



Contract start: September 2002
Duration: 24 months
Coordinator: Universidade de Aveiro – IEETA
Partners: UPM, ISC, Linku, Genomica

READER'S GUIDE

ROADMAP TO PROJECT DELIVERABLES

IEETA

Abstract: This report provides a comprehensive overview of the project work packages and contractual deliverables. It further describes the deliverables selected to the 1st Annual Technical Review and can be used as a guide to the project Deliverables.

Version: 1.0

Reporting period: n.a.

Contributing partners: IEETA

Due on: n.a.

Preparation date: 2003-08-08

Quality Assurance

Document title	Roadmap to project deliverables
Version	1.0
Filename	INFOGENMED_RoadmapToDeliverables_v1.0_IEETA.doc
Lead partner	IEETA
Reviewers	IEETA / PM
Circulation	INFOGENMED Consortium
Copyright notice	Copyright © INFOGENMED
Comments	

Document history

Version	Date	Main changes
1.0	2003-09-10	Initial version.

Coordinator

Universidade de Aveiro

IEETA - Instituto de Engenharia Electrónica e Telemática de Aveiro

Address: Campus Universitário de Santiago
3820-193 Aveiro — PORTUGAL

Telephone: + 351 234 370 500

Fax: + 351 234 370 545

More at: www.infogenmed.net

Table of contents

1	Introduction	4
1.1	Purpose.....	4
1.2	Brief project description.....	4
1.3	Participants list.....	6
1.4	Overview of work-packages	6
1.5	Deliverables listing	7
2	Deliverables reading list for the 1st Annual Technical Review	9
2.1	Management and progress monitoring reports.....	9
2.2	Technical reports.....	10
2.3	Additional documents	11
3	Project glossary	12

1 Introduction

1.1 Purpose

This report was prepared to assist the project Reviewers and other readers to quickly understand the project work structure and produced deliverables.

A comprehensive list of deliverables to the 1st Technical Review is provided in Section 2, along with a description of their goals.

Although in general each report contains an initial acronym section, this report includes a transversal project glossary for the convenience of the reader.

1.2 Brief project description

Project data sheet

Acronym	INFOGENMED
Full title	INFOGENMED: A virtual laboratory for accessing and integrating genetic and medical information for health applications
Number	IST-2001-39013
Project web site	http://www.infofenmed.net
Budget	Total cost: 1,454,264 EC contribution: 1,008,000
Timeframe	Start date: 01-Sep-2002 Duration: 24 months
Keywords	Biomedical informatics, biomedical information hubs, rare genetic diseases, distributed information integration, health knowledge management.

Table 1: Project data sheet.

Project objectives

Medical informaticians have long been providing physicians with computing aids for patient care and management, while bioinformatics experts have more recently been building and managing large databases of genetic information, as the Human Genome Project exemplifies.

The increasing availability of genetic and clinical information, both in public and in-house databases, demands for practical information technology tools, able to give clinicians integrated and comprehensive access to previously scattered and unconnected data.

INFOGENMED aims to build methods and integrated tools that enable the access to and use of dispersed, heterogeneous databases, to improve clinical and research practice. Medical information regarding diseases will be easily accessible and related genetic information, a type of valuable data the clinicians are mostly unfamiliar with, will also be located, retrieved and presented in a unified, user-friendly way.

This will empower medical practice, research, knowledge development and the collaboration between bioinformatics and medical informatics, paving the way for individualized medicine: as the understanding of the genetics base of drug action increases and genetic information is added to Patient Records, both patient-tailoring of medical action and population studies of genetic epidemiology will be improved.

The project motivation is related to both informatics and medical innovations. We propose for the former the integration of different original methods and tools to solve the problem of data integration from multiple heterogeneous sources. For the latter, we believe that integrated medical and genetic information will be increasingly needed in routine patient care. Both dimensions are included in the current (and future) EC research agenda.

Expected results

Having started in September 2002, INFOGENMED will reach conclusion in September 2004, producing the following results:

- Design and implementation of a software system to support the search and linkage of the contents of scattered databases, including:
 - o Development of methods and tools to locate, access and integrate distributed medical and genetic data;
 - o Construction of a vocabulary server to bridge and relate different medical and genetic terminologies;
 - o Development of an “assistant” to help users in resorting to such methods and tools for the benefit of their practice (e.g.: flowchart representation for clinical pathway visualization);
- The field-validation of the entire system in the field of rare genetic diseases.

1.3 Participants list

Partic. no.	Participant short name	Participant name	Country	Partic. Role*
1	IEETA	Universidade de Aveiro	Portugal	C
2	UPM	Universidad Politécnica de Madrid	Spain	P
3	ISC	Instituto de Salud Carlos III	Spain	P
4	LinkU	Linkoeplings Universitet	Sweden	P
5	GENOMICA	STAB VIDA, Investigação e Serviços em Ciências Biológicas, Lda.	Portugal	P

* C: Coordinator; P: Principal contractor.

Table 2: Participants list.

1.4 Overview of work-packages

The INFOGENMED project is structured in eight Work Packages (Table 3).

WP no.	Work package title	Lead contractor	Deliverables No.	Key milestones
1	Project Management and Coordination	IEETA	1.1, 1.2; 1.31 to 1.35; 1.4, 1.5; 1.61 to 1.64; 1.71 to 1.73; 1.81 to 1.82; 1.9	Project presentation. Progress monitoring and reporting. Dissemination and use plans. Quality assurance.
2	Analysis of state of the art	Linku	2.1	Analysis of the state of the art report.
3	Functional analysis and user requirements	IEETA	3.1 and 3.2	Functional specification of INFOGENMED software environment.
4	Database creation and validation	ISC	4.1 and 4.2	Prototype of a virtual database for rare genetic diseases.
5	Design and implementation of the informatics tools	UPM	5.1 to 5.8	Vocabulary server prototype; Heterogeneous databases integration tools; protocol visualization tool; complete integrated system.
6	Clinical testing and evaluation	ISC	6.1 and 6.2	Evaluation methodology; Clinical evaluation report.
7	Analysis of confidentiality and security issues	Linku	7.1	Review of confidentiality and security and an access control model.
8	Dissemination and exploitation	Genomica	8.1	Exploitation and business plan for project results.

Table 3: Description of Work packages.

The Gant chart for the execution of project work packages is depicted in Figure 1.

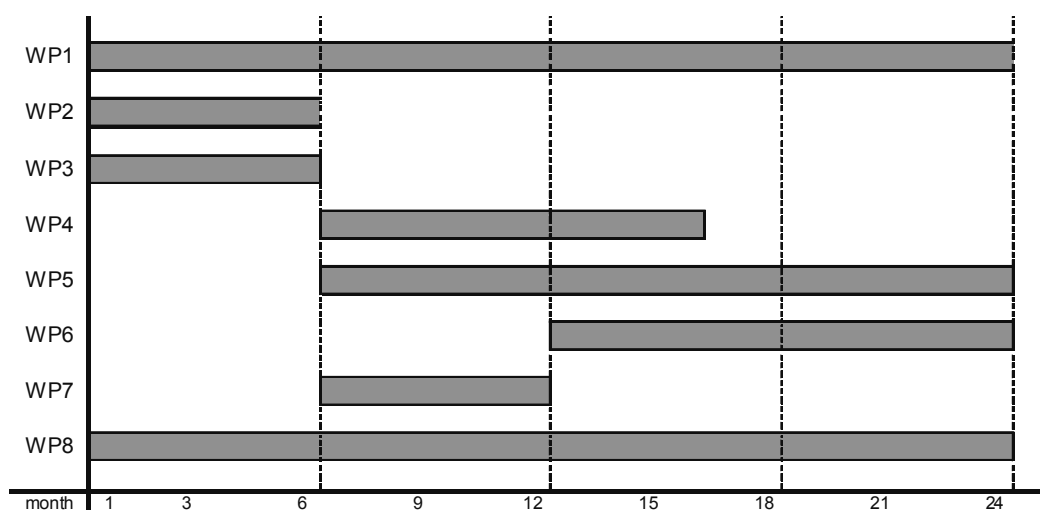


Figure 1: Overall Gantt chart.

1.5 Deliverables listing

Del. no.	Deliverable name	WP	Lead	Type ¹	Security ¹	Circulation ²	Month
1.1	Intra Consortium Agreement	1	IEETA	R	INT	All	1
1.2	Project (web) Presentation	1	IEETA	O	PU	All	3
1.31	Quarterly Management Report	1	IEETA	R	RE	All	3
1.4	Project Handbook - Quality Assurance	1	IEETA	R	FP5	All	3
2.1	Analysis of the state of the art	2	Linku	R	FP5	Linku	6
3.1	Functional analysis	3	IEETA	R	RE	All	6
3.2	User requirement	3	IEETA	R	RE	All	6
1.5	Dissemination and Use Plan	1	IEETA	R	FP5	All	6
1.61	Periodic Progress Report (6 months)	1	IEETA	R	RE	All	6
1.71	Cost Statements (6 months)	1	IEETA	R	RE	All	6
5.1	Technical specifications	5	UPM	R	RE	IEETA,UPM	9
1.32	Quarterly Management Report	1	IEETA	R	RE	All	9
5.2	Vocabulary server prototype	5	UPM	P		IEETA,UPM	12
7.1	Security and confidentiality report	7	Linku	R	FP5	IEETA,Linku	12
1.62	Periodic Progress Report (6 months)	1	IEETA	R	RE	All	12
1.72	Cost Statements (6 months)	1	IEETA	R	RE	All	12
5.3	Validation of the vocabulary server	5	UPM	R	RE	IEETA,UPM	15
5.4	Tools for heterogeneous DB integration	5	UPM	P	RE	IEETA,UPM	15

¹ R: Report; P: Prototype; O: others; PU: Public; RE: Restricted; INT: Internal; FP5: Circulation within Framework Programme participants.

² Besides Project Manager and CE contract definitions.

1.33	Quarterly Management Report	1	IEETA	R	RE	All	15
6.1	Evaluation methodology	6	ISC	R	RE	ISC, Genomica	15
4.1	Virtual Database rare genetic diseases	4	ISC	P	RE	ISC, Genomica	16
4.2	Database validation criteria	4	ISC	R	RE	ISC, Genomica	16
5.5	Validation of DB integration	5	UPM	R	RE	IEETA, UPM	18
5.6	Protocol visualization tool	5	UPM	P	RE	IEETA, UPM	18
1.34	Quarterly Management Report	1	IEETA	R	RE	All	18
5.7	Validation of protocol visualization tool	5	UPM	R	RE	IEETA, UPM	20
5.8	Integrated system	5	UPM	P	RE	IEETA, UPM	21
1.35	Quarterly Management Report	1	IEETA	R	RE	All	21
6.2	Clinical evaluation report	6	ISC	R	RE	ISC, Genomica	24
1.63	Periodic Progress Report (12 months)	1	IEETA	R	RE	All	24
1.73	Final Cost Statements (12 months)	1	IEETA	R	RE	All	24
8.1	Exploitation and Business Plan	8	Genomica	R	RE	All	24
1.81	Final Report	1	IEETA	R	RE	All	24
1.82	Edited Final Report	1	IEETA	R	PU	All	24
1.9	Technology Implementation Plan	1	IEETA	R	RE	All	24

■ Project management and progress monitoring report (under WP1).

Table 4: Project deliverables.

The periodicity of the management tools (Quarterly Management Reports, Periodic Progress Report and Cost Statements) is clarified in Table 5. The reporting period of each document is graphically delimited in the table.

Year 1		
Month	QMR	PPR/CS
1	QMR1	PPR1 + CS1
2		
3		
4	n.a.	
5		
6		
7	QMR2	PPR2 + CS3
8		
9		
10	n.a.	
11		
12		

Year 2		
Month	QMR	PPR/CS
13	QMR3	PPR3 + CS3
14		
15		
16	QMR4	
17		
18		
19	QMR5	
20		
21		
22	n.a.	
23		
24		

Table 5: Management reporting schema.

2 Deliverables reading list for the 1st Annual Technical Review

2.1 Management and progress monitoring reports

Technical Annex

The technical description of work, annexed to the project proposal.

D 1.3x – Quarterly Management Reports

Deliverables **1.31** (months 1-3) and **1.32** (months 7-9) provide project progress monitoring information. They should be read in respect to the Periodic Progress Reports.

See also Table 5 for an overview of the progress reporting strategy.

D1.4 – Project Handbook – Quality assurance

The purpose of this document is to introduce a Project Manual for the INFOGENMED project

This manual describes the internal regulations, management structures and procedures to be used by the Consortium in the INFOGENMED project. It includes information regarding the organization, project plans, controls, best practices, risk analysis and project agenda.

Del 1.5 – Dissemination and Use Plan

This deliverable describes the dissemination and use plan for the products and services developed in the INFOGENMED project. Its scope encompasses the consortium's knowledge and ideas about the target market and on how developed methods and tools can be exploited. Tools under developed will support practitioners and researchers in the exploitation of modern information resources, spread over multiple networks and information systems, regarding molecular, genetic and clinical information for health applications. The ideas presented in this document are based on the current status of the products/services development and knowledge. A more elaborated view on dissemination and exploitation will be released later in the project (planned Deliverable 8.1).

Del 1.6x – Periodic Progress Reports

Deliverables **1.61** (months 1-6) and **1.62** (months 7-12) provide cumulative project progress monitoring details, for each semester. The activities herein reported include and surpass the contents of Quarterly Management Reports.

See also Table 5 for an overview of the progress reporting strategy.

Del 1.7x – Cost Statements

Deliverables **1.71** (months 1-6) and **1.72** (months 7-12) provides formal cost statements.

2.2 Technical reports

Del 2.1 – Analysis of the State of the Art

This report addresses the state of the art in the scientific fields related to INFOGENMED, using the perspective of the intersection between Bioinformatics and Medical Informatics. Special attention is given to the field of rare genetic diseases, which is the focus of the test bed for the tools to be implemented. Enabling technologies (standards, programming tools, knowledge representation, etc.) and initiative related to the project are also reviewed.

Del 3.1 – Functional Analysis

The report identifies and describes the features to be implemented in INFOGENMED software environment for the successful integration of distributed genetic and medical information sources. It discusses the envisaged overall architecture and the requirements of the main components.

Del 3.2 – User Requirements

This document defines the end-user requirements with respect to the INFOGENMED environment. It describes the methodology used for the identification of the user requirements and the results obtained. A use case model of the future tools is provide, along with a description of user profiles.

Del 5.1 – Technical specification

This deliverable describes the system's components in respect to their functions and software implementation strategy. A detailed description of developed prototypes is included along with the planned approach for further development of tools. This report can be read in reference to the previous deliverables 3.1 and 3.2.

Del 5.2 – Vocabulary server prototype

This deliverable is a software tool to enable the storage of terminology systems in a machine-readable form, to be used by integration systems in semantic reasoning and mediation. The tool also includes a user interface for query and update the terminologies repository. This vocabulary server will capture concepts from medical and bio informatics fields, and semantic links between them.

Del 7.1 – Security and confidentiality report

Deliverable 7.1 analyses the security risks to patient data in new networked IT environments. It studies the threats raised by unauthorized disclosure of clinical and genetic data. While the former has been addressed in the past, the last is still under an intense debate. An extensive

review of security and privacy enhancing techniques and practices is conducted. The report also proposes a model for the deployment of secure biomedical, internet-based systems, and managing patient data, including genetic level information. Finally, the security provisions envisaged for the deployment of the INFOGENMED environment are presented.

2.3 Additional documents

In addition to those contractual deliverables already presented, the Consortium is also submitting the following results:

Prospective study on GRID technologies

This report presents the results of the study on GRID technology and applications undertaken by the Consortium. The document provides an initial revision of concepts, relevant GRID research and exploitation initiatives, main technologies and tools, and tries to put in perspective the application of GRIDs in the life science domain. Finally, the report establishes links to the INFOGENMED project, pointing out ways on how the project can benefit from GRID technology.

3 Project glossary

BI	Bioinformatics. A field of biology concerned with the development of techniques for the collection and manipulation of biological data, and the use of such data to make biological discoveries or predictions. This field encompasses all computational methods and theories applicable to molecular biology and areas of computer-based techniques for solving biological problems including manipulation of models and datasets (MeSH definition). It is an interdisciplinary field, which harnesses computer science, mathematics, physics, and biology
BMI	Biomedical Informatics. BMI is the field that studies biomedical information and knowledge: their structure, acquisition, integration, management, and optimal use. The field involves multidisciplinary research in, application development for, and administrative approaches to all aspects of health care delivery, biomedicine, and public health. BMI adopts, applies, evaluates, modifies, and expands results from a variety of disciplines including Information Science, Computer Science, Library Science, Cognitive Science, Business management and Organization, Statistics and Biometrics, Mathematics, Artificial Intelligence, Operations Research, Economics, and of course, basic and clinical Health Sciences. (Vanderbilt University)
CORBA	Common Object Request Broker Architecture. Middleware specification, part of the OMG standards.
CORBA-IIOP	Common Object Request Broker Architecture - Internet Inter-ORB Protocol. See also: CORBA.
DBMS	Database management system.
DBMS	Database Management Systems
DNA	Deoxyribonucleic acid.
DW	Data warehouse
EC	European Community.
EHR	Electronic Health Records
FDBS	Federated Database System.
Genomics	The systematic study of the complete DNA sequences (GENOME) of organisms (MeSH definition).
HGP	Human Genome Project.
HTTP	Hypertext Transfer Protocol.
ICT	Information and Communication Technology.
IDL	Interface Definition Language. Part of the OMG standards; see also: CORBA.
IQL	INFOGENMED query language: an interrogation language to formulate queries over the INFOGENMED system's global model.
IST	Information Society Technologies.
IT	Information Technology.
JDBC	Java Database Connectivity.
JVM	Java Virtual Machine.
MeSH	Medical Subject Headings

MI	Medical Informatics. "The field of information science concerned with the analysis and dissemination of medical data through the application of computers to various aspects of health care and medicine" (MeSH definition).
Molecular Medicine	Medical research and practice focusing on the understanding of the basic molecular biology and the analysis of disease mechanisms at the level of cells and molecules and its translation into diagnosis, prevention, treatment and cure of human diseases.
OMG	Object Management Group; a consortium aimed at setting standards in object-oriented programming.
OMG	Object Management Group
OQL	Object Query Language
ORB	Object Request Broker; part of the OMG standards. See also: CORBA.
Proteomics	The study of the full set of proteins encoded by a genome (HGP)
RMI	Remote Method Invocation. Part of the Java programming language library which enables a Java program running on one computer to access the objects and methods of another Java program running on a different computer.
SQL	Structured Query Language
UML	Unified Modeling Language
VDB	Virtual Database
VRM	Virtual Repository – Mapping.
VRU	Virtual Repository – Unification.
WP	Work package.