Title:	*Specification of mapping to Provenance architecture, and domain specific Provenance handling*
Author:	*Work Package 8 ("Application 2: organ transplant management")*
Editor:	*Javier Vázquez-Salceda and Steve Willmott (UPC), Kifor Tamás and László Zs. Varga (MTA SZTAKI)*
Reviewers:	*All project partners involved*
Identifier:	*D8.1.1*
Type:	*Deliverable*
Version:	*v1.0*
Date:	*September 20h , 2005*
Status:	*Final*
Class:	*Confidential*

Summary

This document describes the relationship between the Organ Transplant Management (OTM) -application adopted as a demonstration case in the Provenance project-, the Electronic Health Care Record Application (ECHR) -used by the OTM application to store medical records- and the Provenance logical architecture. The mapping will be used as the basis for the implementation of the OTM demonstration system in months 12-24 of the project

The PROVENANCE project receives research funding from the European Community's Sixth Framework Programme

Members of the PROVENANCE Consortium:

IBM United Kingdom Limited	United Kingdom
University of Southampton	United Kingdom
University of Wales, Cardiff	United Kingdom
Deutsches Zentrum für Luft- und Raumfahrt e. V.	Germany
Universitat Politecnica de Catalunya	Spain
Magyar Tudomanyos Akademia Szamitastechnikai es Automatizalasi Kutato Intezet	Hungary

Foreword

This document describes the relationship between the Organ Transplant Management (OTM) -application adopted as a demonstration case in the Provenance project-, the Electronic Health Care Record Application (ECHR) -used by the OTM application to store medical records- and the Provenance logical architecture. The mapping will be used as the basis for the implementation of the OTM demonstration system in months 12-24 of the project . The document provides:

- Detailed descriptions of the OTM domain.
- High level mapping rules between the OTM application and the Provenance Logical Architecture.
- High level mapping rules between the ECHR application and the Provenance Logical Architecture.
- Instantiations of these mapping rules for the planned project demonstration deployment.

The primary audience of this document includes: A) Grid computing practitioners seeking to understand how Provenance technologies might be applied and, B) Information technology practitioners in the health care domain interested in applying Provenance to their own medical systems.

Table of Contents

List of Acronyms

Acronym	*Description*
DNI	Spanish National Identity Number
EHCR	Electronic Health Care Record
EHCRS	Electronic Health Care Record Store
GMPID	Global Medical Patient Identifier
GP	General Practitioner
GUI	Graphical User Interface.
HCR	Health Care Record
HLA	Human Leukocyte Antigens test
IT	Information Technology
LMPID	Local Medical Patient Identifier
OCATT	Organització Catalana de Transplantaments (Catalan Transplant Organisation)
ONT	Organización Nacional de Transplantes (Spanish National Transplant Organisation)
OTA	Organ Transplant Authority
OTM	Organ Transplant Management
PDF	Portable Document Format, Adobe Inc.
PID	Patient Identifier
PS	Provenance Store

1 Introduction

This document describes the relationship between the Organ Transplant Management (OTM) application adopted as a demonstration case in the Provenance project and the Provenance architecture (WP3 of the project). The mapping will be used as the basis for the implementation of the OTM demonstration system in months 12-24 of the project. Throughout the document the mapping is split into two separate elements: 1) the OTM application itself and 2) the underlying Electronic Health Care Record (EHCR) application which is treated as a distinct Provenance problem.

The primary audience of this document includes: A) Grid computing practitioners seeking to understand how Provenance technologies might be applied and, B) Information technology practitioners in the health care domain interested in applying Provenance to their own medical systems.

1.1 Purpose of the Document and Overview

The purpose of the document acts as an initial specification for two interconnected demonstration applications: OTM over Provenance and EHCR over Provenance. A detailed demonstration specification taking into account these mappings will subsequently be developed for each application. Secondary goals include:

- Providing an example for others to build upon in future uses of Provenance technology.
- Generating feedback on the current Provenance architecture.

In order to achieve these goals, the document aims to:

- Detail the Organ Transplant Management (OTM) application modeled in the project as well as the associated EHCR application (Section 2 of the document).
- Describe the architecture of the OTM application system implementation being developed by the CARREL FIS project in Catalunya Spain and the planned EHCR architecture (Section 3 of the document).
- Describe a set of high level mapping rules / conventions applied to the two applications in order to guide detailed mapping to the Provenance architecture at implementation time (Section 4 of the document).
- Provide a detailed guide as to how process documentation for Provenance purposes will be stored, managed and queried in the two (OTM and EHCR) application systems (Section 5 of the document).

1.2 Links of other Provenance Documents

The contents of this document are based on the following existing Provenance project documents:

- Requirements expressed for the OTM application in the WP2 Requirements deliverables D2.1.1 and D2.2.1.
- The Provenance project frozen architecture document D3.1.1, preliminary version.
- Project internal note on “Representing Provenance in the OTM application” [Miles05].

Further supporting documents are provided in the references section.

2 Application Overview

This chapter describes the distributed organ transplant process according to current practice. The presentation covers major actors, the description of the process and the decision criteria used during the transplantation workflow.

2.1 The Organ and Tissue Allocation Problem

Organ transplantation from human donors is the only treatment option available in medical cases where there is a major damage or malfunction in an organ. Other approaches have been unable to fully provide the full range of functionalities of human organs (e.g., artificial hearts) or they imply complex follow-up procedures to keep the organs working within acceptable parameters (e.g. Xenotranplants, which are organs from genetically-altered animals). Additionally, human organ transplantation not only has a very positive impact in the patients quality of life, but also it is also very important for the health care authorities in cost terms – in Spain for example, the transplantation of one kidney compared with dialysis would save between 186,400 and 240,530 Euro for a given patient. over their lifetime. Furthermore, a successful transplant often leads to a higher quality of life for a patient.

For these reasons, organ transplantation represents an important element of most advanced health care systems. Over the years, transplant techniques have evolved; knowledge of donor-recipient compatibility has improved and so have immunosuppressant drug regimes, leading to an increase in the number of organs that can be transplanted, but also in the range of transplants, moving beyond organs (heart, liver, lungs, kidney, pancreas) to tissues (bones, skin, corneas, tendons). Since 1980 the number of requests for the application of transplant techniques has risen so much[1] that transplant coordinators — the medical staff who act as the interfaces between the surgeons internally and the organ transplant organisations and tissue banks externally — now face significant challenges in dealing with the volume of work involved in the management of requests and piece assignment and distribution. In addition, organ transplant organisations which act as central authorities responsible for the allocation of organs, are facing similar problems in growth. In particular, the need to optimise allocation of a very scarce resource (donated human organs) and the logistics of the transportation, from the donor's hospital to the recipient's hospital in the short period of time in which an organ can be kept outside a human body for a growing volume of transplants means new methods are needed to handle larger volumes with better outcomes.

For these reasons there is strong demand in the sector for Information Technology solutions which would assist medical staff in the management of data related to transplantation cases, support the execution of workflow and provide efficient communication channels between the many staff involved in a transplantation activity.

2.2 Classes of Transplants

Transplantation operations are divided into two broad classes:

1. *Live organ transplants (heart, lung, intestine, liver, pancreas, kidney)*: In this case, the item being implanted is a live internal organ. Organs are complex structures with a wide range of cell types with different optimal preservation temperatures. As it is impossible to find the optimal preservation conditions for the whole structure, such organs deteriorate rapidly between the moment they are extracted from the donor's body and the moment they are implanted (becoming useless in less than 24 hours in some cases[2]) – creating significant time pressure on

1 The continuous raise in requests is due, among other factors, to the introduction of new immunosuppressors which have significantly decreased rejection in recipients' clinical evolution.

2 More specifically heart, lung, pancreas, intestine can be kept only for 4-6 hours between retrieval and implant before they become useless, kidneys maybe kept longer (24-36 hours). However in all cases the lower the

transplantation. As organs cannot be stored in banks, the allocation process is triggered when a donor becomes available. The allocation process then takes the form of a search for a suitable recipient in some number of hospitals. Patients waiting for a suitable organ are registered in waiting lists. When a donation is made at a given moment in time, the organ must be assigned to one of the patients in the waiting lists (or no one if no good matches are found).

2. *Tissue transplants:* In this type of transplant, the item being transplanted is a tissue such as a cornea, skin, bone or marrow. Tissues are clusters of essentially homogeneous cells, so the optimal temperature for preservation of all the cells composing the tissue is almost the same. Thus, tissues can be preserved for several days or more (from six days in the case of corneas to years in the case of bones) in "tissue banks", which are special centres that receive, check, catalogue and preserve large collections of tissues. For tissues, the allocation process is triggered when there is a recipient with a need for a certain tissue, at which time some number of tissue banks are searched for a suitable piece of tissue. This search is carried out by matching the requirements of the incoming recipient with the catalogued tissues – making decision making a "one recipient to one of many possible donors" matching problem. It is important to note that blood is actually a special type of "liquid tissue", that is separated in its different components (plasma, leucocites, etc) to be optimally stored in blood banks.

Problems in these distinct cases are therefore significantly different in structure and furthermore the Provenance issues may be somewhat different. For the remainder of the document presentation will focus on the organ transplantation case since:

- This is the application in which new IT technologies are predicted to have the largest beneficial impact.
- Through partner UPC, the project has access to an ongoing Spanish Government funded collaboration project with Hospital St. Pau in Barcelona, Spain.

2.3 Coordination Structure: OTM in Spain

At the time of writing, more than one million people in the world have successfully received an organ, and thereafter, in most cases, lead normal lives. Spain has on average 33 cadaveric organ donors per million population (pmp) in 1999 (37 pmp in Catalonia). This success places the Spanish Organización Nacional de Transplantes (ONT) and the Catalan Organització CATalana de Trasplantaments (OCATT) transplant organisations among the most effective and demonstrate the highest global volume of transplants per head of population. Both the ONT and OCATT act as technical organisations under the authority of the Spanish Ministry of Health and Consumer Affairs. These agencies deliver a service to the National Health System which provides for the most appropriate and correct distribution of organs according to transplant related legislation.

2.3.1 Coordination Levels

The success of OTM processes in Spain arguably comes from a re-structuring and optimisation of the process at two levels:

- *Intra-hospital level*: with the introduction of the role of Hospital Transplant Coordinator to improve the coordination of all the individuals involved in any step of the donor procurement, allocation and transplantation process.

- *Inter-hospital level*: where an intermediary organisation—OCATT for Catalonia, ONT for the whole of Spain—was created to improve the communication and coordination of all the participating health-care organisations (namely hospitals and tissue banks).

time outside the body is the more likely ultimate success is.

Coordination	• Inter-hospital Coordination of all multiorgan retrieval procedures. • Up-date and maintenance of liver, heart and lung transplant waiting lists. • Cooperation in kidney exchanges. • Coordination of air/land transportation of transplant teams and organs for transplantation. • Cooperation in patients transfer if needed. Channeling of patients reports for pre-transplant evaluation. • Channeling of requests for bone pieces or other tissues. • Channeling of ”Bone Marrow Donor Searches”.
Regulations and reports	• Elaboration of technical reports related directly or indirectly with the organs, tissues and haematopoietic progenitors transplantation. • Promotion of Agreements and Consensus Reports.
Studies	• Collection of data on procurement and transplantation activities. • Data analysis. Publications. • Evaluation of health requirements: legal, human and material. • Promotion and coordination of multi-center studies and research projects.
Information	• Information covering donation and transplantation activities as well as health related topics to i) Health Administration, ii) transplant coordinators, iii) transplant professionals, iv) international transplant organisations. v) patients associations. • Information to the general public by means of i) public campaigns for social sensitisation, ii) issuing of donor cards, iii) management of a telephone line to provide information about any question related with donation and transplantation • Spreading of informative, didactic and working material.
Others	• International Cooperation. • Promotion of Specific Training Courses. • Development of the Spanish Society of Donation and Transplantation of Organs and Tissues.

Table 1– Tasks of the Spanish National Transplant Coordinator [Source: ONT].

The ONT is now structured as a network system established at three basic levels: national, regional and local:

- *National Coordination*: The National Transplant Coordinator is the head of the ONT, and has the mission to act as a nexus among a) local, national and European health authorities, b) health professionals, c) the different social agents involved in organ donation and transplantation and d) the general population. The National Transplant Coordinator works at the Central Coordination Office, which functions are listed in Table 1 above.
- *Regional Coordination*: Each of the seventeen Spanish Autonomous Communities has a Regional Transplant Coordinator. These individuals are responsible for the coordination of resources, tasks relating to information, circulation and promotion at regional level. These regional coordinators also are members of the Organ and Tissue Transplant Commission, where any subject related with transplantation that affects more than one Autonomous Community is discussed.

- *Hospital Coordination*: Within a region, hospital coordinators catalyse the detection of donors. A full review of the hospital coordinator's role and the difficulties faced is presented by López-Navidad [LopezNavidad97]. In most cases this work is combined with their daily health care work so that the coordinators still maintain contact with the actual hospital life.

2.3.2 Primary Actors and Roles

A transplant *case* is defined as a single episode of organ transplantation (or attempted organ transplantation) including all processes from the arrival of the donor to the completion of surgery and after care of the recipient. In a given case the major actors / types of actors and their roles are:

- *Donor*: person donating the organ or organs in a particular case. The individual must be associated with an available medical history or a quite complete medical history should be created by carrying out several laboratory tests. (Without such a record, transplants cannot take place.)
- *Recipient*: person or persons being operated upon (successfully or unsuccessfully) to implant the donated organ. Associated with a particular medical history and a particular possible implantation center.
- *Recipient Waiting List*: ordered list of individuals who may act as potential recipients if an organ becomes available (grouped by the type of organ they required). A subgroup of the patients which are in the most severe condition are highlighted as "urgency 0" (a special code for the most urgent cases) patients in these lists, having the highest priority.
- *Retrieval Team*: medical personnel (surgeon, nurses, technicians, etc.) carrying out the retrieval of an organ from the donor. In Spain, the personnel in this team is not associated with the retrieval center, but with the implantation center.
- *Implant Team*: medical personnel (surgeon, nurses, technicians, etc.) carrying out the implantation of an organ in the recipient. Associated with a particular implantation center.
- *Duty Transplant Surgeon*: individual physician/surgeon on duty at the retrieval or implantation center.
- *Consultant Transplant Surgeons (experts)*: individual(s) other than the duty surgeon who may be consulted by the duty surgeon during any given case. Associated with one or more retrieval/implantation centers.
- *Remote retrieval site*: location where the retrieval takes place if this location is not a hospital or suitably equipped retrieval center.
- *Retrieval center*: hospital coordinating / carrying out the retrieval of an organ – either at the hospital itself or at a remote retrieval site.
- *Implantation center*: hospital carrying out the implantation of an organ.

- *Post operation care center*: hospital or medical center looking after the patient in post-operation care.
- *Test Laboratories*: specialist medical units or even separate centers performing blood and other analyzes of organs in order to determine matches in key indicators (HLA analysis or crossmatching[3] for example).
- *Regional Organ Transplant Authority (OTA)*: regulatory and oversight body for all transplants in a given region. Associated with a number of retrieval / implantation centers. The OTA center also acts as the coordinating point to find recipients if local recipients are not available. The OCATT is the Catalan OTA which coordinates the area of Catalonia, Valencia and the Balearic Islands. For the rest of Spain, ONT acts as OTA.

Figure 1 illustrates the communication paths between these primary actors and Figure 2 illustrates the primary power relationships. In a standard incident, the duty transplant physician of the retrieval centre is alerted to the availability of a possible donor, this individual sets in motion processes for assessing the donor.

Information on which organs may be donated is then passed to local transplant teams (in the same center) and the Organ Transplant Authority (OTA) to find potential donors. Depending on the type of organ, conditions and the protocols for the situation the duty physician and OTA mediate to find an appropriate recipient. Once a recipient has been found, a two part transplant team from the potential implant center takes charge – a retrieval team is sent to the location where the donor is and an implant team is readied at the implant center, the leader of the transplant team (comprising both parts) takes charge of the proceedings.

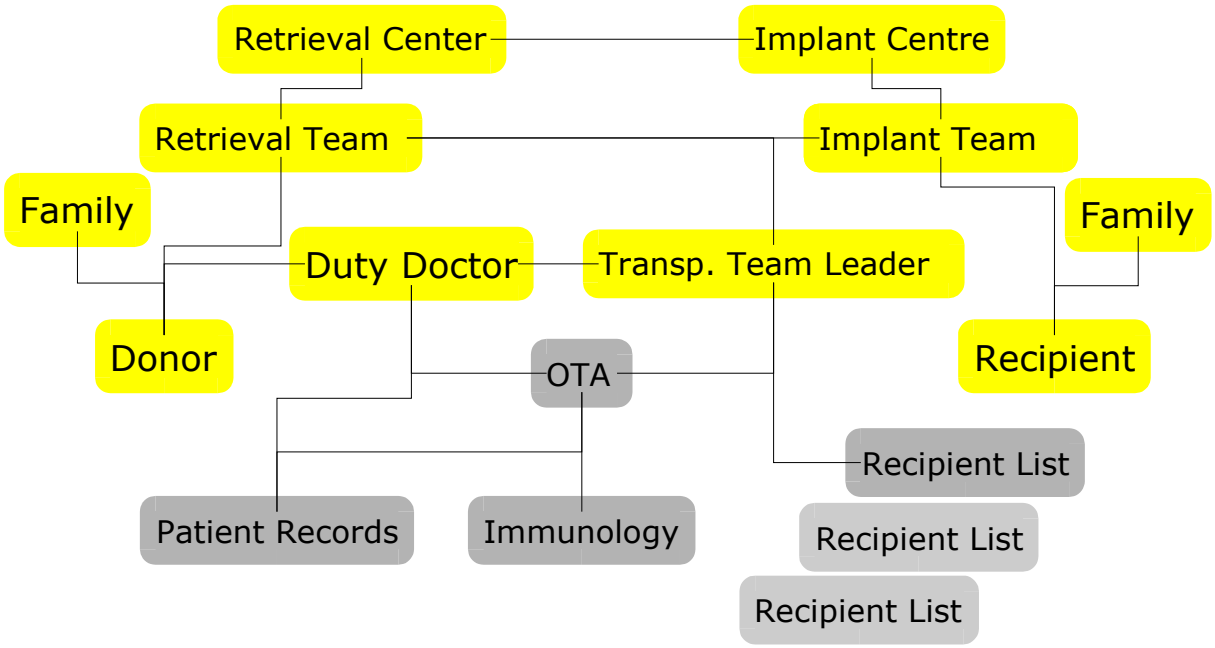

Figure 1– Direct communication during a transplant case (post-operation care centre not shown).

The physical distribution of the actors shown in Figure 1 are approximately as follows:

- Retrieval center and Implant center are both medical centers, each with its own physical location. In certain cases the same medical center may play both roles.

3 HLA Analysis and crossmatching are two important blood analysis techniques required to determine whether organs received are likely to be compatible with a particular potential recipient.

- The duty transplant physician is always located at the retrieval center site.
- If the transplant is from an accident the retrieval team may be at an arbitrary location which is not a medical center (however this case is rare – almost always the donor is moved to the nearest medical center.)
- Patient records are stored at each medical center patients are registered with but can be treated as a single distinct site (data is accessible from all sites).
- The immunology center is a distinct medical center per region – it may or may not be in the same place as the retrieval or implant centers.
- Experts may be at one of the previously mentioned medical centers, at another medical center entirely or in an arbitrary place (reachable by phone). They are generally members of transplant teams currently not on duty but may have additional experience of special situations.
- The organ transplant authority (OTA in the diagram) is located at a single further geographic site and is on 24 hour call.
- The post-operation care center may be one of the retrieval or implantation centers or another physically located center.

2.4 Description of the Transplant Process and Workflow

This section provides a more detailed description of the organ transplant process, including a precise model of the process by means of workflow diagrams.

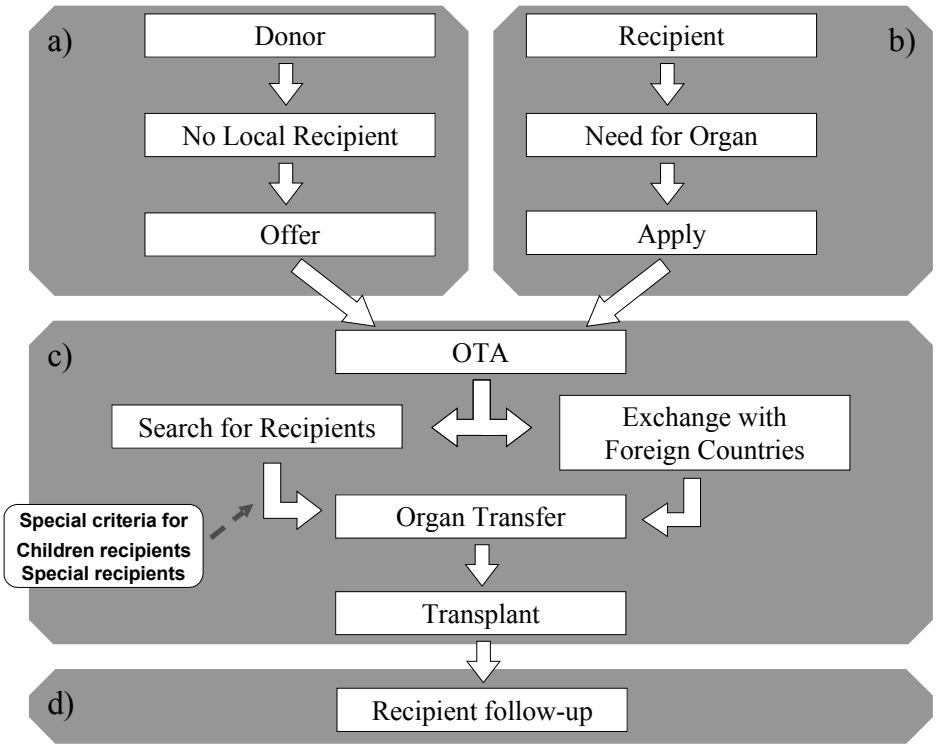

Figure 2– Overview of the transplant process: a) donor detection and evaluation, b) recipient detection and evaluation, c) allocation, d) follow-up.

Figure 2 shows an outline of the process which is divided in 4 stages or sub-processes: donor detection and evaluation process, recipient detection and allocation process, the allocation process and a recipient follow-up process. Each of these stages is further described in the following sections by including detailed workflow diagrams. The meaning of the different elements in the diagrams is as follows:

- *White Boxes*: represent a relevant event in the execution of the service, to be recorded in Provenance stores.
- *White Diamonds*: are minor conditional points in the workflow. The diamond is labelled with the condition to be checked. Lines that exit the diamond are labelled with the response needed to follow each line.
- *Black Diamonds*: are major decision points always taken by a human expert.
- *Full lines*: sequential workflows between white boxes and/or diamonds.
- *Dashed Line*: Parallel workflows between elements (e.g., in the diagram, the requests for the patient records and the lab tests are not done sequentially but all at the same time).

2.4.1 Recipient Detection and Evaluation Process (Pre-allocation)

This process starts in a hospital when a patient is diagnosed with a pathology or malfunction in one or several organs, and the only possible treatment is the transplant of one or several organs to substitute the one(s) that are not in good condition. From this moment onward, the patient is registered in the *recipient waiting list* of the hospital, and a preservation process is started to keep the patient in a stable condition until the moment that the transplant is performed. This process may be lightweight (such as regular check ups) or heavyweight (such as 24 hour medical care for patients in a critical condition).

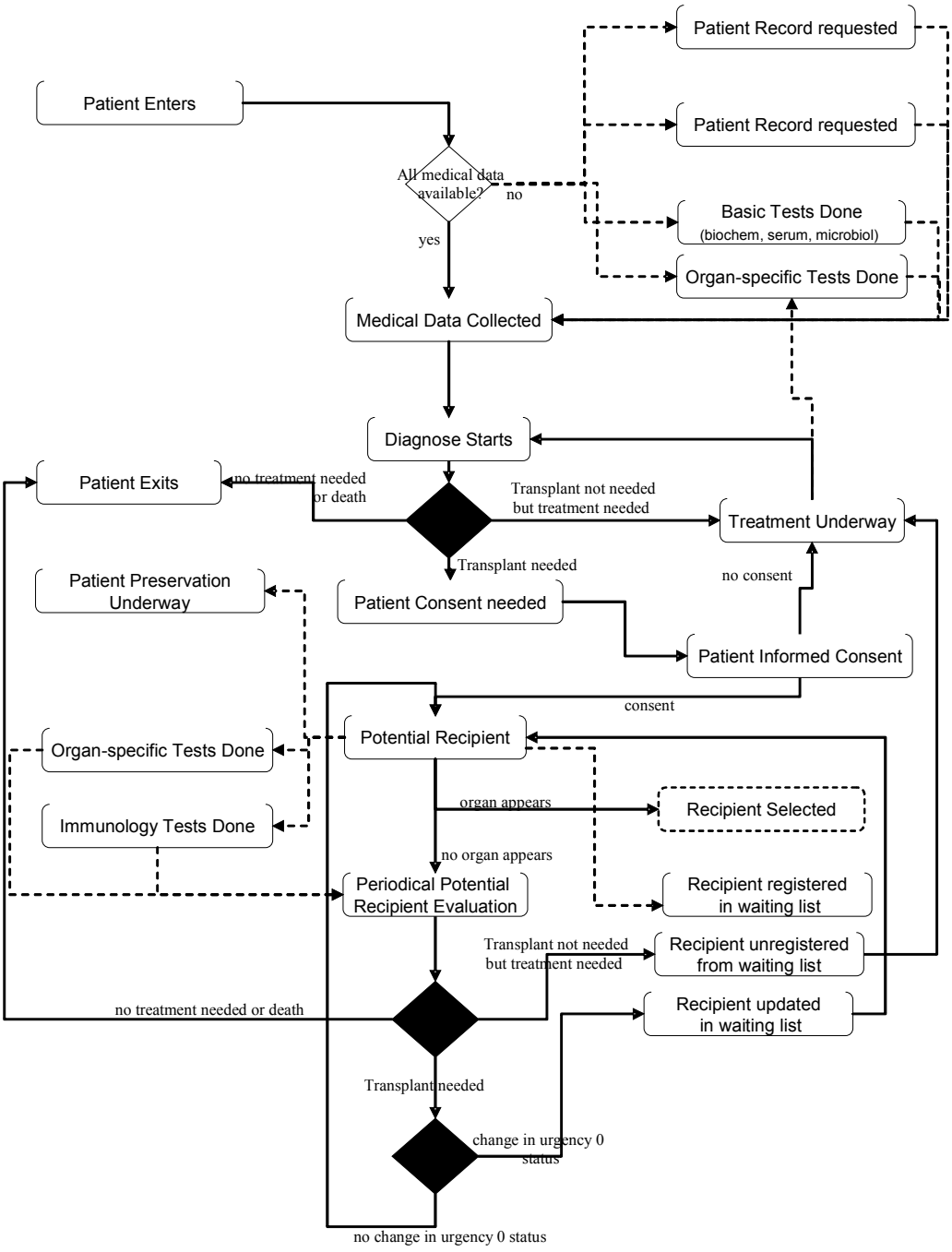

Figure 3 – Workflow for recipient detection and evaluation process.

Additionally, hospitals and the OTA maintain a list of urgent cases (called urgency 0) containing all the recipients whose condition is life-threatening. If there is a suitable recipient for an organ in this list then they are accorded higher priority in the assignment process over all other recipients.

Figure 3 shows the complete workflow for the recipient detection and evaluation process. It includes all the events (white boxes) and decisions (diamonds) that are important in such process. At the top of the diagram a process provides for the collection of the patient data relevant for the patient diagnosis. After the diagnoses, it is either considered that the patient needs no treatment, or that he/she needs a treatment which does not involve an organ transplant, or finally that a transplant is needed. In such

cases, the patient (or a legal representative) is informed about the procedures to be taken and their possible positive and negative consequences.

If and only if the consent of the patient is given, the patient is considered a potential recipient and included in the waiting lists. The bottom of the diagram shows the periodic cycle of evaluation of the patient while waiting for a transplant: important changes in the patient health status may need an update of the waiting list, or the removal of a patient in case of death, terminal conditions or recovery. If an organ appears during this cycle for a particular patient, then the patient is selected as recipient, which leads to the "recipient selected" event in Figure 3.

2.4.2 Donor Detection and Evaluation Process (Pre-allocation)

This process starts when the members of the transplant coordination team inside a certain hospital are made aware of a potential donor by the coordinator of one of the hospital units. A donor alarm is then sent to the OTA. This alarm is signalled by telephone, and a human member of the staff lists the basic attributes of the donor, including the results of clinical analysis, and a first evaluation of the organs and tissues that could be extracted is carried out. This first call is carried out as early as possible, usually when brain death of the potential donor is diagnosed and no problems are foreseen in getting the relatives and legal consent. Enough time is then available to organize the supply infrastructure and transport. At the time of the first call, basic clinical, analytical and anthropometric donor data is provided, which facilitate the subsequent assessment of the possible use of the organs, as well as donor/recipient compatibility. After carefully recording the donor's data, a dossier is opened for each case including an incident sheet used to record all the steps being taken and the time when each of then occurs.

Figure 4 shows the complete workflow for the donor detection and evaluation process. It again includes all the events (white boxes) and decisions (diamonds) that are important in such process. The top of the diagram shows the collection of the patient. If a potential donor is found, then the consent of relatives (and in cases of non-natural death, a legal ruling) or, in live donor transplants, the patient themselves is requested. If consent is given, then the patient becomes an available donor. By analysing the results of the different tests, experts evaluate the possible state of each of the donor's organs, and sends the offers to the OTA for those organs that are evaluated as suitable for potential transplant.

In the case that there is a potential recipient in the same hospital as the donor, current practice in Spain states that such patient would be likely the one to receive the organ, as in this case the time between organ extraction (from the donor) to the organ implantation (to the recipient) will be minimal, increasing the quality of the organ to be implanted. The primary exception to this local assignment is the existence of a national priority case (a recipient suitable for that organ which is in critical clinical condition, see section 2.4.3). Therefore, the hospital must always send the offer to the OTA along with extra information of a recipient available in the hospital, so the OTA will confirm the allocation (if there is no national priority recipient).

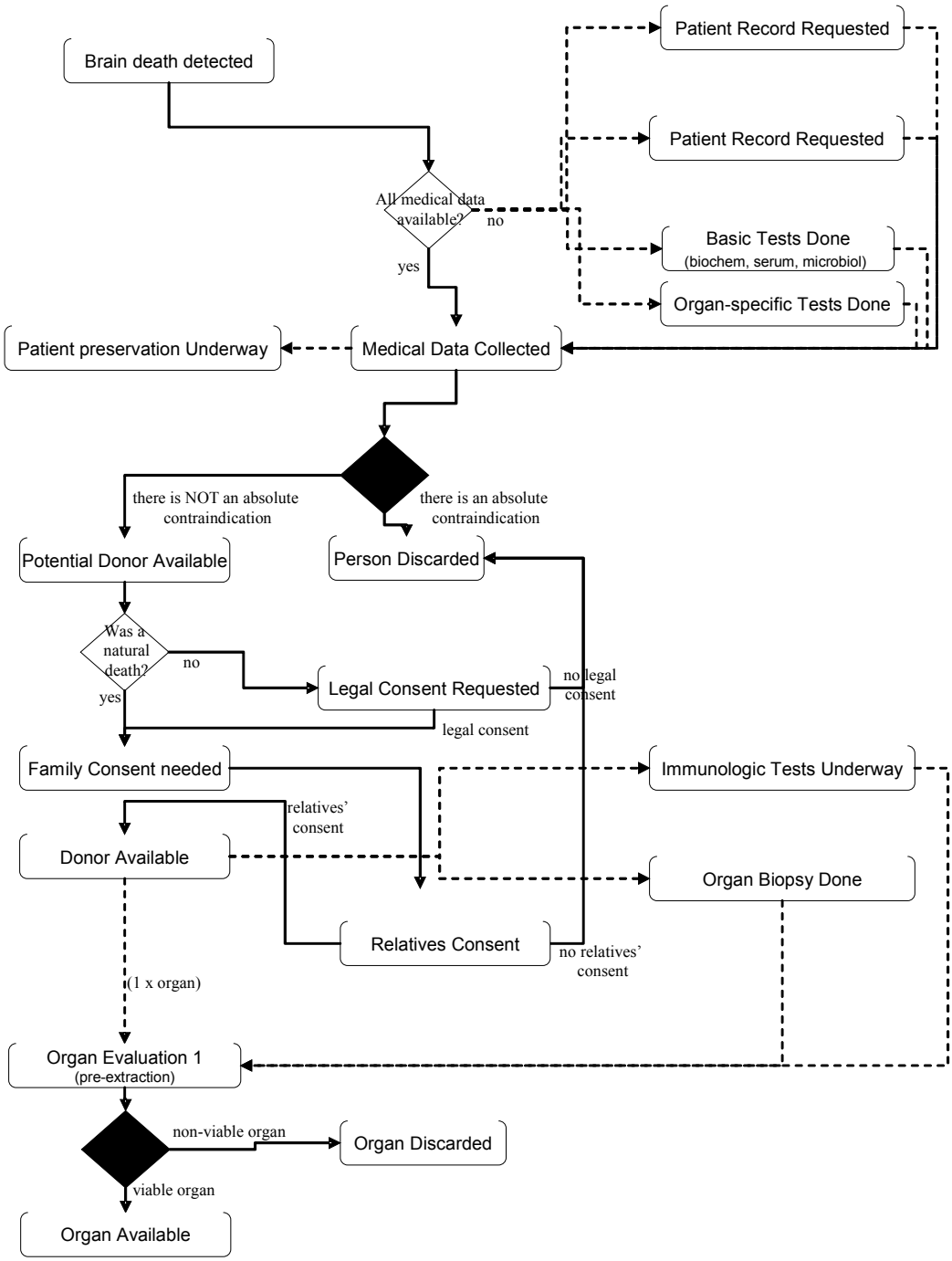

Figure 4 - Workflow for the donor detection and evaluation process.

2.4.3 Allocation Process

The next step in the overall process is to search for suitable recipients. The OTA carries out a recipient search for each organ that may be available by calling all hospitals with information about the organs. To speed up this search process, each organ is assessed separately with reference to well-defined the distribution criteria. These criteria are divided into clinical and geographical criteria. The clinical criteria are established and reviewed every year by all the transplant teams and ONT representatives, whereas the geographical distribution criteria are established by the Interterritorial Council of the National Health System. In this way, Spain is divided into six areas (Figure 5).

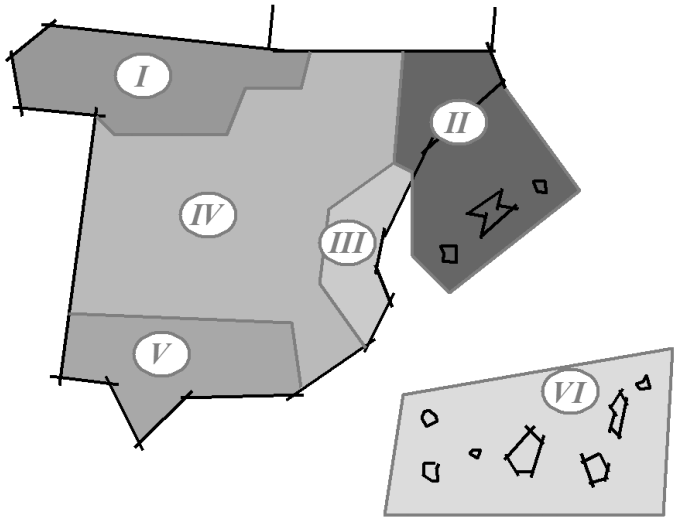

Figure 5– Zones defined in Spain.

As organs are a very scarce resource, recipient waiting lists use to be quite extensive in all centres. This also leads to the fact that for a given organ, almost any centre has a potential recipient. Therefore distribution criteria have been defined in order to enforce a territorial equilibrium in the distribution (if possible). Current distribution criteria can be summarized in the following steps:

1. *Urgency-0 cases:* The OTA and the centres keep an updated list of maintain a list of highly urgent cases, which become a national priority over any other recipient. If there are urgency zero cases with sufficient compatibility (blood group) then the assignment is made to those.
2. *Back to the extraction centre*: If there are no suitable urgency zero recipients, the OTA offers the organ back to the extraction centre – at which point the transplant team head must accept or decline based on his/her list of patients.
3. *Local, regional and national turns:* If the extraction centre refuses, a multi-level round robin system is used to call other possible implant centre to find a potential recipient: first the other hospitals in the same metropolitan area, then hospitals in the same Autonomous Community, in the same zone or in the whole country (see Figure 6). For each level, a turn rotation mechanism is used: each time a team transplants an organ using his "turn" for it, the team goes to the last place of the rotation.
4. *International offer:* if no centre accepts an organ, search goes international, by contacting other transplant organizations.

Each time a centre receives an organ offer, all the donor's data is provided together with the conditions established by the generating hospital, particularly with regards to the time of removal and other possible requirements. The team that is to perform the implant makes an assessment, analysing the information available about the organ, and decides whether the removal and implant can in fact be performed. The informational background for such assessment is derived from the patient care record, immunology analyses of the donor / donated organ and the physician's own knowledge (of any of the patients). If the offer is turned down, it goes to the next centre according to the distribution criteria.

If the offer is accepted then the delivery stage starts: the donor's hospital is informed, transport is organized (ambulance or helicopter from one hospital to one nearby, to the airport or to a train station; plane and/or train from one city to the other) and the main process schedule is defined, including the delivery plan (which must take into account transportation system schedules). How the OTA is involved in the organisation of transport depends on the distance to be covered and if it a removal team has to be mobilised or just the organ has to be carried. In any event there always exists a minimal involvement even if it is only as a point of reference between the generating hospital and the implanting one regarding transport details.

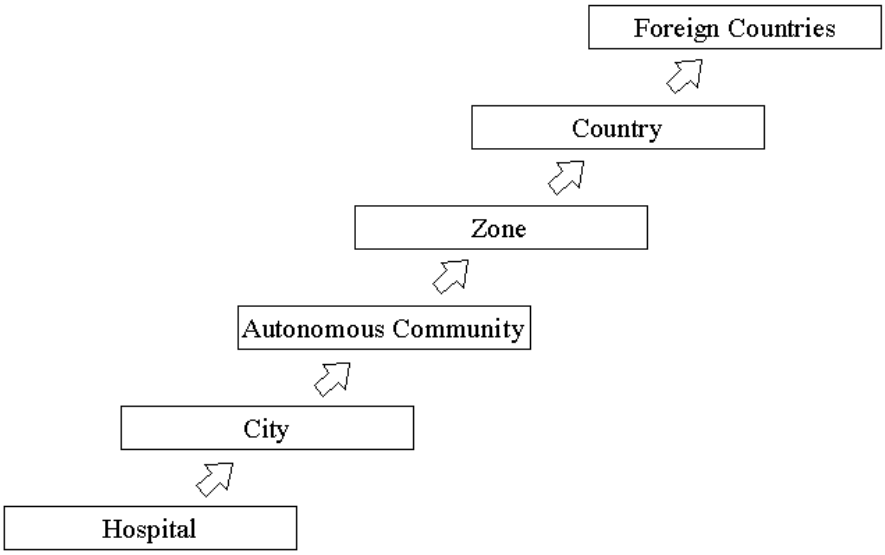

Figure 6– Distribution criteria for organs in Spain since 1996.

- *Local donor:* A donor who is in the same city as the extracting/implanting team, but in another hospital. In this case the hospital coordinator organizes the transport according to the internal agreements among both hospitals.

- *Non-local donor:* If the donor and the extracting team are in different cities, ONT staff will organize the transfer and that in general becomes a more complex operation. There are two scenarios:

    - *Short distances*: in distances less than 200 km., teams are usually carried in ambulances or helicopters. If necessary the collaboration of the police is requested to open the way. On occasions military helicopters / landing facilities may be used and on other occasions they are civilian, normally belonging to the civil protection services provided by the Autonomous Communities. These means of transport are used as long as the climate and schedule permit.

    - *Long distances*: given the short period of physical ischemia that is tolerated by the organs, private aircraft are used for this type of distances and occasionally the help of the Air Force is required. At this point, it should be taken into account that the preparation of a flight requires at least two hours, (checking the aircraft, calling the crew, establishing the flight plan, etc), which is why it is so important to advise the ONT of the existence of a donor as soon as possible. When the flight plan is ready both hospitals are informed of the schedule, the company and the flight number. It is very important to inform of the number of people who are travelling to arrange for sufficient air and land transport from the airport to the hospital.

Once the removal team arrives at the hospital, the ONT staff waits to be informed of the implant so that the patient is immediately removed from the waiting list. The hospital coordinator from the generating hospital will later send the ONT the donor's registration sheet duly completed with all the transplant performed.

Figure 7 shows the complete workflow for the allocation process. It is important to note here that it only includes events and decisions that are important form a medical perspective (all the details about how transportation of the organ is organized from one hospital to another have been omitted). When

an organ is available the offer is registered by the OTA. At this point the search for recipients begins, taking into account the distribution criteria mentioned above. According to such distribution, the organ is offered to one or several extraction teams (one at a time), which will evaluate the offer and decide if they will accept or reject the organ. If the organ is rejected, then an offer is sent to the next extract team, following the distribution criteria.

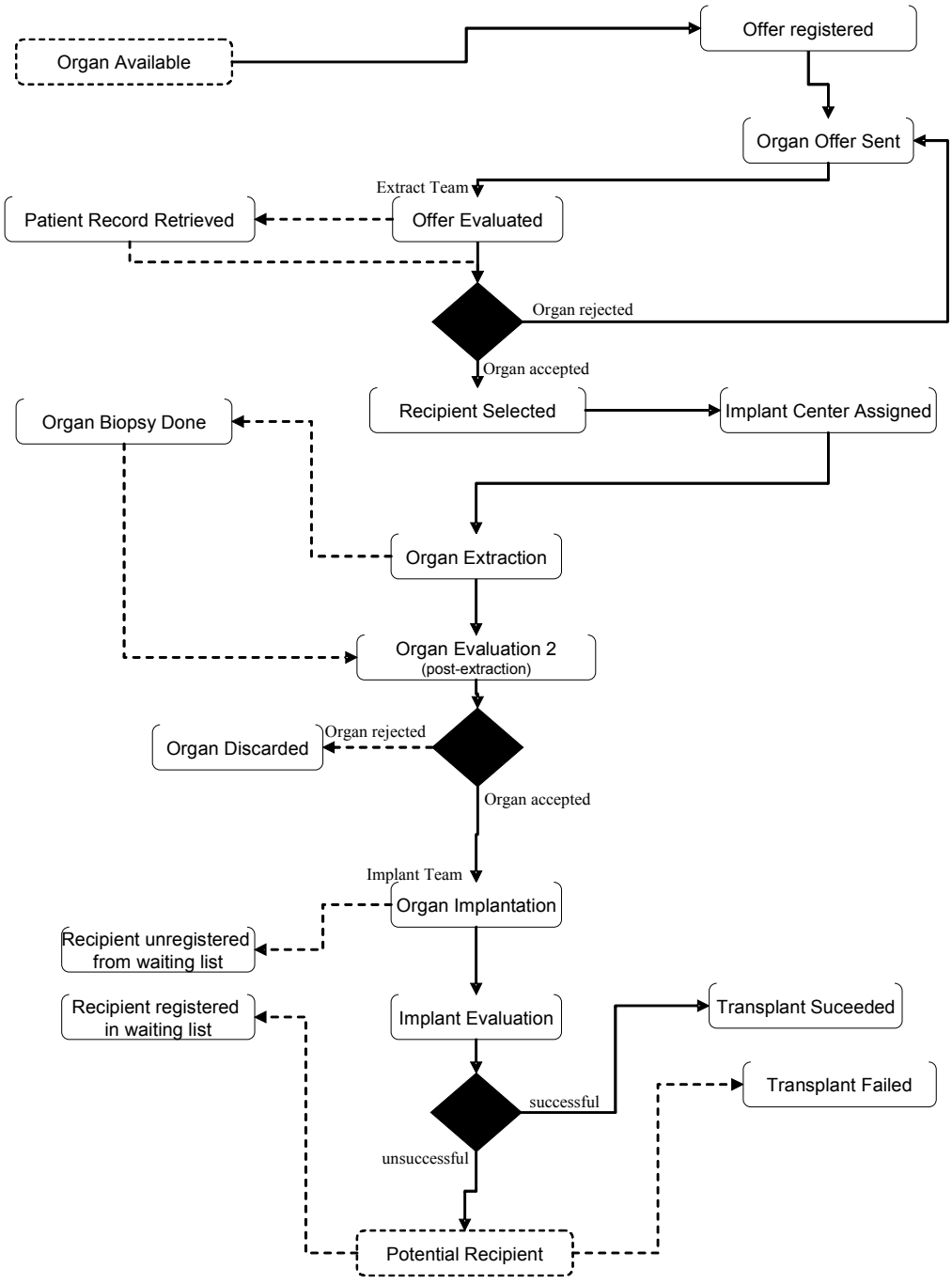

Figure 7- Workflow for the allocation process.

The decision that the extraction team should take each time an offer arrives is one of the more complex ones, and can be summarized as:

> *"Given an available organ **x**, which patient **y** from the set of potential recipients **Y** should be selected as the recipient?"*

Important factors which impact on this decision are:

- *Which recipient has the best medical chance of successfully accepting a given organ?*
  - How good is the clinical match of **x** to each **y** in **Y**? In terms of major and minor factors such as ABO blood type, age of donor and recipient etc.
  - Are there additional compatibility issues? (For example, the donor is infected with a given virus such as HIV or Hepatitis B/C – in which case recipients also infected with this virus may be able to receive it whereas for those not carrying these viruses implantation would carry a risk of transmission).
  - Are there additional surgical, logistical etc. issues which would worsen / improve the chances for one or other of the potential recipients? (For example, if one of the recipients is immediately available and another is not or one of the recipients is located a great distance from the donor.)
- *Which recipient is in most urgent need of a transplant?*
  - Is any of the potential recipients in danger of imminent death if they do not receive a transplant?
  - Which recipient's quality of life stands to improved by the greatest margin by a given organ? For example, are there familial / social circumstances (for example, financial hardship caused by inability to work) which constitute extenuating circumstances.
- *Which recipient has been waiting for the longest period of time for a given organ?*
  - Which potential recipients were added onto the waiting list earliest for a particular transplant / transplant center?
  - Where is the potential recipient registered? In other words, is the recipient on the waiting list of the retrieval center? Of a center in the locality? Further afield?

If the extraction team accepts the organ, then both the recipient and the implant center are automatically selected. The extraction team then travels to the hospital to extract the organ, and makes a further evaluation of the organ once it has been extracted. At this point damage or other pathologies may be detected that lead to the organ being discarded. If the organ passes this second evaluation, the extraction team carries the organ back to their hospital, where the organ will be implanted. If the implant is successful, then the transplant has succeeded, if not, the recipient becomes a potential recipient (usually at urgency-0 level) and he is re-registered in the appropriate waiting lists.

2.4.4 Recipient follow-up

In Spain all recipients of an organs are regularly assessed to determine their general health and the functional state of the organ(s) transplanted. Figure 8 shows the workflow for this process. On the left hand side, the periodic evaluation process is depicted. After tests are carried out, an expert assesses if the organ is being rejected or not, and whether it is working within acceptable boundaries. If not, the recipient is again registered as a potential recipient (and connects with the “Potential Recipient” state in Figure 3.

The right hand side of the diagram in Figure 8 shows the follow-up carried out by a general practitioner, who in principle only prescribes the treatment (following the hospital's evaluation) rather than making new diagnoses.

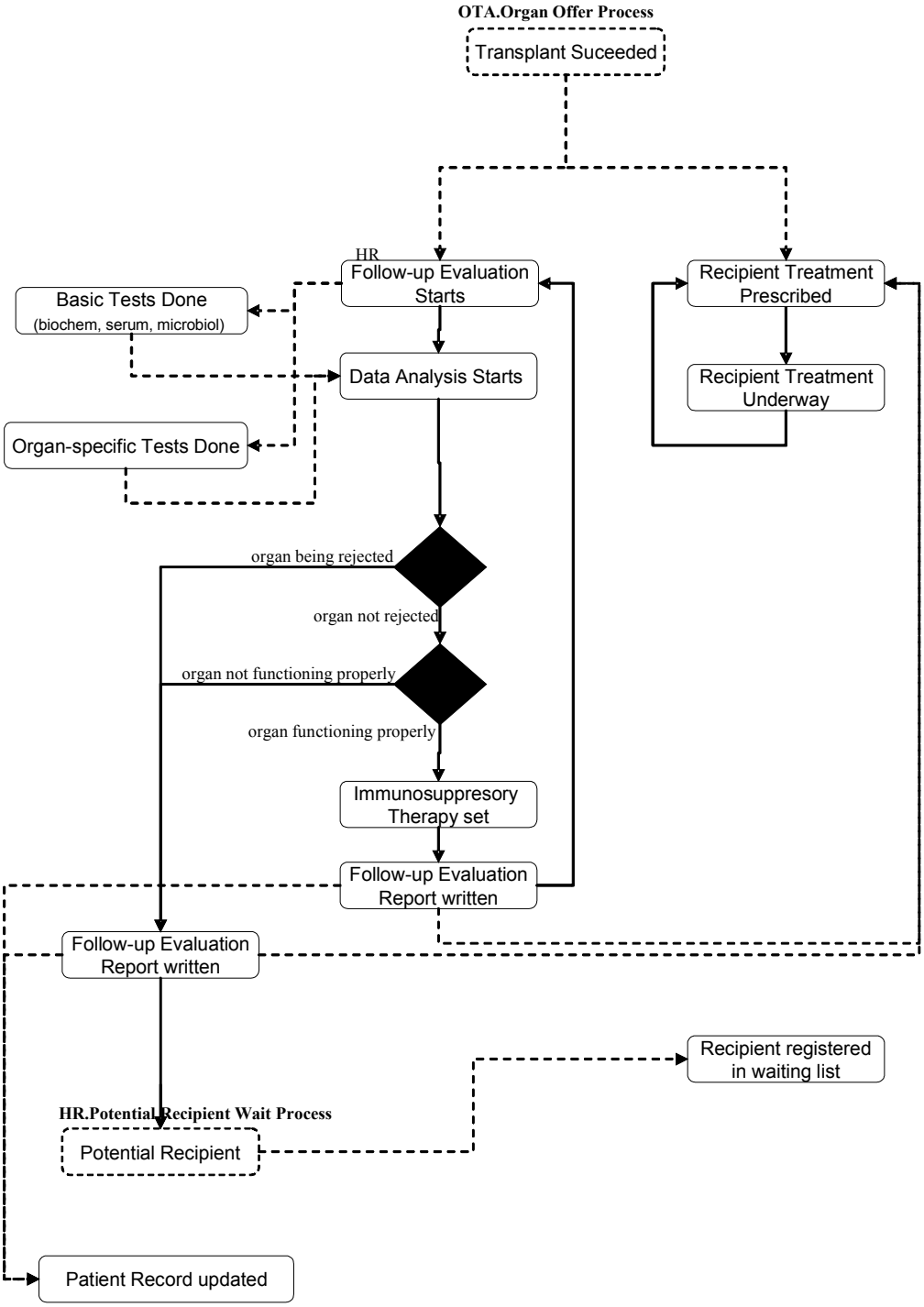

Figure 8 --- Workflow for the recipient follow-up process.

2.5 Summary

This chapter provided an introduction to the organ allocation process problem in Spain. Although procedures vary across national boundaries many of the issues addressed are similar elsewhere and results transferable. Description here covers the process itself, the actors and the important events and decisions to be made, represented in detailed workflows. These workflows will now be used as base model of the allocation problem for the OTM application, described on Chapters 3 and 4.

3 Application System Architecture

As discussed in the previous chapter, treatment of patients through the transplantation of organs or tissue is one of the most complex medical processes currently carried out. This complexity arises not only from the difficulty of the surgery itself but also from a wide range of associated processes, rules and decision making which accompany any such surgery. Depending on the country where a transplant is being carried out, procedures and the level of electronic automation of information / decision making may vary significantly. However, it is recognized worldwide that ICT solutions which increase the speed and accuracy of decision making could have a very significant positive impact on patient care outcomes.

Electronic systems that might be implemented for transplant management can be divided into two main types:

1. *Transplantation Management*: information systems used by medical staff during the process of a transplant incident (a single patient receiving an organ or tissue) to access existing case or background data, share it with colleagues, carry out matchmaking and/or otherwise provide decision support. This also includes long term, post incident data analysis techniques able to extract aggregate information such as general trends over large sets of previous transplant case records.

2. *Medical Record management*: the storage, access and modification of medical patient care records for patients in a given geographic region. Gathering, access and modification of such data is regulated by European, national and regional laws and forms an underlying information system for any treatment process management system.

Each of these types is considered separately in this section, however in the final demonstration application both will function together, with the OTM application directly accessing and making use of the EHCR functionality.

3.1 Transplantation Management and Post-processing: The OTM Application

The OTM application involves a large number of individuals, units and administrative domains – each of which provides different services which must be combined to carry out the whole procedure. Figure 9 summarizes the different administrative domains (solid boxes) and units (dashed boxes) that are involved during a transplantation management scenario. Each of these interact with each other through Web Service interfaces (circles)[4] that send or receive messages. The Organ Transplant Authority (OTA) is an administrative domain with no internal units. In a transplantation management scenario, one or more hospital units may be involved: the hospital transplant unit, one or several units that provide laboratory tests and the unit that is responsible for the patient records (which will use the EHCR application services). The diagram also shows some of the data stores that are involved: apart of the patient records, these include stores for the transplant units and the OTA recipient waiting lists (WL). Hospitals that are the origin of a donation also keep records of the donations performed, while hospitals that are recipients of the donation may include such information in the recipient's patient record. The OTA has its own records of each donation, stored case by case.

More specifically, Figure 9 shows that a transplant management scenario starts with a potential donor in Hospital A's transplant unit. In order to evaluate the donor, this unit may request the patient records inside the hospital and order a number of tests, some of them to internal laboratory units and others to some specialized external laboratories. Once the donor is evaluated and, if valid, the transplant unit contacts the OTA, which sends first the offer to hospital C. As the transplant unit in hospital C rejects

4 See section 3.1.4 for more details on the architecture of the Web Service interfaces.

the donation, the OTA sends the offer to hospital B, which has a potential recipient for the organ offer (as in the case of Hospital A, all the medical data needed for the recipient was previously collected by hospital B by interacting with the ECHR application and the testing laboratories). During extraction and implantation, direct communication between hospital A and hospital B and also between the OTA and the hospitals occurs.

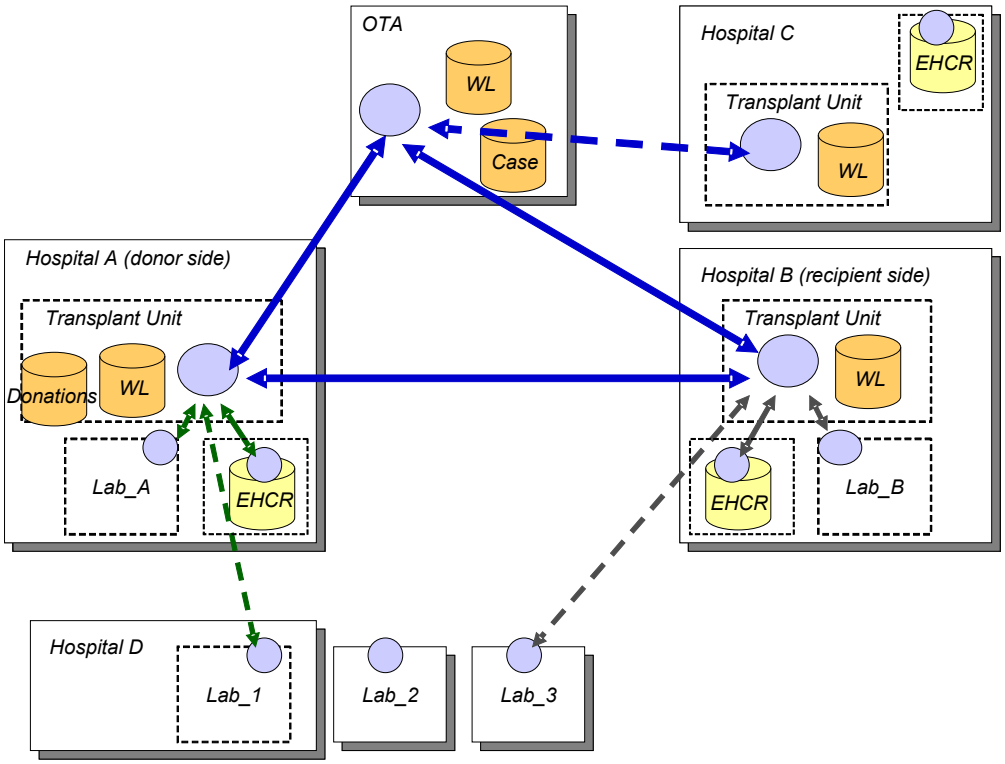

Figure 9 -- Mapping of the recipient detection and evaluation process workflow into Web Services.

In the following sections we will provide a detailed description on how transplantation management is distributed through the actors and services in the OTM application. Section 3.1.1 Identifies the actors and services provided, as well as mapping them to the workflows presented in Chapter 2. Section 3.1.2 covers the data stores that each actor may use. Section 3.1.3 defines the two kinds of cases to be managed during transplantation management. In Section 3.1.4 the architecture of the Web Services is described, and Section 3.1.5 provides details of the deployment of the OTM demonstrator.

3.1.1 System Services

In this section, the administrative domains, units and individuals in the OTM application are mapped into an Information Technology (IT) infrastructure such that:

- Each real world process is mirrored in the IT world by one or more services, which carry out the mirrored process. If the real world process involves the activity of an individual or a team, one person involved in the real world process will interact with the service(s) in the IT world in order to introduce or query information, by means of a Graphical User interface (GUI). Each of the services also serve as a representation of the individual/team in the IT world.

- These services are grouped together into units that intuitively reflect the organisational units in the Hospital / Health care system.

- These services are distributed across all possible participants in the Hospital / health care system (that is, each Hospital with a certain type of laboratory is modelled as having its own IT service to represent this laboratory).

Figure 10 shows a section of the modelling diagrams used in this section. Throughout this description, the meanings of the different elements of the diagram are as follows:

- *Dark Boxes*: correspond to Web Services representing individuals/teams in the OTM application). Each Web Service is identifiable by a name that includes the administrative domain or unit the service belongs to.

- *White Boxes*: represent a relevant event in the execution of the service, to be recorded in Provenance stores.

- *White Diamonds*: are minor conditional points in the workflow. The diamond is labelled with the condition to be checked. Lines that exit the diamond are labelled with the response needed to follow each line.

- *Black Diamonds*: are major decision points always taken by a human expert. All decision points are vital in the process and all of them are to be recorded in the Provenance store.

- *Full Lines*: sequential workflows between white boxes and/or diamonds.

- *Dashed Lines*: Parallel workflows between elements (e.g., in the diagram, the requests for the patient records and the lab tests are not done sequentially but all at the same time).

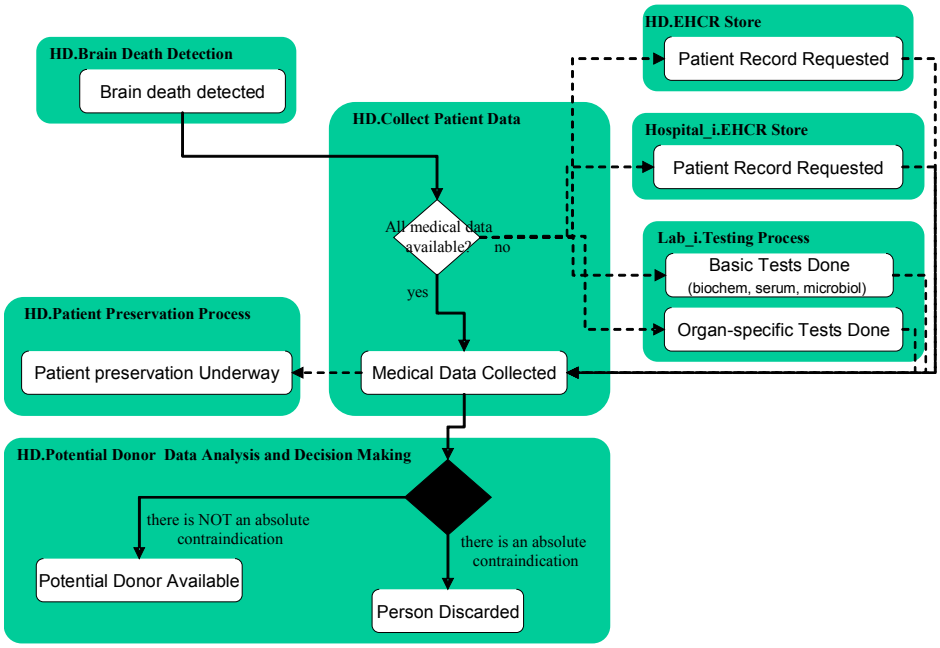

Figure 10-- Service Diagram showing Dark Boxes, White Boxes, Diamonds, Lines and Labels.

The following section lists the individual Web Services in the system which represent participants. The figure in brackets after the name indicates the expected number of such representatives in any

one deployment (1 for a unique instance, n for multiple instances).[5] Some of the services are informational in nature (such as patient care record retrieval) and some are real-world / medical in nature (such as patient preservation). For real-world / medical processes an proxy information / electronic process is created to support and record data from the real-world process.

It is important to note that the actors models do not generally represent individual *human beings* but *processes managed by teams of humans* (e.g. a surgery team). Modelling individuals is important for data access / rights / responsibility, however this is managed at a different level to the IT actors/services themselves (for example an actor may record the humans responsible for a particular execution run of the service).

Per Hospital with transplant facilities [n] (donor side)
- Brain death detection [1] (this is the process which kick starts all the others – note that "heart first" death cases significantly change work flow and are not currently considered here.)
- Electronic Health Care Record (EHCR) store [n][6]
- Testing lab / testing process [n]
- Potential donor data analysis & decision making process [1]
- Patient preservation process [1] (run once per potential donor)
- Legal consent checker [1]
- Family consent checker [1]
- Pre-extraction organ evaluation process [1] (run once per potential donated organ)
- Collect Patient Data process [1]

Per External Lab [n]
- Testing process [n]
- Biopsy Process [1] (optional on request – per organ)

Per Organ Transplant Authority [1]
- Organ availability/offer registration [1]
- Organ offer process [1] (offer generation and tracking – 1 process per organ)
- Potential recipient waiting list [1]

Per Hospital with transplant facilities [n] (recipient side)
- Organ offer process [1] (evaluation)
- Potential recipient waiting list [n]
- Electronic Patient Care Record store [n]
- Organ extraction process [1] (run once per potential donated organ at the extraction site)
- Post-extraction organ evaluation process [1] (run once per potential donated organ)
- Implantation process [1]
- Implantation evaluation process [1]
- Implantation post-operation medical evaluation process [1]
- Patient treatment process [1]
- Patient preservation process [1]
- Patient consent checker [1]
- Patient care record store [1]
- Patient registration [1]
- Collect patient data [1]
- Potential recipient data analysis & decision making process [1]
- Potential recipient wait process [1]

5 Note that replication for robustness is not considered here – for replication purposes arbitrary numbers of actual service instances may be deployed.

6 EHCR stores serve as the primary data store for patient medical records in the system and are therefore modelled as an independent service in this document – a description of the EHCR application is given in Section 3.2.

General Practice Post-Care [N] (recipient side)

- Patient post care process [1][7]

In the sake of brevity, each of the services listed above is described by mapping them to the OTM process workflows described in section 2.4. Figures 11 to 14 show such mapping.

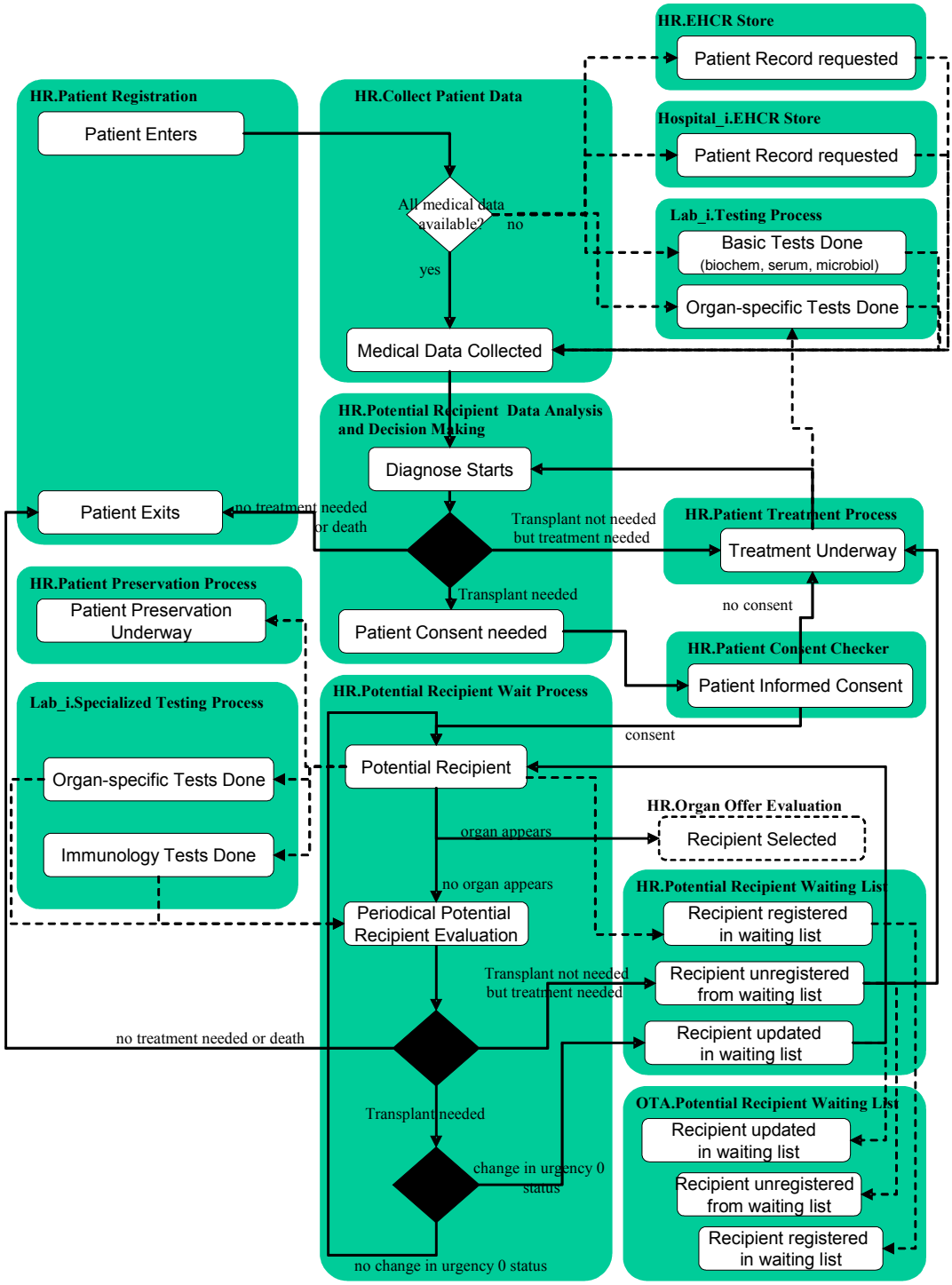

Figure 11 -- Mapping of the recipient detection and evaluation process workflow into Web Services.

7 Note that the hospital post-operation medical evaluation process and the GP post-care processes run in parallel but are not the same. The former is a regular check up on progress, the second is the implementation of post-operation prescribed treatment.

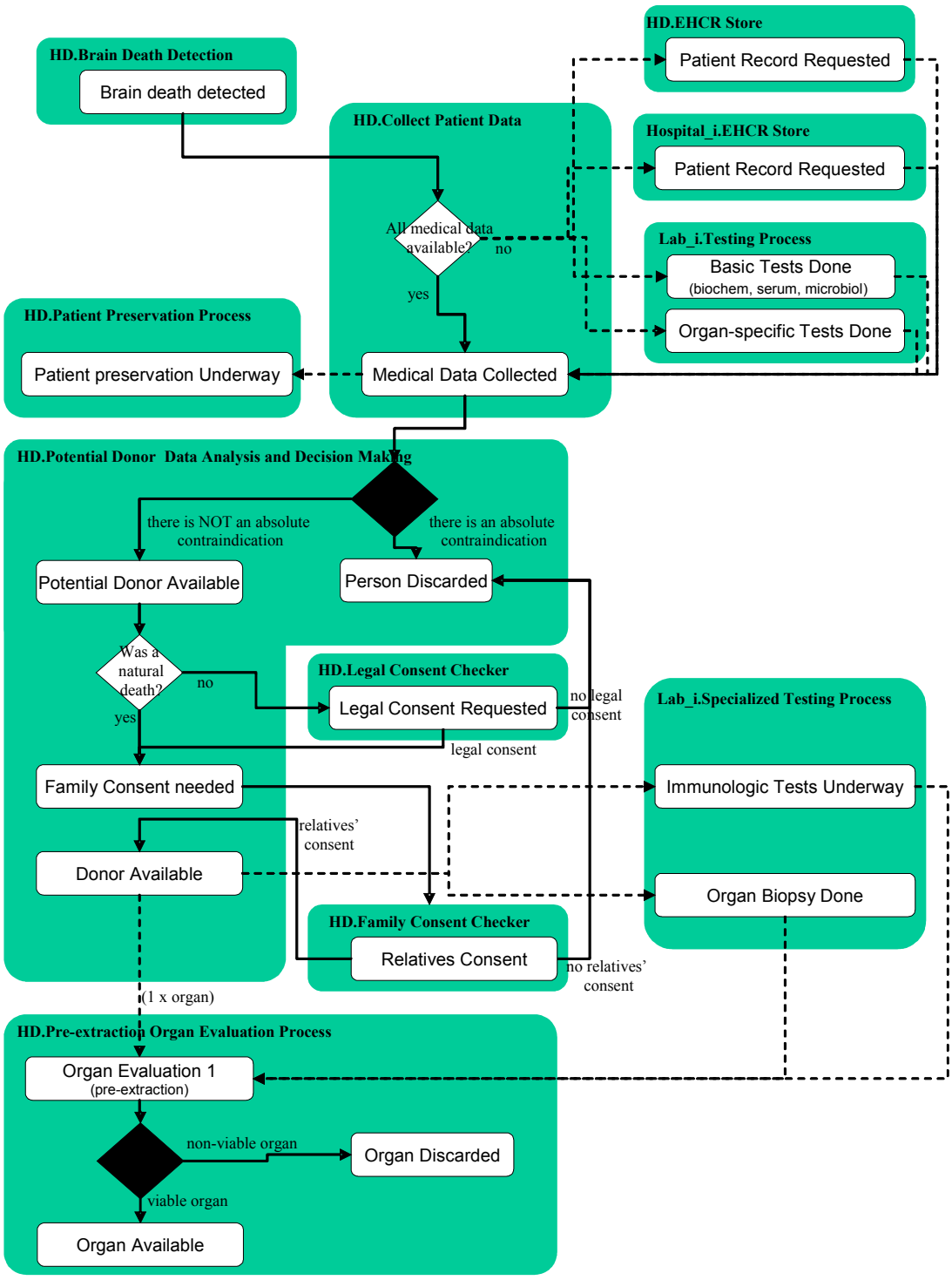

Figure 12 - Mapping of the donor detection and evaluation process workflow into Web Services.

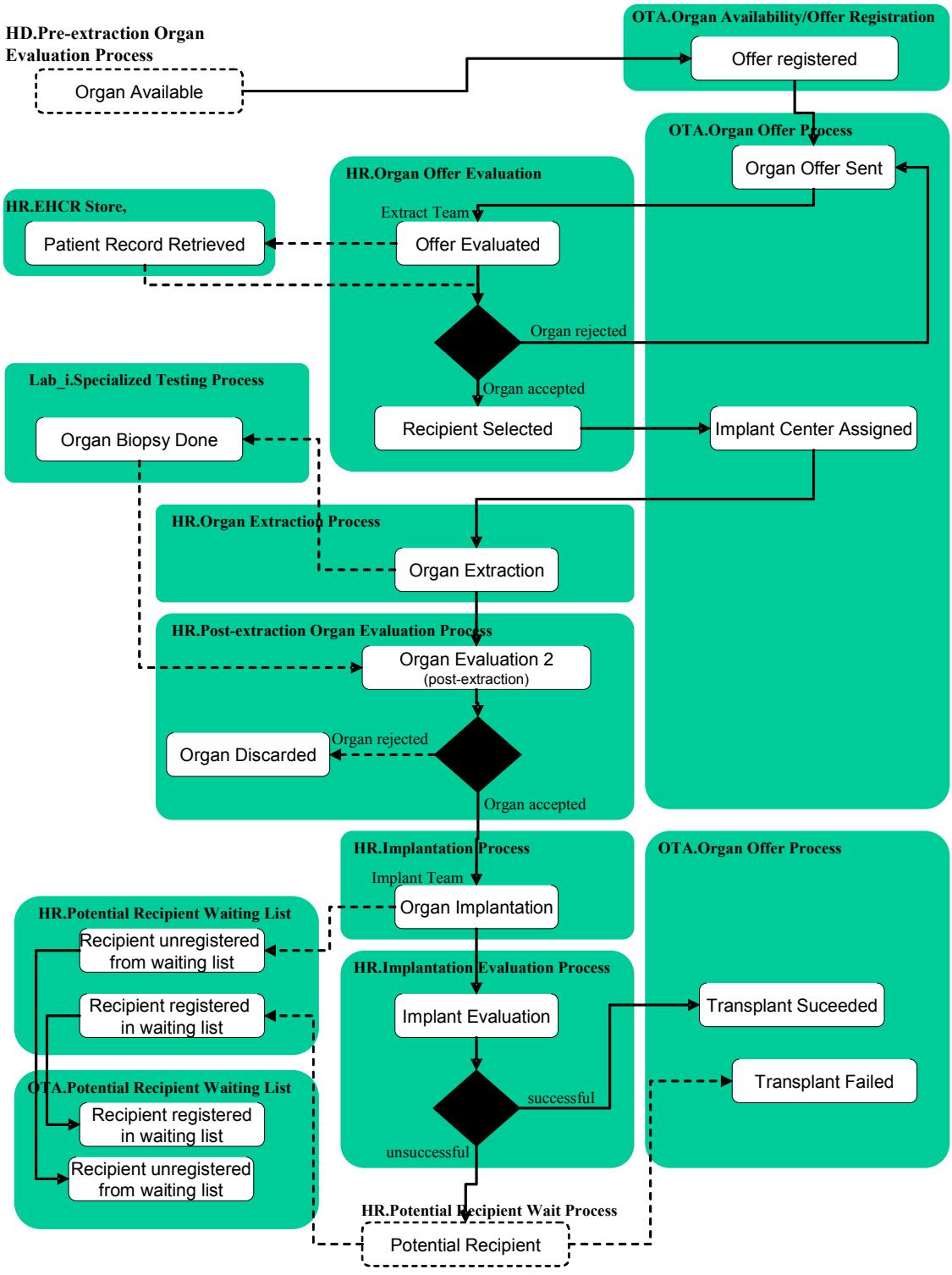

Figure 13- Mapping of the allocation process workflow into Web Services.

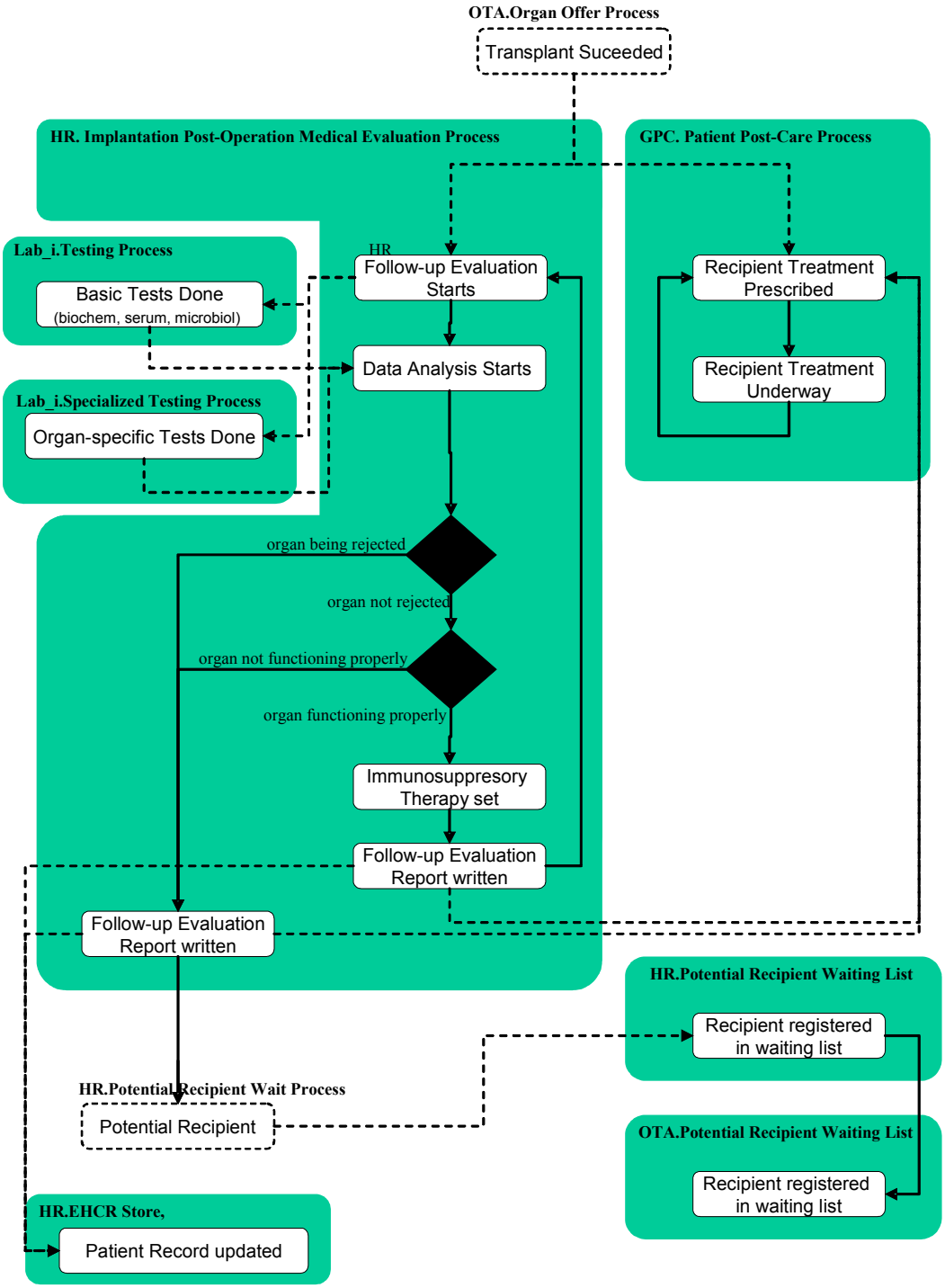

Figure 14- Mapping of the recipient follow-up process workflow into Web Services.

The organisational entities which own these services are divided into several levels: 1) individual laboratories/units within hospitals, 2) individual hospitals, GPs or the transplant authority as legal entities, 3) the Catalan Health Authority (CatSalut) and ultimately 4) the Spanish national health authority. Each service, when executed, runs in a specific local context that:

- Captures which health care staff are currently present and/or responsible for operation.
- Captures date / time / location as appropriate.
- Captures the *case* a particular execution is associated with.
- Captures any data associated with the case that is necessary for the execution of the service.

During the execution of the system, services are activated to carry out their function and retain a kind of "virtual token", which, together with the presence of the medical staff involved in the real world process,[8] denotes that the service is "active". As the medical procedures progress the token is moved between station to indicate the current status of the workflow – activating services in turn. Once they are activated, almost all the services in the OTM application are driven by the user. Staff sign-on / sign-off procedures move the token along the work flow to represent the current active point in the process (note that where processes run in parallel several services may be active at the same time). This virtual token may also be used as a case identifier label in Provenance stores (see Section 5.1.3.1).

3.1.2 Data Stores

As a rule of thumb, every instance of a service described in the previous section (e.g. every hospital biopsy testing centre instance) will have *its own data store* which:

- Formally archives the reports / decisions / permissions documents generated by the actions the service carries out.
- Is considered the authoritative location for the master copy of the original record datum (even if other copies exist elsewhere).
- May be a database containing structured records, a database containing (for example) PDF electronic documents or in some cases a database containing only administrative data about the existence of an off-line document.[9]
- Has access rights and permissions set by the entity which owns the service (a particular hospital, the medical authority, etc.), often in response to legal regulations.

In some cases, these data stores will be large complex systems (e.g. a hospital patient record system). In other cases they may be simple local activity logs. In order to save space the following presentation presents just several of the major data stores the system is concerned with, others may be added on a per-need basis.

Per Hospital with transplant facilities [N] (donor side)

- Patient Care Record Store [n]
  - Owned and managed by the hospital
  - Access from outside the hospital X by another hospital / medical centre Y only if the patient is currently physically being treated at Y.
  - Access only by authorised medical personnel. Access is logged. Certain data (such as HIV status) requires additional checks.
  - Extensive structured database capturing different types of reports, test results, medical attributes.
  - Currently do not follow EU standards
  - [A particular patient may have records or fragments of records in multiple places – see section on patient care records.]
  - The OTM application will use the ECHR application for storage of all medical data (see Section 3.2.2.1.).

- Testing Lab [n]
  - Depending on the type of lab [Microbiology, immunology, ...]
  - Test databases which record the result of each test carried out and result delivered.

8 An example would be duty surgeons for an operation logging into a hospital theatre console using coded keys (dongles) to activate the surgery recording process.
9 Example: a signed family consent form.

- Often currently recorded as PDF documents which are archived electronically and on paper.[10]

- Medical Data Analysis / Decision making [1]
  - Record keeping for the decisions taken for a case – in particular, the duty surgeon's report about decisions with respect to a particular donor.
  - Often currently recorded as PDF documents which are archived electronically and on paper.

- Family consent checker [1]
  - Process for checking with family whether organs may be reused.
  - Archive of responses (positive or negative) including archive of signed consent form (paper).

Per Organ Transplant Authority [1]

- Organ availability/offer registration [1]
  - Once decisions on a donor have been made a donation dossier containing test results, available organs, patient details etc. is sent to OCATT. These arriving dossiers are archived as inputs.
  - Organ transplant dossier currently recorded as PDF documents which are archived electronically and on paper.

- Organ offer process [1] (offer generation)
  - On the basis of each dossier an offer process passes the dossier to hospitals using a round robin sequence. Records are kept of responses and in particular of the allocation made.
  - Currently recorded as PDF documents which are archived electronically and on paper.

Per Hospital with transplant facilities [N] (recipient side)

- Organ offer process [1] (evaluation)
  - Counter-part to the OTA offer process which stores incoming dossiers and decisions.
  - Currently recorded as PDF documents which are archived electronically and on paper.

- Potential recipient waiting list [n]
  - Ranked list of potential recipients for each organ type along with links to updated medical records. Regularly refreshed using a different work flow. Updated if treatment is carried out on any individual.
  - Electronic database.

- Implantation process [1]
  - Surgical reports database archiving team recorded results for surgery.
  - Currently recorded as PDF documents which are archived electronically and on paper.

General Practice Post-Care [N] (recipient side)

- Patient post care process [1]
  - Additions to patient medical record and (in some instances), database to record post care evaluation results for statistical purposes.
  - Electronic database for patient records, additional reports – variable formats.

More details on the clinical data that is stored in the OTM application can be found in Appendix A.

10 Note that the description here refers to current practice. The demo system will automate this with electronic storage and structured data.

3.1.3 Notion of Case in the OTM Architecture

In medical records currently kept for transplants in Catalonia there is no overarching "case file" that captures all details of a transplantation. The different elements of the process (recipient waiting lists, donor decision making, transplantation and post-care) all run separately. However, there at least two useful notions of *case*:

1. The process starting with the brain death report of a potential donor and terminating in a single or multiple organ implant in a single recipient (that is only two patients involved). (Labeled the *recipient perspective*.)
2. The process starting with the brain death report of a potential donor and terminating in all potentially transplantable organs of the donor being evaluated and/or implanted (that is potentially n patients involved). (Labeled the *donor perspective*.)

In the OTM Provenance architecture the word *case* is important since it is strongly linked to the need for identifiers in Provenance stores (see Section 5.1.1.4). In order to clarify the relationship between donors and recipients the current application will make use of 2 different notions of case:

1. *A donation case*: corresponding to the second of the above definitions – starting from the donation and including all possible recipients. This is managed by generating a donation case label which is carried forward with every new activity.
2. *A recipient case*: corresponding to the time between organ extraction and the final Health care outcome for one recipient. Thus for each organ reused from a donor, one new label is generated which propagates forward only from then one.

In general, these two case labels will be used in parallel once both are available, such that p-assertions on a recipient can be associated both with the donation and the recipient case. See Section 5.1.1.4 on the naming convention for cases.

3.1.4 Architecture of the Web Services

Via the Spanish national FIS CARREL project, core Web Services functionalities, database access and client facing services have already been developed with a combination of open source Web Services tools. In particular these include: Apache Web Servers, Jakarta Tomcat, the AXIS SOAP toolkit, and MySQL/PostGres databases (see Figure 15).

The work done for the Provenance OTM demonstration will extend this existing base using similar technologies and adding a second layer of services based on a generic agent like model [FIPA02a, W3C05] which has been adopted in FIS CARREL. Concretely this means:

- Participating Web Services will be modeled and implemented as having persistence (and a persistent identity in the application).
- Communication will be over SOAP/XML and HTTP as normal, however a limited generic interface corresponding roughly to the Agent Communication Language FIPA-ACL [FIPA02b, FIPA02c] will be used for all messages to all services. Using this device all messages will be generated using general classes of message such as *Inform, Request, Agree* etc. which characterize the type of message being sent.
- Complex application data to be transmitted will (where possible) be encoded in RDF (XML serialisation) and linked to associated domain ontologies in order to increase future re-use.

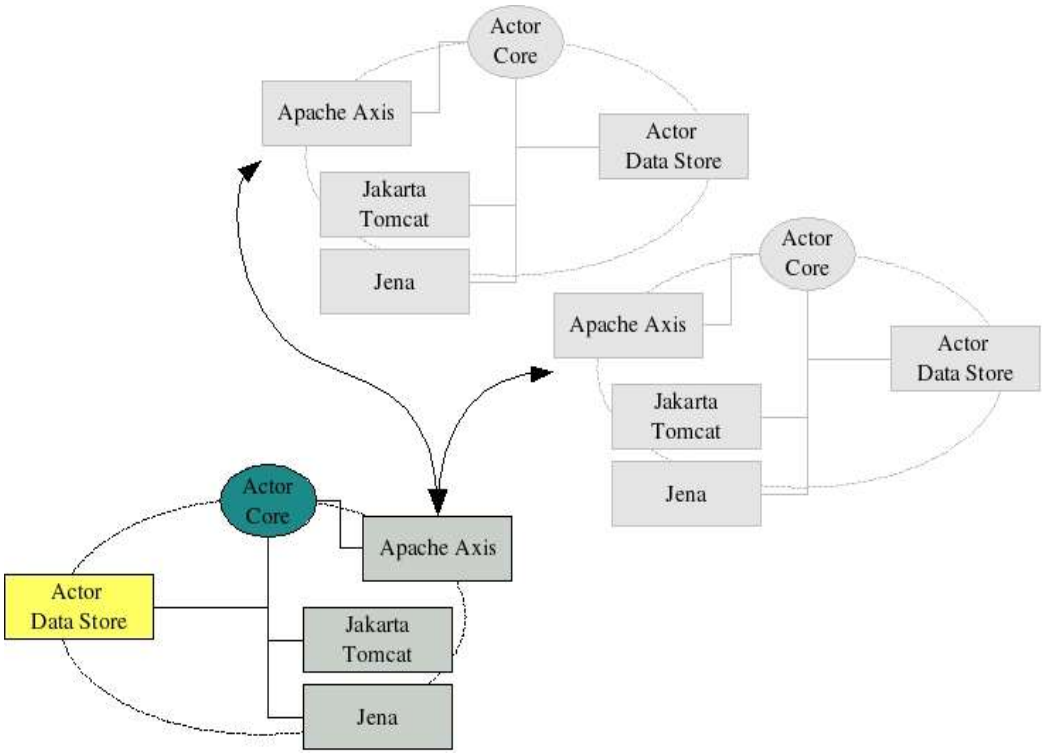

Figure 15: Generic Application Architecture

3.1.5 Deployment Details

The OTM over Provenance application is being developed in conjunction with the Spanish national FIS CARREL project and will therefore benefit from a joint demonstration deployment. The demonstration scenario will consider a system approximately the size of the Catalan OTM problem with:

- 1 health Authority (CatSalut).
- 1 Organ Transplant Authority (OCCATT).
- 3-4 Major hospitals (Each with internal transplant facilities in laboratories).
- 3-4 Minor hospitals (With no transplant facilities but with general patient care facilities).
- 3-4 external laboratories.

The demonstration to be shown will simulate the majority of these actors and use realistic but anonymous data. The final demonstration is expected to span at least two, possible 3 or 4 remote sites with 8-10 machines hosting different simulated services.

3.2 Medical Record Management: the EHCR Application

While European standards for the interchange of patient care data do exist they have not been widely adopted to date. Currently the standards have also not been used in the Catalan health care system which currently works as follows:

- Storing master copies data about individual medical interventions on a patient at the place where interventions are carried out.

- Most commonly a single GP oversees a patient's medical history and thus integrates interventions not carried out under his/her own supervision post event.
- However there is no standard process for forwarding medical details which might form part of the record to a central registry or a master copy of a particular patient's record.
- Information is retrieved from different health care providers on the basis of either the patients National Identity Number (DNI) or the brand-new CatSalut Identity Number, with more and more organizations moving recently to the latter.
- A health care provider A may only ask for record information from another provider B for a patient X if the patient X is physically being treated at A.
- Further, the database schemas used by different hospitals / health care providers are potentially different, although in Catalonia there is a plan to converge towards a common Catalan standard, still to be defined.
- Patients do have an entry in the CatSalut database if they are legal inhabitants of Catalonia. This record however, only contains the CatSalut Identity Number, and basic information of the patient such as name, date of birth and current address. It is important to note that no medical information is contained in these records. The Medical records are stored in the health institutions the patient is assigned to or has visited in an emergency case.

For the purposes of organ transplants, potential recipients are identified in waiting lists and depending upon their condition assessed via a variety of tests to ensure their medical records are A) up to date and B) held by the transplant hospital they are registered with. This ensures that if a donor becomes available, information on all potential recipients is readily available.

For the Provenance project there are essentially three options available in order do model the health care record element of the application:

1. Deploy a system mirroring the current one based on fragments of records in different places which can be pulled together to produce a unified view on demand (depending on the permissions of the viewer).
2. Deploy a system of a more centralized nature with a CatSalut based master record which can be read and written to by authorized health care providers (in a controlled fashion) and possible cached at a particular health care provider.
3. Deploy a hybrid system which stores fragments of data with providers but records high level events in a central master record.

In each case, the interchange protocol could be one of the new European standards, the approach chosen would then study the Provenance questions which arise for the health care records themselves. In order to fully explore these issues the EHCR application will pursue the 3rd option – distributed storage of records with some aggregation of information. This approach also best matches current best practice in many countries including other regions inside Spain.

We can observe that the application has two major parts: one which mainly involves individual health care related activities (the health care domain), and one which mainly involves the assembly of the full health care record of a patient. Activities in the health care domain produce and store data in a distributed way: master copies of data are stored at the place where a particular medical interventions is carried out. Fragments of records stored in different places are then subsequently pulled together on demand to produce a unified view. The formation of this unified view can be seen as similar to the creation of a Provenance document reflecting the health care status of the patient. As a result, the EHCR application and the Provenance system can be represented as shown in Figure 16.

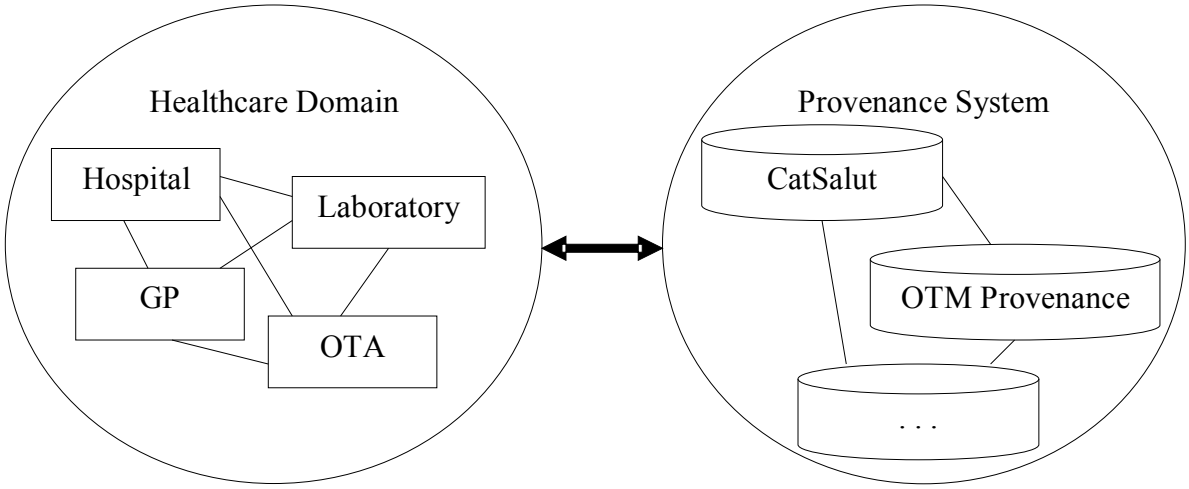

Figure 16– EHCR architecture for the OTM application using the Provenance system for pulling together fragments of the EHCR.

Every EHCR system stores all fragments of records that were queried before or are changed by actors associated with it. The Provenance system is then used to keep track of the location of other, related, fragments of records for query purposes. This scheme is based on the observation that the full health care record of the patient is the same as the Provenance document of the health care status of the patient (Figure 16). Essentially:

- Different medical procedures, tests, outcomes can be seen as intermediate results on the path to the full current health care status of the patient.
- Causal relationships exist between these fragments, such as a test result triggering an operation, further tests or a referral to a different medical centre.

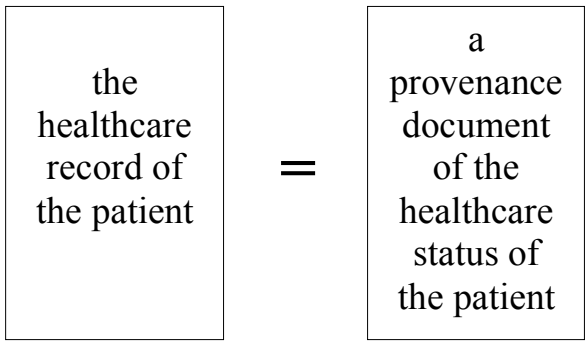

Figure 17– General mapping of EHCR to the Provenance architecture.

3.2.1 Overview of Standards for Patient Care Records

This section briefly introduces the European pre-standard for health care data exchange, called ENV 13606 [ENV13606]. While the format is not currently widely adopted it is one of the leading candidates for potential European wide use. The full specification is provided in the form of an XML DTD, however this section covers just subsets of the definitions used.

3.2.1.1 Messages

The standard defines three type of messages: `request`, `provide` and `notification`. All of these contain the fields *identification of message, issue date and time of message, EHCR source/destination* (see health care agent in this document), *urgency of message, patient matching information* (subject of the message), *message receipt acknowledgment request*. Besides these pieces of information, any message *may* contain *EHCR message related agents* (important health care agent (s) other than EHCR source or destination), *comments on message, language* (of the requested/provided EHCR), *health care agents directory* (see later in this document), *message references* (to a related message).

`request` and `provide` EHCR messages must (`notification` may) contain a *specification of EHCR information* (*language* and *completeness* of content of the EHCR information requested or provided). This information may include an identification of the nature of the *enterprise environment* and/or *communicating community* of the party sending the message (e.g. organ transplant management application). In addition:

- A request also contains a *reason for request*.
- A notification also contains *type and comment of the notification*.
- A provide also contains *distribution rule* directory and an EHCR.

3.2.1.2 EHCR

Every message is about one and only one patient and his/her EHCR. An EHCR consists of record components. The *simplest* instance of an EHCR consists of an `EHCR extract` (root component of the EHCR) class containing a single *text data item* with the *component role* "Narrative Text":

- The main types of the record components are *EHCR extract* (root component of the EHCR), `folder`, `composition`, `headed section`, `cluster`, `link set item` and the `data item`. There is exactly one `EHCR extract` in a provide EHCR message and zero in any other message. EHCR extract contains all the other record components in the message.
- A folder *shall not be a member of* any other record components.
- An *original complex* component is a collection of record components.
- The contents of an original component complex are collected at a given time and in a given situation and added to the EHCR.
- The original complex component is an abstract data type and can subsequently be specialized to form types such as `folder`, `composition`, `headed section` and `cluster`.

The `folder` is therefore one of the original component complex, the contents which may be data collected by different people, at different times and places (e.g. nursing notes, specialist departmental record). A `folder` *shall be a member of* one and only one of an EHCR extract or another folder. An EHCR extract and `folder` *may contain* folder, composition, link set item, text data item and empty record item. EHCR extract and folder *shall not contain* `headed section`, `cluster` and `data items` other than `text data items`.

A `composition` contains a set of record components relating to one time and place of care delivery, a single session of recording or a single document included in the EHCR (e.g. operation note, laboratory report). A composition *shall be a member of* one and only one of an EHCR extract or a folder. `Compositions` *may contain* headed `section`, `cluster`, `link set item`, `data items` and `empty record items`. Compositions *shall not contain* `folders` or `other compositions.`

`Headed sections` represent a sub-division within a composition, the contents of which have a common theme or are derived through the same health care process (e.g. examination, treatment). A headed section *shall be a member of* one and only one of a `composition` or another `headed section`. `Headed sections` *may contain* headed section, cluster, data item and empty record item. `Headed sections` *shall not contain* `folder`, `composition`, and `link set items`.

`Clusters` are used where it is necessary or desirable to group data items into logically or clinical associated collections. The data is collected by the same person in the same time and place (e.g. blood pressure measurement consisting of two data items (one for systolic and the other diastolic pressure)). A `cluster` *shall be a member of* one and only one of a composition, a headed section or another cluster. `Clusters` further, *may contain* `cluster`, `data item` and `empty record item` types. `Clusters` *shall not contain* `folder`, `composition`, `headed section`, `link set items`.

A `Data item` represents the smallest structural unit into which the content of the EHCR can be broken down without losing its meaning. A data item may aggregate information that cannot be safely disaggregated and can express the following things: person identification, person name, telecom data, address, external digital data reference, physical entity (e.g. a paper file, a sample) reference, structured coded data (machine readable form) (e.g. marital status, sex), medication, event (e.g. phone call, consultation, date of birth/death), language, patient related party (person or organization who has a role in relation too the care of a patient other than as a health care agent) information, result of a quantifiable observation (e.g. result of a laboratory investigation), text and other community defined information. A `Data item` *shall be a member of* one and only one of a `composition`, a `headed section` or a `cluster`. `Text data items` with the component role „Narrative Text" can also be a member of the EHCR extract or a `folder`. Data items *shall not contain* any other record component.

`Link set items` provides labeled links between one EHCR message component (nominated as the source component) and one or more other EHCR message components (nominated as target components). This structure is used to indicate relationships between record components other than those relationships that are determined by the original information context. Some examples of the nature of these relationships include:

1. Something is derived from / is source for something.
2. Something has caused / is caused of something.
3. Something has goal / is goal of something.

The source component and the target components *shall not be* `link set items`. A `link set item` furthermore *shall be a member of* one and only one of the EHCR extract, a `folder` or a `composition`.

A `selected component complex` is used to provide an alternate view of record components. It represents an aggregation of record components arranged in a manner not determined by the time and situation in which they were originally added to the EHCR.

There is one more record component, the `empty record item`. This record component is used to communicate the deletion of existing record component in update messages.

3.2.1.3 Health Care Agent

A *health care agent* is a health care person, a health care organization, a health care device (e.g. x ray machine, ECG machine), or a health care software component that performs a role in a health care activity. Health care agents are, for example, the sender / recipient of an EHCR message, the requester / provider of an EHCR, a person signing of a message or record entry, originator or author of a record entry. Relationships between two health care agents (such as employee / employer) can be defined.

The same health care agent can exist in different contexts (e.g. the same doctor working in different hospitals). A *health care agent in context* has a unique identifier, a reference to a health care agent, function (e.g. duty doctor, locum) and relationships to other health care agents. In terms of implementation, a *health care agents directory* may contain several health care agents (in context). Using this directory, the sender of a message need only include the full details of any health care agents (in context) once. A health care agents directory:

1. Can be part of each EHCR message.
2. May be communicated in a separate message (used for maintaining and aligning health care agents (in context) information).
3. Can be a shared or distributed directory.

People and organizations (called health care parties) have a name, address, telecommunication data, languages' details, medical specialties and types (types of people for example doctor, dentist, nurse, pharmacist; types of organization for example university hospital, private clinic). People also have positions (e.g. head of department), qualifications (e.g. MD, M.Sc.), military ranks etc.

Devices have type (e.g. computer, ECG machine), manufacturer, model name, version, serial number and location.

Software has product name, manufacturer, internal name, filename (of the file to start the program), version and date (creation or last modification).

3.2.1.4 Distribution Rules

With *distribution rules*, the provider of the EHCR (or somebody else) can define who, when, where, how and with what type of access can access a part of the EHCR. Rules can also be applied to add/invalidate distribution rules to part of an EHCR. In addition to this information, a distribution rule contains the necessary data to be able to identify the author of the distribution rule. A provide EHCR message may contain a *distribution rule directory* in order to keep distribution rules together. This directory can be both in the message and a shared directory.

To attach a distribution rule to a message component they must be used in conjunction with a *distribution rule reference*. These references contains information about by whom and when the rules should be applied to the message component, the interval of the validation of the rule, the country where the rule is valid and the reference to the health care agent in context who invalidated the rule within the period of time originally applied. If a distribution rule reference acts as a health care person's demonstration of consent according to an existing, applied distribution rule, then the distribution rule reference *shall contain a reference* to that distribution rule.

There are two functional flags in the distribution rule reference which are used as follows:

- A *Negation statement* informs the information system that the contents of the rule of which is a part shall both be interpreted as a disabling mechanism and take precedence over all enabling rules containing any of the same distribution rule components applied to this EHCR message component.
- A *Basic distribution rule* represents the fact that the distribution rule is applied as a basic distribution rule to the EHCR message component. If one or more basic distribution rules are applied to an EHCR message component, it shall not be possible to apply any non-basic distribution rule if its contents do not comply with one or more of the basic distribution rules.

3.2.2 The EHCR Store Application

The Electronic Health Care Record (EHCR) store application is intended to be not only the application to store medical records for the needs of the OTM application, but a generic system for

storing and collating health care records across multiple health care providers. In essence, fragments of records are kept at each site a particular patient has visited and Provenance techniques are used to aggregate and determine the origin of data.

Note that there is a basic difference between typical workflow applications and the medical applications including the EHCR application. In a typical workflow application the actors participating in the workflow are in contact, while in medical applications the actors are not in contact. For example when the manufacturer of a bolt sends the bolt to the manufacturer of the aircraft, the aircraft manufacturer knows the identity of the manufacturer of the bolt and the identity if the bolt is given by the bolt manufacturer to the aircraft manufacturer (they both know which bolt they are talking about). In the EHCR application the actors are not in direct contact. The doctors do not always send the patient to each other. The patient may be treated by one doctor, then the patient may be healthy for a while, and then the patient may go to another doctor with another disease which is a consequence of the previous disease. In this case the second doctor is not in contact with the first one, they do not know each other's identity and that they are treating the same patient, as long as the patient or some other actor matches the different identities. The lack of direct contact between the EHCR application actors necessitates the introduction of unique identifiers.

The application will be used by the OTM application as its primary store of patient care data.

3.2.2.1 The Application

The EHCRS consists of two parts:

1. A Web Service that receives and sends messages in format ENV 13606 for remote medical applications.
2. A Java API for local medical applications that can be used to access the EHCRS directly.

The application is deployed on top of a database to store data of the patients, messages sent/received by the Web Service, data of health care parties (with their public key or username-encrypted_password) and possible other data.

The application uses Provenance services in the following way:

- To log the messages sent/received by the Web Service.
- To query whether the EHCR is up to date or not.
- To query where newer parts of the EHCR can be found.

The application also uses an authentication Web Service (ws_auth) to authorize request messages from remote health care parties. The application uses an encryption Web Service (ws_crypt) to get global anonymous ID of a patient (GMPID=Global Medical Patient ID).

GMPID is used to identify patients in Provenance and plays the same role in the Provenance system as the national health care identifier in the health care system, but the the GMPID is anonymised to hide the real identity of patients in the Provenance system. The GMPID is important to allow the Provenance system and the EHCR application to link the different Provenance information related to the same patient. Provenance information cannot be linked in other ways, because very often there are no direct contacts between the medical actors.

The system also requires the two other Web Services mentioned in the previous paragraph: ws_auth and ws_crypt. The former Web Service must be able to decide whether a health care agent is allowed to access a piece of an EHCR or not. The latter service is used to make the real data anonymous for Provenance store.

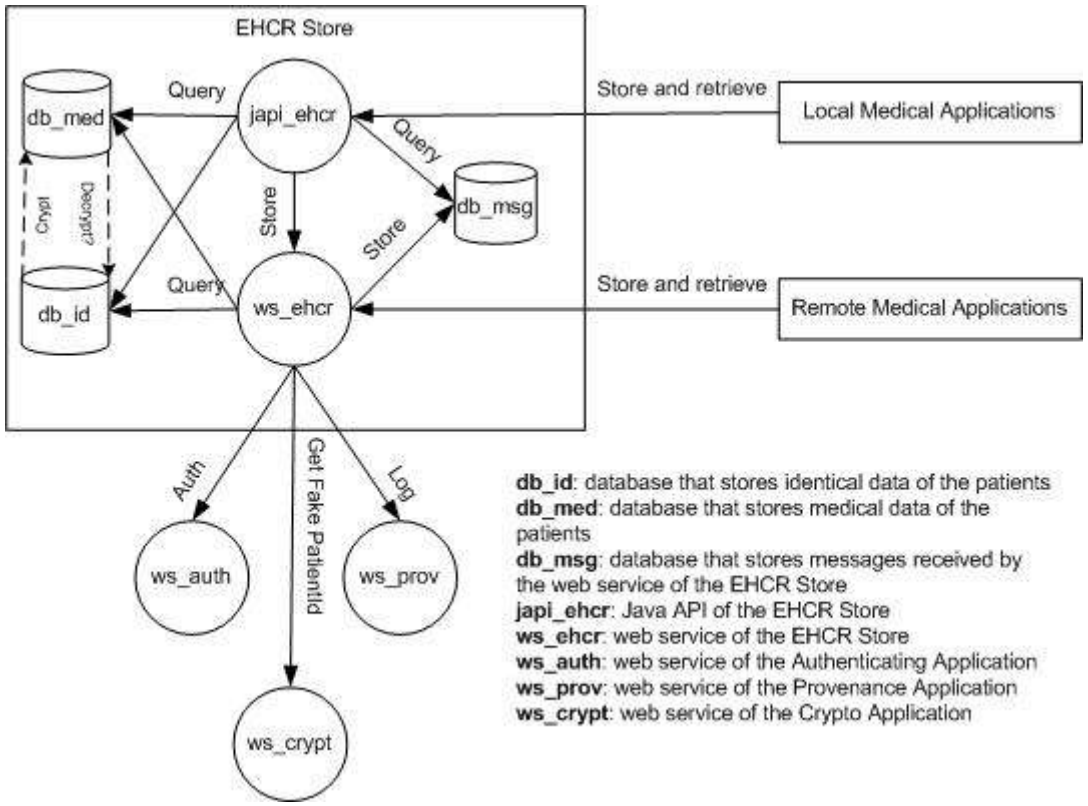

Figure 18 – EHCR Store.[11]

3.2.2.2 Database

For legal reasons, in most countries medical data of the patient must be separated from the identification data of the patient and the latter must be inaccessible from the former without sufficient authorisation. Whether or not medical data is directly accessible from identification data in this level is not decided yet (see Figure 18), however the architecture provides the means to separate these two types of data as necessary.

To implement this idea, an interface is defined in the Java API of the EHCRS which can give the medical data id (LMPID – Local Medical Patient ID) for the identification data id (PID – Patient ID).

Rules of medical data treatment require that:

- patient identification data (PID) and medical data of the patient must be kept in different databases,
- the medical database should not include the real patient identification,
- the mapping from the patient identification (PID) to the patient identification used in the medical database (LMPID) should be irreversible, i.e. nobody should be able to find out the real person identity (PID) from the patient identification used in the medical database (LMPID).

The LMPID may (and probably is) different in each EHCR store.

11 In the case of the OTM application, OTM services may access to EHCR stores as local or remote medical applications, depending if there is local network access between the OTM service and the EHCR store or not.

Identification data of the patient, such as national insurance number, name, mother's name, birth place and time, birth name, etc. are described in Section 3.2.2.4. Medical data in database level is one blob (a large binary object in the database) for each patient. This is a simple solution (and enough in this case). The part of the database - where the received/sent messages are - can be queried directly by the medical application.

3.2.2.3 Java API

This API serves local medical applications that have direct access to the local ECHCR store. Authentication in this component will provisionally use the same auth. application as the Web Service interface or only username/password.

This component provides a set of useful functions to search, list, create or modify identification data of patients; to query a set of medical data of unknown patients for statistical purpose; to query the EHCR of a single patient; to update an EHCR of a single patient locally or remotely (the API calls his own Web Service or a remote one to do this); to query a part of an EHCR from a remote EHCRS; to send notification messages to remote medical applications (e.g. request of EHCR is rejected)

Medical data in programming level is a Java object representing an EHCR extract (this expression is from ENV 13606). This object could be easily transformed into XSD or the type a WSDL specification.

3.2.2.4 Data Types

This section details the data types which will be used by the application.

a. NationalInsuranceNumber

- National insurance number of a patient or a encrypted form of it.

b. Timestamp

c. NotificationType

d. health careParty

- Data of a health care party; health care party is a human, institute, software or hardware that can send or receive EHCR message.

e. Authentication

- Digital signature or username-password.

f. EHCRExtract

- Representing an EHCR of a patient.

g. StateOfUpdateEHCR

- Messages sent during an update process and the answers of them; the state of an update process: under processing, succeeded, failed, ...

h. Message

- A Web Service message sent/received by the Web Service interface of the EHCRS.

i. Messages

- An array of Web Service messages sent/received by the Web Service interface of the EHCRS.

j. MessageReference

- A message id.

k. DistributionRule

- Who, where, when, why and how can update an EHCR.

3.2.2.5 Interfaces

The following interfaces are defined for the application.

a. **constructor of EHCRS(**
health careParty health careParty,
Authentication authentication)

- Introduce and authenticate the health care party using EHCRS.

b. **void setEHCR(**
NationalInsuranceNumber nationalInsuranceNumberOfPatient,
EHCRExtract ehcrExtract)

- Store an EHCR.

c. **EHCRExtract getEHCR(**
NationalInsuranceNumber nationalInsuranceNumberOfPatient)

- Query the whole EHCR stored locally.

d. **boolean isEHCRUpToDate(**
NationalInsuranceNumber nationalInsuranceNumberOfPatient)

e. **void updateEHCR(**
NationalInsuranceNumber nationalInsuranceNumberOfPatient,
DistributionRule distributionRule)

- Starts an update process.

f. **StateOfUpdateEHCR stateOfUpdateEHCR(**
NationalInsuranceNumber nationalInsuranceNumberOfPatient)

- Queries the states of the last update process.

g. **void sendEHCR(**
health careParty destination,
NationalInsuranceNumber nationalInsuranceNumberOfPatient,
EHCRExtract ehcrExtract

- send an EHCR to a destination (when there was no request to it)

h. **void sendEHCR(**
MessageReference messageReferenceOtRequest,
EHCRExtract ehcrExtract

- Send an EHCR to a destination in response to a request.

i. **void sendNotification(**
MessageReference messageReferenceOtRequest,
NotificationType notificationType
...

- Send a notification to a destination in response to a request.

j. **Message getMessage(**
MessageReference messageReference)

k. **Messages getMessages()**

- Request all the messages received or sent by the Web Service interface of the EHCRS.

l. **Messages getInputMessages()**

- Request all the messages received by the Web Service interface of the EHCRS.

m. **Messages getOutputMessages()**

- Request all the messages sent by the Web Service interface of the EHCRS.

3.2.2.6 Web Services Interface

In addition to the Java API used to access the EHCRS locally, there will be a Web Service interface to access the EHCRS remotely. When a message arrives, the Web Service authenticates it using a separate application and logs it using the Provenance service (see Figure 19) . Information about the authentication and logging application can be inside the message as a third party agent. There can be also a digital signature of the sender in the message for the authorization. After authentication and logging, the message is stored in a database from where local medical applications can query it with the Java API of the EHCRS.

If the type of the message is provide (see ENV 13606), the EHCR extract, included in the message, is also stored in a database (in db_med in Figure 18).

If the type of the message is request, then:

1) If the message can be answered automatically (without human interaction).
   The application (a Web Service client) sends a provide message. Human interaction is not needed if the whole EHCR is requested or the application is smart enough to decide which parts of the EHCR is requested.
2) If the message can be answered only with human interaction.
   The application sends a notification that the answer is being processed.
   A doctor using a medical application selects the parts of the EHCR record that must be sent and call the Java API of the EHCRS to send the answer.

3.2.2.7 WSDL

Three operations will be described in EHCR WSDL (following ENV 13606):

- `Request`
- `RequestAll`
- `Provide`
- `Notification`

The `request` and `requestAll` operations receive an EHCR query. The latter requests the whole EHCR that the EHCRS owns. The `provide` operation is the EHCRS response to a successful query. The `notification` operation encodes status reports about the EHCR query (e.g. request received, request accepted/rejected etc.) or about the response (e.g. EHCR accepted/rejected).

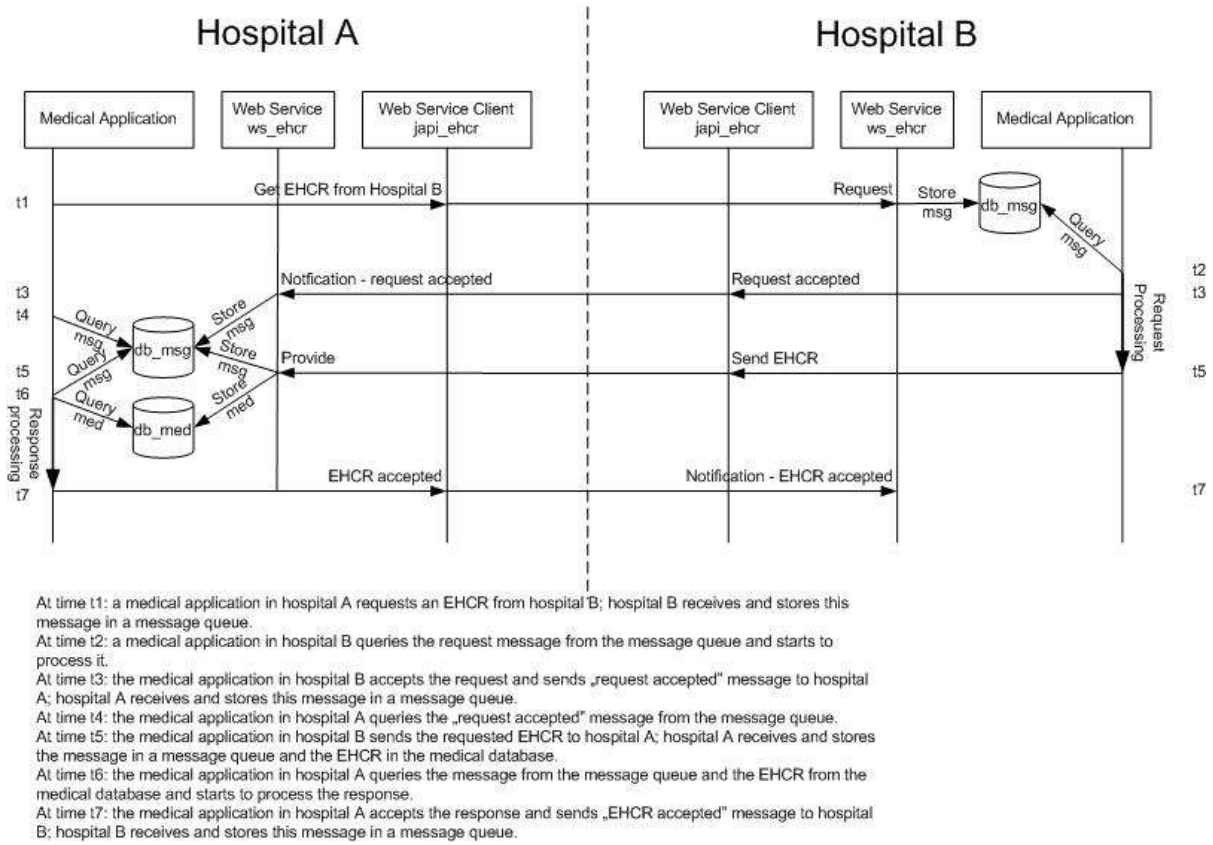

Figure 19 – Interaction diagram of requesting an EHCR.

Communication in the system is asynchronous; the service only receives messages sent by the Web Service client. Every message contains an id and can be identified by this id (and optionally with the time and the sender of the message). Also we will describe in the WSDL the domain information model of ENV 13606-4.

3.2.2.8 Medical Applications

Medical applications can access the remote EHCRS via EHCR Web Service to store and retrieve the data of a single patient. To retrieve the whole data of the patient they have to ask information about the location of the fragments from the Provenance application (see Figure 20).

When a doctor uses the medical application (see Figure 20), he/she can query the messages from the local message queue. It can be seen whether new messages arrived. New messages can be requests or can be responses (provide or notification) to previous messages. If a request message is not answered yet, the doctor can make the answer and send it using the Java API EHCRS. This answer can be a notification (e.g. request accepted/rejected) or a provide message (which includes an EHCR).

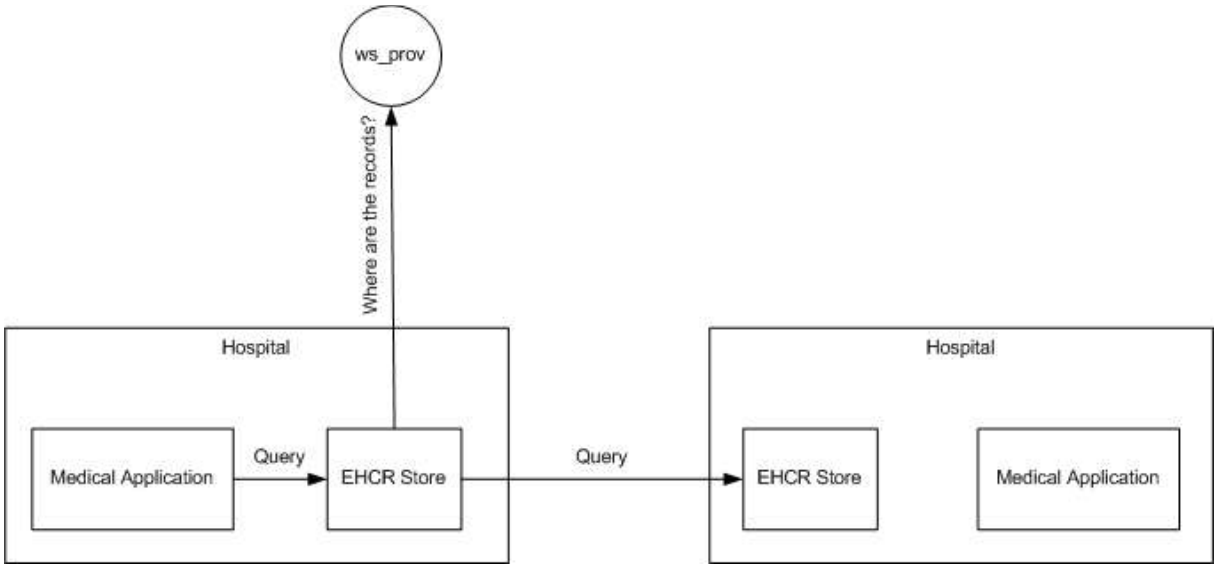

Figure 20 – Pulling together a full EHCR.

3.3 Summary

In summary, the OTM application is divided into two parts: an underlying health care record management element (Section 3.2) and the OTM application itself (Section 3.1). Important conclusions from the description of the application given include:

- The application spans multiple jurisdictions: numerous hospitals, an organ transplant authority, external labs and individual units within each of these entities.
- All actors/services in the system save data in their own individual data archive (the type of which varies with the type of the actor).
- Copies of full medical reports or decisions should in general never be logged in a Provenance store.
- A default Provenance deployment to support an OTM deployment is expected to use one Provenance store to shadow each actor/service.

4 Logical Application – Provenance Mapping

This section provides a high level view of how the OTM application maps to the Provenance architecture. Description is given in terms of general principles, identification of components, and mapping rules. Chapter 5 further details domain specific encoding and other issues required for the mapping.

Presentation is divided into two major sections: Section 4.1 on the OTM application itself and Section 4.2 on the EHCR system. Section 4.3 provides a summary of the logical mapping.

4.1 Mapping to the OTM Application

The OTM application's use of the Provenance infrastructure is based upon the following main principles:

- All major Web Service components in the OTM/EHCR application will be coupled to their own local Provenance store.
- Sensitive medical data is not retained in Provenance stores but referenced using pointers from Provenance stores to secure application data stores.
- Provenance stores are subsequently linked together to answer queries in a form of overlay infrastructure spanning all relevant OTM/EHCR services.

4.1.1 Provenance and Application Data Mapping

While the precise objectives of storing certain items of data are dependent on the queries finally supported (see Chapter 5) the general aim of data storage in the system is:

1. To provide a coherent way of tracking/locating interim results related to a case across all records / documents / reports.
2. Provide (without additional security clearance) a small amount of anonymized, low security risk meta-data to characterize the result types, decisions and outcomes without being exposed to the details of each decision.
3. [Optional] Provide a skeleton / framework for a more detailed probe which is able to apply security clearance to retrieve detailed records as needed in conjunction with Provenance meta-data.

It is important to note that all the data referred to previously in Chapter 3 is *data stored by the application itself for the purposes of the application* (no Provenance involved). The data referred to in this section however, relates to the *process documentation for Provenance purposes* which might be stored at runtime *in addition* to this data.

4.1.1.1 General Rules

Before beginning, several rules of thumb are helpful to establish default practice:

- Provenance stores are considered less secure than the medical systems/services themselves and the grade/type of data must in general be non-medically sensitive only.[12] [13]

12 Whether or not the Provenance systems are less secure of course is another issue, however the jurisdictions / control of the Provenance systems is distinct (different people may access them) – this alone makes it necessary to keep more sensitive data out. The point of this note is to state that there is a distinction between the *security features implemented* and the management/configuration of a particular deployment. i.e. the subsets of people who have access to a medical data store and a Provenance store may not be the same - this alone is reason to treat the Provenance stores differently.

- Any data relating to specific patients or health care individuals must use an anonymous ID only. Management of these IDs must use security controlled naming system for mapping back to identifiable individuals that is separate from the Provenance system itself.
- Copies of reports, diagnoses, full decisions etc. should not appear in the Provenance store. Only very limited factual / summary data may appear and be de-referencable via a unique ID to the full datum in the relevant data store.[14] (An example would be a medical report on the extraction operation – part of a specific transplant case – would be stored in a hospital surgery registry, but what is stored in the Provenance store would not be the report but a unique ID to the report in the secure database + a small amount of summary meta-data.).
- As a consequence of the previous item, full messages between actors/services will generally not be stored in the Provenance-stores (see schemas below).

In general, access to Provenance stores will be available to persons across multiple hospitals / entities – whereas internal data stores, generally are not accessible to anybody but local teams. If absolutely necessary however in certain cases data could be included. Such needs will be evaluated on a case-by-case basis.

Lastly one of the most important concerns is whether or not patient data can be linked to a particular patient:

- Identifiers which would allow a non-authorized viewer of a medical datum to recognize an individual patient must be removed.
- The system must be provably safe (patients cannot be identified) even across arbitrary access to all data stores in the system – i.e. it must not be possible to identify patients even by correlation across multiple storage sites.

➔ ***General Rule OTM.1****: Anonymisation of patient identifiers must be carried out before any data is stored in Provenance stores. A system wide anonymisation mechanism is required. (Instantiated in Section 5.3.1.)*

➔ ***General Rule OTM.2****: Source medical data is never stored in Provenance stores but only referenced therein. A system wide medical data referencing scheme is required. (Instantiated in Section 5.3.1.)*

4.1.1.2 Deployment of Provenance Stores

Each actor/service is expected to keep its own authoritative records of activities (and is generally responsible for one data type – such as family consent forms for example). Although the number of Provenance stores could be fine tuned and optimized by aggregating them, a default expected deployment is expected to be *one Provenance store per actor/service/data-store*. The reasons for this are primarily:

- Ensuring a logical architecture match between Provenance stores and deployed OTM components.
- Jurisdiction / ownership issues for Provenance stores will mirror those for the components themselves.
- Ease of synchronisation of Provenance store with local naming / identifiers used in a particular service / data store.[15]

13 This decision may be revisited if it can be shown that in a restricted case a Provenance service with a "higher level of clearance" can provide high value.

14 Appendix A shows in detail, for the different kinds of clinical data to be exchanged and stored in the OTM application, the subset that can appear in the p-assertions stored in the Provenance Stores.

15 This means that since a Provenance store will generally contain records which reference records in the service archival store the naming scheme can be more easily synchronised if Provenance stores are always twinned with services.

- Reduced delay in writing "actor" and interaction Provenance for local actors.

Provenance stores for some actors/services may be dropped if the actor/service is not deemed important to recording. Additional Provenance stores may be added for system wide elements of Provenance not related to a particular actor. However the prototypical deployment is likely to be primarily a "shadowing" of all OTM Actors/Services/Data Stores with Provenance stores.

➔ ***General Rule OTM.3****: Each application component (actor/service) is shadowed by its own local Provenance store. (Instantiated in Section 4.1.2 "Mapping to Logical Architecture".)*

➔ ***General Rule OTM.4****: Each such Provenance store is managed by the entity responsible for the application component the store is associated with. (Instantiated in Section 4.1.2 "Mapping to Logical Architecture".)*

➔ ***General Rule OTM.5****: Provenance stores are interlinked and communicate with one another, they are considered to be in one single-sign-on domain for security purposes, even though the application components will generally not be. (Instantiated in Section 5.3.3.)*

4.1.1.3 Process Documentation for Provenance Purposes

Description of process documentation is divided into two subsections: schemas for actions and events and schemas for messages.

4.1.1.3.1 Schemas: Actions/Events

Although the precise p-assertions stored by each service for each type of event / action / occurrence could be fine tuned (and would need to be in a full deployment) we begin with a set of generic template schemas for what data is stored:

1. *Decision Event*: whenever a medical decision is taken within a service (such as accepting a donor) the actor involved makes an actor state p-assertion in the Provenance store that:
   1. A decision was taken.
   2. The ID (and version number) of the corresponding report / data in the database.
   3. Standard elements of the decision including the primary outcome [Yes / No / Referral, ...], medical staff involved, data, time, medical warnings (special conditions which must be watched due to a particular reading / finding) and so forth.
   4. A case ID if available.

   In parallel, this actor-state p-assertion should be linked with the p-assertions of all the information inputs for the decision. These links are added to the Provenance store as relationship p-assertions.
2. *Condition Event*: whenever a important conditional point should be recorded within the execution of a service, the actor involved in checking the condition makes an actor state p-assertion in the Provenance store that:
   1. A check on the condition was made.
   2. The ID (and version number) of the corresponding report / data in the database.
   3. Standard elements of the condition including the result of the check, the time the check was performed, and any other information about the condition (depending on the condition being tested.
   4. A case ID if available.

   In parallel, this actor-state p-assertion should be linked with the p-assertions of all the information inputs for the condition. These links are added to the Provenance store as relationship p-assertions.

3. *Test Result Event (Datum Generated)*:[16] whenever a service responsible for producing/retrieving a datum generates a result (e.g. a blood analysis outcome, a patient care record retrieval, ...), the actor representing the whole service makes actor state p-assertions in the Provenance store on:
   1. A result was generated.
   2. The ID (and version number) of the corresponding report/data in the database.
   3. A number of standard elements similar to those under decisions.
   4. A case ID if available.

   In parallel, this actor-state p-assertion should be linked with the interaction p-assertion of the message which requested the test. This link is added to the Provenance store as relationship p-assertions.
4. *Edit / Update Event (Datum changed):* whenever an actor responsible for a particular datum (test result) or an actor involved in a decision makes a change to the datum (changing version, adding something, removing something), the actor responsible logs in the Provenance store that: [17]
   1. A record was changed.
   2. The ID (and version number[s]) of the corresponding report/data in the database.
   3. A number of standard elements similar to those under results/decisions incl. time and date of change.
   4. A case ID if available.

   In parallel, this actor-state p-assertion should be linked with the p-assertions which caused the update. These links are added to the Provenance store as relationship p-assertions.
5. *Consult Action (Datum / Decision Read):* whenever an actor provides a datum or record to anybody (e.g. in response to a query). the corresponding event, who invoked it, with what permissions etc. must be logged – not clear it needs to be logged in the Provenance store though.

These recorded elements provide the major state changes which affect a particular outcome – recording the major steps in the work/data flow and associating them with the archived actual detailed records. *The default is that all decision, result and edit (possibly consult) actions defined in the workflow are recorded.*

It is important to note that data retention policies for the OTM application differ per data type and it will not always be the case that all data needed for a particular Provenance query will be available at any time in the future. That is, once certain data has been removed (as per a policy) it will no longer be available to Provenance and a query requiring it will no longer work.

➔ ***General Rule OTM.6****: All decision, result and edit (and possibly consult) actions/events are recorded in the system by the actor responsible for carrying them out. (Instantiated in Section 5.1.3.1.)*

➔ ***General Rule OTM.7****: All such generated process documentation is expected to be available indefinitely, however the underlying medical data may be removed over time due to the application of data retention policies. (Rule not further instantiated – applied as is.)*

16 Potentially this is not a separate class to decision but separate for now.

17 Note that the application databases are expected to track versions if edits are allowed – we do not expect Provenance to be able to do this. Provenance should be given enough information to always get back the actual version used in a particular workflow.

4.1.1.3.2 Schemas: Messages

While it is not clear they are strictly necessary in all cases (primary use is in work flow reconstruction), the actors may also record a *reduced/restricted form of some (or all) of the messages they exchange*. Typically we expect that:

- All messages to be asynchronous (that is document style calls rather than RPC style calls).
- Message types to be of one of the following types:[18] [19]
  1. A REQUEST: asking for an action to be taken.
  2. A QUERY: asking for a result / test result.
  3. A RESPONSE to a query or request.
  4. A FAILURE: generally stating a reason (such as non-availability of data).
  5. A REFUSAL : generally stating a reason (such as lack of credentials to access data).
  6. An AGREEMENT: generally to carry on and try to fulfill a request.
  7. A INFORM-RESULT: generally containing a datum (e.g. test result, decision)
- All messages are between exactly two actors (sender and receiver).
- A flow of REQUESTS for action (followed by RESPONSES) is used to drive the central workflow, as steps are taken REQUESTS are generated for the next responsible Actors to carry on the process (passing on tokens as to which is the active service).
- The messages will contain data items which are copies of original data items stored at the Actor that produced them (the actor retains the master copy).

In terms of schemas for Provenance recording of messages in the general, the default is that *all messages referenced in the work flow are stored by the Actor sending and the Actor receiving them* in a reduced form which does not include a full copy of the content datum but logging that:

- A message was sent / received.
- The type of the message (request, ... etc.).
- The sender and receiver.
- Time, Date etc.
- In the case of RESPONSE, FAILURE, REFUSAL, AGRREEMENT and INFORM-RESULT, a reference to the original message that trigered the current message.
- A pointer to the archived local copy of the medical data contained in the content.
- Possibly a small amount of meta data (administrative data) about the meaning of the content (e.g. nature of a decision).

➔ ***General Rule OTM.8****: Messages stored in the Provenance system MAY NOT be the complete messages originally sent, but a reduced form removing sensitive medical data. (Instantiated in Section 5.1.3.2.)*

➔ ***General Rule OTM.9****: All messages sent in the system are stored by BOTH the sender and the receiver. (Instantiated in Section 5.1.3.2.)*

18 The semantics for each type are loosely based on the FIPA ACL semantics [FIPA02b] but not strictly adhered to – these are to illustrate general message types used in control flow. It is an open question as to whether it is worth exploring the use of formal message semantics for answering certain types of Provenance questions.

19 The set of types may be extended.

4.1.2 Mapping to Logical Architecture

This section identifies the major components of the OTM application in terms of the architectural notions defined in the Provenance Logical Architecture D3.1.1. Concrete instantiations of the rules generated are given in Section 5 of the document. The overall mapping is shown in Figure 21.

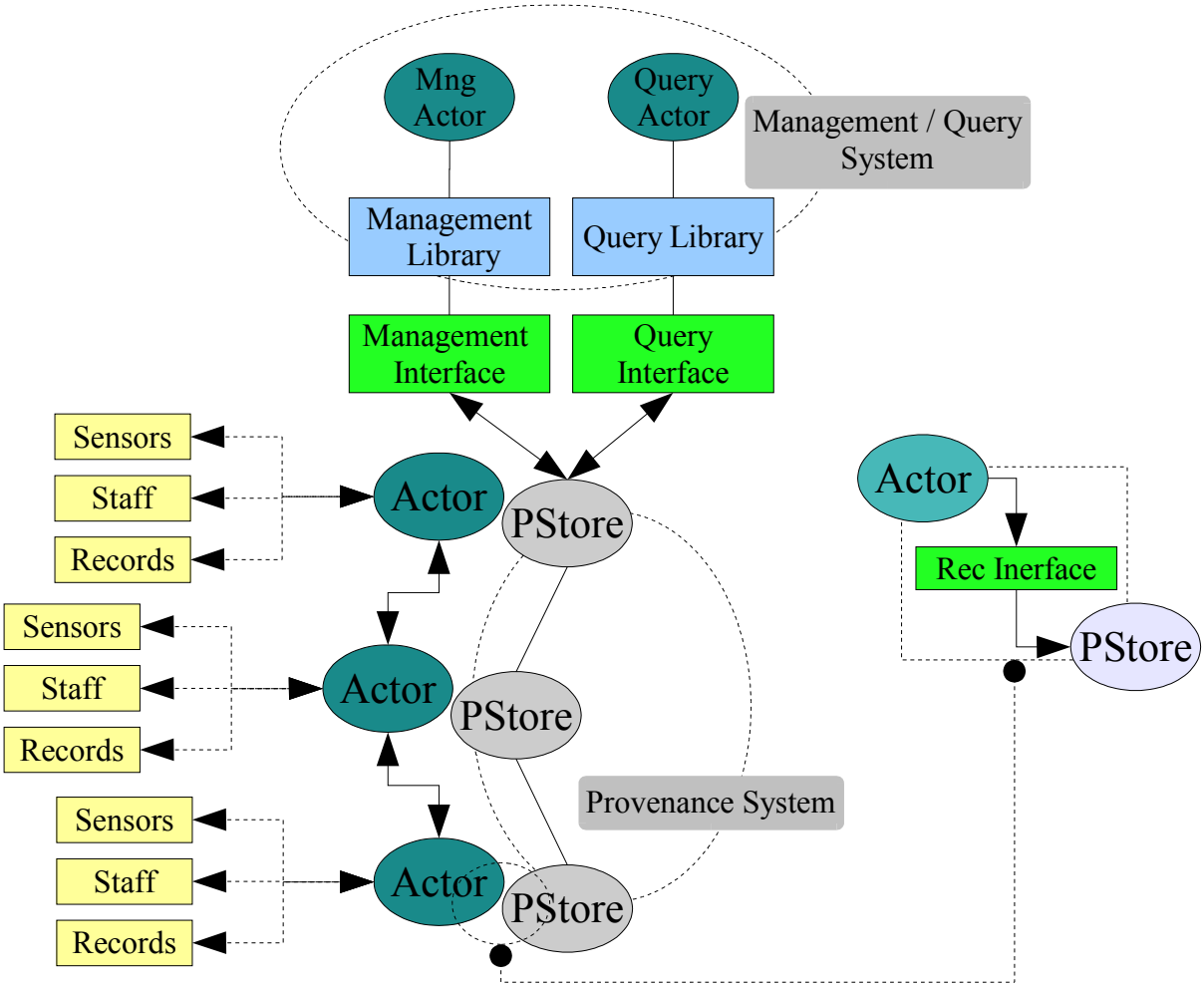

Figure 21: Overall OTM Mapping to Provenance Logical Architecture

Provenance Roles

Application Actors (responsible for carrying out the application's business logic)

➔ ***General Rule OTM.10****: Each of the services identified in Section 3.1.1. is represented by a special Application Actor (called the Controller). Apart from the controller, services may also be composed by some Application Actors (called Internal Actors)*[20] *which perform a limited number of functions within the service. (Instantiated in Section 5.1.3.1.)*

➔ ***General Rule OTM.11****: For each service, the following mapping is carried out:*

20 The terms *Controller* and *Internal Actor* are introduced for convenience of discussion and to ease the mapping to the Provenance Logical Architecture.

1. *Each service has a lead application actor, called the Controller, which receives and sends the information or requests from/to other service controllers.*
2. *An Event occurring within a service corresponds to an Internal Actor receiving information and then sending information.*
3. *For each information item the Event takes as input, it must receive it from an actor that has previously received it, the Controller or preceding Event for example, or otherwise know it.*
4. *A Decision Point or a Condition Point (in a particular instance) having been made within a service are considered as an Event occurring. Therefore each Decision Point and Condition Point corresponds to an Internal Actor receiving the inputs and then producing the corresponding output.*
5. *As in the OTM scenario a Decision depends on a human making the decision, the actor state (of the Application actor mapping the decision) may contain further information on why the particular decision was made and, if available), the name(s) of the team members involved in the decision.*
6. *For those services that keep activity logs, or results' logs or that create any other specific documentation about the results of the process carried out, an extra event "Documentation being recorded" should be added (an Event occurring after the Controller has sent out all the resulting information to other Controllers and before the service execution ends).*
7. *For each Event occurring (including Decisions having been made or Documentation having been recorded), a causal relationship is recorded between the outputs of the Internal Actor (the effects) and inputs of that actor (the causes), so that the process can be retraced.*

(Instantiated in Section 5.1.3.1.)

It is important to note that the individual humans (doctors, nurses) that participate in the real world process will not be mapped as application actors, as:

- Most of the activities and decisions in the real world are carried out / made by the "team", not individuals.
- Reports are usually completed at the end of one step in the allocation process by any member of the team that has the proper credentials and time to do it.[21]

Therefore, any member of the team will send the information about what happened in the real world through a GUI interface to the webservice, which will "represent" the team as a whole. Of the humans (or, equipment) involved in the process, these are generally, not represented explicitly in the Provenance. Only the Web Services that govern the processes (dark boxes) that make up the OTM workflow can be explicitly represented.

➔ ***General Rule OTM.12****: Individuals (doctors, nurses, assistants) taking part in a given real-world process are not directly mapped as actors in the Logical Architecture. Instead, the team as a whole is represented by the webservice which is documenting the process in the IT world. (Rule not further instantiated – applied as is.)*

Provenance Stores (responsible for making persistent, managing and providing controlled access to recorded p-assertions)

21 Records of the process are usually done at the end of any step in the allocation process in order to avoid delays in critical steps: for instance, a surgeon should not stop the implantation of an organ in the recipiet to go to the GUI interface and record his last decissions and actions taken. If there is enough personnel in the surgery room, a nurse or an assitant will record the events and decissions in parallel; if not, recording is done after the surgery.

➔ ***General Rule OTM.13****: Each controller is shadowed by a Provenance Store which records Provenance only for this actor. (Rule not further instantiated – applied as is.)*

Asserting Actors (actors that create p-assertions about an execution)

➔ ***General Rule OTM.14****: All application actors are asserting actors. As a minimum they assert statements relating to incoming / outgoing messages. (Rule not further instantiated – applied as is.)*

Recording Actors (actors that submit p-assertions to a Provenance store for recording)

➔ ***General Rule OTM.15****: All asserting actors are recording actors. All recording actors by default record to their local Provenance Store unless otherwise specified. (Rule not further instantiated – applied as is.)*

Querying Actors (actors that issue Provenance queries to a Provenance store)

➔ ***General Rule OTM.16****: Provenance store queries are not expected at application execution time but occur as a separate process. Queries are carried out by one or more designated query actors not included in the list of application actors in Section 3.1.1. (Instantiation in Section 5.1.2.)*

Managing Actors (actors that interact with the Provenance store for management purposes)

➔ ***General Rule OTM.17****: Provenance management interactions are not expected at application execution time but occur as a separate process. Management actions are carried out by designated management actors not included in the list in Section 3.1.1. (Instantiation in Section 5.1.2.)*

Libraries and Interfaces

Actor Side Libraries

➔ ***General Rule OTM.18****: Actor side libraries for Provenance recording are embedded in all application actors. Actor side libraries for management and querying are added only in designated additional actors. (Instantiated in Section 5.1.3.1 for all actors, in Section 5.1.2 for query / management actors.)*

P-header (Provenance-related context information, sent along with the interaction's message. tracers)

➔ ***General Rule OTM.19****: A P-header is included with every application message interchanged between any 2 Application Actors (see Section 5.1.3.1).*

Recording Interface

➔ ***General Rule OTM.20****: Recording Interfaces are used by Application Actors only, typically an Application Actor will use only the recording interface of the Provenance Store directly associated with itself. (Rule not further instantiated – applied as is.)*

Query interfaces

➔ ***General Rule OTM.21****: Query interfaces will be used only by specialized query and management actors – not Application Actors. (Instantiated in Section 5.1.2.)*

Management interface

➔ ***General Rule OTM.22****: Query interfaces will be used only by specialized query and management actors – not Application Actors. (Instantiated in Section 5.1.2.)*

Processing Services and Presentation User Interfaces

➔ ***General Rule OTM.23****: These interfaces are instantiated by the Management and Query Actors only . (Instantiated in Section 5.1.2.)*

Policies

Use of explicit policies in the OTM application is currently limited only to definition of naming schemes to be used by deployed services. These schemes are defined in Section 5.1.1.

4.2 Mapping in the ECHR Application

This section describes the mapping of the EHCR application part of OTM to the Provenance architecture. The EHCR application is a generic component of OTM. In principle it could be part of any medical application, although sometimes we use OTM specific information. The EHCR application part of OTM focuses on the data management part of the OTM application and the OTM specific Provenance issues are mapped in the previous parts of this document. The EHCR application mapping to the Provenance architecture focuses on the Provenance of patient health care status from the information stored in the EHCR stores.

4.2.1 Provenance and Application Data Mapping

Here we follow the general rules that health care data and patient information are not stored directly in the Provenance system, only through references. The reference to patient information is via the Global Medical Patient ID (GMPID) and the reference to the health care data uses the system wide medical referencing scheme.

4.2.1.1 Objectives of Provenance Stores for the EHCR Application

As discussed earlier, the full EHCR of a patient is basically the documentation of the Provenance of the health care status of the patient. Because EHCR fragments are scattered through health care organisations and the the full EHCR has to be pulled together from different places, the EHCR application needs to track somehow the creation of EHCR data and be able to find the EHCR traces of the development of the health care history of the patient. This is a typical Provenance problem, therefore the main objective of the Provenance stores for the EHCR application is the assembly of the full EHCR of the patient.

In particular Provenance stores should tell to an EHCRS:

- Whether EHCRS owns an up to date EHCR of a patient or not.
- Where EHCRS can find the missing fragments of an EHCR of a patient.

In order to be able to provide this data we expect that Provenance stores log sent/received EHCR messages and the change of an EHCR. Provenance stores also should receive information about the success of an EHCR update process.

➔ ***General Rule EHCR.1****: In order to assembly the full EHCR of the patient, the EHCR application uses the Provenance information returned from the Provenance store. (Rule not further instantiated – applied as is.)*

4.2.1.2 Deployment of Provenance Stores

In the current health care systems there are some solutions to pull together the full EHCR of patient, although these solutions do not give perfect results and doctors are satisfied with almost complete EHCR as well. In the case of Catalonia, CatSalut has no system to help the collection of the different pieces of the EHCR, as CatSalut records only keep information about which General Practitioner centre each patient is assigned to (based in geographical criteria). There is no track done by CatSalut of all the health institutions that have pieces of a patient's ECHR, and no plans for a system to do that in the near future. Therefore, to save time, doctors usually make their own laboratory investigations to collect the relevant data needed from the patient. Other EU regions have some kind of central authority. In this centralised approach, all medical applications submit relevant information to the central authority.

The Centralised approach is not the best solution from informatics point of view, because of its decreased fault tolerance and scalability. A distributed approach could provide better solutions to these problems, therefore the EHCR application will use the *one Provenance store per actor/service/data-store* approach as described in section 4.1.1.3. It is up to the Provenance system to connect these distributed Provenance stores and answer Provenance questions as if the distributed Provenance stores were logically a single centralised Provenance store.

➔ ***General Rule EHCR.2****: The EHCR application will use the one Provenance store per actor/service/data-store approach to store p-assertions. (Instantiated in Section 3.2.2.1.)*

➔ ***General Rule EHCR.3****: The Provenance system connects the distributed Provenance stores and answers Provenance questions as if the distributed Provenance stores were logically a single centralized Provenance store. (Instantiated in Section 3.2.2.1.)*

4.2.1.3 Process Documentation for Provenance Purposes

The process documentation of the EHCR application follows the general rules described in section 4.1.1.1. However, as noted earlier, the identification of the patients is a special problem, because healthcare actors are not always in contact with each other and do not give the patient identification to each other, which means that they cannot directly link the different pieces of the patient healthcare history.

The health care data storage regulations specify that the identification data of the patient and the health care data of the patient have to be kept separate, and different patient identifications must be used in the two databases. Moreover the health care data cannot directly be retrievable by knowing the identification of the patient in the patient identification database. As a consequence, the identification of the patient used in the health care database is an anonymous patient identification derivable from the public patient identification (usually the national social security number) in a secure and encrypted way. Usually each EHCR database in the health care application has a different method to derive the anonymous patient identification from the public patient identification. Although there is a real global identification for each patient, the anonymous patient identification is different in each EHCR database.

The Provenance system must be able to tell that EHCR pieces coming from different EHCR databases belong to the same patient or not in order to be able to pull together the full EHCR of the patient. Because the healthcare actors cannot link the different patient identifications, there must be a global patient identifier in the Provenance system. However this global Provenance patient identifier cannot be the real global patient identifier, because the Provenance system should not be able to infer health care information of real patients. If the Provenance system knew the real identification of the patient, then it could tell for example whether HIV test was carried out on the patient or not. Therefore the Provenance global patient identification must be an anonymous global identification which means that the anonymous patient identification of each EHCR store must be mapped to the anonymous global patient identification. The EHCR stores should not store the mapping information between the local and the global anonymous patient identification, rather they should ask for the mapping each time they submit p-assertions to the Provenance store as shown in the following figure:

PID is a public identifier of the patient, such as national insurance number.

LMPID is private (used only in EHCR store) identifier of the patient counted from PID by a cryptography module of the EHCR store. Different EHCR stores can use different algorithms to count LMPID. If somebody can steal the database he can't join the identification and the medical data of the patient without knowing the algorithm.

Local OTMA (and other local medical applications) identify the patient in EHCR store with PID for non medical data of the patient (such as name, birth date, mother name) and with LMPID for medical data of the patient.

EHCR store identifies the patient in other EHCR stores with PID. The information channel must be secure in this case (for example using https protocol).

GMPID is private (used only in provenance store) identifier of the patient counted from PID by a central service (for example from CatSalut). We don't care how this service maps PID to GMPID, but this mapping must have the same properties as the PID -> LMPID mapping, i.e. not reversible and should not allow anyone to identify the patient of some medical data/provenance entry by knowing the GMPID.

EHCR stores identify the patient in provenance stores with GMPID that can be queried from a central service (for example from CatSalut) after authorization.. The EHCR store doesn't store any identification or medical data of the patient in provenance store, only a fake id of the patient and references to medical data. This fake id known only by authorized applications.

OTMA (and other medical applications) identify the patient in provenance stores with anything they want.

Registration and authorization is required for EHCR store, provenance store and CatSalut.

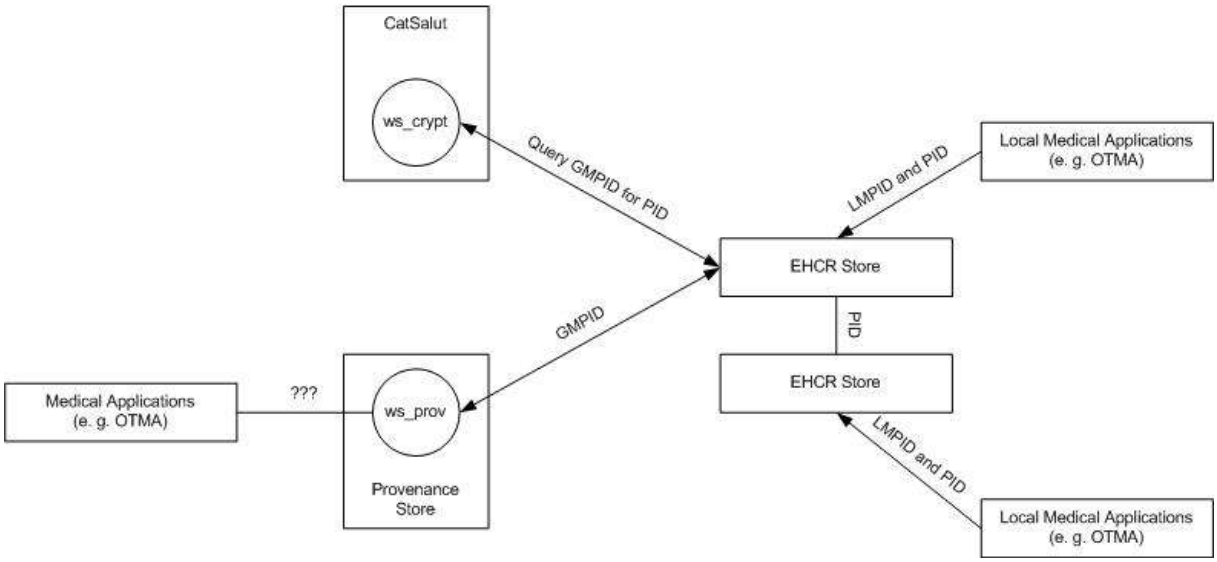

Figure 22 – Assigning global anonymous patient identification to p-assertions.

The PID -> GMPID mapping (provided by ws_crypt) must have the same properties as the PID to LMPID mapping, i.e. not reversible and should not allow anyone to identify the patient of some medical data/Provenance entry by knowing the GMPID. This is why a given EHCR store is not allowed to keep the PID to GMPID mapping. The GMPID is used only in communication with the Provenance store. The EHCR store sends the PID of the patient, which is then translated to the GMPID by the ws_crypt application and used in the Provenance store. The Provenance store does not

know the PID, only the GMPID. When the EHCR asks Provenance queries from the Provenance store, then formulates the query with the PID, which is again translated to the GMPID by the ws_crypt application.

This way the Provenance store does not know the patient and does not know the medical data. Even if a series of medical interventions can be reconstructed from the Provenance store, then it can only be known that the EHCR of some person (not known who) was updated/modified in certain medical institutions, but it is not known what these updates/modifications were. From this information nothing can be inferred for the patient, one can only infer which medical institutions were involved in a case. This includes information on the medical practice, but nothing particular for individual patients or doctors.

The contents of the p-assertions documenting the communication between the Provenance store and EHCRS are the following:

- A global anonymous patient identification (GMPID).
- A timestamp.
- The type of EHCR message: request, requestAll, response, notification.
- The health care party: data of a human, institute, software or hardware that can send or receive EHCR message.
- The address of an EHCRS.
- Authentication information: digital signature or username-password.

➔ ***General Rule EHCR.4****: Patient information are stored in the Provenance system only through references using the Global Medical Patient ID (GMPID). (Instantiated in Section 5.3.1.)*

➔ ***General Rule EHCR.5****: Health care data are stored in the Provenance system only through references using the system wide medical referencing scheme. (Instantiated in Section 5.3.2.)*

•

4.2.2 Mapping to Logical Architecture

As it is shown in Figures 16 and 18 in Section 3.2, the ECHR application directly talks to the Provenance system and all Provenance activities related to the assembly of the full EHCR is in the control of the EHCR store. The Provenance related activities will be hidden for the medical applications and the medical applications will see the EHCR store as a store with complete information on the full EHCR of the patient. Therefore principles of the mapping of the EHCR store to the Provenance Logical Architecture is simple: the actors are the EHCR stores. The p-assertions sent from the EHCRS to the Provenance system are composed of the activities described in Section 4.2.1.1 and the data described in Section 4.2.1.3.

➔ ***General Rule EHCR.6****: Provenance activities related to the assembly of the full EHCR is in the control of the EHCR store and are hidden from medical applications. (Rule not further instantiated – applied as is.)*

Provenance Roles

Application Actors (responsible for carrying out the application's business logic)

➔ ***General Rule EHCR.7****: japi_ehcr and ws_ehcr (see Figure 18) are application actors. (Rule no further instantiated – applied as it is).*

Provenance Stores (responsible for making persistent, managing and providing controlled access to recorded p-assertions)

➔ ***General Rule EHCR.8****: The EHCR application will use the one Provenance store per actor/service/data-store approach to store p-assertions.(Rule not further instantiated – applied as is.)*

Asserting Actors (actors that create p-assertions about an execution)

➔ ***General Rule EHCR.9****: ws_ehcr is the only asserting actor (see Figure 18). (Rule not further instantiated – applied as is.)*

Recording Actors (actors that submit p-assertions to a Provenance store for recording)

➔ ***General Rule EHCR.10****: ws_ehcr is the only recording actor (see Figure 18).(Rule not further instantiated – applied as is.)*

Querying Actors (actors that issue Provenance queries to a Provenance store)

➔ ***General Rule EHCR.11****: ws_ehcr is the only querying actor (see Figure 18). (Rule not further instantiated – applied as is.)*

Managing Actors (actors that interact with the Provenance store for management purposes)

➔ ***General Rule EHCR.12****: Provenance management interactions are not expected at application execution time but occur as a separate process. Management actions are carried out by designated management actors not included in Section 3.2.(Rule not further instantiated – applied as is.)*

Internal Actors (actors added in the p-structure to reflect the activities of real actors in the world or to ease queries on events and decisions)

As the EHCR application has to mirror real actors that are not made within the computational system, the p-structure has to be extended with extra actors, called internal actors. These real actors stands behind the medical applications (such as the OTM Application) and must be described by these applications.

Libraries and Interfaces

Actor Side Libraries

➔ ***General Rule EHCR.13****: Actor side libraries for Provenance recording are embedded in only the recording actors, for management in only the managing actors and for querying in only querying actors. (Instantiation in Section 5.2.)*

P-header (Provenance-related context information, sent along with the interaction's message. tracers)

➔ ***General Rule EHCR.14****: A P-header is included with every message interchanged between any 2 ws_ehcr.(Rule not further instantiated – applied as is.)*

Recording Interface

➔ ***General Rule EHCR.15****: Recording Interfaces are used by Recording Actors only.(Rule not further instantiated – applied as is.)*

Query interfaces

➔ ***General Rule EHCR.16****: Query interfaces will be used only by specialized query actors. (Instantiation in Section 5.2.)*

Processing Services and Presentation User Interfaces

➔ ***General Rule EHCR.17****: There will be no presentation user interface in EHCR application at all. (Rule not further instantiated – applied as is.)*

4.3 Summary

In summary, the two elements of the application (OTM and EHCR) both use a similar mapping to the Provenance architecture:

- Application Components are mapped 1:1 with they own local Provenance Store.
- Both will apply generic systems for hiding patient data from users able to access only the Provenance stores and rely on application data stores to hold sensitive data.
- Mappings to anonymized Identifiers are used throughout the Provenance recording and querying procedure.

Chapter 5 instantiates / extends the general rules defined in this section to concrete mappings to be used in application deployment.

5 Domain Specific Provenance Handling

The general rules defined in Section 4 provide an overall view of how Provenance is to be applied in the OTM / EHCR application. This section goes on to defined specific solutions / decisions for individual elements of the mapping – instantiating these rules. As previously, presentation is divided up between the OTM application (Section 5.1) and the EHCR application (Section 5.2).

5.1 Provenance Handling in the OTM Application

Presentation of Provenance handling in the OTM application is divided into four areas: Section 5.1.1 covers technical decisions on naming, and name spaces, Section 5.1.2 covers management and query functions, Section 5.1.3 covers process documentation recording, Section 5.1.4 covers Provenance queries.

5.1.1 Naming and Namespaces

The following sections describe mechanisms for the identification of elements in the application.

5.1.1.1 Preliminaries

The identification of items of different types in the system a hierarchical naming scheme. All names for which there is not another convention already existing, have the following form (based on the IETF DNS Specification for Domain names [RFC1035]):

```
<name>          ::= <subname> | " "
<name>          ::= <label> | <subname> "." <label>
<label>         ::= <letter> [ [ <ldh-str> ] <let-dig> ]
<ldh-str>       ::= <let-dig-hyp> | <let-dig-hyp> <ldh-str>
<let-dig-hyp>   ::= <let-dig> | "-"
<let-dig>       ::= <letter> | <digit>
<letter>        ::= any one of the 52 alphabetic
                    characters A through Z in upper case and a
                    through z in lower case
<digit>         ::= any one of the ten digits 0 through 9
```

Further, we allow for the possibility of adding an arbitrary qualifying number to the end of a standard name:

```
<ext-name>      ::= <name> ":" <number>
<number>        ::= <digit> | <number> [ <digit> ]
<digit>         ::= any one of the ten digits 0 through 9
```

In addition:

- The first subname is used to denote the type of the entity being named, e.g. `ACTOR.CatSalut.hospitalhsp.ward37`. The reserved keywords for this purpose are defined in the subsequent section.
- A URN qualifier can be added to make the names compatible with [RFC2141] as a logical namespace. The qualifier adopted is "URN:X-OTM:", giving fully qualified names such as `URN:X-OTM.ACTOR.CatSalut.hospitalhsp.ward37.` (Note that through the remainder of this document this prefix is not used. The X- prefic indicates the prefix is experimental and not registered with IANA [RFC2114])
- The qualifying identifier is *ACTOR* etc. is typically dropped if the type is obvious.

5.1.1.2 Actor Identity Management [Prefix ACTOR]

The identification of Actors in the system follows the hierarchical naming scheme defined in the previous section. More specifically in the OTM application, Actor names prefixed by the identifier `ACTOR` are allocated according to a set of levels of domains as shown in Table 2.

Level	***Meaning***	***Allowed Values***
0	Health Authority	*"CatSalut"*
1	Health care Organisation (e.g. hospital, OTA, Surgery Practice)	Individual values for all organisations identified in Section 3.
2	Department / Unit	Individual values for all departments/units identified in Section 3.
3	Service (Dark Boxes in Section 3 diagrams)	Individual values for all services identified in Section 3.
4	Internal Actor (involved in an Event, Decision or Condition in Section 3 diagrams)	Individual values for all events identified in Section 3.

Table 2 -- Specific levels of Actors applied in the OTM application. Hence the fully qualified name `CatSalut.hsp.lab-i` might be applied to a particular laboratory I at hospital HSP in catalunya. The name `CatSalut.hsp.lab-i.test-complete` may refer to a particular event within the Lab.

Each name in the hierarchy refers to a notional actor which may or may not be separately instantiated in the system. Hence, the `ACTOR.CatSalut.hsp.lab-i` is logically part of the actors `CatSalut.hsp` and `CatSalut`, even if neither of these are actually instantiated as a Web Service themselves. Further down the hierarchy, although individual events are represented in the namespace it is not necessarily the case that they are represented by a individual Web Service in the real world. Instead an individual Web Service representing an actor such as `CatSalut.hsp.lab_1` may use the names `CatSalut.hsp.lab_1.in` and `CatSalut.hsp.lab_1.out` as different events occur.

Lastly, it is assumed that each high-level actor so identified manages their own namespace, such that the actor `CatSalut` for example is associated with a list of assigned names in the namespace `CatSalut.*`. These assignments are made at design time but may be managed dynamically at run time with a DNS style name assignment service.

5.1.1.3 Medical Data Identity Management [Prefix DATA]

Apart from the Actors and patients in the system, provision also needs to be made for the logical identification of the various pieces of medical data which are stored during application execution. As defined in general rules **OTM.13** and **OTM.14**, each actor responsible for storing medical data is logically associated with their own medical data store which may impose its own rules on structure, access, security etc. For implementation purposes however it is assumed that the actor responsible for a data store:

- Generates a identifier for every new data item stored.
- Retains a mechanism which makes it possible to identify a particular data item given such an identifier (and vice versa). (This mechanism may for example be adding extra labels to the data store, retaining a mapping table or using a conversion function etc.)

In most cases one and only one actor is responsible for a data store and each actor generally has at most one data store associated with it. However in order to allow for more generality a separate namespace is defined for data stores and data items. Specifically:

- All names are prefixed with the label DATA.
- The unique identifier for a data item in a given store is added as the numeric label in the extended name defined in Section 5.1.1.1.
- Similarly to Actors, data stores are labeled using the conventions defined in Table 3.

Level	***Meaning***	***Allowed Values***
0	Health Authority	*"CatSalut"*
1	Health care Organisation (e.g. hospital, OTA, Surgery Practice)	Individual values for all organisations identified in Section 3.
2	Department / Unit	Individual values for all departments/units identified in Section 3.
3	Data Store	

Table 3: Specific levels applied to naming of Data stores in the OTM application.

Examples of data store names would therefore be: `DATA.CatSalut.hsp.immunology.d1` and `DATA.CatSalut.hsp.immunology.d2`. An example of a data item in the first database would be `DATA.CatSalut.hsp.immunology.d1:452348`. Lastly it is important to note that the names spaces of Actors and Data stores are deliberately independent therefore:

- An actor `ACTOR.CatSalut.hsp.immunology`, could well be responsible for a data store `DATA.CatSalut.hsp.general.d1`.
- More than one actor (from different places) could write to a data store `DATA.CatSalut.hsp.general.d2`.
- The unique identifier mechanism associated with a data store must be shared between Actors where more than one actor is at liberty to create data items in the store.

5.1.1.4 Case Identity Management (Tracers) [Prefix: CASE]

In a similar manner to the issuance of identifiers for data stores, organ transplant cases are also issued issued with hierarchical identifiers annotated with a case number. However, in this case it is assumed that the new identifier is generated by combining:

- The fully qualified name of the Actor creating the case (see Section 5.1.1.2).
- A unique identification number generated by this actor.

Hence cases typically have names such as `CASE.CatSalut.hsp.otm-unit:3289`.

It is noted that the case identifier is distinct from the data record identifier of an case file which may exist. In general a case may be associated with many data items.

5.1.1.5 Patient Identity Management

The identification of Patients in the OTM system is handled by a separate process from application data since it must be consistent across all health care tasks rather than just organ transplantation. These identifiers are used:

- Within medical data stored out of reach of the Provenance system
- Within the Provenance stores deployed for the application.

System wide patient identifier and data anonymisation is discussed in Section 5.3.1.

5.1.2 Management and Query Services

The majority of the mapping presented here is concerned with the run-time operation of the OTM application, that is, the activities relevant to the *Application Actors*. As described in general rules **OTM.16** and **OTM.18** however, in addition to these however, the application will include a number of management / query services which make use of Provenance protocols to monitor application activity and provide the interfaces needed to answer Provenance questions. Following general rule **OTM.23**, such services are mapped to *Management* and *Query Actors*.

The management and query services are provided by two separate components:

- *Management and Monitoring Service*: A system which provides a console with two major features: 1) simple configuration of access controls to the single sign on domain covering the deployed Provenance stores, 2) monitoring and statistical information over the operation of the Provenance stores. The service is not expected to carry out significant on-the-fly reconfiguration of Provenance stores.
- *Query Service*: A system which combines two main features: 1) single sign on to the Provenance system and (where possible) sign on to associated authentication domains within the OTM application itself, 2) a set of standard query types taken from Section 5.1.4.

In both cases these services will be developed on top of tools provided by WP6 of the Provenance project.

5.1.3 Run-time Provenance Storage

This section defines how data will actually be stored by the OTM application.

5.1.3.1 Storing Events and States

Provisions for storing data and events are set in accordance with the following general rules defined in Chapter 4: **OTM.6** (recording of all decision, result and edit or consult actions), **OTM.10** (subdivision of actors), **OTM.18** (use of actor side libraries) and **OTM.19** (inclusion of `p-headers`). In the case of the OTM application therefore, *all the events, decision and conditional points* that appear in the process workflows (Figures 11 to 14) are relevant and, therefore, *should be recorded properly in the Provenance stores*.

To illustrate how events, decisions and conditions are stored, we present an example for a single OTM service: HR.Organ Offer Evaluation.[22] The following figure isolates that OTM service retaining only its internal detail and information flow with other OTM services. Initially, an organ offer comes from OTA.Organ Offer Process, this is offer is evaluated by the Offer Evaluated event, a decision is made about the organ based on the Offer Evaluated event results and the patient record obtained from HR.Patient Care Record Store. If the organ is rejected, the OTA.Organ Offer Process is notified, otherwise the Recipient Selected event occurs and the OTA.Organ Offer Process again notified.

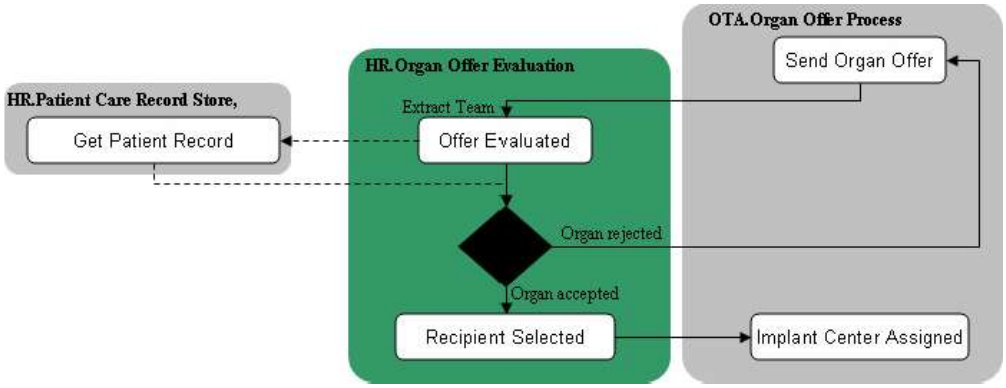

Figure 23 -- The Hr.Organ Offer Evaluation service and services it interacts with during execution.

Following General Rule **OTM.11**, we identify five application actors corresponding to the HR.Organ Offer Evaluation OTM service:

- The Controller (*URN:X-OTM.ACTOR.CatSalut.HR.otm-unit.OrganOfferEvaluation)*, which appears in the diagrams as **HR.Organ Offer Evaluation**
- An Internal Actor 1 (*URN:X-OTM.ACTOR.CatSalut.HR.otm-unit.OrganOfferEvaluation. involved-in-OfferEvaluated)*, who is involved in the Offer Evaluated event and appears in the diagrams as **1**.
- An Internal Actor 2 (*URN:X-OTM.ACTOR.CatSalut.HR.otm-unit.OrganOfferEvaluation. reports-OrganDecision)*, who reports the Organ Decision and appears in the diagrams as **2**.
- An Internal Actor 3 (*URN:X-OTM.ACTOR.CatSalut.HR.otm-unit.OrganOfferEvaluation. involved-in-RecipientSelect)*, who is involved in the Recipient Selected event and appears in the diagrams as **3**.
- An Internal Actor 4 (*URN:X-OTM.ACTOR.CatSalut.HR.otm-unit.OrganOfferEvaluation.*

22 Please note that, for the sake of clarity, we use in the example reduced versions of the service names. The full name for HR.Organ Offer Evaluation would be *URN:X-OTM.ACTOR.CatSalut.HR.otm-unit.OrganOfferEvaluation*, while the full name for HR.Patient Care Record Store would be *URN:X-OTM.ACTOR.CatSalut.HR.general.PatientCareRecordStore*, and for OTA.Organ Offer Process it would be *URN:X-OTM.ACTOR.CatSalut.OCATT.OrganOfferProcess*.

`reporting-Results)`, who is involved in the creation of the documents/reports for humans about the process[23] and appears in the diagrams as **4**.

As remarked in the 5th point of Generic Rule **OTM.11**, in some cases Internal Actors may record (as an actor p-state assertion) information about who made a decision (for instance, the chief of the team evaluating the offer), if such information is available. In addition, we identify two external actors corresponding to the Controllers of the OTM services with which this OTM service interacts.

- The Controller of OTA.Organ Offer Process.
- The Controller of HR.Patient Care Record Store.

We assume for this mapping that all interactions between OTM services go via the Controllers of those services, for example if these Controllers are Web Services, all information is communicated via Web Services and possibly then to GUIs for the use of medics.

Figures 24 to 28 show the evolution of the contents of a Provenance store while the execution of HR.Organ Offer Evaluation is taking place. Application Actors are shown as ovals; interactions are shown by solid, arrowed, horizontal lines between actors; causal relations are shown as dotted, vertical lines between interactions; and, actor state information are shown in boxes attached to actors with dashed horizontal lines. First, we show how the OTM service is activated. As mentioned in Section 3.1.1, services are activated to carry out their function by receiving a "virtual token", denoting that the service is active. This token is part of the `p-header` of the message that triggers the service, and includes any tracer(s) needed for tagging properly all the p-assertions in the Provenance store. In our example, the Controller of the HR.Organ Offer Evaluation receives an Organ Offer message (including the activation token) from the Controller of the OTA.Organ Offer Process, and passes it to the internal actor 1, who is involved in the Offer Evaluated event.

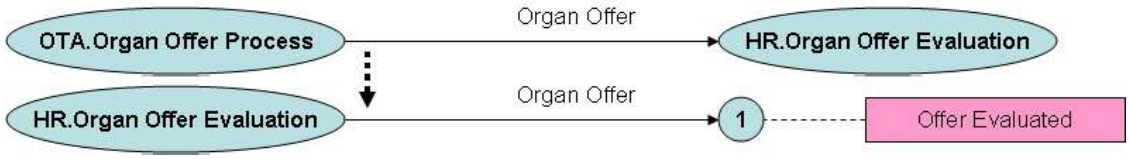

Figure 24 -- Partial content of the Provenance Store when the HR.Organ Offer Evaluation service is triggered.

The Offer Evaluated event causes a patient record to be retrieved from the HR.Patient Care Record Store and this record sent to the actor making the Organ Decision, with all communication between OTM services being conducted via the Controllers.

23 These documents/reports are for human use and should not be confused with the p-assertions all application actors do in the Provenance store to document the Provenance of the result.

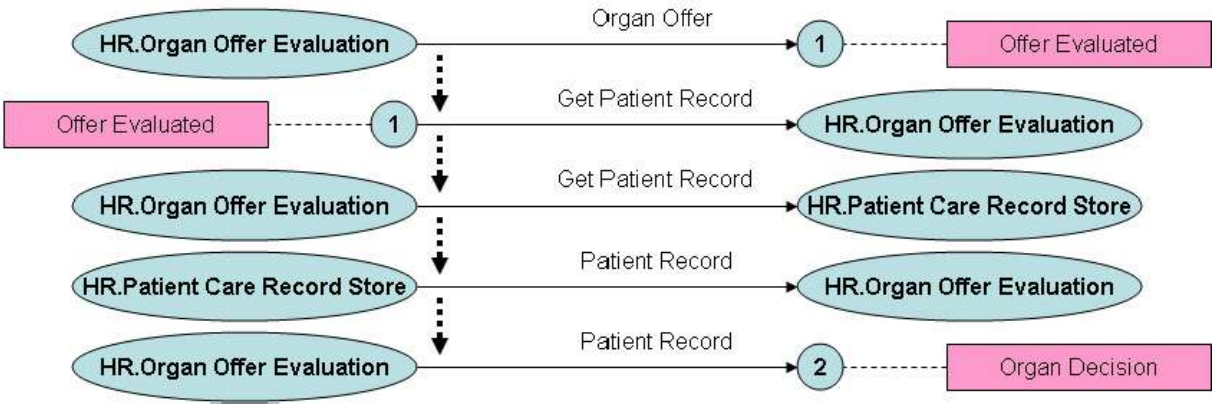

Figure 25 -- Partial content of the Provenance Store after the patient record is retrieved.

As shown in the original OTM service figure, the Organ Decision uses both the patient record and the result of the Offer Evaluated event to decide whether to accept or reject an organ. Here, for brevity, we assume rejection and so this rejection is sent back to the OTA.Organ Offer Process.

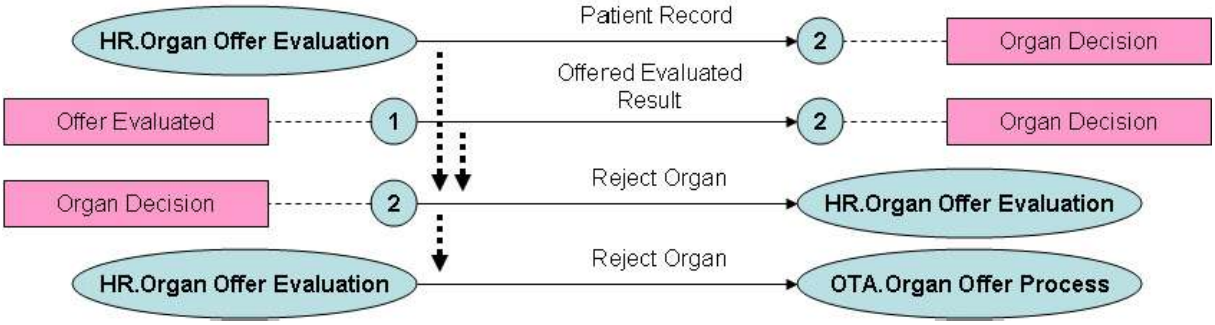

Figure 26 -- Partial content of the Provenance Store after the decision (rejection) is taken.

Finally, after the HR.Organ Offer Evaluation has completed, standard medical documentation is produced based on the events that have occurred and decisions made. Therefore, each actor that performed an event, including the initial trigger of the OTM service, or made a decision must send that documentation to the actor documenting the process. For each event or decision, the documentation of its outcome is caused by the interaction that triggered the event or decision (nothing would be documented if nothing was triggered to occur).

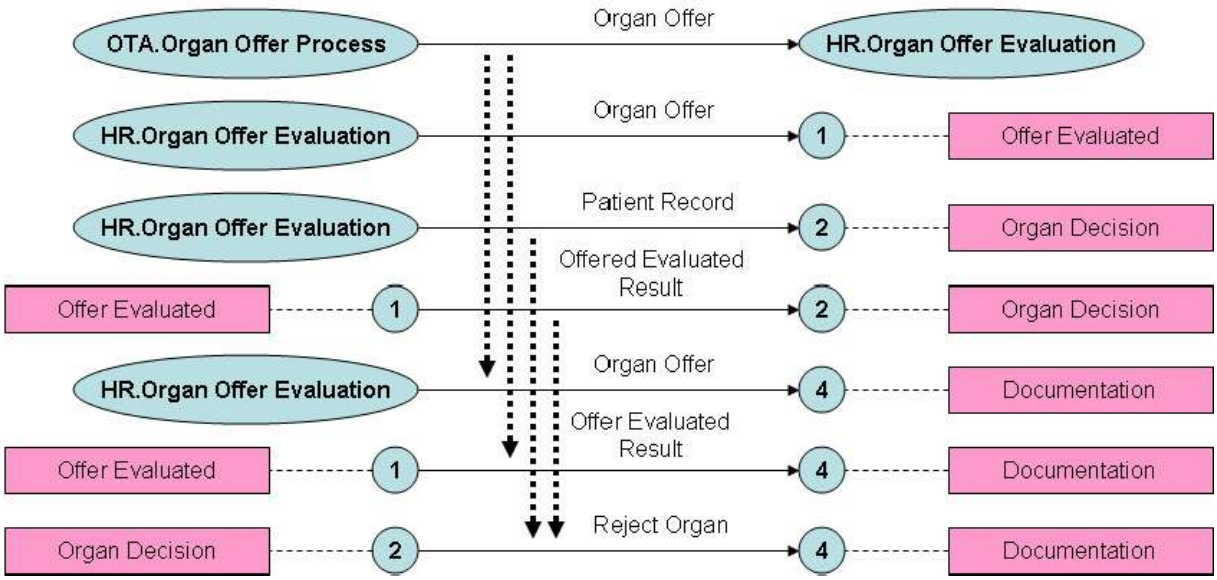

Figure 27 -- Partial content of the Provenance Store when Internal Actor 4 documents the process.

Figure 28 shows the full contents of the Provenance store after an organ has been rejected. It is important to note that the figure is an schematic representation of the contents of the Provenance store, as:

- interactions between controllers of the different services will be stored following the message schemas in Section 4.1.1.3.2.;
- interactions between the controller and the Internal Actors and between two Internal Actors are stored as standard interaction p-assertions;
- causal relations in the diagram (dotted, vertical lines between interactions) are recorded as relationship p-assertions;
- actor state information from the internal actors is stored as actor state p-assertions following the schemas in Section 4.1.1.3.1.

In this example, all information is recorded directly by the actors involved in the events or decisions into Provenance stores. But in some cases, when special processes, decisions, conditions or events occur in the real world by individuals or teams which are not represented in the system (e.g, a lawyer or judge involved in the legal consent for non-natural deaths; a private laboratory which is not connected to the OTM application performing an specific test), the information may be inferred from the standard medical/legal documentation produced at the end of the process (the legal consent, the laboratory results).

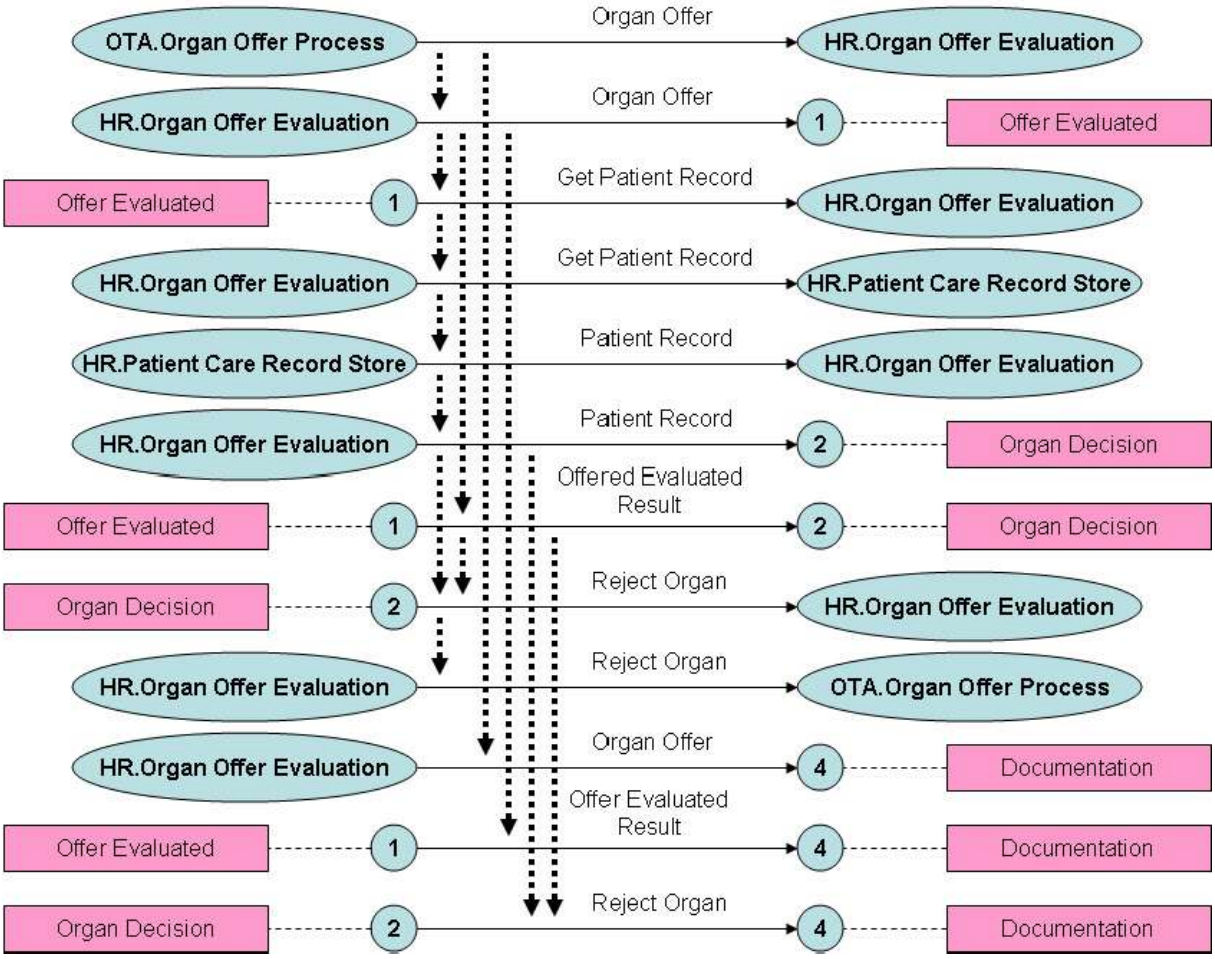

Figure 28 -- Content of the Provenance Store after the execution of the HR.Organ Offer Evaluation service.

5.1.3.2 Storing Interactions

As described in Section 4.1.1.3.2, the following general rules apply to interaction storage:

- In general, all messages between Application Actors are stored in Provenance traces by both the sender and the receiver (***General rule OTM.9***).
- However, they will often be stored in reduced form to remove sensitive medical information (***General rule OTM.8***).

The application uses asynchronous, document style message passing over Web Services SOAP interfaces to exchange data between services using an XML encoding and semantics derived from the FIPA ACL Agent Standard. Messages will therefore have the following characteristics (see Section 4.1.1.3.2 for a high level view):

- Interchanges will be asynchronous, one-way messages with invocations passing data in one direction but not expecting an immediate response.
- Interactions which require more than one message are identified using a conversation identifier generated by one of the participants to allow Web Services to track which messages belong together.
- All messages will have a number of common *meta data elements*: a message type (Request, Agree, etc.), named `sender` and `receiver`, `protocol` in use, `content type` in use, and

`content`. This data also includes the `p-header data` required by the Provenance architecture.

- Message content will be split into three possible types, which can co-exist in the same message:
  - *XML encoded data*: explicitly capturing the data which the Web Service receiving the data may be able to process. This comes in two forms: *open data* – which may be stored in the Provenance store and *closed data* – which for security reasons may not be.
  - *Attachments*: *data items* such as medical data items which are generally NOT stored in the Provenance store. Attachments are always accompanied by a reference in the XML encoded data which indicates the nature of the attachment and where the original is stored (its identifier in a data store). In some cases the attachments themselves may not be sent – simply referred to.

A set of pseudo-code messages is given in Figures 29 to 31. These messages show an example of a serology test (i.e. a blood test to search for infectious diseases) that is requested by the transplant unit to the immunology laboratory (Figure 29) and the full form of the message returned (all the information of the test, which in the example detected Hepatitis C antibodies) and the reduced form of the message (to be stored in a Provenance store).

The reduced form of a messages to be stored in a Provenance store is defined as being: the message *meta data* and the *open XML encoded data*. Further, the open XML encoded data is expected to include references to any closed XML encoded data or private attachments which are associated with the message. Appendix A shows, for each piece of clinical data, the part that can be stored in the Provenance Stores. In the example (Figure 30) we can see that the reduced form of the message only contains unclosed information about the date and time of the test and that the test was positive, without further details. Only an actor with the proper clearance level will be able to see the closed content to get more information about which serology tests were positive. The Provenance store may also capture invocation information for the message at the Web Service interfaces (e.g. precise method call parameters etc.).

```
<Message>
 <Header>
      <Type>QUERY</Type>
      <Sender>ACTOR.CatSalut.hsp.otm-unit</Sender>
      <Receiver>ACTOR.CatSalut.hsp.immunology.dutylab</Receiver>
      <Message-ID>E987324</Message-ID>
      <In-Response-To>NA</In-Response-To>
      <Conv-ID>ACTOR.CatSalut.hsp.otm-unit:832</Conv-ID>
      <Ontologies>OTM-Ontology</Ontologies>
      <Content-Language>OTM-ContentLanguage</Content Language>
      <Protocol>Query</Protocol>
      <Pheader> --- </Pheader>
 </Header>
 <Body>
      <Open Content>
            <TestResult>
                  <Type>Serology</Type>
                  <Patient-ID>EU384802-FC</Patient-ID>
            </TestResult>
      </Open Content>
 </Body>
</Message>
```

Figure 29: Example Actor – Actor message, requesting the results of a serology test for a particular patient.

```
<Message>
   <Header>
        <Type>INFORM</Type>
        <Sender>ACTOR.CatSalut.hsp.immunology.dutylab </Sender>
        <Receiver>ACTOR.CatSalut.hsp.otm-unit</Receiver>
        <Message-ID>E987323</Message-ID>
        <In-Response-To>E987324 </In-Response-To>
        <Conv-ID>ACTOR.CatSalut.hsp.otm-unit:832</Conv-ID>
        <Ontologies>OTM-Ontology</Ontologies>
        <Content-Language>OTM-ContentLanguage</Content Language>
        <Protocol>Query</Protocol>
        <Pheader> --- </Pheader>
   </Header>
   <Body>
        <Open Content>
              <TestResult>
                    <Type>Serology</Type>
                    <Patient-ID>EU384802-FC</Patient-ID>
                    <Date>2005.08.12</Date>
                    <Time>22:35.15(GMT+1)</Time>
                    <Status>Completed</Status>
                    <Diagnose>POSITIVE</Diagnose>
              </TestResult>
              <Record>
                    <Type>Serology.Record.XML</Type>
                    <DataStore>DATA.CatSalut.hsp.x1</DataStore>
                    <DataItem>DATA.CatSalut.hsp.x1:657</DataItem>
              </Record>
              <Record>
                    <Type>Serology.Record.PDF</Type>
                    <DataStore>DATA.CatSalut.hsp.x2</DataStore>
                    <DataItem>DATA.CatSalut.hsp.x2:657</DataItem>
              </Record>
        </Open Content>
        <Closed Content>
              <TestResult>
                    <Result1>
                          <Test>Serology.HbsAg</Test>
                          <Value>Negative</Value>
                    </Result1>
                    <Result2>
                          <Test>Serology.antiCoreHBV</Test>
                          <Value>Negative</Value>
                    </Result2>
                    <Result3>
                          <Test>Serology.antiHCV</Test>
                          <Value>Positive</Value>
                    </Result3>
                    <Result4>
```

```
                        <Test>Serology.antiHIV1</Test>
                        <Value>Negative</Value>
                    </Result4>
                    <Result5>
                        <Test>Serology.antiHIV2</Test>
                        <Value>Negative</Value>
                    </Result5>
                    <Result6>
                        <Test>Serology.HIV1p24</Test>
                        <Value>Negative</Value>
                    </Result6>
                    <Result7>
                        <Test>Serology.antiCMV</Test>
                        <Value>Negative</Value>
                    </Result7>
                    <Result8>
                        <Test>Serology.RPR</Test>
                        <Value>Negative</Value>
                    </Result8>
                    <Result9>
                        <Test>Serology.HATP</Test>
                        <Value>Negative</Value>
                    </Result9>
                    <Result10>
                        <Test>Serology.EBVIgG</Test>
                        <Value>Negative</Value>
                    </Result10>
                    <Result11>
                        <Test>Serology.ToxoplasmIgG</Test>
                        <Value>Negative</Value>
                    </Result11>
                    <Result12>
                        <Test>Serology.SHVIgG</Test>
                        <Value>Negative</Value>
                    </Result12>
                    <Comments>Antibodies of Hepatitis C found in test
                     (antiHCV test is clearly positive).</Comments>
                </TestResult>
            </Closed Content>
        </Body>
    </Message>
```

Figure 30: Example Actor – Actor message, providing a response to the message in Figure 29. Included are a high meta data, a high level view of the result, references to two forms of stored data (an XML form which might contain the closed-content element of the message and a PDF form) and a closed content which gives more details of the test result.

```
<Message>
 <Header>
      <Type>INFORM</Type>
      <Sender>ACTOR.CatSalut.hsp.immunology.dutylab </Sender>
      <Receiver>ACTOR.CatSalut.hsp.otm-unit</Receiver>
      <Message-ID>E987323</Message-ID>
      <In-Response-To>E987324 </In-Response-To>
      <Conv-ID>ACTOR.CatSalut.hsp.otm-unit:832</Conv-ID>
      <Ontologies>OTM-Ontology</Ontologies>
      <Content-Language>OTM-ContentLanguage</Content Language>
      <Protocol>Query</Protocol>
      <Pheader> --- </Pheader>
 </Header>
 <Body>
      <Open Content>
            <TestResult>
                  <Type>Serology</Type>
                  <Patient-ID>EU384802-FC</Patient-ID>
                  <Date>2005.08.12</Date>
                  <Time>22:35.15(GMT+1)</Time>
                  <Status>Completed</Status>
                  <Diagnose>POSITIVE</Diagnose>
            </TestResult>
            <Record>
                  <Type>Serology.Record.XML</Type>
                  <DataStore>DATA.CatSalut.hsp.x1</DataStore>
                  <DataItem>DATA.CatSalut.hsp.x1:657</DataItem>
            </Record>
            <Record>
                  <Type>Serology.Record.PDF</Type>
                  <DataStore>DATA.CatSalut.hsp.x2</DataStore>
                  <DataItem>DATA.CatSalut.hsp.x2:657</DataItem>
            </Record>
      </Open Content>
      <Closed Content>Omitted</Closed Content>
 </Body>
</Message>
```

Figure 31 – Example of reduced Actor – Actor message stored in a Provenance store. This message is identical to that in Figure 30 apart from the removal of the closed content.

Whilst the figures in the messages are presented in XML pseudo-code, in the application they are encoded as WSDL interfaces / SOAP method calls such that the top level elements in the meta data and content are individual parameters in a standard method call. Hence a typical method call would take the following form:

- `messageDeliver(type, sender, receiver, ...).`
- or `messageDeliverInform(sender, receiver, ...)`.

The `p-header` will be instantiated using the standard pattern provided for by the Provenance architecture specification and implementations thereof. Assertions for messages exchanged can be made in several forms: either in the form of the WSDL method call, in the form of the complete XML message body or as a collection of individual assertions for each element of the message.

5.1.4 Provenance Queries

5.1.4.1 The Objects of Provenance Queries

Before treating suggested Provenance queries themselves, it is useful to consider what the objects of those queries may be (specifically what the queries are about) since the transplant process does not produce a single datum or result.

While other possible query objects may be added, the main possible query objects to be considered include:

- *Recipient Perspective – Post-Surgery*: The clinical outcome of implantation surgery on a particular recipient at stabilization after surgery (rejection, acceptance of the organ). Represented by / captured in a surgery report.
- *Recipient Perspective – After Care*: The clinical outcome of implantation surgery on a particular recipient at some interval of time after surgery and in post-care – typically 1month, 1 year, 3 years of 5 years after implantation (rejection, full versus partial functionality, complications). Represented by /captured in a surgery report.
- *Organ Perspective*: The fate of a particular potential donor organ (extracted? extracted and rejected? donated? rejected? accepted?). Represented in / captured in one of several surgery reports / decision documents depending on what the final result was.
- *Donor Perspective*: The fate of the set of potential donated organs from a particular patient (a patient may donate different organs to different recipients). Represented in / Captured in a set of surgery reports depending on the outcome.

Each of these views are potential "outcomes" about which queries on the process which lead to the outcome might be structured. Each of the objects of a query generally has:

1. An *essential nature* (outcome positive, outcome negative etc.).
2. An *internal structure* of standard fields for the result type (for example, for a medical outcome standard indicators such as patient vital life signs, standard potential complications, drug dosages applied etc.). Some of these may be "hidden" and only available in full reports, some may be "visible" - available in the summary data in a Provenance record.
3. A set of *additional administrative attributes (metadata)* which characterize the outcome in various ways such as: entity responsible for the result, data produced etc. Some of these may be "hidden" and only available in full reports, some may be "visible" - available in the summary data in a Provenance record.

Ideally it should be possible to ask Provenance questions each of these types of properties of the object of a query.

5.1.4.2 Example Expected Queries

The following are a set of potential Provenance queries which could usefully be asked over the p-assertions logged in Provenance stores (and by extension over the more detailed medical reports logged in the health care application itself. The queries are classified according to:

Recipient Perspective [Post-Care and Post-Surgery Perspectives]

From the recipient perspective the intrinsic result is the outcome of the surgery, e.g. Success / Failure / Partial Success. Within this there is a significant amount of structure covering for example: acceptance / rejection of the organ, complications incurred, test results, surgical notes, medication used etc.

--- *On the overall result*

1. Retrieve meta-data and references to all actions / events associated with a particular case.
2. Retrieve meta-data and references to all actions / events associated with a particular case as well as medical system data for which the current query issuer(s) has (have) access permission.[24]
3. Determine a decision tree for a particular case (decision nodes only).
4. Determine a medical analysis tree for a particular case (medical data items only).
5. Determine whether a standard work flow was followed or if there were deviations / unusual events.

--- *On aspects of the result*

6. Determine the likely contributing factors to a element/aspect of the internal structure of a result. [Carried out in a number of possible ways – potentially by a search pattern which identified known possible patterns in supporting medical reports to see if they occurred.] For example if the final medical report notes the presence of a certain type of pathology – determine whether there are prior indicators as to why this might be the case. Example possibilities:
   i. A known side effect of a drug used in one of the operations.
   ii. A prior condition of the patient.
   iii. A combination of two separate results (for example, a certain type of blood test result and a particular incident in surgery).

--- *Result Meta-Data*

7. Determine the evolution / reasons for composition of one or more of the meta-data parameters of a result such as:
   i. The medical staff members participating in an outcome or responsible for one or more steps of the outcome.
   ii. The set of institutions involved in decision making (potentially including those which turned down an organ).
   iii. The time taken for the result to be reached (breaking down in time taken for each step in the process).
   iv. The generation of medical warning flags associated with an outcome (medical warnings - special conditions which must be watched due to a particular reading / finding that can be added by a step in the process) – which tests/decisions resulted in which flags and which staff were responsible.

--- On aggregate data

8. Deriving aggregate data across many transplant incidents for any one of the above queries.
9. Establishing whether a particular case is a statistical outlier in terms of decisions made, results obtained etc. with respect to the aggregate.

24 Note that this should include tracking version numbers of results actually used at the time of execution.

Organ Perspective

For an organ the intrinsic result falls into a broad category such as: not considered, patient rejected, organ rejected before extraction, rejected upon post-extraction examination, implanted but unsuccessfully, implanted successfully but then rejected, implanted successfully and accepted. Beyond this data available would cover: dimensions of the organ, blood types, functioning (if implanted), surgical notes etc.

Questions are likely to be similar to those for recipients but with a different perspective.

--- *On the overall result*

1. Retrieve meta-data and references to all actions / events associated with a particular organ.
2. Retrieve meta-data and references to all actions / events associated with a particular organ as well as medical system data for which the current query issuer(s) has (have) access permission.[25]
3. Determine a decision tree for a particular organ (decision nodes only).
4. Determine a medical analysis tree for a particular organ (medical data items only).
5. Determine whether a standard work flow was followed or if there were deviations / unusual events.

--- *On aspects of the result*

6. Determine the likely contributing factors to a element/aspect of the internal structure of a result. For example if the final result is a rejection and this is noted to be due to a certain type of organ damage – determine whether there are prior indicators as to why this type of deformity may be present. Example possibilities:
    1. A known side effect of a condition the donor had.
    2. A surgical error / problem.
    3. A combination of two separate results (e.g. a certain type of donor, a certain type of complication in surgery).[26]

--- *Result Meta-Data*

7. Determine the evolution / reasons for composition of one or more of the meta-data parameters of a result such as:
    i. Similar to recipient cases.

--- On aggregate data

8. Deriving aggregate data across many transplant incidents for any one of the above queries (particular matching organ types / rejections against background context).
9. Establishing whether a particular case is a statistical outlier in terms of decisions made, results obtained etc. with respect to the aggregate.

Donor Perspective

The questions here are very similar to those posted for the recipient perspective, however they are carried out over a greater number of process executions (since one donor may give rise to several implantations).

25 Note that this should include tracking version numbers of results actually used at the time of execution.

26 Again ideal here would be the ability to look for know patterns for particular outcomes in organ outcomes – explaining why a particular outcome was the way it was. This may be in the medics notes explicitly but may not be.

5.1.4.3 Rules/Mappings/Guidelines for Extracting Responses to Provenance Queries from the Data Stored.

While determining the exact algorithm / query pattern to use to extract results for each of the queries described in the previous section is an extensive job, a number of examples are provided here as to how results may be obtained. In each case queries need to be made in terms of sub selects of increasing specificity over data stored.

Example 1: *Meta-data and references to all actions / events associated with a particular case.*

This information is easily available in a one step query if the case identifier is known (since it appears in the p-header) or in two steps if the case identifier is not known:

1. Query registry of case identifiers (one or more) with case details to return possible matches as case identifiers.
2. Use the case identifier selected to return all related meta-data associated with the case-id which is used as a tracer and stored: in the p-header of messages and forms part of the assertion for events / states.

Example 2: *Determine a decision tree for a particular case (decision nodes only).*

Given access to the data related to a particular case, a view of the decisions taken in a case can be generated by:

1. Extracting all decision events recorded (each of which has an identifier)
2. Extracting all conversations communicating a decision (each of which has an identifier and refers to decision identifiers).
3. Constructing a temporally tree graph of the decision by assigning decisions and their decision makers responsible for them together as nodes and conversations as linking decisions as arcs in the tree.[27]

This provides a view across the meta data of decision items, but not necessarily details of each decision and the reasons for taking them. This type of data can be extracted at the next level.

Example 3: *The medical staff members participating in an outcome or responsible for one or more steps of the outcome.*

Given a decision tree, data tree or other view on a case two types of further extraction can be carried out:

- Gathering more meta-data associated with the tree
- Drilling down into restricted data.

An example of the former is iterating over the tree to expand upon all the names reasons / events given for a decision being made at each node. Further iterations may reveal deeper chains of causality for decisions. However this view can only work with high level events stored as public in the Provenance stores.

For the second type of query a user must have clearance to access one or more medical data stores in addition to the Provenance stores. In this case the query mechanisms may also access the data stores the user has clearance for to retrieve:

- XML encoded closed data: which provides more detail on an event/decision and may be used to explore more parts of the tree.

27 Note that due to clock synchronisation issues the graph may not be in perfect temporal order – but the flow of activity should be correct.

- PDF / Other formatted data which cannot be further machine processed but provide the user with a detailed account of an event/decision.

5.2 Provenance Handling in the EHCR Application

Provenance handling in the EHCR application is divided into two sections: the runtime storage of process documentation (Section 5.2.1) and queries over process documentation (Section 5.2.2).

5.2.1 Run-time Provenance Storage

This section defines how data will actually be stored by the EHCR application. The overall data flow for the application is defined in Section 3.2.2.

5.2.1.1 Storing Events and States

The data stored by the application is based on the pre-standard adopted in the application and is used to generate separate p-assertions for both states and events.

The EHCR store makes the following two types of p-assertions:

1. A given part of an ECHR of a patient (identified by the GMPID) was modified at a given time at a given EHCR store.
2. The EHCR of a patient (identified by the GMPID) was updated at a given time at a given EHCR store.

This makes the p-assertions secure, because from these p-assertions one can only know that the EHCR of someone was modified somewhere or was up-to-date somewhere at sometime. The actual medical information is moved only between EHCR stores based on these Provenance information using ENV 13606 rules.

The parameters of the p-assertions are the following:

P-assertion Events:

s1. ehcrChanged: EHCR changed

parameters:
gmpid:String,
timestamp:Timestamp,
ehcrs:EHCRS
ehcrComponents:Set

P-assertion States:

s2. updateEHCRSuccessful: whether the result of an update process is a full and up to date EHCR or not

parameters:
gmpid:String,
timestamp:Timestamp,
ehcrs:EHCRS

The submissions are always made by EHCRS actors as defined in Section 3.2.2.

5.2.1.2 Storing Interactions

The following pseudo-code messages represent the types of messages exchanged by services in the system. A similar division will be carried out as detailed in Section 5.1.3.2 in order to separate sensitive data from open data in the Provenance store.

Interactions:

s3. ehcrMessageSent: EHCR message sent

parameters:
gmpid:String,
timestamp:Timestamp,
sender:EHCRS,
receiver:EHCRS,
type:String[Request|RequestAll|Provide|Notification],
senderHealtcareAgent:HealtcareAgent,
receiverHealtcareAgent:HealtcareAgent,
ehcrComponents:Set
...

s4. ehcrMessageReceived: EHCR message received

parameters:
gmpid:String,
timestamp:Timestamp,
sender:EHCRS,
receiver:EHCRS,
type:String[Request|RequestAll|Provide|Notification],
senderHealtcareAgent:HealtcareAgent,
receiverHealtcareAgent:HealtcareAgent,
ehcrComponents:Set
...

If there is no sender of the message, then the message comes from a healtcareAgent without EHCRS. In this case the EHCRS will also send an ehcrChanged p-assertion to Provenance.

5.2.2 Provenance Queries

Provenance queries in the EHCR system correspond to questions about patient care records which must draw together data from several sources.

5.2.2.1 The Objects of Provenance Queries

The application allows for scenarios which are both user activated (an external entity asking for a complete EHCR for a patient) or internal to the application. In this case EHCRS systems issue Provenance queries to one another to generate information updates.

Scenario1

An EHCRS wants to know whether it owns an up to date EHCR of a patient. The identification of patient is based on GMPID.

The EHCRS owns an up-to-date EHCR of the patient if, after the last successful update process of the asking EHCRS, the EHCR was not changed in other EHCRSs and other EHCRSs did not receive any provide messages in which the sender field was not filled in.

Scenario2

An EHCRS wants to update the EHCR of a patient. It needs information where the missing fragments of that EHCR can be found. The identification of the patient is based on GMPID.

Provenance gives a list of places (and timestamps of the update or change of EHCR) where the EHCR fragments can be found. The first item in the list is the place where the last successful update process happened (let us to call the time of the last successful update process as t1). The following items in the list are the places where the EHCR was changed before t1 in decreasing time order (i.e. newer updates first, followed by older updates).

5.2.2.2 List of Expected Queries

The types of queries expected based on the scenarios from the previous section are therefore:

q1. boolean isEHCRUpToDate: whether EHCR of a patient is up to date in an EHCRS or not.
return: true if the EHCRS owns up to date EHCR of the patient, no otherwise .

parameters: gmpid:String, ehcrs:EHCRS

q2. list getUpdatePlaces
return: a list of places (and timestamps of the update or change of EHCR) of where the missing fragments of an EHCR can be found.
parameters: gmpid:String, ehcrs:EHCRS

5.2.2.3 Rules/Mappings/Guidelines for Extracting Responses to Provenance Queries from the Data Stored.

The answers to queries q1 and q2 are generated from the information of s1 and s2 as it was described in scenario1 and 2 in 5.2.2.1. To answer the queries, the EHCR will assemble the latest full EHCR from the response of the Provenance store as shown in the following figure:

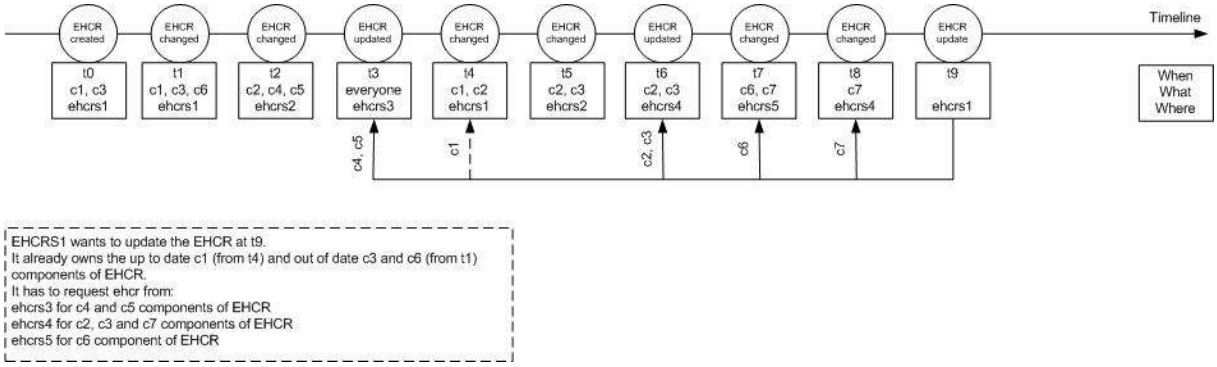

Figure 32 – Assembling a full EHCR from Provenance information.

5.2.2.4 Accessing Confidential Information from the Provenance Query: Patient Identities, Clinical Data

The EHCRS application does not store or query any data of identities of patient in the Provenance Store. Instead, it uses the GMPID to identify patients. The connection between GMPID and identities of patient comes from a separate application that we called ws_crypto application. This application can be run for example in CatSalut and the identification of patient could be based for example on national insurance number. It means that the patient identities and data from Provenance could be assigned only with Provenance and ws_crypto application together. Provenance alone is safe from that point of view. Further, clinical data are not stored in the Provenance system. Clinical data is stored and moved only between EHCR stores which provide these data only through proper authorisation. In order to deal with this the EHCR application therefore makes the following two assumptions;

- A name mapping application such as the one shown in Figure 22 (Section 4.2.1.3) will be present in the environment (operated by an Actor such as CatSalut for example).
- Authentication / Sign-on services exist for each domain which may hold sensitive data that may be linked to a medical record.

These systems are discussed in Section 5.3.

5.3 Patient Identifier Anonymisation and Access to Sensitive Health Care Data Items

One of the most important recurring themes in the modeling of the OTM and EHCR application is the need to protect sensitive patient data at all times. This is reflected in five general rules arising in different parts of the specification in Chapter 4:

- **OTM.1**: Anonymisation of patient identifiers must be carried out before any data is stored in Provenance stores. A system wide anonymisation mechanism is required.
- **OTM.2**: Source medical data is never stored in Provenance stores but only referenced therein. A system wide medical data referencing scheme is required.
- **OTM.5**: Provenance stores are interlinked and communicate with one another, they are considered to be in one single-sign-on domain for security purposes, even though the application components will generally not be.
- **EHCR.4**: Patient information are stored in the Provenance system only through references using the Global Medical Patient ID (GMPID).
- **EHCR.5**: Health care data are stored in the Provenance system only through references using the system wide medical referencing scheme.

The following sections describe security / authentication mechanisms which will be put in place on behalf of the health authority in order to approximate the required data safety measures. These measures are expected to be used across both the OTM and EHCR applications.

5.3.1 Protected Identities for Patients

One of the most important principles of medical data storage is that patients should not be identifiable to anybody other than someone with the correct authorization. As described in Section 4.1.2.3, an important step to combating this is the use of substitute / temporary identifiers which can only be mapped to public IDs such as the Spanish national identification number (DNI) or regional health service number (the CatSalit Patient Identifier) through restricted naming services.

In the case of Catalunya the regional health authority (CatSalut) is the primary body responsible for keeping a list of registered patients (i.e. Citizens that are covered by CatSalut Health Services) . It is CatSalut the body that assigns the patient identifiers (The CatSalut Patient Identifier). Health Care Centers under the authority of CatSalut should check, for each patient, if the patient is registered in CatSalut, in order to know if the costs of the health services will be fully, partially or not-covered at all by CatSalut. Apart from that mandatory check, there is right now no obligation for hospitals to use the CatSalut Patient Identifier as part of the health care record of the patient (some still use the Spanish DNI as patient identifier), although some centers are moving towards the use of CatSalut ID's.

Currently there is no single, unified system to anonymize patient data. Each Health Care center implements EU ans Spanish regulations on data protection following their own procedures. However, in the OTM application demonstrator we propose to deploy a unified patient ID anonymization system as follows:

- CatSalut will provide a naming service which relates a medical service internal ("substitute") identifier for a given patient to public identifiers such as the DNI and CatSalut number.
- This identifier is named the GMPID as defined in Section 4.2.1.3.[28]

28 It is important to note the difference between the GMPID (to be provided by a service in CatSalut) used by both the OTM and the ECHR application to anonymously identify the p-assertions related to a given patient,

- Access to the mapping service provided by CatSalut is restricted only to authorized medical staff.
- The OTM application uses the GMPID provided by CatSalut's namig service in order to tag all messages between the services related to a given patient, and to tag all p-assertions stored in various Provenance stores about that patient. In this way only those who have the proper credentials to get the link from the real identity of the patient to the GMPID will be able to cross such information.

These provisions provide basic anonymization for data in the system. However, the OTM application creates the possibility for the generation of a large amount of correlated data about a patient even if the original identity of the patient is not known. One element of the solution to this is to restrict the amount of data which can be correlated (see next section). However, in addition to this, a number of additional mechanisms are considered for the OTM application (not all of which may be used in the demonstration system).

- *Local Identifiers*: even if not related to a patient's real identifier a single token used throughout storage of all Provenance assertions provides a simple way to retrieve all data related to one individual. To combat this phenomenon, an additional mechanism of identifier dereferencing could be put in place in which individual organizations assign arbitrary local names to patients. The use of LMPID's for local clinical data storage (introduced in 3.2.2.2) is an implementation of this mechanism.
- *Local Patient and Case Identifiers*: although the use of different identifiers reduces one possible correlating element, other tokens such as the CASE identifier (tracer) or certain dates may still provide a strong correlating identifier. We will study if further randomization and renaming should be carried out for these features only.

As a result of these measures, an individual with unauthorized access to the Provenance stores would not be able to identify which assertions related to the same case – with the contents of the Provenance stores essentially scrambled. These two solutions however clearly also have an effect on the Provenance system since they remove identifiers (tracers) which are relied upon to formulate queries. A simple way to deal with this however would be to require users of the Provenance system to sign into the OTM naming service prior to query execution and allow the query mechanism to retrieve records based on the related identifiers. Access to the naming service would therefore allow unscrambling of Provenance store content.

5.3.2 Storage and Retrieval of Sensitive Medical Data

As the example messages in Section 5.1.3.2 show, certain types of data may not be generally released and are accessible only to specialists treating a patient. Further, other types of data may compromise the anonymity of a patient. The general rule that source medical data should not be stored in the Provenance stores must therefore be enforced. However, this is interpreted in different ways in the OTM and EHCR application:

- In OTM the Provenance stores are considered external to the application data stored – with a lower clearance level. For this reason, a combination reduction/summarization and referencing is used for such data. Reductions are used as described in Section 5.1.3.2 to summarize information for the Provenance store such that only high-level non-sensitive material remains – an analysis for what this entails for a range of test types can be found in Annex A of this document. Referencing is used as described in Section 5.1.3.2 by identifying records stored in secure medical databases by reference rather than including the data.
- In the EHCR application, the Provenance stores are used as the basis of EHCRS services. Without the accessibility of the Provenance stores, the EHCRS services can only provide limited services, because they cannot collect a complete EHCR. However medical data is not stored in the Provenance stores and medical data travels only between EHCR stores. In the EHCR case

and the LMPID, generated by the ECHR application only for local storage of patient data in a way that separates the identification data of the patient and the medical data.

therefore: 1) the availability of Provenance stores is a prerequisite of the EHCRS services and 2) provisions from the ENV 13606 pre-standard are used to govern access to medical data.

In each of these cases an analysis will be carried out at build time for each data type in order to carry out the approach to be used in a given case.

5.3.3 Single Sign-On for Provenance Stores

Although authority for access to medical data stores is dependent on individual institutions the Provenance stores in the application are deployed as a single security domain at the regional level (potentially with federated domains at national / international level). This measure:

- Significantly simplifies deployment of the Provenance stores.
- Implies (as described in Chapter 4) that Provenance stores are considered less secure than individual institution medical data stores.

Provisionally however be separate domains for the OTM and EHCR systems. Hence access to the former is no guarantee of access to the later and vice versa. This provision is primarily taken since A) the Provenance stores in each part of the application play a significantly different role and B) in a real world deployment EHCR access would likely be much more widespread than OTM access.

In terms of deployment/implementation of the sign-on services, the systems developed will follow standard guidelines provided by WP3 and WP4 of the project.

5.4 Summary

The OTM and EHCR applications present a significant and complex use-case for Provenance deployment. The demonstration system plan is likely to involve several tens of individual services and several tens of Provenance stores capturing a wide range of data. Furthermore the mapping provided here describes:

- How sensitive medical data is separated from “reduced” views in Provenance stores.
- How such data is subsequently references in Provenance assertions.
- Example naming conventions which can be used in scenarios as complex as this.
- Example queries which can be used in a medical domain.

The next step in the development of the OTM over Provenance application development is the completion of a detailed implementation and deployment plan. It is expected that a number of mapping decisions described in this document may need to be revised as a result of the new experience generated by this activity. The deliverable will also be revised one or more times to remain in line with:

- Updates of the Architecture specifications (D3.1.1).
- Updates of the Tool specifications (D6.1.1).
- Updates of the application mapping provided in WP7 (D7.1.1).

Appendix A Clinical Data to be Recorded

The following data items are an example of the types of elements contained in medical data items which may be exchanged and stored in the OTM or EHCR applications. A more extensive list will be provided with the demonstration description. For each type of data, an indication is given as to what may be recorded in the Provenance Store as unclosed content.

A.1 Data about Donors

A.1.1 Data for all donors

Anthropometric Data (from Patient Exploration)
- Height (cm)
- Weight (Kg)
- Chest circumference (cm)
- Waist circumference (cm)
- Sternum length (cm)

Allowable in Provenance Stores: all values (none of them is enough to identify the patient).

Analitical tests (see section A.3.1)
- Haematology
- Blood Biochemistry
- Urine Biochemistry
- Gasometry

Allowable in Provenance Stores: see section A.3.1 for details on each of them.

Microbiological and Immunological tests (see section A.3.2)
- Human Leukocyte Antigens Test [HLA]
- Serology
- Microbiological Cultures
- Urine sedimentation

Allowable in Provenance Stores: see section A.3.2 for details on each of them.

Medical Imaging (see section A.3.3)
- Chest Radiography
- Abdominal Echography
- Echocardiography
- Electrocardiogram

Allowable in Provenance Stores: see section A.3.1 for details on each of them.

A.1.2 Data about donor preservation

The following data is only gathered for brain-dead donors.

Haemodynamics
- Test Date
- Test Time
- Systolic arterial pressure (60-340 mmHg)
- Diastolic arterial pressure (0-160 mmHg)
- Hipotension > 30min (Sí / No)
- Premature Ventricular Complex [PVC] (3-30 mmHg)
- Previous cardiac arrest (yes + minutes / No)
- Preliminary analysis: free text pointing out relevant findings.

Allowable in Provenance Stores: Test Date, Test Time, and the extra observations made on the analysis.

Preservation Medications
(list for each date and time the medications given to the donor and the dosage)
- Medication Date
- Medication Time
- Dopamine (0-40 ug/kg.min)
- Dobutamine (0-40 ug/kg.min)
- Noradrenaline (0-13 ug/kg.min)
- Adrenaline (0-8 ug/kg.min)
- Desmopresine (0-10 ug/kg.min)
- Other (name + dosage in ug/kg.min)

Allowable in Provenance Stores: Medication Date and Time, list of medications given.

Transfusions
(list for each date and time the fluids used in the transfusions done before extraction)
- Transfusion Date
- Transfusion Time
- Red Blood Cells (0-15 units)
- Platelets (0-15 units)
- Plasm (0-15 units)

Allowable in Provenance Stores: Transfusion Date and Time, list of fluids.

A.1.3 Data about the donor organs

Organ anatomical description
- Organ structure (normal/pathological)
- Anomalies (yes/no + free text describing them)
- Number of cyst and local lesions
- Cyst/lesion description (for each one: size, location, type, free-text description)
- Number of tumours
- Tumour description (for each tumour: size, location, type, free-text description)

Allowable in Provenance Stores: Organ structure, anomalies, Number of cysts and lessions, number of tumours.

Organ Extraction Report
- Clamping date (i.e. instant when the organ is disconnected from the blood stream)
- Clamping time
- Perfussion liquids (Winsconsi, Eurocollins, other)
- Extracted (yes/no)
- Anatomical anomalies (yes/no + free text describing them)

Allowable in Provenance Stores: All of them, as none can be used to identify the donor.

Organ evaluation
- Valid (yes/no)
- Argumantation for no-valid (free text arguing the causes for the organ being non-valid, linking it to specific values in the tests performed)

Allowable in Provenance Stores: All of them, as none can be used to identify the donor.

A.2 Data about Recipients

A.2.1 Data for all recipients

Waiting List Data
- Recipient type (organ or organs required)
- Date of inclusion in the waiting list
- Urgency-0 status (yes/no)
- Birth date
- Gender
- Blood type (O A B AB)
- Rh (+/-)

Allowable in Provenance Stores: all except birth date and gender.

Anthropometric Data (from Patient Exploration)
- Height (cm)
- Weight (Kg)
- Chest circumference (cm)
- Waist circumference (cm)
- Sternum length (cm)

Allowable in Provenance Stores: all values (none of them is enough to identify the patient).

Analitical tests (see section A.3.1)
- Haematology
- Blood Biochemistry
- Urine Biochemistry
- Gasometry

Allowable in Provenance Stores: see section A.3.1 for details on each of them.

Microbiological and Immunological tests (see section A.3.2)
- Human Leukocyte Antigens Test [HLA]
- Serology
- Microbiological Cultures
- Urine sedimentation

Allowable in Provenance Stores: see section A.3.2 for details on each of them.

Medical Imaging (see section A.3.3)
- Chest Radiography
- Abdominal Echography
- Echocardiography
- Electrocardiogram

Allowable in Provenance Stores: see section A.3.1 for details on each of them.

A.3 Data from tests

A.3.1 Analitical tests

(all tests have date and time; there can be time series)

Haematology
- Test Date
- Test Time
- Haemoglobin (10-20 g/dL)
- Haematocrit (30-50 %)
- Mean Cell Volume [MCV] (80-1000 fL)
- Eritrocite sedimentation rate [ESR] (0-20 mm/1h)
- Leucocytes (4500-10000 /mL)
- Basophil granulocytes (0-300 /mL)
- Percentage basophil (0.0-2.5 %)
- Eosinophil granulocytes (0-1000 /mL)
- Percentage eosinophil granulocytes (0.0-14 %)
- Neutrophil granulocytes (0-1000 /mL)
- Percentage neutrophil granulocytes (0.0-10 %)
- Lymphocytes (1000-4000 /mL)
- Percentage lymphocytes (10-60 %)
- Monocytes (0-2000 /mL)
- Percentage monocytes (1-16 %)
- Platelet Count (100000-400000 /mL)
- Corrected Prothrombin Ratio [INR] (0-10)
- Cephalin Ratio (0-10)
- Fibrinogen (1.5-4 g/L)
- D-Dímer (300-3000 umol/L)
- Preliminary analysis: free text pointing out relevant deviations in the values.

Allowable in Provenance Stores: Test Date, Test Time and the extra observations made on the analysis.

Blood Biochemistry
- Test Date
- Test Time
- GOT (10-5000 U/L)
- GPT (10-5000 U/L)
- Alkaline Phosphatase (10-500 U/L)
- Gamma Glutanyl Transferase [GGT] (10-5000 U/L)
- Lactic Dehidrogenase [LDH] (0-1000 U/L)
- Total Bilirubin (5-200 umol/L)
- Direct Bilirubin (5-200 umol/L)
- Ammonia (5-200 umol/L)
- Sodium [Na] (100-200 mmol/L)
- Potassium [K] (2.0-8.0 mmol/L)
- Amylase (10-200 U/L)
- Lipase (0-200 U/L)
- Glucose (2.5-9.9 mmol/L)
- Glycohemoglobin (0-50 %)

- Calcium [Ca] (1-4 mmol/L)
- Creatine Kinase [CK] (0-10 U/L)
- Troponin T (0-3 ug/L)
- Phosphate (0.5-3.0 mmol/L)
- Magnesium (0.0-2.0 mmol/L)
- Urea (1.0-9.9 mmol/L)
- Creatinine (0-200 umol/L)
- Creatinine Clearance (20-150 mL/min)
- Myoglobin (80-140 ug/L)
- Lactic Acid (0.4-3.0 mmol/L)
- Proteins (20-200 g/L)
- Albumine (30.0-60.0 g/L)
- Cholesterol (1.0-9.9 mmol/L)
- Triglycerides (0.0-5.0 mmol/L)
- Total Acid Phosphatase (0.0-9.9 U/L)
- Prostatic Acid Phosphatase (0.0-9.9 U/L)
- HCG (0-10 U/L)
- Prostate Specific Antigen [PSA] (0-10 ug/L)
- Carcinoembryonic Antigen [CEA] (0-10 ug/L)
- Alpha Fetoprotein [AFP] (0-20000 U/L)
- Preliminary analysis: free text pointing out relevant deviations in the values.

Allowable in Provenance Stores: Test Date, Test Time and the extra observations made on the analysis.

Urine Biochemistry
- Test Date
- Test Time
- Albumine (Negative – Positive)
- Red Blood Cells (Negative – Positive)
- Specific Gravity (Negative – Positive)
- Amylase (10-2000 U/L)
- Liver bilis pigments (Negative – Positive)
- Preliminary analysis: free text pointing out relevant deviations in the values.

Allowable in Provenance Stores: Test Date, Test Time and the extra observations made on the analysis.

Gasometry
- Test Date
- Test Time
- Fraction of oxygen in inspired gas [FiO2] (0-100)
- Oxygen Saturation [Sa02]
- Time (min)
- Alveolar Oxygen Tension [PaO2] (40-700 mmHg)
- PaCO2 (15-110 mmHg)
- pH (7.11-7.50)
- Bicarbonate [HCO3] (10-40 mmol/L)
- Base Excess (-20 to +20)
- Preliminary analysis: free text pointing out relevant deviations in the values.

Allowable in Provenance Stores: Test Date, Test Time and the extra observations made on the anlysis.

A.3.2 Microbiological and Immunological tests

Human Leukocyte Antigens Test [HLA]
- Test Date
- Test Time
- A (two values between 0-99)
- B (two values between 0-99)
- C (two values between 0-99)
- DR (two values between 0-99)
- DW (two values between 0-99)
- DRW (two values between 0-99)

Allowable in Provenance Stores: in this case information is so important for histocompatibility donor-recipient, and the values indicate nothing sensitive about the patient identity or the patient health, that all values may appear in the Provenance store.

Serology
(all values are positive, negative or undetermined)
- Test Date
- Test Time
- HbsAg (test for Hepatitis B)
- anti-Core HBV (another test for Hepatitis B)
- anti-HCV (test for Hepatitis C)
- anti-HIV-I (test for HIV)
- anti-HIV-II (test for HIV)
- HIV-1 p24 antigen (test for HIV)
- anti-CMV (test for citomegalovirus)
- Rapid Plasmin Reagin [RPR] (test for sífilis)
- HATP (another test for sífilis)
- EBV IgG (test for Epstein-Barr virus)
- Toxoplasm IgG
- SHV IgG (test for Simple Herpes)
- Others (name and value)
- Preliminary analysis: free text pointing out the positive tests.

Allowable in Provenance Stores: in this case there is very sensible data. In the message recorded in the PS only will appear a “positive” (if any test was positive) “negative” (if all went negative) or “unknown” (if there is no information to certify a full negative).

Microbiological Cultures
(There is a long list of microorganisms)

- Starting Test Date
- Starting Test Time
- Test Duration
- Blood culture (list of organisms found)
- Urine culture (list of organisms found)
- Respiratory secretions (list of organisms)
- Preliminary analysis: free text pointing out relevant findings.

Allowable in Provenance Stores: Test Date, Test Time, Duration and the extra observations made on the analysis.

Urine Sedimentation
- Gran Stain (microorganisms observed yes/no)
- Red Blood Cell Count [RBC] (units / field)
- White Blood Cell Count [WBC] (units / field)
- Preliminary analysis: free text pointing out relevant findings.

Allowable in Provenance Stores: Test Date, Test Time, Duration and the extra observations made on the analysis.

A.3.3 Medical Imaging

Chest Radiography
- Test Date
- Test Time
- Images (set of images)
- Length Pulmonary vertex – left diaphragm (10-120 cm)
- Length Pulmonary vertex – right diaphragm (10-120 cm)
- Chest diameter (10-120 cm)
- Preliminary analysis: free text pointing out relevant findings.

Allowable in Provenance Stores: Test Date, Test Time, and the extra observations made on the analysis.

Abdominal Echography
- Test Date
- Test Time
- Images (set of images)
- Liver structure (homogeneous / non homogeneous)
- Live brightness (normal / pathological)
- Left kidney cortex thickness (mm)
- Left kidney medulla (normal / pathological)
- Left kidney cortex/medulla ratio [L C/M]
- Left kidney size (cm)
- Right kidney cortex thickness (mm)
- Right kidney medulla (normal / pathological)
- Right kidney cortex/medulla ratio [R C/M]
- Right kidney size (cm)
- Pancreas (normal / pathological)
- Preliminary analysis: free text pointing out relevant findings.

Allowable in Provenance Stores: Test Date, Test Time, and the extra observations made on the analysis.

Echocardiography
- Test Date
- Test Time
- Images (set of images)

- Septum thickness (6-14 mm)
- Rear wall thickness (6-14 mm)
- Global contractibility (preserved / decreased)
- Segmented contractibility (preserved / decreased)
- Sistolic left ventricular diameter 8-30 mm)
- Diastolic left ventricular diameter (20-50 mm)
- Eyection Fraction [EF] (10-90 %)
- Left atrium diameter (10-40 mm)
- Aortic root diameter [AR] (10-40 mm)
- Mitral valve (normal / pathological)
- Tricuspid valve (normal / pathological)
- Aortic valve (normal / pathological)
- Pulmonic valve (normal / pathological)
- Preliminary analysis: free text pointing out relevant findings.

Allowable in Provenance Stores: Test Date, Test Time, and the extra observations made on the analysis.

Electrocardiogram
- Test Date
- Test Time
- Images (set of images)
- Cardiac Rhythm (sinusoidal / pathological)
- Heart rate (beats/sec)
- Repolarisation Irregularities (Yes / No)
- PR interval (number)
- DQRS (number)
- QT interval (number)
- QT corrected interval [Qtc] (number)
- P wave (number)
- QRS (number)
- T wave (number)
- Preliminary analysis: free text pointing out relevant findings.

Allowable in Provenance Stores: Test Date, Test Time, and the extra observations made on the analysis.

References

[D2.1.1] Project Deliverable D2.1.1, "User Requirements Document". Version 1.0, 2005 . http://twiki.gridprovenance.org/pub/Restricted/DeliverableD2dot1dot1/GRID_PROVENANCE-STD-M3-UserRequirementsDocument-WP2-Input-final.pdf (Project Internal web site).

[D2.2.1] Project Deliverable D2.2, "Software Requirements Document". Version 1.0, 2005, http://twiki.gridprovenance.org/pub/Restricted/DeliverableD2dot2dot1/SRD-v1-ofr.sxw (Project internal web site).

[D3.1.1] Project Deliverable D3.1.1, "An Architecture for Provenance Systems", Provenance project.

[ENV13606] CEN/TC251 WG I.: Health Informatics-Electronic Healthcare Record Communication- Part 1-4, Final Draft prENV13606-1 (1999-2000).

[FIPA02a] The Foundation for Intelligent Physical Agents "FIPA Abstract Architecture Specification", Specification Number SC00001, December, 2002, available online at http://www.fipa.org/specs/fipa00001/ .

[FIPA02b] The Foundation for Intelligent Physical Agents "FIPA Communicative Act Library Specification", Specification Number SC00037, December, 2002, available online at http://www.fipa.org/specs/fipa00037/ .

[FIPA02c] The Foundation for Intelligent Physical Agents "FIPA Message Structure Specification", Specification Number SC00061, December, 2002, available online at http://www.fipa.org/specs/fipa00061/ .

[LopezNavidad97] A. Lopez-Navidad. Professional characteristics of the transplant coordinator. Transplantation Proceedings, (23):1607–1613, 1997.

[LopezNavidad97b] A. Lopez-Navidad, J. Kulisevsky, and F. Caballero, editors. El donante de ´organos y tejidos: Evaluación y manejo. Springer-Verlag Ibérica, 1st edition, 1997.

[Miles 05] Project Internal document "Representing Provenance Data in the OTM application", Simon Miles, Javier Vazquez, July, 2005. http://twiki.gridprovenance.org/bin/viewauth/Restricted/OrganTransplantPStructure (project internal web site).

[RFC1035] P. Mockapetris, "Domain Names – Implementation and Specification", Internet Engineering Task Force RCF 1035, November 1987, available online at http://www.ietf.org/rfc/rfc1035.txt .

[RFC2141] R. Moats, "URN Syntax", Internet Engineering Task Force RFC 2141, May 1997. available online at http://www.ietf.org/rfc/rfc2141.txt .

[W3C05] D. Booth, H. Haas, F. McCabe et. al. "Web Services Architecture, W3C Working Group Note", 11 February 2004, available online at: http://www.w3.org/TR/2004/NOTE-ws-arch-20040211/ .