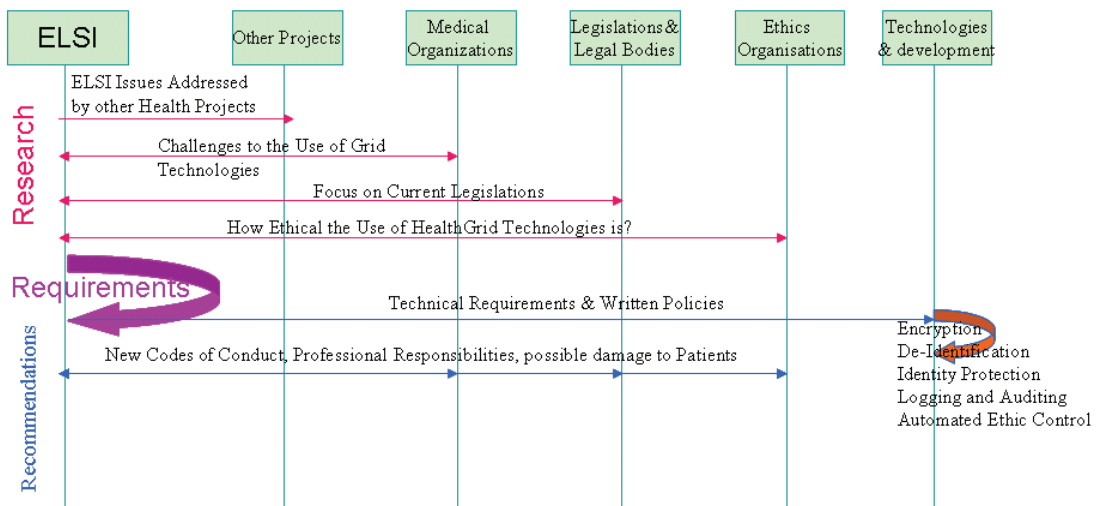
These will cover data processing and transfer issues such as legitimacy, accessing the minimum data required, the ethical transfer of data, compliance with confidentiality rules and limiting the period of data storage.

We suggest three policy milestones. Each one should start some time before the corresponding deployment milestone as requirements are expected to differ when changing from a computational grid to a data grid or to a knowledge grid.



ELSE Milestones Diagram

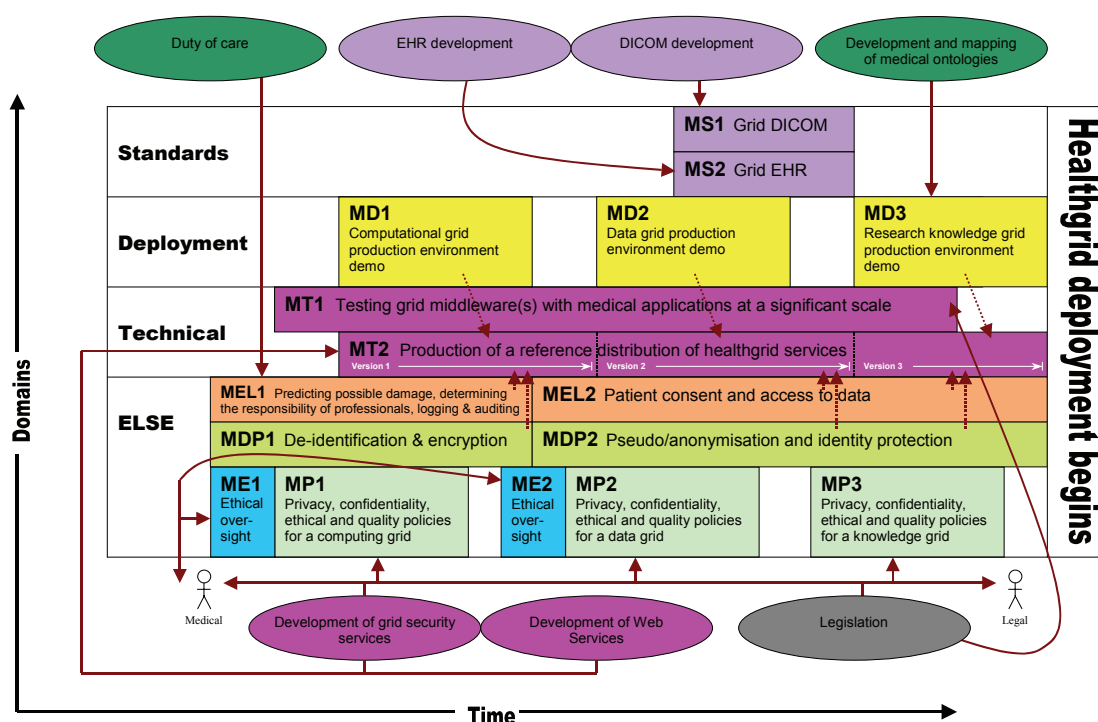
The necessary interactions in addressing various ethical, legal and socio-economic issues in relation to healthgrid are set out in an interaction diagram below. In the Appendix, we display an equivalent interaction diagram for each milestone.



ELSE Roadmap Actions Diagram

### 3.3. THE INTEGRATED ROAD MAP I

We now have all the necessary elements to compose our integrated map. The main task is to describe and justify the interactions between the technical road map and its milestones with the proposed actions in the ethical, legal and socio-economic (ELSE) road map. The integrated map may be read as a series of extended actions from left to right. Technical development is shown as its middle track, with deployment and ELSE tracks on either side. Standards that straddle healthgrid and medical informatics in general make up another track. Influences that may rightly be considered external are shown on either side as ovals. From these to the tracks and from outer tracks inwards are lines of influence. These depict the impacts on technology development of every other aspect in the realisation of healthgrids.



#### 3.3.1. Phasing ELSE Requirements into the Technical Roadmap

We now turn to a more practical aspect of our analysis and note that all the suggested milestones need to go through different phases of development in order to be achieved as specified in the ELSE road-map diagram. The different phases of development are research, requirements and recommendations.

##### Research Phase

During this phase an analysis about how other projects addressed similar issues to the ones with which we are concerned needs to be done, as well as how efficient their solutions were. Discussions with legal, ethical and medical bodies could help with the prioritisation of issues to be addressed. Understanding real-world medical requirements through scenarios that clinicians could provide can help both in understanding the issues and resolving potential conflicts before they occur. These stories could also inspire healthgrid researchers to develop future stories and scenarios that could occur when healthgrid services are put in practice.

### ***Requirements Phase***

This phase may start by carefully mapping typical use case scenarios for each issue and analysing them in order to specify the requirements needed in order to solve the problem. The requirements are then prioritised and classified into technical (in case we think the solution could be offered by an automated system) and non technical requirements. Some ELSE requirements may be ‘technical’ in the sense that there exists potential for at least the majority of cases to be handled on an automatic basis.

### ***Recommendations Phase***

The outcomes from the previous phases could be used to produce some recommendations to address the various ethical, legal and social issues. Technical recommendations could be made to developers responsible for building the healthgrid infrastructure, the development of tools and technologies to provide de-identification, encryption, anonymisation, pseudonymisation, automated means of ethical control and the management of patient consent.

Legislation, such as privacy and confidentiality through data protection, could be mapped in order to be implemented within the grid infrastructure. Other issues might have to be sorted out without relying on automated systems such as liability issues. In this case we might recommend that a new code of conduct needs to be published outlining the responsibility of each healthgrid professional along with risk analyses of possible damaging impacts on patients.

Recommendations for new legislation could also be made especially for issues such as liability. We have seen in previous sections that some EU legislation could have an impact on data protection issues within the healthgrid context. In particular, it appears to exacerbate liability issues. It may be thought that the most appropriate way to address these issues is to treat healthgrid products and services like any other product in the market and to apply equivalent rules. There is a difference both in perception, when resources and data are geographically distributed and shared across the member states, and in the reality of interprofessional responsibilities. These are matters that will require extended negotiation between different healthgrid participants and relevant authorities, not least to decide on the importance and benefit of seeking new legislation to address healthgrid related legal and ethical issues. Once the problem has been recognised and a decision to support automated transactions has been made, a framework will need to be developed perhaps similar to those devised to support e-commerce. It would be necessary to present scenarios showing when current legislation does not ideally cover these issues.

### **3.3.2. Cross-Domain Impacts**

With the golden goal of free deployment of healthgrids, the project has identified three distinct milestones as deployment of (a) a computational grid, (b) a data grid, and (c) a knowledge grid. In order to fix ideas somewhat, we consider a computational grid to be used mainly in diagnostic or therapy planning activities, thus dealing with only one patient at any given time. This may encompass services such as simulation for surgical reconstruction, computer-aided detection (CADe) in radiology or radiotherapy plans. A data grid is taken to mean the inclusion of an electronic patient or health record (EPR/EHR), the latter being understood to be some sort of summary of the former. This may be used both in individual and public health interventions, for example in seeking case-based evidence or analysing aggregate phenomena. Genomic medicine, even in cases in which only consultation of large data banks would be required, may be included here. The final goal of a knowledge grid is significantly more ambitious, encompassing the incorporation and provision of at least some medical knowledge among the services offered. This may include second opinion based on ‘similar’ cases, thus dynamically growing the evidence base, or checking that some combination of treatments the patient is to undergo is not problematic.

Beginning with the lower tracks in the road map, we anticipate that external developments, i.e. developments outside the healthgrid domain, in grid security and web services would be impacted to some extent by the requirements of the medical and legal domains. This feeds into one of the major tracks of activity, the production of a reference distribution of healthgrid services.

Medical and legal requirements are also expressed through policies in the next track up, where also, anticipating the first two major milestones in deployment, the requirement for direct ethical oversight and monitoring is included explicitly. By the time of the third deployment milestone, that of a knowledge grid, it is proposed that oversight and monitoring be included as a policy requirement which is expressed internally in the system and automated at least to some extent.

The policies for privacy, confidentiality, ethics and quality are interpreted in more directly applicable terms as ‘data protection’ actions, with the realisation of de-identification and encryption policies as the first milestone in time for the deployment of a computational grid. Correspondingly, in time for data grids, we require at least robust pseudo/anonymisation or more sophisticated identity protection methods.

On the basis of current understanding, for the purposes of a computational grid, ethical concerns would be addressed by an appropriate interpretation of ‘duty of care’ and anticipating any possible harm to the patient, ensuring that interventions are by appropriately authorised individuals with due diligence as to their professional standing and role in the process, and maintaining a precise and complete record of what is undertaken. In the deployment of a data grid and beyond, we must ensure that patient data is manipulated according to explicitly given permissions, subject to the considerations outlined above (and in greater detail in Deliverable 4.2).

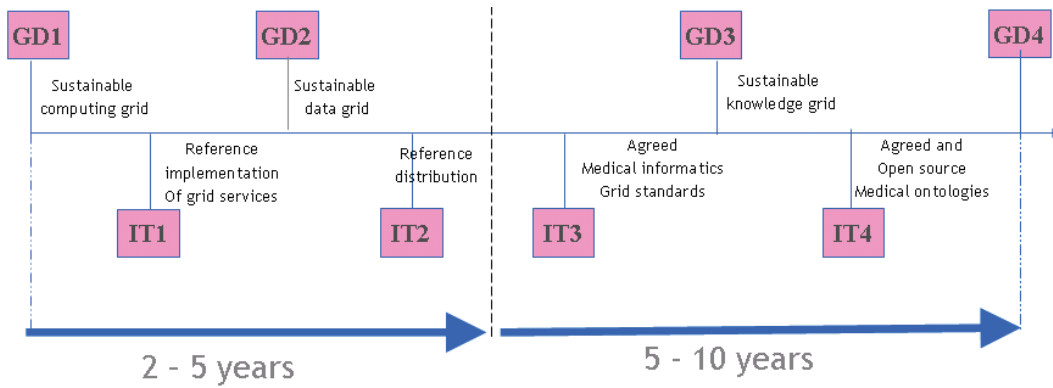
Both legislation and long-established standard practice in engineering require extensive validation, verification and testing of systems, especially those deployed in safety-critical applications. Validation and verification are sometimes distinguished by the questions they are said to ask, “*Are we building the right product?*” – is the specification correct? – contrasted to, “*Are we building the product right?*” – i.e. assuming we have the correct specification, does what we are building satisfy its requirements? Testing is the traditional engineering approach to verification, in which as many representative cases as possible are explicitly tried out in the real system. This activity deserves its own track running between the development and the deployment tracks.

Finally, we observe that significant developments from within and outside the healthgrid community are anticipated in standards for electronic patient/health records, for images as well as in ontologies. Grid interpretations or adaptations of the standards will be necessary, especially in domains such as radiology where proprietary extension of standards by equipment manufacturers can make a common approach to, say, annotation of images very difficult. In relation to ontologies, we note the multiplicity, almost cacophony, of ontologies currently in development or use and anticipate that some standardisation will also be necessary here before it is safe to approach medical knowledge in an ontology-driven manner.

#### 4. HISTORICAL DEVELOPMENT OF THE ROAD MAP

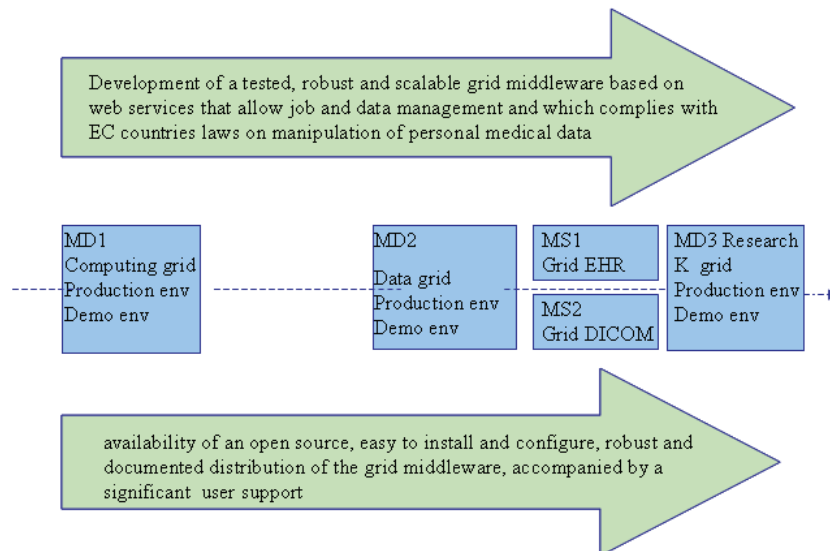
It is undeniable that the road map presented above is rather complex and may therefore have less visual impact than a simpler representation. We present here two historical precedents to help explain where it comes from. Both are technical road maps with minimal acknowledgement of ELSE issues.

The first technical road map was presented at IST2006 in Helsinki and at e-Science 2006 in Amsterdam. It cleanly divides up the milestones into ‘technology’ and ‘deployment’, moves from research to healthcare, from a single disease to generic solutions and has what was, by common consent in discussion, an optimistic time scale. (One respected authority commented that a ‘0’ may be safely added after each projected number of years!)



**Technical Roadmap Version 1**

We display the second road map in juxtaposition for comparison:

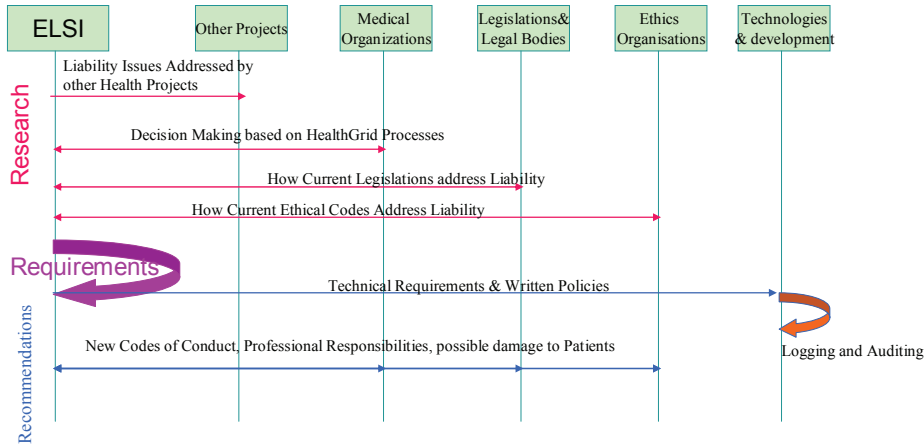


**Technical Roadmap version2**

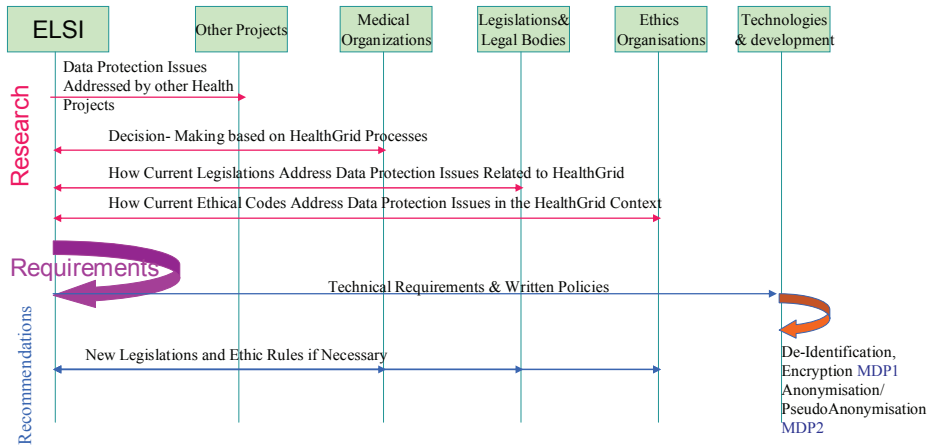


Despite the simplicity of the first road map, this has the greater virtue of separating concurrent ‘grid world’ developments in the form of two parallel arrows outside the main healthgrid strand. However, it has the potential to mislead, inasmuch as it appears to make the transition from a data grid to a knowledge grid merely a matter of standards, and especially with no mention of ontologies. None the less, this was precisely our starting point in the integrative attempt to bring the technical road map and the ELSE conceptual map together.

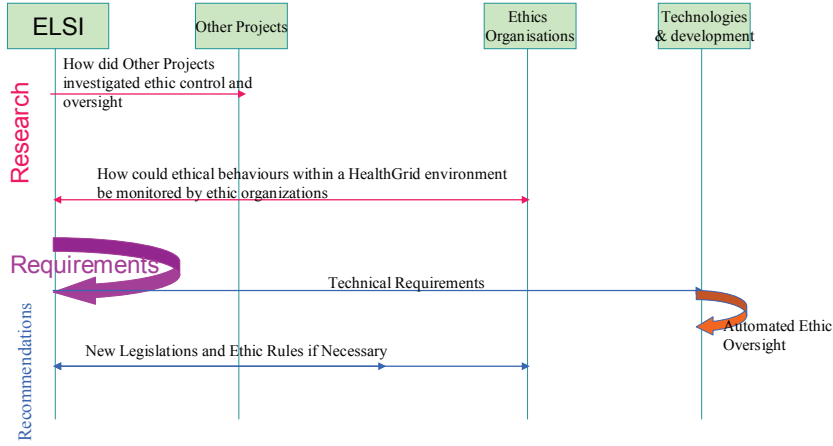
5. APPENDIX



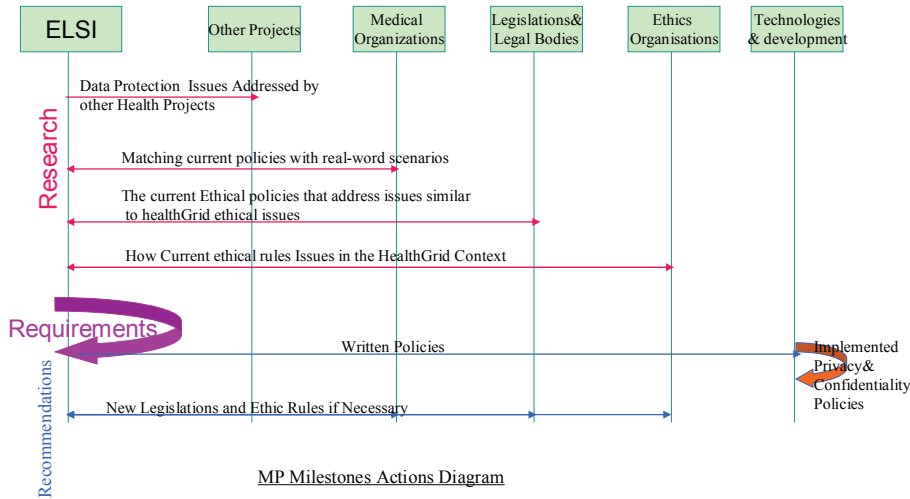
MEL1 Interaction Diagram



MDP Milestones Actions Diagram



ME Milestones Actions Diagram



MP Milestones Actions Diagram