



Universitat de les  
Illes Balears



Treball de Fi de Grau

GRAU D'ENGINYERIA INFORMÀTICA

# Magatzem de dades aplicat a l'àmbit sanitari. Elaboració d'un cub OLAP per al LIS.

Álamo Consulting

JAVIER HERNÁNDEZ LLINÀS

**Tutor**

*Gabriel Fontanet Nadal*

Escola Politècnica Superior

Universitat de les Illes Balears

Palma, 14 de Juliol del 2017

---

# Índex

---

<b>ÍNDEX</b> .....	<b>3</b>
<b>LLISTA D'IL·LUSTRACIONS</b> .....	<b>5</b>
<b>ACRÒNIMS I TERMES</b> .....	<b>7</b>
<b>SOBRE EL DOCUMENT</b> .....	<b>9</b>
RESUM.....	9
SITUACIÓ ACTUAL A L'EMPRESA.....	10
<b>1. INTRODUCCIÓ</b> .....	<b>11</b>
1.1. TEMA PRINCIPAL DEL PROJECTE.....	11
1.2. MOTIVACIÓ.....	11
1.3. DESCRIPCIÓ DEL PRODUCTE.....	12
1.3.1 FUNCIONALITATS DEL PRODUCTE.....	12
1.3.2 LIMITACIONS DEL PRODUCTE.....	13
<b>2. ESTAT DE L'ART</b> .....	<b>15</b>
2.1. ENFOCAMENT.....	15
2.1.1 <i>Emmagatzematge de les Dades</i> .....	16
2.1.2 <i>Descàrrega de les Dades</i> .....	17
2.1.3 <i>Tractament de les Dades</i> .....	18
2.1.4 <i>Creació del Cub</i> .....	19
2.1.5 <i>Visualització del Cub</i> .....	19
<b>3. MODEL DE NEGOCI ÁLAMO CONSULTING</b> .....	<b>21</b>
3.1 ÁLAMO CONSULTING.....	21
3.2 ÁLAMO CONSULTING (MALLORCA).....	22
3.2.1 <i>La Sanitat en l'actualitat</i> .....	22
3.2.1.1 <i>Un nou concepte de salut (Salut-e)</i> .....	24
3.2.2 <i>Contractació d'Álamo Consulting per part de l'Ib-Salut</i> .....	24
3.2.2.1 <i>Àrea de Tecnologia i Sistemes de Salut de l'IB-Salut</i> .....	25
3.2.3 <i>Estructura Tecnològica d'Álamo Consulting</i> .....	26
3.2.3.1 <i>Magatzem de Dades (Data warehouse)</i> .....	28
3.2.4 <i>Activitat de consultoria per a l'Ib-Salut</i> .....	30
3.2.4.1 <i>Extracció de dades dels orígens</i> .....	31
3.2.4.2 <i>Tractament de les dades</i> .....	33
3.2.4.2.1 <i>Integració de dades: ETL</i> .....	34
3.2.4.2.2 <i>Ús de l'integració de dades</i> .....	37
3.2.4.2.3 <i>ETL en el context de Pentaho</i> .....	38
3.2.4.3 <i>Cubs OLAP</i> .....	40
3.2.4.3.1 <i>Elements OLAP</i> .....	42
3.2.4.3.2 <i>Creació Cub OLAP</i> .....	43
3.2.4.4 <i>Visualització de les dades</i> .....	45
3.2.4.4.1 <i>Informes</i> .....	45
3.2.4.4.2 <i>Quadre de comandament</i> .....	45
<b>4 SOLUCIÓ ELEGIDA</b> .....	<b>47</b>
4.1 CARACTERÍSTIQUES DE LA SOLUCIÓ ELEGIDA.....	48

<b>5</b>	<b>LABORATORY INFORMATION SYSTEM (LIS)</b> .....	<b>49</b>
5.1	CARACTERÍSTIQUES GENERALS .....	50
5.1.1	<i>Estructuració de la informació</i> .....	50
5.1.2	<i>Criticitat</i> .....	51
5.1.3	<i>Traçabilitat</i> .....	51
5.1.4	<i>Altres característiques del LIS</i> .....	52
5.2	FLUXE DE TREBALL DELS LABORATORIS .....	52
5.3	EL LABORATORI DELS HOSPITALS DE LES ILLES .....	53
5.3.1	<i>Dades generades per l'activitat dels laboratoris</i> .....	54
<b>6</b>	<b>CREACIÓ CUB PER AL LIS</b> .....	<b>55</b>
6.1	DADES DELS LIS GENERADES ALS HOSPITALS.....	55
6.2	PROCÉS GENERAL .....	57
6.2.1	<i>Descàrrega de les dades</i> .....	57
6.2.2	<i>Tractament de les dades</i> .....	58
6.2.3	<i>Càrrega de les dimensions</i> .....	59
6.2.4	<i>Associació amb la taula de fets</i> .....	59
6.2.5	<i>Creació del cub (Gestió de peticions de Laboratori)</i> .....	60
6.2.6	<i>Visualització de les dades del cub</i> .....	61
<b>7</b>	<b>CONCLUSIONS</b> .....	<b>65</b>
	<b>REFERÈNCIES</b> .....	<b>67</b>
	AGRAÏMENTS .....	69

## Llista d'il·lustracions

---

Il·lustració 1: Mòduls de les tecnologies a estudi per tal de donar resposta al projecte.....	16
Il·lustració 2: Tipus de Bases de Dades segons la seva llicència.....	17
Il·lustració 3: Definició gràfica d'un enllaç Oracle de tipus DataBase Link (DBLink).....	18
Il·lustració 4: Logotip de Pentaho Data Integration. Eina d'ETL.....	18
Il·lustració 5: Logotip de Schema Workbench.....	19
Il·lustració 6: Logotip de Saiku Business Analytics.....	19
Il·lustració 7: Exemple d'alguna de les empreses a les que Álamo ofereix els seus serveis.....	21
Il·lustració 8: Imatge del canvi de paradigma de la gestió mèdica.....	22
Il·lustració 9: Implantació de les TIC dins l'àmbit sanitari.....	23
Il·lustració 10: Imatge corporativa de l'Ib-Salut.....	24
Il·lustració 11: Organigrama simplificat de l'Ib-Salut.....	25
Il·lustració 12: Etapes d'implantació d'un model BI ( <i>Business Intelligence</i> ).....	27
Il·lustració 13: Arquitectura d'un Data Warehouse.....	30
Il·lustració 14: Estructuració general de l'Ib-Salut com a servei de salut.....	31
Il·lustració 15: Composició del Hospital Information System i altres subsistemes.....	32
Il·lustració 16: Esquema de comunicació entre la FIC i els orígens de dades.....	33
Il·lustració 17: Esquema d'integració de dades ETL.....	34
Il·lustració 18: Funcionalitats eines d'integració de dades ETL.....	35
Il·lustració 19: Imatge Pentaho Data Integration (PDI).....	38
Il·lustració 20: Procés general PDI.....	38
Il·lustració 21: Procés exemple d'Integració de dades en PDI.....	39
Il·lustració 22: Exemple genèric d'un Cub OLAP.....	41
Il·lustració 23: Opcions disponibles de Pentaho Schema Workbench.....	43
Il·lustració 24: Exemple d'estructura tipus estrella de les taules d'un cub OLAP.....	44
Il·lustració 25: Imatge genèrica d'un Quadre de Comandaments.....	46
Il·lustració 26: Imatge marc tecnològic elegit.....	47
Il·lustració 27: El HIS i els mòduls que el formen.....	50
Il·lustració 28: Procés habitual de laboratori.....	52
Il·lustració 29: Hospitals a considerar l'activitat del LIS.....	54
Il·lustració 30: Hospitals a considerar l'activitat del LIS.....	54
Il·lustració 31: Motius de cancel·lació de les proves de LIS.....	56
Il·lustració 32: Procés ETL general del Laboratori.....	57
Il·lustració 33: Descàrrega de les dades [Carrega dades Orígens].....	57
Il·lustració 34: Càrrega del model d'exploració [Carrega dades Exploració].....	58
Il·lustració 35: Càrrega de les dimensions del Cub [Carrega dades Anàlisi].....	59
Il·lustració 36: Associació de les dimensions a la taula de fets [Carrega dades Anàlisi].....	59
Il·lustració 37: Estructuració de les dimensions i la taula de fets del cub.....	60
Il·lustració 38: Llistat dels cubs disponibles.....	61
Il·lustració 39: Selecció de les dades a mostrar del cub.....	62
Il·lustració 40: Filtrar la informació del cub per any.....	63
Il·lustració 41: Resultat de executar la consulta sobre el cub.....	63



# Acrònims i Termes

---

En aquest apartat es destaquen els acrònims i termes utilitzats en el projecte i serviran per aclarir i facilitar la comprensió del lector. Son els següents:

**ATB:** Antibiòtic.

**BI:** *Business Intelligence*. (veure definició a l'apartat 3.2.3).

**BIMM:** *Business Intelligence Maturity Model*.

**CDC:** *Change Data Capture*.

**CDI:** *Customer Data Integration*.

**CPM:** *Corporate Performance Management*.

**Data profiling:** procés d'examinació d'informació d'origens i recollida de dades i d'estadístiques.

**DBLink:** tipus d'objecte que permet realitzar una connexió des d'una base de dades a una altra (*Database Link*).

**DOIP:** Departament d'Orientació i Inserció Professional.

**EAI:** *Enterprise Application Integration*.

**EII:** *Enterprise Information Integration*.

**EIM:** *Enterprise Information Management*.

**EPS:** Escola Politècnica Superior.

**eSIAP:** aplicació centralitzada d'Atenció Primària.

**ETL:** *Extract, Transform and Load*.

**FIC:** Factoria d'Informació Corporativa.

**FUEIB:** Fundació Universitat-Empresa de les Illes Balears.

**GUI:** *Graphical User Interface*.

**HIS:** *Hospital Information System*. (veure definició a l'apartat 3.2.4.1).

**LIS:** *Laboratory Information System* (veure definició al capítol 5).

**Magatzem de Dades:** *Data Warehouse* (veure definició a l'apartat 3.2.3.1).

**MDM:** *Master Data Management*.

**MDX:** *Multidimensional Expressions*.

**ODBC:** *Open Database Connectivity*.

**OLAP:** *On-Line Analytical Processing*. (veure definició a l'apartat 3.2.4.3).

**PDI:** Pentaho Data Integration. Eina d'integració de dades de Pentaho (també conegut com Spoon o Kettle).

**PACS:** *Picture Archiving and Communication System*.

**Pentaho suite:** *software* lliure d'intel·ligència de negoci.

**PIM:** *Product Information Management.*

**PSW:** software propietat de Pentaho per la creació de cubs (*Pentaho Schema Workbench*).

**SaaS:** *Software as a Service.*

**SCD:** *Slowly Changing Dimensions.*

**Saiku Business Analytics:** software propietat de Meteorite.bi.



# Sobre el document

---

## Resum

Aquest document recull la memòria del treball de fi de Grau d'Enginyeria Informàtica, realitzat durant el curs 2016-2017.

La idea va sorgir quan vaig unir-me a per *Álamo Consulting* començar les pràctiques i, després d'un temps treballant a l'empresa, varem parlar per fer una proposta de treball de fi de Grau de manera col·laborativa.

Les dades tractades en el present document i en tota documentació consultada o generada, són de caràcter personal fent referència a dades mèdiques de pacients. Per tant, com dictamina la Llei Orgànica 15/1999, del 13 de desembre, de Protecció de Dades de Caràcter Personal, no es fa pública cap dada o informació referent a pacients o qualsevol dada que pugui comprometre el contracte de confidencialitat signat amb la pròpia empresa (*Álamo Consulting*).

En aquest document s'introdueix el concepte de **Business Intelligence** (veure definició a l'apartat 3.2.3) o Intel·ligència de negoci i com aquest tipus de paradigma és de cada vegada més aplicat per obtenir informació de les dades de l'activitat empresarial d'una organització. A més, es descriu la necessitat d'un **Data warehouse** o Magatzem de dades (veure definició a l'apartat 3.2.3.1) veient el marc tecnològic necessari per tal de treballar amb un conjunt de dades consistents i fiables. La descàrrega i el tractament de les dades a través de **Processos ETL**, abans de ser emmagatzemades al *Data warehouse* corporatiu, amb eines que ofereixen diverses funcionalitats per realitzar el refinament de les dades. Es descriu el concepte **OLAP** (*On-Line Analytical Processing*) (veure definició a l'apartat 3.2.4.3) i les seves característiques principals. Finalment, el concepte de **Quadre de comandament** (veure definició a l'apartat 3.2.4.4.2) i com es centralitza aquest servei en un aplicatiu web que permet la visualització de l'informació tenint en compte els permisos amb els que conta cada un d'ells usuaris que hi accedeixen.

Els apartats que trobarem al documents són els següents: Al **Capítol 1** s'introdueix el tema principal del projecte, la motivació que apareix per realitzar aquest treball i la descripció del producte resultant. Al **Capítol 2** es presenten els diferents aspectes a tractar, les tecnologies que actualment s'usen i els aspectes que caracteritzen cada una d'elles. Al **Capítol 3** es fa una introducció de l'empresa (*Álamo Consulting*), el model de negoci i l'entorn empresarial en el que opera. A més es fa una breu introducció a l'entorn de la sanitat i com als darrers anys ha canviat el paradigma del tractament de les dades generades als hospitals i centres de salut. Al **Capítol 4** es descriu la solució elegida i els motius que han propiciat aquesta elecció. Al **Capítol 5** s'introdueix el concepte de *Laboratory Information System* (veure definició al capítol 5) i les seves característiques generals. Seguidament, al **Capítol 6** es descriu el procés per la creació del cub *OLAP* i quina informació es necessita. Finalment, al **Capítol 7** s'extreuen les conclusions que s'obtenen després d'haver realitzat el treball.

## **Situació actual a l'empresa**

*Álamo Consulting* és una empresa espanyola de consultoria, que opera a nivell nacional, especialitzada en Tecnologies de la Informació, Estratègia i Gestió que ofereix un portfoli integral de serveis i solucions.

Actualment faig feina com a consultor en pràctiques, a *Álamo Consulting*, realitzant diferents tasques relacionades amb el *Business Intelligence*.

# 1. Introducció

---

En aquest apartat es descriu **el tema principal del projecte** que és el de crear un cub *OLAP* específic que doni resposta a un conjunt de requisits concrets per a la subdivisió de Laboratori. Abans, es contextualitza l'entorn operacional i l'ecosistema en el que el cub coexistirà.

## 1.1. Tema principal del projecte

L'objectiu principal del projecte és la creació d'un cub *OLAP* (veure definició a l'apartat 3.2.4.3) que permet la consulta de la informació generada durant l'activitat de la secció de Laboratori dels diferents hospitals de les Illes Balears.

Abans d'adentrar-se en la creació del cub *OLAP*, es dona una visió general sobre l'estat actual de la sanitat i com l'introducció de la tecnologia a tots els nivells ha impulsat un canvi de tendència sobre la gestió de la informació i la importància de la interpretació de les dades generades.

Es defineix el model de negoci establert per *Álamo Consulting*, l'estructura necessària i les metodologies emprades, quines eines i tecnologies es fan servir a l'organització per tal d'oferir un servei de qualitat a l'IB-Salut.

La jerarquia dels sistemes d'informació que coexisteixen als hospitals i més concretament entrant en detall al sistema d'informació del laboratori (*Laboratory Information System* o *LIS*).

Després d'haver fet aquesta contextualització que permeti conèixer tot l'entorn operatiu, s'explica el procés que s'ha de dur a terme per poder obtenir les dades, tractar-les i gestionar-les de tal manera que permeti construir el cub *OLAP* que doni resposta a les necessitats dels directius i personal autoritzat per la visualització de les dades.

## 1.2. Motivació

La idea de realitzar aquest treball de fi de grau va sorgir arrel de la conversa amb el responsable d'*Álamo Consulting* a Mallorca, Guillem Pérez, que hem va demanar si tenia alguna idea per realitzar el Treball de Fi de Grau (TFG) per als meus estudis d'Informàtica.

La meva idea inicial era la de realitzar el Treball de Fi de Grau de manera conjunta amb l'empresa i poder desenvolupar algun producte que em pogués servir tant a mi, per poder presentar com a Treball de Fi de Grau a l'Escola Politècnica Superior (EPS) de la Universitat de les Illes Balears (UIB), com a la pròpia empresa.

Després de parlar amb el propi Guillem Pérez per tal de consensuar la temàtica elegida sobre la qual realitzar el TFG, i després de diverses propostes, varem acordar que el Treball de Fi de Grau seria la creació d'un cub específic per al Sistema d'Informació per al Laboratori (*LIS*).

### 1.3. Descripció del producte

El projecte consisteix en oferir una visió del terme Business Intelligence (veure definició a l'apartat 3.2.3), l'entorn tecnològic que implica aplicar aquesta intel·ligència de negoci dins una organització i com *Álamo Consulting* ofereix un servei de consultoria en l'àmbit sanitari que permet transformar un conjunt de dades en informació útil per donar resposta així a les necessitats de l'IB-Salut.

Es descriu tota la infraestructura desplegada i el procés d'obtenció de les dades dels diferents orígens, el tractament d'aquestes dades per processos d'integració i, finalment, l'emmagatzematge d'aquestes dades al repositori de l'organització.

Després de fer aquesta explicació, desenvolupar un cub *OLAP* per al *LIS* el qual pot ser consultat a través d'un navegador web que permet obtenir dades dels diferents Laboratoris dels hospitals de les Illes Balears.

Aquest cub reflecteix una sèrie d'indicadors que ofereixen una visió general de l'activitat dels diferents Laboratoris.

#### 1.3.1 Funcionalitats del producte

El cub *OLAP* permet consultar les dades dels Laboratoris de cada un dels hospitals de les Illes Balears. Aquestes dades hauran estat tractades previament i permetran extreure una sèrie d'indicadors que aportaran informació agregada per poder gestionar els diferents sistemes de laboratori i poder obtenir una visió general de l'activitat dels Laboratoris segons els orígens dels que es disposa (hospitals des d'on es descarreguen les dades).

Es té en compte la següent informació dels laboratoris:

- Nombre de peticions totals
- Nombre de peticions realitzades
- Nombre de determinacions realitzades
- Nombre de proves derivades
- Nombre de proves cancel·lades
- Nombre de pacients diferents
- Temps de demora de laboratori ordinari
- Temps de demora de laboratori urgent
- Nombre d'antibiogrames

### **1.3.2 Limitacions del producte**

El cub està centrat en oferir dades d'una sèrie d'indicadors concrets. Per aquest motiu, únicament permetrà oferir les dades dels indicadors que s'han sol·licitat de Laboratori. Tan sols es podran consultar aquest indicadors ja que un centre hospitalari compta amb altres àrees diferents que poc tenen a veure amb l'àmbit de Laboratori.



# 2. Estat de l'art

---

En aquest apartat s'exposen quins son els principals factors a tenir en compte per solucionar el problema plantejat. A més, s'expliquen les principals eines disponibles o que s'estan utilitzant en l'actualitat tant dins l'empresa com a l'entorn tecnològic. Es tracta de cercar i arribar a una solució que compleixi amb les necessitats dels interessates per aconseguir l'objectiu plantejat.

## 2.1. Enfocament

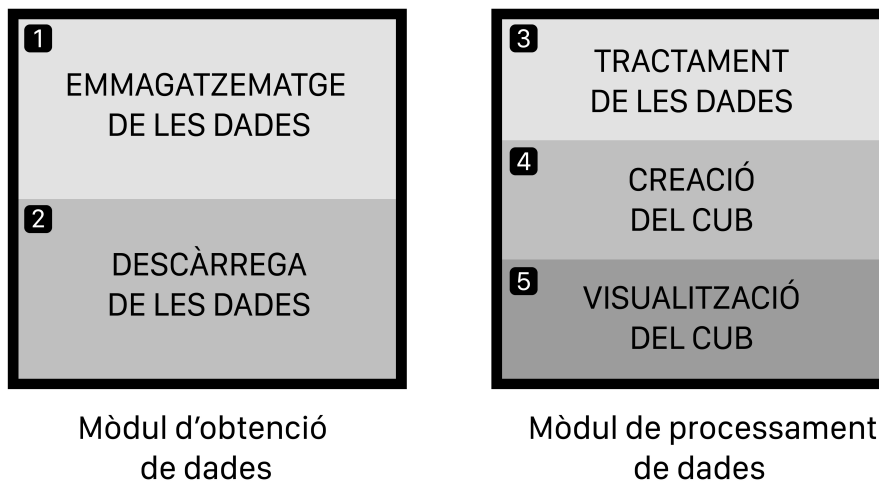
Als darrers anys, ha pres molta força dins el mon de l'informàtica el concepte de *software* de codi obert o *Open Source* que s'ha convertit en tota una filosofia de desenvolupament de software que compleix amb els següents principis:

- Obert: La comunitat te lliure accés, ús i participació del codi font, així com la possibilitat de fer ús de foros per proporcionar feedback.
- Transparència: La comunitat te accés al *full de ruta*, documentació, defectes i agenda dels objectius marcats.
- Early & Often: La informació es publica de manera freqüent a través de repositoris públics (incloent el codi font).

Per aquest motiu és un factor a tenir molt present a l'hora d'elegir entre les diferents opcions i tecnologies que actualment estan dins el mercat.

Podem trobar 5 aspectes claus que afectaran de manera directa al projecte actual:

1. Emmagatzematge de les dades
2. Descàrrega de les dades
3. Tractament de les dades
4. Creació del Cub
5. Visualització del Cub



Il·lustració 1: Mòduls de les tecnologies a estudi per tal de donar resposta al projecte.

### 2.1.1 Emmagatzematge de les Dades

Actualment existeixen nombroses opcions a l'hora de crear i gestionar una base de dades. Podem trobar nombroses opcions de Codi Obert (*Open Source*) o opcions comercials en les que s'ha d'adquirir una llicència per poder fer-ne ús.

· Opcions de Codi Obert: *MySQL, Cassandra, PostgreSQL, Firebird, MongoDB, MariaDB o SQLite* entre d'altres.

· Opcions Comercials: *Oracle, IBM DB2, Microsoft SQL Server, Redis, Teradata o SAP Sybase* entre d'altres.



## Bases de Dades Open Source

## Bases de Dades comercials



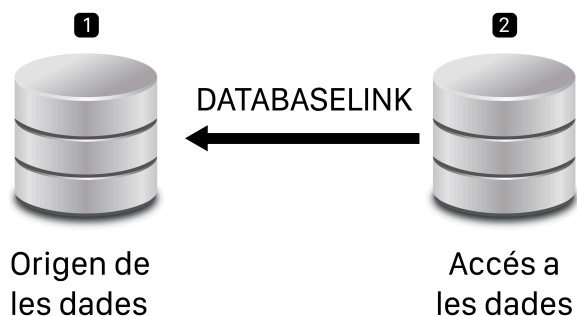
II-lustració 2: Tipus de Bases de Dades segons la seva llicència.

*Álamo Consulting* treballa amb tecnologia Oracle la qual ofereix un sistema de gestió de base de dades de tipus objecte-relacional que permet poder establir el model de *Data warehouse* que permet, a partir de tot l'entorn tecnològic desplegat, poder oferir unes dades consistents que proporcionen la informació que es sol·licita des de l'IB-Salut.

### 2.1.2 Descàrrega de les Dades

El fet que *Álamo Consulting* faci servir *Oracle* per la gestió de la base de dades que li permet operar amb la informació que rep, fa que les connexions entre les estacions es realitzin a partir de *Database Links (DBLinks)*.

Un *Database Link* o *DBLink*, com el seu propi nom indica, és un tipus d'objecte que permet realitzar una connexió desde una base de dades a una altra. El seu principal objectiu es ocultar el detall dels paràmetres de connexió necessaris, facilitant l'accés als recursos disponibles entre bases de dades, independentment de que aquestes bases de dades siguin o no al mateix servidor.



Il·lustració 3: Definició gràfica d'un enllaç Oracle de tipus DataBase Link (DBLink).

### 2.1.3 Tractament de les Dades

Un magatzem de dades o repositori de dades es caracteritza per emmagatzemar les dades de forma íntegra, consistent, no volàtil i actualitzades.

Per tal d'assegurar les característiques intrínseques d'un *Data warehouse*, abans d'emmagatzemar les dades, s'han de tractar les dades abans. Per dur a terme aquest tractament es fa ús d'eines *ETL* (*Extract, Transform and Load*) que venen a explicar el procediment que defineix l'extracció de les dades des de el seu origen o des d'altres *Data warehouse*, la transformació de les dades obtingudes i la càrrega final al repositori de l'organització.



Pentaho Data Integration

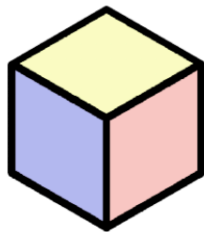
Il·lustració 4: Logotip de Pentaho Data Integration. Eina d'ETL.

Per poder dur a terme aquest tractament farem ús de *Pentaho Data Integration (PDI)* o també conegut com *Kettle* o *Spoon* (veure [1]).

És un *software* de tractament de dades, propietat de *Pentaho*, que de forma gràfica i, a través de *drag and drop*, permet crear un conjunt de processos relacionats entre si per poder dur a terme la integració de les dades.

#### 2.1.4 Creació del Cub

Per a la creació del cub farem ús de l'eina *Schema Workbench*, també de *Pentaho*, que una vegada disposem de la informació a la base de dades ens permetrà definir l'estructura del cub i les dependències entre les diferents taules.



# Schema Workbench

Il·lustració 5: Logotip de Schema Workbench.

#### 2.1.5 Visualització del Cub

Per poder seleccionar i visualitzar les dades que es sol·licitin del cub es farà ús de *Saiku Business Analytics* (veure [2]).



# Saiku Business Analytics

Il·lustració 6: Logotip de Saiku Business Analytics.

Permet explorar, des d'un navegador web, un origen de dades complet fent ús d'una interfície *drag and drop* que facilita la consulta de les dades del cub.



# 3. Model de negoci Álamo Consulting

En primer lloc, i abans d'aprofundir en el procés de creació del cub, es tracta de donar una visió general de l'entorn operacional sobre el qual treballa *Álamo Consulting* per poder oferir servei i quin model de negoci te establert per tal d'assolir les necessitats del client.

## 3.1 Álamo Consulting

*Álamo Consulting* (veure [3]) és una empresa de consultoria, amb seu central a Madrid i fundada al 2007, especialitzada en Tecnologies de la Informació, Estrategia y Gestió que ofereix un portfoli integral de serveis i solucions.



Il·lustració 7: Exemple d'alguna de les empreses a les que Álamo ofereix els seus serveis.

*Álamo Consulting* ha desenvolupat solucions i ofert els seus serveis de consultoria a empreses com *Prisa*, *Banco Gallego*, *Sanitas*, *RSI*, *Seguros RGA*, *Liberty Seguros*, *Euro 6000*, *Parque Científico de Madrid*, *Telefonica*, *Govern de les Illes Balears*, *Mapfre*, *Huawei*, *Bankia* i el *Ministeri d'Igualtat* entre d'altres empreses.

## 3.2 Àlamo Consulting (Mallorca)

A nivell local, a l'arxipèlag Balear, treballa amb el Govern de les Illes Balears per donar servei a la Conselleria de Salut (IB-Salut).

Abans d'entrar en detalls de l'activitat que realitza *Àlamo Consulting* per donar resposta a les necessitats de l'IB-Salut, és necessària una contextualització per conèixer l'estat actual dels serveis sanitaris.

### 3.2.1 La Sanitat en l'actualitat

A Espanya, tret de algunes excepcions, les *TIC* es varen incorporar als hospitals a finals dels anys 70 i a principis dels 80. Aquesta incorporació es va produir mitjançant dues vies d'accés.

Canvi paradigme gestió mèdica (finals dels anys 70)



Il·lustració 8: Imatge del canvi de paradigma de la gestió mèdica.

Les primeres incorporacions de les *TIC* varen venir de la mà de proves com el *TAC* o els incipients sistemes d'informació de laboratori. Aquests sistemes, que formaven part d'equips mèdics, quedaven fora del control de la alta direcció i marcaren un camí per posteriors incorporacions, moltes vegades també sense el control de l'alta direcció.

L'altre via d'entrada va ser més convencional, vinculada a la gestió econòmico-financera i logística dels hospitals.

Amb posteritat, les *TIC* es varen incorporar als sistemes clínics administratius, com la gestió de llits i cita de consultes.

Als inicis dels anys 90 es va iniciar la informatització de l'atenció primària amb la creació de la base de dades de població (tarjeta individual sanitaria) i la implantació de la cita prèvia.



### IMPLANTACIÓ DE LES TIC DINS L'ÀMBIT SANITARI

#### Il·lustració 9: Implantació de les TIC dins l'àmbit sanitari.

Aquest fet respon a la incorporació de TIC als sistemes de salut, denominada *Salut-e*, que obeeix a una decisió per millorar l'efectivitat i la eficiència del sector. Els seus objectius es basen en dos fets:

- La realitat social marcada per les necessitats dels ciutadans, amb unes expectatives de cada vegada majors perquè disposen de l'accés a gran quantitat d'informació.
- Garantir la sostenibilitat dels sistemes de salut controlant costos creixents, optimitzar processos i la reassignació de recursos són reptes permanents de qualsevol sistema sanitari.

La plena incorporació de les *TIC* al nucli d'operacions del sistema de salut es produeix quan aquestes es constitueixen com un instrument clínic formant part de l'activitat assistencial. Aquesta nova etapa ha estat marcada per la irrupció de la història clínica electrònica i de totes les aplicacions clíniques departamentals (laboratoris, imatge i farmàcia, entre d'altres). Aquesta situació ha convertit a les *TIC* en imprescindibles en la funció assistencial i, per tant, en estratègiques per al compliment dels objectius de la organització. Tot això fa que la bona gestió de la funció *TIC* sigui un dels requisits que tot directiu de serveis i sistemes de salut ha de complir.

La gestió de la funció *TIC* en el nostre sistema de salut encara no ha assolit la maduresa suficient per considerar-se consolidada. Pensem que fa poc més de 20 anys que les *TIC* formen part de l'activitat clínica, que moltes consultes dels hospitals espanyols encara utilitzen paper

per gestionar la seva història clínica i que molts sistemes de la informació encara no s'han integrat. Per altra banda, a les dificultats diàries de la gestió s'afegeixen les pròpies dels temps de crisi econòmica, amb greus restriccions presupostaries i major exigència d'eficiència i efectivitat. A més, la generació que administra els recursos dels serveis i sistemes de salut encara no és una "generació digital". És una generació que ha hagut d'aprendre a utilitzar les TIC en la seva vida professional i privada, amb alts costos personals i d'aprenentatge.

### 3.2.1.1 Un nou concepte de salut (Salut-e)

La *Salut-e* té multitud d'implicacions relacionades amb la prevenció, diagnòstic, tractament i monitorització de pacients així com la planificació i control de gestió dels serveis i sistemes de salut. En aquest àmbit te cabuda aplicacions tan diverses com la història clínica electrònica, la gestió clínico-administrativa, la imatge mèdica digital, els sistemes d'informació departamentals, els distints serveis de telemedicina, els sistemes de vigilància de salut pública o laboral, o els programes d'educació a distància de salut, entre d'altres.

A aquest serveis s'ha d'afegir varies necessitats bàsiques, com la infraestructura tecnològica sobre la que han de funcionar, la interoperabilitat que permeti l'intercanvi de dades entre sistemes i les mesures de seguretat i protecció de la informació.

### 3.2.2 Contractació d'Álamo Consulting per part de l'Ib-Salut

L'IB-Salut precisa d'un servei de *Business Intelligence (BI)* per tractar les dades que es generen als hospitals. Es publica un contracte i és presentat a concurs públic. Després de que l'IB-Salut rebí una sèrie de plegs amb les diferents propostes i les característiques pertinents de cada un d'ells se'n tria un de les opcions que té disponibles i s'otorga el contracte per poder dur a terme el servei de consultoria.



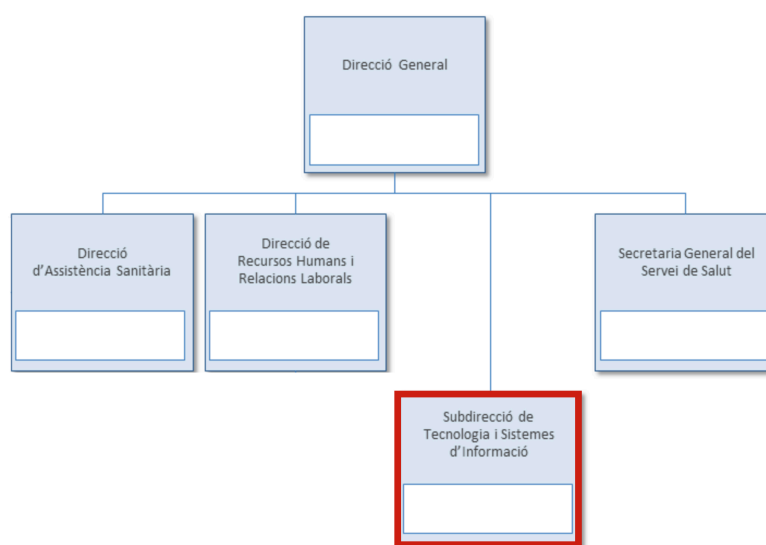
Il·lustració 10: Imatge corporativa de l'Ib-Salut.



Aquest servei permet conèixer l'estat dels diferents centres de salut de les Illes Balears a través del tractament de les dades sanitàries que es van generant degut a l'activitat que es porta a terme a hospitals i centres d'atenció primària.

### 3.2.2.1 Àrea de Tecnologia i Sistemes de Salut de l'IB-Salut

L'IB-Salut està estructurat seguint una organització geràrquica amb un departament específic de Tecnologia i Sistemes d'Informació:



Il·lustració 11: Organigrama simplificat de l'IB-Salut.

És des d'aquesta àrea, la de Tecnologia i Sistemes d'Informació, on es valida tota la informació generada durant l'activitat sanitària.

### 3.2.3 Estructura Tecnològica d'Álamo Consulting

Degut als canvis soferts durant els darrers 20 anys, ha estat una necessitat quasi obligatòria, replantejar la situació i l'evolució dels sistemes sanitaris. Per aquest motiu, esdevé una gran avantatge poder disposar de les dades generades durant l'activitat sanitària per poder fer-ne ús i oferir un millor servei als ciutadans.

En la majoria de les situacions, i per tal d'aprofitar de manera efectiva la informació es convergeix en disposar d'un repositori privat de dades, propietat de l'organització, que permet disposar de les dades per poder realitzar les mesures necessàries per poder oferir un millor servei.

Aquest nou paradigma o pràctica és conegut com a Intel·ligència de negoci o *Business Intelligence (BI)*.

S'entén per **Business Intelligence** al conjunt de metodologies, aplicacions, pràctiques i capacitats enfocades a la creació i administració d'informació que permet prendre millors decisions als usuaris d'una organització.

El context de la societat de la informació ha propiciat la necessitat de tenir millors, més ràpids i més eficients mètodes per extreure i transformar les dades d'una organització en informació i distribuir-la al llarg de la cadena de valor.

La intel·ligència de negoci (o *Business Intelligence*) respon a la necessitat descrita, i podem entendre, en una primera aproximació, que és una evolució dels sistemes de suport a les decisions.

Algunes de les tecnologies que formen part de *Business Intelligence* son:

- *Data warehouse*
- *Reporting*
- Anàlisi *OLAP (On-Line Analytical Processing)*
- Anàlisi visual
- Anàlisi predictiu
- Quadre de comandament
- Minería de dades
- Gestió del rendiment
- Previsions

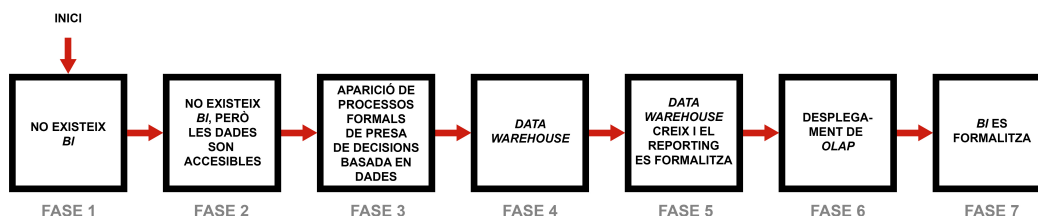
- Regles de negoci
- Integració de dades (que inclou *ETL, Extract, Transform and Load*)

En definitiva, els sistemes de Business Intelligence cerquen donar resposta a les preguntes:

- Què va passar?
- Què passa ara?
- Per quin motiu va passar?
- Què passarà?

Es pot identificar quin és el grau de maduresa del sistema *BI* d'una organització en el que es refereix a la intel·ligència de negoci.

El *BIMM (Business Intelligence Maturity Model)* és un model de maduresa que permet classificar una organització des del punt de vista del grau de maduresa d'implantació de sistemes *Business Intelligence*.



Il·lustració 12: Etapes d'implantació d'un model *BI (Business Intelligence)*.

### 3.2.3.1 Magatzem de Dades (*Data warehouse*)

El concepte i l'enfoc de la intel·ligència de negoci ha evolucionat en gran manera els darrers anys. Un dels conceptes que més ha evolucionat ha estat el repositori de dades o magatzem de dades, també conegut com *Data warehouse*.

En els darrers anys han aparegut molts d'enfocaments de *Data warehouse* tant a nivell tecnològic com estratègic:

- Aparició de bases de dades especialitzades en el desplegament de *Data warehouse* en forma de *software* o *appliances*.
- Desenvolupament de múltiples metodologies que cerquen cobrir el creixement exponencial de dades en una organització.
- Tendències *SaaS* (*Software as a Service*) o *in-memory* que cerquen reduir el temps de creació d'estructures de dades.

Un sistema d'intel·ligència de negoci està format per diferents elements: *ETL*, *OLAP*, *reporting* i el *Data warehouse* que és la peça més rellevant.

Un ***Data warehouse*** és un repositori de dades que proporciona una visió global, comú i integrada de les dades d'una organització – independentment de com s'utilitzin posteriorment per els consumidors o usuaris –, amb les propietats següents: estable, coherent, fiable i amb informació històrica. A l'abastar un àmbit global de l'organització i amb un ampli abast històric, el volum de dades pot ser molt gran.

Les bases de dades relacionals son el suport tècnic més comunament usat per emmagatzemar les estructures de aquestes dades i els seus grans volums.

En resum, presenta les següents característiques:

- Orientat a un tema: organitza una col·lecció d'informació al voltant d'un tema central.
- Integrat: inclou dades de múltiples orígens i presenta consistència de dades.
- Variable en el temps: es realitzen fotos de les dades basades en dates o fets.
- No volàtil: només de lectura per usuaris finals.

Hem de tenir en compte que existeixen altres elements en el context d'un *Data warehouse*:

· *Data warehousing*: és el procés d'extreure i filtrar dades de les operacions comuns de l'organització, procedents dels diferents sistemes d'informació operacionals o sistemes externs, per transformar-los, integrar-los i emmagatzemar-los en un magatzem de dades amb la finalitat d'accedir a ells per donar suport en el procés de la presa de decisions d'una organització.

· *Data mart*: és un subconjunt de les dades del *Data warehouse* que té com objectiu respondre a un determinat anàlisi, funció o necessitat, amb una població d'usuaris específica. Igual que un *Data warehouse*, les dades estan estructurades en models d'estrella o floc de neu, i un *Data mart* pot ser dependent o independent d'un *Data warehouse*.

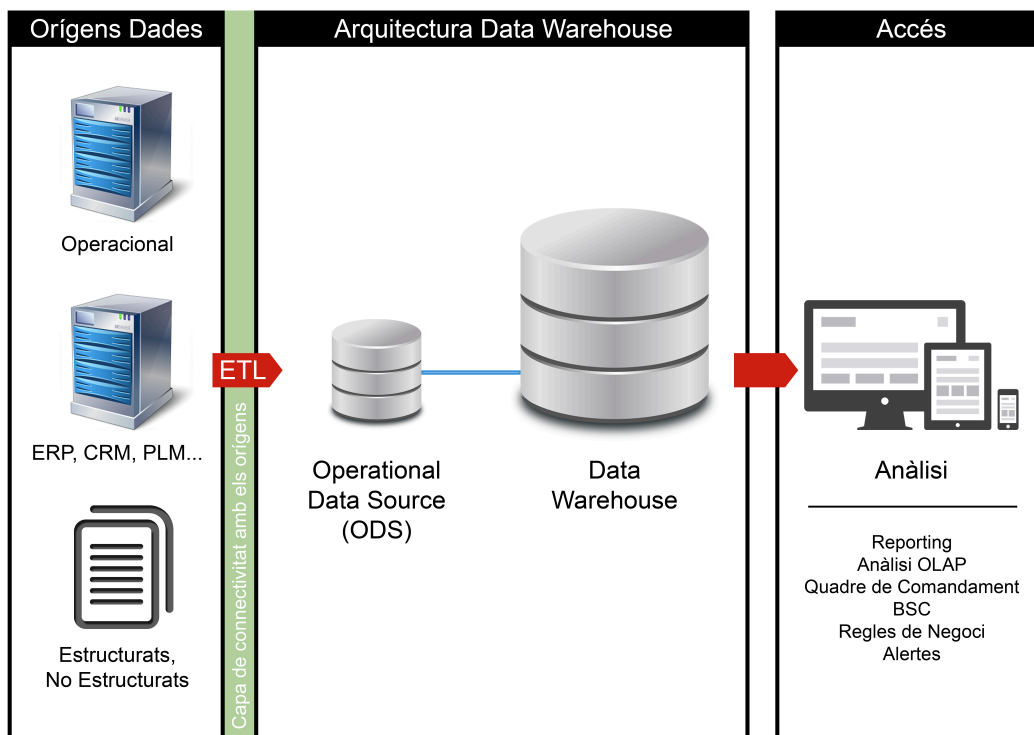
· *Operational Data Store*: és un tipus de magatzem de dades que proporciona només els darrers valors de les dades i no el seu historial. A més, generalment admet un petit desfasament o retràs sobre les dades operacionals.

· *Staging Area*: és el sistema que roman entre les fonts de dades i el *Data warehouse* amb l'objectiu de:

- Facilitar l'extracció de dades desde fonts d'origen amb una heterogeneïtat i complexitat gran.
- Millorar la qualitat de dades.
- Ser usat com *cache* de dades operacionals amb el que posteriorment es realitza el procés de *Data warehousing*.
- Ús de la mateixa *Staging area* per accedir en detall a informació no continguda en el *Data warehouse*.

· *Processos ETL*: tecnologia d'integració de dades basada en la consolidació de dades que s'usa tradicionalment per alimentar el *Data warehouse*, *Data mart*, *Staging area* i *ODS*. Usualment es combina amb altres tècniques de consolidació de dades.

· *Metadades*: dades estructurades i codificades que descriuen les característiques d'instàncies. Aporten informacions per ajudar a identificar, descobrir, valorar i administrar les instàncies descrites.

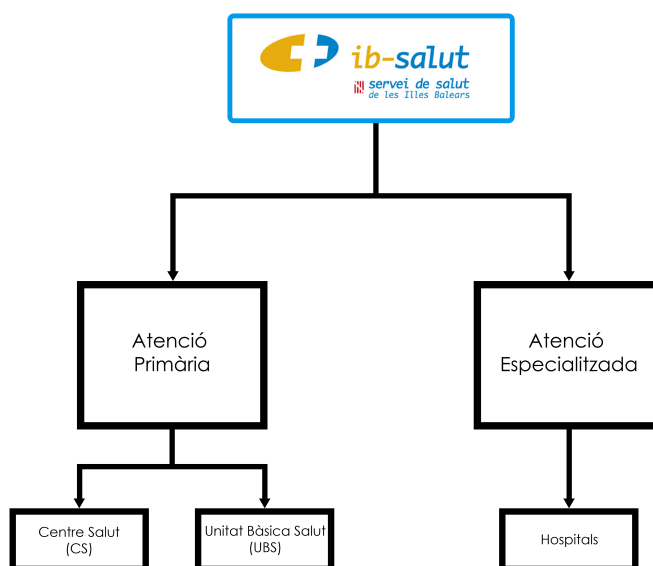


II-Il·lustració 13: Arquitectura d'un Data Warehouse

### 3.2.4 Activitat de consultoria per a l'IB-Salut

Álamo Consulting gestiona la FIC (Factoria d'Informació Corporativa) que és un nou concepte de Sistema d'Informació (*Data warehouse*) que unifica, homogenitza i estandaritza tota la informació assistencial del Servei de Salut de les Illes Balears, tant d'Atenció Primària (Centres de Salut i Unitats Bàsiques de Salut) com d'Atenció Especialitzada (Hospitals).

A grans trets, l'IB-Salut es pot subdividir en dos grans subapartats que són l'Atenció Primària i l'Atenció Especialitzada.



Il·lustració 14: Estructuració general de l'Ib-Salut com a servei de salut.

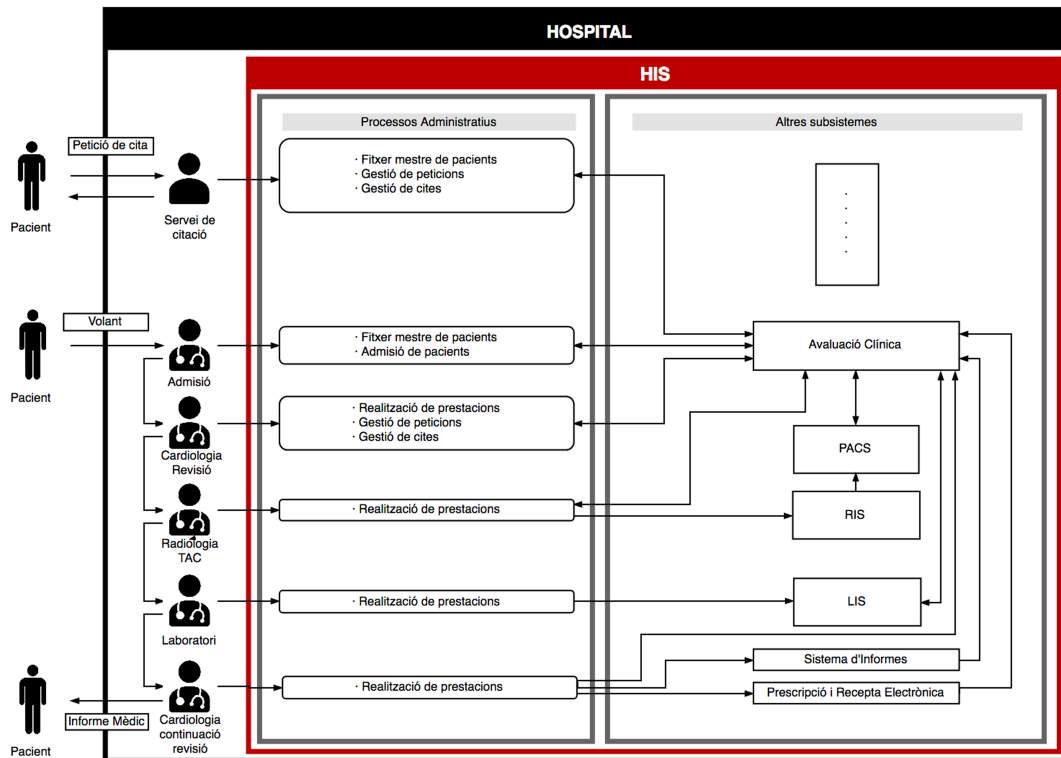
La informació emmagatzemada és accessible de forma segura i controlada, mitjançant les aplicacions de la Factoria d'Informació Corporativa (FIC) des de qualsevol dels centres que conformen la xarxa assistencial de les Illes Balears.

### 3.2.4.1 Extracció de dades dels orígens

La FIC es connecta a 7 sistemes *HIS* (*Hospital Information System*) diferents que es troben a ubicacions distintes i a l'aplicació centralitzada del Servei de Salut d'atenció primària (l'eSIAP), mantenint la base de dades actualitzada al dia anterior. També es connecta a *SAP* (RRHH i Farmàcia) i rep informació via fitxers externs.

El **HIS** (acrònim de *Hospital Information System* o sistema d'informació hospitalari) és un sistema integrat d'informació dissenyat per gestionar tots els aspectes clínics, administratius i financers d'un hospital.

A més, permet obtenir estadístiques generals de pacients, dades epidemiològiques, de salut laboral i salut pública, entre d'altres.



Il·lustració 15: Composició del Hospital Information System i altres subsistemes.

El HIS pot estar format per un o varis components de *software* i una gran varietat de subsistemes d'especialitats mèdiques com el RIS (Sistema d'Informació de Radiologia), LIS (Sistema d'Informació de Laboratori), sistemes per anatomia patològica i altres.

El fet que el sistema estigui estructurat en mòduls permet escollir aquells que son necessaris en funció dels requeriments del centre sanitari i assegurar que futures incorporacions mantindran la integració global del sistema.

Cada nit, la FIC recull, normalitza i tradueix la informació de més de 50.000 actes assistencials, aproximadament uns 38.000 procedents dels serveis i centres d'atenció primària (AP) i uns 12.000 d'atenció especialitzada (AE).





Il·lustració 16: Esquema de comunicació entre la FIC i els orígens de dades.

Quan es descarreguen les dades, i abans de ser emmagatzemades dins el propi *Data warehouse*, han de passar per un o més processos d'integració de dades per tal de tenir finalment, al *Data warehouse*, dades consistents i fiables.

### 3.2.4.2 Tractament de les dades

En un context empresarial, la integració pot donar-se en quatre grans àrees:

- Integració de dades: proporciona una visió única de totes les dades de negoci, es trobin on es trobin. Aquest és l'àmbit del present document i, en particular, en el context de la intel·ligència de negoci.

- Integració d'aplicacions: proporciona una visió unificada de totes les aplicacions tant internes com externes a l'empresa. Aquesta integració s'aconsegueix mitjançant la coordinació de fluxes d'events (transaccions, missatges o dades) entre aplicacions.

- Integració de processos de negoci: proporciona una visió unificada de tots els processos de negoci. La seva principal avantatge és que les consideracions de disseny de l'anàlisi i implementació dels processos de negoci estan aïllats del desenvolupament de les aplicacions.

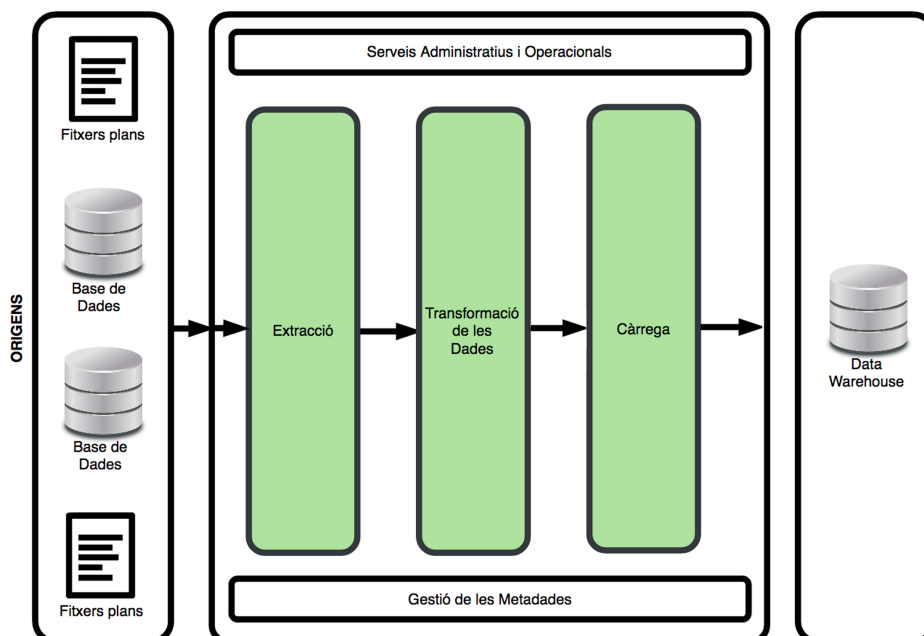
· Integració de la interacció dels usuaris: proporciona una interfície segura i personalitzada a l'usuari del negoci (dades, aplicacions i processos de negoci).

Aquest document es centrarà en l'integració de dades i en els processos *ETL* (Extracció, Transformació i Càrrega) en particular, que és una de les tecnologies d'integració de dades que s'usa en els projectes de implantació de *Business Intelligence*.

### 3.2.4.2.1 Integració de dades: ETL

En el context de la intel·ligència de negoci, les eines *ETL* han estat l'opció més usual per alimentar el *Data warehouse*. La funcionalitat bàsica de aquestes eines està composta per:

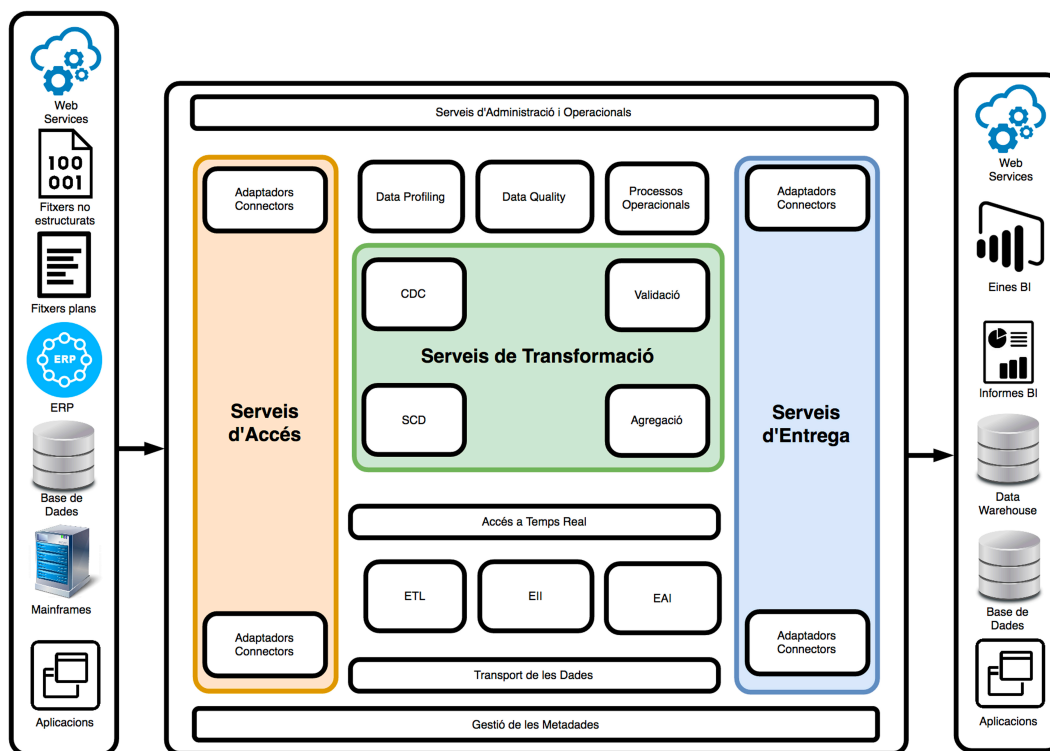
- Gestió i administració de serveis.
- Extracció de dades.
- Transformació de dades.
- Càrrega de dades.
- Gestió de dades.



Il·lustració 17: Esquema d'integració de dades ETL.

En els darrers anys, aquestes eines han evolucionat incloent més funcionalitats pròpies d'una eina d'integració de dades. Podem destacar:

- Serveis d'accés/entrega de dades (vía adaptadors/connectors).
- Gestió de serveis.
- *Data profiling*.
- *Data quality*.
- Processos operacionals.
- Serveis de transformació: *CDC*, *SCD*, validació, agregació.
- Serveis d'accés a temps real.
- *Extract, Transform and Load (ETL)*.
- *Enterprise Information Integration (EII)*.
- *Enterprise Application Integration (EAI)*.
- Capa de transport de dades.
- Gestió de metadades.



II-lustració 18: Funcionalitats eines d'integració de dades ETL.

Aquesta evolució és conseqüència de diversos motius, entre els que podem destacar els diferents tipus de dades existents:

- Estructurats: continguts en bases de dades.
- Semiestructurats: en formats llegibles per màquines, si bé no estan completament estructurats: *HTML*, tabulat, Excel, *CSV*, entre d'altres que poden obtenir-se mitjançant tècniques estàndard d'extracció de dades.
- No estructurats: en formats llegibles per humans, però no per a màquina: *Word*, *HTML*, no tabulat, *PDF*, entre d'altres que poden obtenir-se mitjançant tècniques avançades com text mining o altres.

S'entén per **Integració de dades** al conjunt d'aplicacions, productes, tècniques i tecnologies que permeten una visió única consistent de les dades de negoci.

Respecte a la definició:

- Les aplicacions són solucions a mida que permeten l'integració de dades en base a l'ús de productes d'integració.
- Els productes comercials desenvolupats per tercers capaciten l'integració mitjançant l'ús de tecnologies d'integració.
- Les tecnologies d'integració són solucions per realitzar l'integració de dades.

Existeixen diferents tecnologies d'integració de dades basades en les tècniques presentades:

- ETL: permet extreure dades de l'entorn d'origen, transformar-los segons unes necessitats de negoci per integració de dades i carregar aquestes dades en els entorns destí. Els entorns d'origen i destí són usualment bases de dades i/o fitxers, però en ocasions també poden ser coes de missatges d'un determinat middleware, així com fitxers o altres fonts estructurades, semiestructurades o no estructurades.

A més, planifiquen el transport de dades, log d'errors, log de canvis i estadístiques associades als processos de moviments de dades. Aquest tipus d'eines solen tenir una interfície d'usuari de tipus *GUI* i permeten dissenyar, administrar i controlar cada un dels processos de l'entorn *ETL*.

- EII: l'objectiu d'aquesta tecnologia és permetre a les aplicacions l'accés a dades disperses (des d'un *Data mart* fins a fitxers de text o fins i tot *Web services*) com si estiguessin tots residint en una base de dades comú. Per tant se basa en la federació. L'accés a dades disperses implica la descomposició de la consulta inicial (habitualment *SQL*) direccionalada contra la vista virtual federada en subcomponents, que seran processades en cada un dels entorns on resideixen les dades. Es recullen els resultats individuals de cada un dels subcomponents de la consulta, es combinen adequadament i es torna el resultat a l'aplicació que va llançar la consulta. Els productes de *EII* han evolucionat des de dos entorns d'origen diferenciats: les bases de dades relacionals i les bases de dades *XML*. Actualment, la tendència en productes *EII* és que soporten ambdues interfícies de dades, *SQL (ODBC i JDBC)* i *XML (XQuery i XPath)*.
- EDR: té com objectiu detectar els canvis que succeeixen les fonts d'origen. Està suportada per les tècniques d'integració de dades de *CDC (Change Data Capture)* i per la tècnica de propagació de dades.

### 3.2.4.2.2 Ús de l'integració de dades

Els processos d'integració de dades s'usen en múltiples tipologies de projectes. Podem destacar els següents:

- Migració de dades.
- Processos de qualitat de dades.
- *Corporate Performance Management (CPM)*.
- *Master Data Management (MDM)*.
- *Customer Data Integration (CDI)*.
- *Product Information Management (PIM)*.
- *Enterprise Information Management (EIM)*.
- *Data warehousing*.
- *Business Intelligence (BI)*.

En el cas del present document, aquests dos darrers punts (*Data warehousing* i *Business Intelligence*) seran els dos punts que es tractaran amb més detall.

### 3.2.4.2.3 ETL en el context de Pentaho

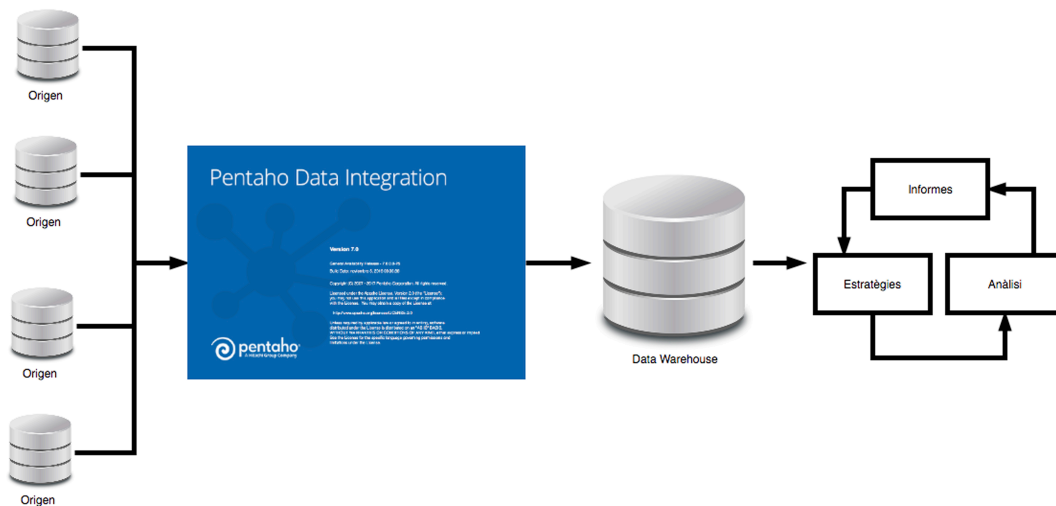
*Pentaho Data Integration (PDI)*, anteriorment anomenat *Kettle*, va ser iniciat el 2001 per Matt Casters. A l'any 2006, *Pentaho* va adquirir *Kettle* i va canviar el nom després de que aquest passés a ser *Open Source*. D'aquesta forma continuava amb la política de crear una *suite* completa d'intel·ligència de negoci *Open Source*.

*PDI* és una solució d'integració de dades programada en *Java* orientada completament a l'usuari i basada en un enfocament de metadades. Els processos *ETL* s'encapsulen en metadades que s'executen a través del motor *ETL*.



Il·lustració 19: Imatge Pentaho Data Integration (PDI).



Aquesta eina permet carregar dades de múltiples fonts d'origen, carregar aquestes dades en un *Data warehouse* perquè posteriorment la informació consolidada sigui utilitzada a nivell operatiu, tàctic i estratègic.



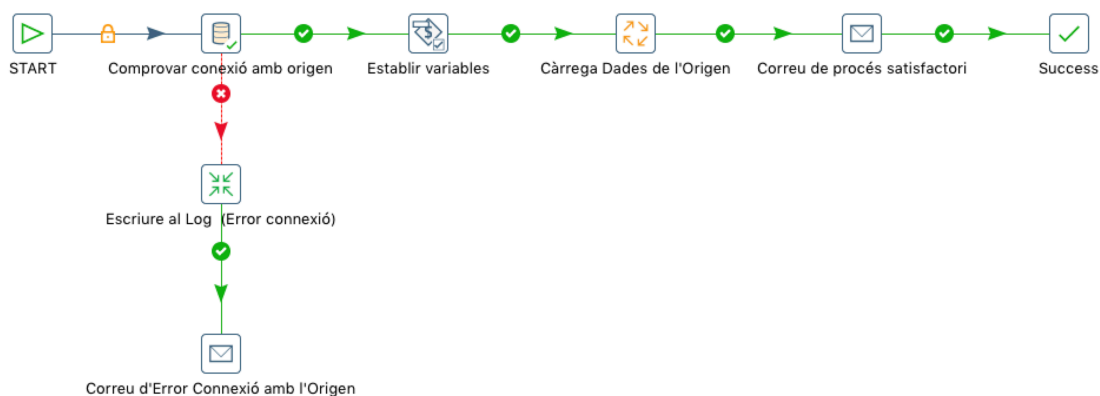
Il·lustració 20: Procés general PDI.

Les principals característiques de *PDI* son:

- Entorn gràfic orientat al desenvolupament ràpid i àgil basat en dues àrees: la de treball i la de disseny/vista.
- Multiplataforma.
- Inclou múltiples connectors a bases de dades, tant propietàries com comercials. Així com connectors a fitxers plans, Excel, XML o altres.
- Arquitectura extensible mitjançant plugins.
- Suporta us de clúster, processos *ETL* en paral·lel i arquitectures servidor mestre-esclau.
- Completament integrat amb la suite de *Pentaho*.
- Basat en el desenvolupament de dos tipus d'objectes:

- Treballs o *Jobs* : permeten gestionar i administrar processos *ETL* a alt nivell.
- Transformacions : permeten definir les operacions de transformació de dades.

Els processos d'integració de dades per norma general segueixen el següent esquema:



Il·lustració 21: Procés exemple d'Integració de dades en PDI.

### 3.2.4.3 Cubs OLAP

És ben sabut que el concepte de *Business Intelligence* engloba múltiples conceptes. Un dels més importants és el concepte *OLAP* (*On-Line Analytical Processing*), denominat així per Edgar F. Codd.

Una manera senzilla d'explicar aquest concepte és dir que es una tecnologia que permet un anàlisi multidimensional a través de taules matricials o pivotants.

*OLAP* forma part del que és coneix com sistemes analítics, que permeten respondre preguntes com '*Per què va passar?*'. Aquests sistemes poden trobar-se integrats en suites de *Business Intelligence* o ser simplement una aplicació independent.

S'entén per **OLAP**, o procés analític en línia, al mètode àgil i flexible per organitzar dades, especialment metadades, sobre un objecte o jerarquia d'objectes com en un sistema o organització multidimensional, i amb l'objectiu de recuperar i manipular dades i combinacions de les mateixes dades a través de consultes o inclús informes.

Una eina *OLAP* està formada per un motor i un visor. El motor és, en realitat, just el concepte que s'ha definit d'*OLAP*. El visor *OLAP* és una interfície que permet consultar, manipular, reordenar i filtrar dades existents en una estructura *OLAP* mitjançant una interfície gràfica d'usuari que disposa funcions de consulta de tipus *MDX* (*Multidimensional Expressions*) i d'altres.

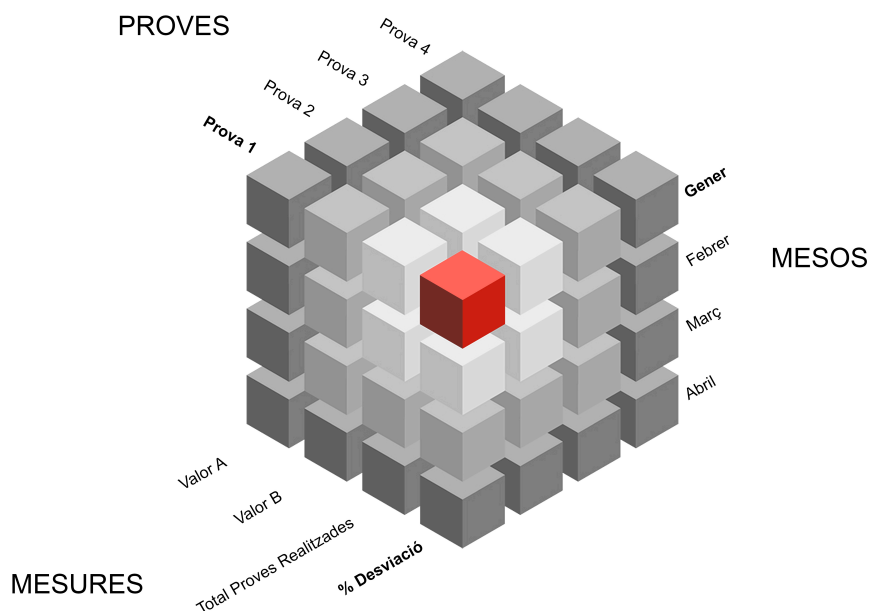
Les estructures *OLAP* permeten realitzar preguntes que serien sumament complexes mitjançant *SQL*.

Un exemple genèric d'un cub *OLAP*, per entendre el concepte i el tipus d'informació que ofereix, que ens permet obtenir informació de les diferents proves realitzades, en un conjunt de mesos concret i amb uns indicadors o mesures seleccionades.



Imaginem que volem respondre a la següent pregunta:

**Quin és el percentatge de desviació de la Prova 1 per al més de Gener?**



**Il·lustració 22: Exemple genèric d'un Cub OLAP.**

Si tenim un cub, com el de l'exemple, format per temps, proves i mesures, la resposta és la intersecció entre els diferents elements (*colorejat en vermell*).

Aquest cub *OLAP* simula una consulta sobre el percentatge de desviació (*% Desviació*) que s'ha obtingut quan s'ha realitzat una prova determinada (*Prova 1*) en un més concret (*Gener*).

Cal observar que una estructura d'aquesta forma permet consultes molt més complexes, com per exemple comparar els diferents valors de les mesures obtingudes de febrer i maig, entre les diferents proves realitzades.

A més, el vistor *OLAP* proporciona llibertat als usuaris finals per realitzar consultes de manera independent al departament de *IT*.

### 3.2.4.3.1 Elements OLAP

*OLAP* permet l'anàlisi multidimensional. Això significa que la informació està estructurada en **eixos** (punts de vista d'anàlisi) i **cel·les** (valors que s'estàn analitzant).

En el context *OLAP* existeixen diferents elements comuns a les diferents tipologies *OLAP*:

- Esquema: un esquema és una col·lecció de cubs, dimensions, taules de fets i rols.
- Cub: és una col·lecció de dimensions associades a una taula de fet. Un cub virtual permet creuar la informació entre taules de fet a partir de les seves dimensions comuns.
- Taula de fet, dimensió i mètrica.
- Jerarquia: és un conjunt de membres organitzats en nivells. En quan a bases de dades, se poden entendre com una ordenació dels atributs d'una dimensió.
- Nivell: és un grup de membres en una jerarquia que tenen els mateixos atributs i nivell de profunditat en la jerarquia.
- Membre: és un punt en la dimensió d'un cub que pertany a un determinat nivell d'una jerarquia. Les mètriques en *OLAP* se consideren un tipus especial de membre que pertany al seu propi tipus de dimensió. Un membre pot tenir propietats associades.
- Rols: permisos associats a un grup d'usuaris.
- MDX: és un acrònim de *Multidimensional eXpressions* (també és conegut com *Multidimensional Query eXpression*). És el llenguatge de consulta d'estructures *OLAP*, si bé no és un llenguatge estàndard, la gran majoria de fabricants d'eines *OLAP* l'han adoptat com estàndard de fet.

### 3.2.4.3.2 Creació Cub OLAP

Els cubs es creen amb *Pentaho Schema Workbench (PSW)* que és una eina multiplataforma de desenvolupament que permet crear, modificar i publicar un esquema de *Mondrian*.

Permet crear tots els objectes que soporta *Mondrian*: esquema, cub, dimensió, jerarquies, nivells, mesures, mesures calculades, elements virtuals, rols i operacions estàndard entre d'altres elements.



Il·lustració 23: Opcions disponibles de Pentaho Schema Workbench.

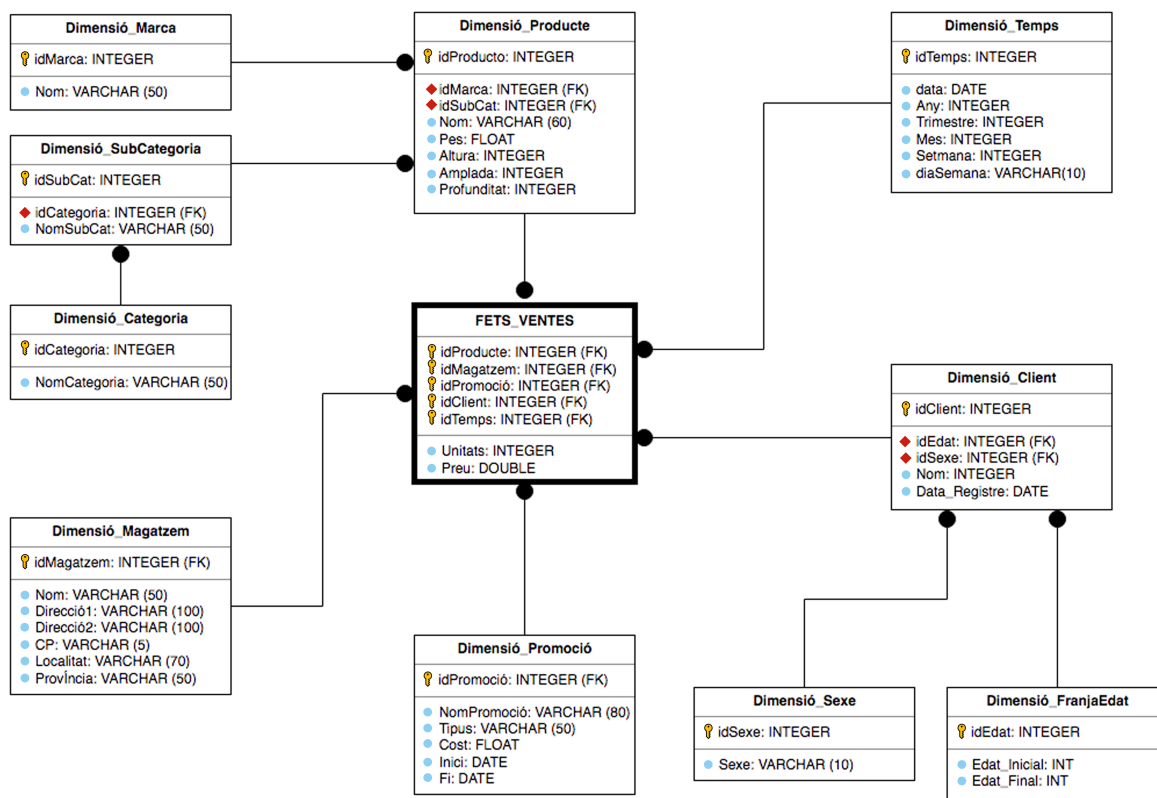
A l'hora de crear un cub s'ha de tenir en compte que el que es fa es mapejar el disseny de la base de dades (taules de fet i dimensions) amb el disseny que es vol realitzar, de forma que:

- És possible crear un esquema a partir d'un subconjunt de les dades emmagatzemades a la base de dades.
- És possible crear un esquema amb la mateixa quantitat d'elements. Es consideren totes les taules de fets i les dimensions.
- És possible crear un esquema amb més elements dels que existeixen a la base de dades. Per exemple, és possible crear dimensions o altres objectes que només existeixen en l'esquema *OLAP* i que es generen en memòria.

Per altre banda, aquesta eina inclou un explorador de la base de dades que, una vegada s'ha creat la connexió a la base de dades, permet explorar l'estructura de les taules per recordar quin és el nom dels camps i atributs que s'usaran.

Les passes en el procés de creació son les següents:

1. Creació d'una connexió al *Data warehouse*.
2. Creació d'un esquema.
3. Creació d'un cub.
4. Creació d'una taula de fets (elegir la taula del *Data warehouse* que té el rol de taula de fets).
5. Creació de  $n$  dimensions amb  $n$  claus foranes (especificar el nom de la taula que conté la informació de la dimensió).
6. Creació d'una o varies jerarquies (definir també la seva clau primaria).
7. Creació d'almenys un nivell per cada jerarquia (definir la taula on es troba la informació, la col·lumna que conté aquesta informació, com s'ordenen els resultats, la tipologia del valor, la tipologia de nivell, si els valors son únics i si és necessari ocultar el nivell en algun cas).
8. Alguns membres contenen informació addicional que es poden mostrar com a propietats. En aquest cas, es pot mostrar la descripció del valor. Afegim una propietat i definim el nom, la col·lumna que conté la informació i la tipologia del seu valor.
9. Una vegada definides les dimensions, és necessari definir les mètriques.
10. Finalment, es guarda l'esquema creat (generarà un fitxer amb extensió .xml)



Il·lustració 24: Exemple d'estructura tipus estrella de les taules d'un cub OLAP.

Un exemple genèric de l'estructuració de les taules en un cub *OLAP* (veure [il·lustració 24](#)) on es pot veure un esquema tipus floc de neu o estrella on hi ha una taula central (taula de fets) que està composta per les *foreign keys* de les taules que l'envolten (taules de dimensions) i que permet accedir a la informació que emmagatzemen aquestes taules que fan el rol de dimensions. Una dimensió associada a una taula de fets pot tenir, al mateix temps, més dimensions enllaçades.

### 3.2.4.4 Visualització de les dades

#### 3.2.4.4.1 Informes

El punt d'entrada tradicional per una eina d'intel·ligència de negoci en el context d'una organització és la necessitat d'informes operacionals.

Durant la vida d'una empresa, la quantitat de dades que es generen per la seva activitat de negoci creix de forma notable, i aquesta informació és guardada tant en bases de dades de les aplicacions de negoci com en fitxers en múltiples formats.

És necessari generar i distribuir informes per conèixer l'estat del negoci i poder prendre decisions a tots els nivells: operatiu, tàctic i estratègic.

Per dur a terme aquesta finalitat, s'usa l'eina *Pentaho Report Designer* que és el motor de reporting de *Pentaho* que està integrat en la suite.

*Pentaho Report Designer* és un editor basat en *eclipse* amb prestacions professionals de customització d'informes destinat a desenvolupadors.

#### 3.2.4.4.2 Quadre de comandament

Tant els informes com els *OLAP* són eines que proporcionen informació als usuaris finals. La gran quantitat d'informació que normalment inclouen aquestes eines pot fer-les inadequades per usuaris que necessiten prendre decisions de forma ràpida a partir d'elles.

S'entén per **Quadre de comandament** o *dashboard* al sistema que informa de l'evolució dels paràmetres fonamentals de negoci d'una organització o d'una àrea del mateix.

El Quadre de comandament prové del concepte francès *tableau du bord*, i permet mostrar informació consolidada a alt nivell.

Es focalitza en:

- Presentar una quantitat reduïda d'aspectes de negoci.
- Ús majoritari d'elements gràfics.
- Inclusió d'elements interactius per potenciar l'anàlisi en profunditat i la comprensió de la informació consultada.



Il·lustració 25: Imatge genèrica d'un Quadre de Comandaments.

A la imatge anterior es pot veure un conjunt de dades agregades i que dona informació general fent ús de gràfics per sintetitzar les dades i oferir una visió global de l'estat del negoci.

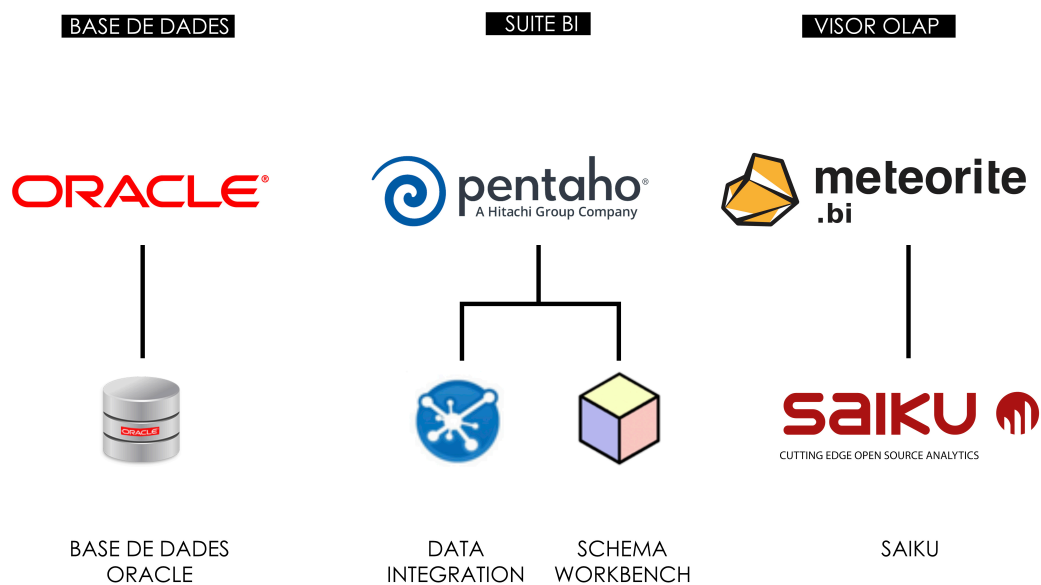
## 4 Solució elegida

---

En aquest apartat es tracta la solució que es considera més adient i s'exposen els motius per els quals s'ha fet aquesta elecció, tenint en compte l'entorn, els marcs teòrics, de negoci i tecnològics actuals.

Per dur a terme el procés de creació del cub per al LIS hem de tenir en compte que *Álamo Consulting* disposa i treballa amb:

- Base de Dades: Oracle
- Suite BI: Pentaho
- Visor OLAP: Saiku



Il·lustració 26: Imatge marc tecnològic elegit.

## 4.1 Característiques de la solució elegida

· Oracle: és el motor de base de dades objecte-relacional més usat a nivell mundial. Un entorn multiplataforma que pot executar-se a qualsevol tipus de ordinador (PC o supercomputador). Alta protecció de les dades, oferint una seguretat completa en l'entorn de producció i de proves. Ofereix la gestió de còpies de seguretat. Llenguatge de disseny de bases de dades molt complet (PL/SQL) que permeten adaptar-se a les necessitats canviants del negoci. Alta disponibilitat. Ofereix la possibilitat d'escalabilitat, protecció i alt rendiment per l'activitat empresarial i la gestió d'usuaris.

· Pentaho: és líder mundial de sistemes d'Intel·ligència de Negoci *Open Source*. *Pentaho BI Suite* ofereix una ampla gama d'eines orientades a la integració i a l'anàlisi intel·ligent de les dades d'una organització. Compta amb potents capacitats per la gestió de processos *ETL* (Extracció, Transformació i Càrrega de dades), informes interactius, anàlisi multidimensional de informació *OLAP* o mineria de dades.

· Saiku: és un excel·lent visor *OLAP*, *Open Source*, que proporciona una eina per realitzar ànlisi de forma fàcil i intuïtiva. Ofereix solucions d'excelent qualitat a la vanguardia de la tecnologia i delicada experiència d'usuari.



## 5 Laboratory Information System (LIS)

---

Com ja s'ha definit, el **HIS** (acrònim de *Hospital Information System* o sistema d'informació hospitalari) és un sistema integrat d'informació dissenyat per gestionar tots els aspectes clínics, administratius i financers d'un hospital.

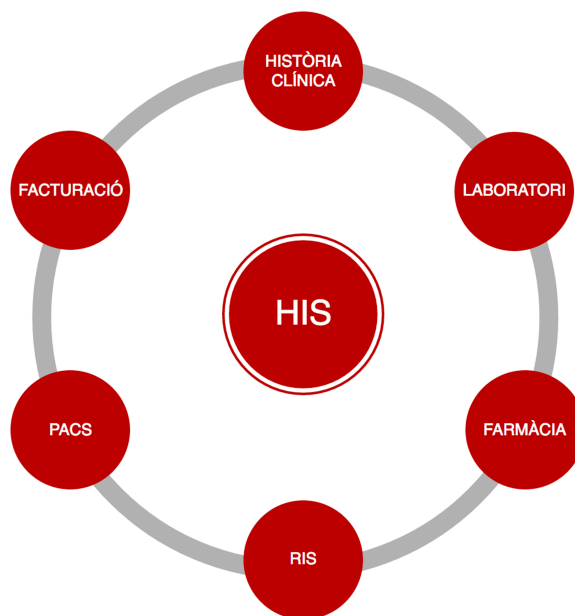
Està format per diferents mòduls com:

- *RIS (Radiological Information System)*
- *PACS (Picture Archiving and Communication System)*
- Facturació
- Història Clínica
- Laboratori, **LIS** (*Laboratory Information System*)
- Farmàcia

El **LIS** és un conjunt de hardware i software que dona suport a l'activitat d'un laboratori clínic.

Els sistemes d'informació del laboratori clínic (*LIS* o *SIL*) constitueix avui en dia una eina indispensable i crítica per l'activitat dels laboratoris clínics.

L'evolució del *LIS* unida al desenvolupament dels grans autoanalitzadors i la robòtica, han possibilitat un augment extraordinari de la capacitat i l'eficiència dels laboratoris amb elevats nivells de qualitat.



Il·lustració 27: El HIS i els mòduls que el formen.

## 5.1 Característiques generals

### 5.1.1 Estructuració de la informació

En general, els laboratoris clínics informen dels resultats de proves analítiques realitzades en mostres procedents d'un pacient en un moment de la seva vida amb una finalitat determinada.

La base de dades principal del LIS hauria de tenir una estructura jeràrquica amb aspectes referents a:

- Pacient: dades demogràfiques i administratives.
- Solicitud o petició: tipus, data, hora, motiu, peticionari o altres.
- Mostra: sang total, sèrum, orina i/o LCR, entre d'altres.
- Prova (amb mètode): per exemple, glucosa, urea, hemograma o d'altres.
- Resultat (amb unitats i interval de referència en el seu cas): numèric, alfanumèric, informe, comentaris o d'altres.

### 5.1.2 Criticitat

El propi desenvolupament dels sistemes d'informació de laboratori (LIS), unit a les possibilitats d'automatització i robotització, ha implicat un gran augment de la capacitat productiva dels laboratoris, que a la seva vegada ha augmentat la seva dependència del LIS. Actualment, un laboratori d'un hospital de 1.000 llits amb assistència primària pot rebre al dia un nombre aproximat de 3.000 sol·licituts d'anàlisi que corresponen amb l'extracció d'uns 10.000 tubs, la realització de 30.000 proves i l'informe de 60.000 resultats.

El LIS està normalment connectat a temps real a molts analitzadors que necessiten una resposta ràpida als seus requeriments de dades, o per contra, s'interromprà la comunicació. La velocitat resposta, per tant, és un requisit imprescindible dels LIS, havent de ser aquesta immediata després de la sol·licitud de l'autoanalitzador. En l'exemple anterior, el LIS pot estar connectat directament a 50 equips analitzadors.

En l'actualitat, en la majoria de casos no existeix una alternativa manual al sistema informàtic i quan aquest falla no hi ha suficient capacitat operativa, lo que genera retrassos en l'entrega de resultats i, en ocasions, deteriorament irreversible de les mostres amb el conseqüent perjudici als pacients.

### 5.1.3 Traçabilitat

Les normes legals i administratives, així com els sistemes de qualitat obliguen a que tot el procés de laboratori sigui rastrejable, de tal manera que el sistema permeti reconstruir tot el que ha passat des de que es realitza la sol·licitud fins que es rebí o es visualitzi l'informe.

Això suposa conèixer quina persona o instrument ha duit a terme qualsevol acció en tot el procés, el moment en que ha ocorregut i el resultat de l'acció. Alguns exemples d'aquest fet són: qui i quan va fer la sol·licitud, qui i quan va obtenir la mostra i quants tubs es varen extreure, qui i quan va realitzar el fraccionament d'una mostra i quantes fraccions es varen obtenir, i quan ha entrat la mostra en un determinat analitzador i quines proves es varen sol·licitar, entre altres coses. En compliment de la normativa legal de seguretat de la informació qualsevol acció realitzada sobre les dades, registre, consulta, validació, informe o altres ha de quedar registrada.

Aquesta ingent quantitat d'informació serveix per delimitar responsabilitats, establir accions de millora i obtenir indicadors de qualitat que permetin marcar objectius i realitzar el seu seguiment.

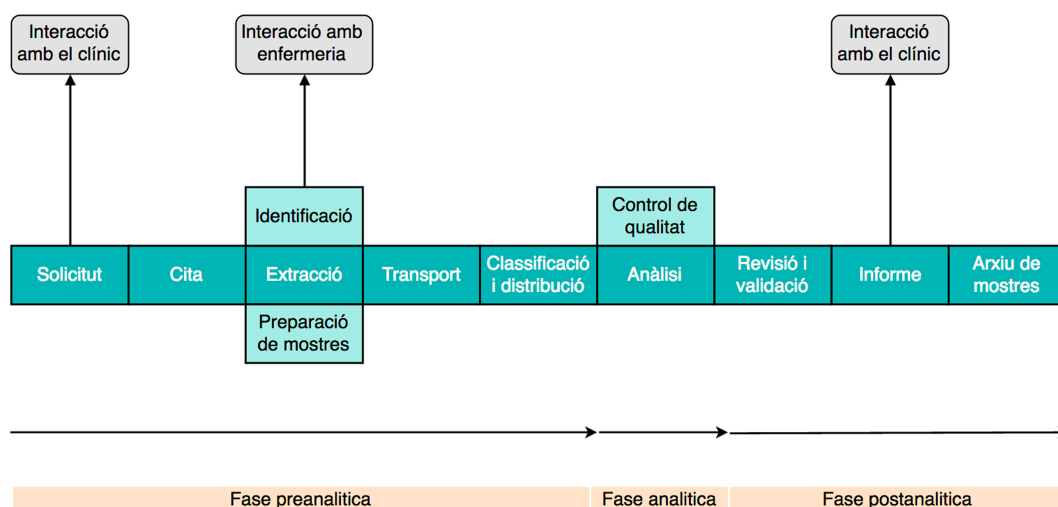
### 5.1.4 Altres característiques del LIS

- Flexibilitat: les característiques dels laboratoris son molt variades i es necessiten sistemes flexibles que s'adaptin a l'organització i necessitats del laboratori, i no a l'inversa.
- Modularitat i escalabilitat: ja que les necessitats dels laboratoris evolucionen contínuament, és important que els LIS permetin el creixement i la incorporació de noves funcionalitats.
- Seguretat i confidencialitat: son de màxima importància ja que el tipus d'informació que es maneja, així com el gran nombre i dispersió de les persones que accedeixen a ella.

## 5.2 Fluxe de treball dels laboratoris

El procés del laboratori normalment comença amb una sol·licitud realitzada per un clínic i finalitza amb el corresponent informe. Entre la sol·licitud i l'informe ocorren una sèrie de fases o subprocessos en els que els sistemes d'informació juguen un paper cada vegada major. seqüència d'aquestes fases (veure *il·lustració 28*), que de forma agrupada solen denominar-se fase preanalítica, fase analítica i fase postanalítica.

En l'apartat assistencial, els LIS ajuden en els **aspectes preanalítics** (sol·licitud, cita, obtenció de mostres, preparació, transport, fraccionament, i distribució), en els **aspectes analítics** (processament, gestió d'equips i rutes, control de qualitat, validació tècnica) i en els **aspectes postanalítics** (validació clínica, edició d'informes, distribució, arxiu de mostres).



Il·lustració 28: Procés habitual de laboratori.

### 5.3 El Laboratori dels Hospitals de les Illes

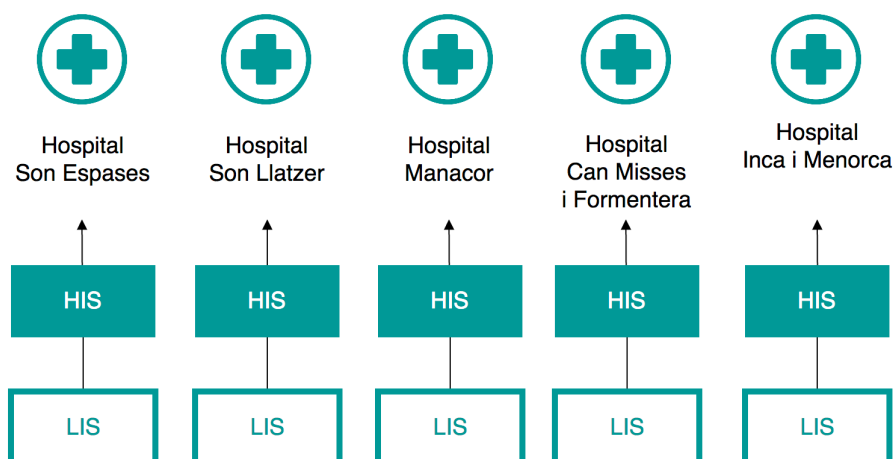
A les Illes Balears, l'IB-Salut compta amb un total de 7 hospitals:

- Hospital Universitari Son Espases
- Hospital Son Llàtzer
- Hospital de Manacor
- Hospital Mateu Orfila
- Hospital Comarcal d'Inca
- Hospital Can Misses
- Hospital de Formentera

Cada un d'aquests hospitals compta amb el seu propi *HIS* que permet gestionar l'activitat hospitalària, però en l'apartat del *LIS* s'han de tenir en compte dues consideracions rellevants:

1. L'activitat de laboratori de l'Hospital Can Misses (Eivissa) i de l'Hospital de Formentera (Formentera), es tractaran conjuntament.
2. L'activitat de laboratori de l'Hospital Mateu Orfila (Menorca) i de l'Hospital d'Inca (Mallorca), es tractaran conjuntament.

Per tant, es tractarà l'activitat del LIS dels següents hospitals:



Il·lustració 29: Hospitals a considerar l'activitat del LIS.

### 5.3.1 Dades generades per l'activitat dels laboratoris

Cada un dels hospitals genera informació per l'activitat duta a terme en els laboratoris, si es té en compte únicament l'activitat d'un dels hospitals es generen dades estructurades en les següents tipologies:

Usuaris	Unitat	Tipus procedència
Pacients	Secció	Procedències
Metges	Àrea	Proced. Servei
Proves	Laboratoris	Centre procés
Tipus de prova	Estat resultat	Centre extracció
Tipus de mostra	Determin. mes	Aïllat
Atb	Antibiogrames	Microorganismes

Il·lustració 30: Hospitals a considerar l'activitat del LIS.

## 6 Creació cub per al LIS

---

En aquest apartat s'explica pas a pas com crear el cub per al LIS i quines dades s'usaran. Com s'obtenen aquestes dades i el procés necessari per poder crear el cub i la seva posterior visualització.

### 6.1 Dades dels LIS generades als Hospitals

Com a conseqüència de l'activitat del laboratori, es pot extreure una gran quantitat d'informació associada. Totes aquestes dades queden emmagatzemades a diferents taules segons els diferents hospitals d'origen esmentats amb anterioritat (veure *il·lustració 30*).

A partir d'aquestes dades extraurem, posteriorment, una sèrie d'indicadors.

Un **indicador** és un valor o dada precalculada que serveix per conèixer o valorar característiques i la intensitat d'un fet o per determinar la seva evolució.

Concretament, en l'àmbit de laboratori seran totes aquelles dades agregades referents a l'activitat dels diferents laboratoris dels hospitals de les Illes Balears que permetran als rols rols indicats avaluar l'activitat d'un període o comparar-la amb activitats de períodes anteriors.

- Nombre de Peticions Totals: S'entén per petició cada una de les ocasions distintes en les que a un pacient se li pren una mostra per ser analitzada posteriorment en el laboratori.
- Nombre de Peticions Realitzades: S'entén per petició realitzada cada uan de les peticions en les que s'ha emès un resultat.
- Nombre de determinacions realitzades: S'entén per determinació o prova analítica el producte final d'una activitat de laboratori que genera un resultat que s'emet al facultatiu sol·licitant.
- Nombre de proves derivades: Sumatori de totes les proves diferents de cada una de les peticions que son transferides a un centre de procés.
- Nombre de proves cancel·lades: Sumatori de totes aquelles proves l'estat del resultat està inclòs (veure *il·lustració 31*)

Estat del resultat	Descripció
10	Resultat importat de l'històric
11	Resultat duplicat
12	Regla programació rec/reflexiva
13	Regla de rebuig
90	Prova anul·lada
91	Prova/Facturació anul·lada
92	Prova pacient duplicada període
93	Prova refusada
94	No calculable
95	Anul·lada per regla
96	Anul·lada per incidència
97	Mostra anul·lada

Il·lustració 31: Motius de cancel·lació de les proves de LIS.

- Nombre de pacients diferents: Sumatori dels pacients que han sol·licitat alguna prova en un període d'estudi (per defecte, un mes). Un pacient que ha sol·licitat diferents proves en el mateix període d'estudi, contarà únicament una sola vegada.

- Temps de demora:

- Laboratori Ordinari: Diferència en hores entre les dates de validació tècnica i la data de registre en el laboratori, dividit per el nombre total de proves ordinaries. Quan la data de validació tècnica no està informada se considerarà la data de validació facultativa. S'exclouen proves derivades i en situació especial.

- Laboratori Urgent: Diferència en hores entre les dates de validació tècnica i la data de registre en el laboratori, dividit per el nombre total de proves urgents. Quan la data de validació tècnica no està informada se considerarà la data de validació facultativa. S'exclouen proves derivades i en situació especial.

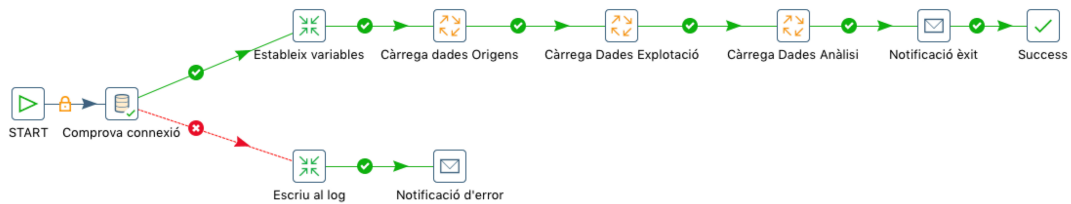
- Nombre d'Antibiogrames: Correspon al nombre de proves d'antibiogrames realitzades per aquest període.



## 6.2 Procés General

Per crear el cub de laboratori s'ha fet ús d'un procediment de Pentaho Data Integration. Aquest procés engloba les etapes de descàrrega de dades, tractament de les dades i la càrrega a les taules adjacents per poder crear el cub.

El procés segueix la següent estructuració:



Il·lustració 32: Procés ETL general del Laboratori.

### 6.2.1 Descàrrega de les dades

La descàrrega es fa mitjançant una connexió a cada un dels orígens al Job anomenat 'Càrrega dades Orígens' (veure *il·lustració 32*).



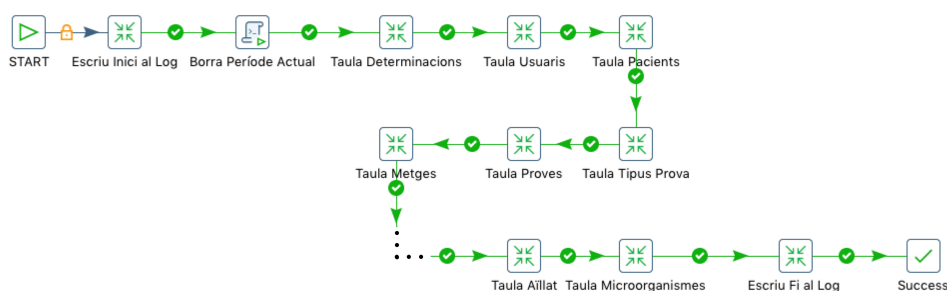
Il·lustració 33: Descàrrega de les dades [Carrega dades Orígens].

A través de *DBLinks* s'accedeixen als orígens i es fa la descàrrega (veure *il·lustració 33*). D'aquesta manera, s'obtenen les dades dels diferents orígens i es copien a taules temporals de la zona de maniobra del del *Data warehouse*.

## 6.2.2 Tractament de les dades

Les dades, emmagatzemades actualment a taules de la zona de maniobra que únicament s'usen com a contenidor per poder realitzar la descàrrega, es carreguen al següent pas a la zona d'exploració al *Job* anomenat '*Càrrega Dades Exploració*' (veure *il·lustració 32*). Aquest pas borra, si existeixen, dades del període actual per evitar duplicats a les taules finals del *Data warehouse*.

En aquest *Job*, una vegada les dades han estat descarregades dels diferents orígens i s'ha assegurat que no existeixen dades per el període a carregar, les dades son tractades segons l'àmbit de les dades (veure *il·lustració 30*) i volcades a taules del model d'exploració de la *FIC*.

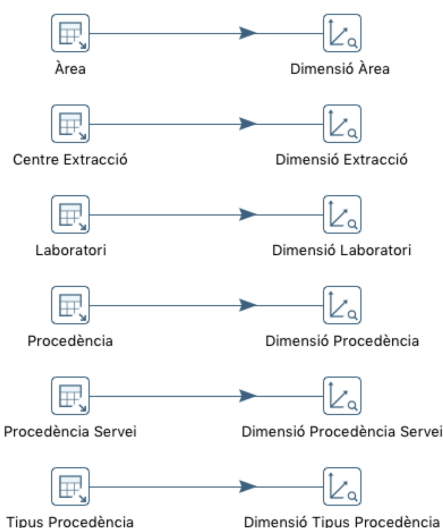


Il·lustració 34: Càrrega del model d'exploració [Carrega dades Exploració].

A cada una de les transformacions es carreguen, a les taules del model d'exploració, les dades pertinents de cada un dels apartats que s'han considerat del laboratori.

### 6.2.3 Càrrega de les dimensions

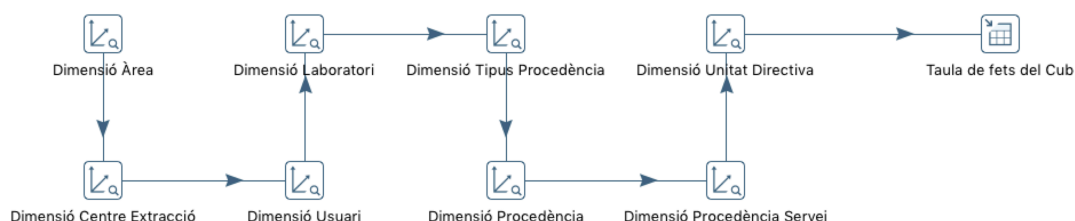
Una vegada es disposa de les dades a taules del model d'exploració, es fa la càrrega de les dimensions del cub. Aquest pas proveirà d'informació a les taules de dimensions que el cub consultarà a l'hora d'efectuar una consulta sobre el cub. Aquesta tasca es porta a terme al Job anomenat 'Càrrega Dades Anàlisi' (veure *il·lustració 32*).



Il·lustració 35: Càrrega de les dimensions del Cub [Carrega dades Anàlisi].

### 6.2.4 Associació amb la taula de fets

Una vegada es disposa de les dades a taules del model d'exploració, es fa la càrrega de les dimensions del cub. Aquest pas proveirà d'informació a les taules de dimensions que el cub usarà a l'hora d'efectuar una consulta sobre el cub. Aquesta tasca es porta a terme al Job anomenat 'Càrrega Dades Anàlisi' (veure *il·lustració 32*).

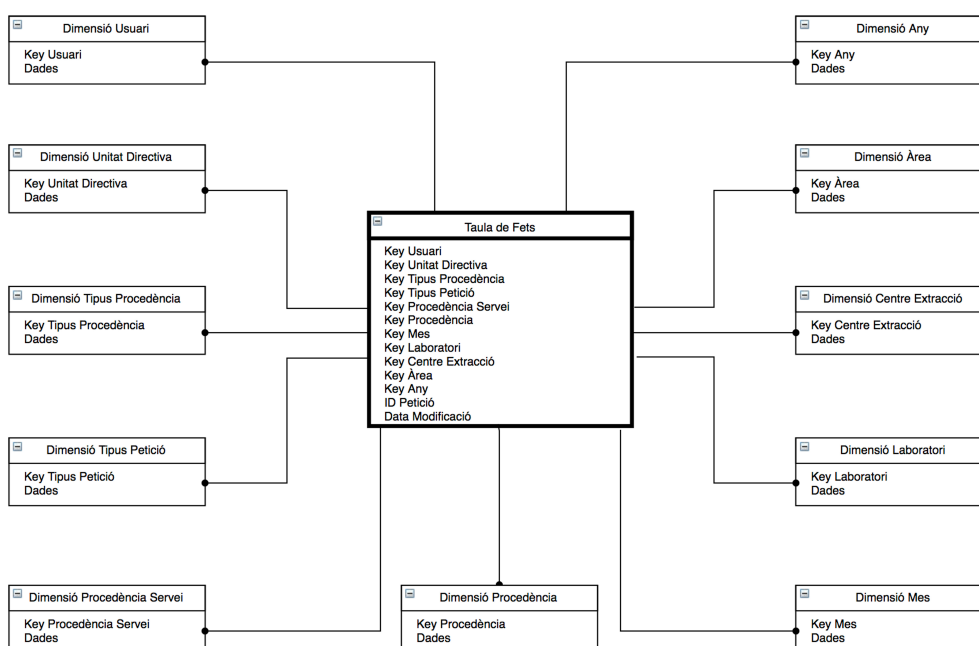


Il·lustració 36: Associació de les dimensions a la taula de fets [Carrega dades Anàlisi].

## 6.2.5 Creació del cub (Gestió de peticions de Laboratori)

A l'hora de crear el cub s'ha fet ús de l'eina *Pentaho Schema Workbench* i es segueix el procés de creació d'un cub i totes les seves dependències (veure apartat [3.2.4.3.2 Creació cub OLAP](#)). Es necessari que la vinculació entre les taules es faci de forma correcta per poder accedir a la informació de manera correcta.

L'enllaç de la taula de fets amb les seves pertinents dimensions ha seguit la següent estructuració:



Il·lustració 37: Estructuració de les dimensions i la taula de fets del cub.

## 6.2.6 Visualització de les dades del cub

Una vegada s'ha creat el cub, obtenint el fitxer amb extensió .xml, es posa a disposició dels usuaris finals per poder realitzar les consultes pertinents.

Es pot accedir al cub visitant el quadre de comandament habilitat per consultar les dades de la *FIC*, a l'apartat *Anàlisi Lliure*.

Nombre	Etiquetas	
[blurred]	ACT FAC BO	Link
[blurred]	GRH BO	Abrir
[blurred]	SCA BO	Abrir
[blurred]	HdS FIC BO	Link
[blurred]	CM BO	Abrir
[blurred]	CM BO	Abrir
[blurred]	MAL USOFIC PENTAHO	Abrir
[blurred]	MAL USOFIC PENTAHO	Abrir
Gestor de Peticiones - LIS	MAL GPE PENTAHO	Abrir
[blurred]	PENTAHO	Abrir
[blurred]	CM GESMA BO	Link
[blurred]	AT BO	Abrir
[blurred]	CON BO	Link

Il·lustració 38: Llistat dels cubs disponibles.

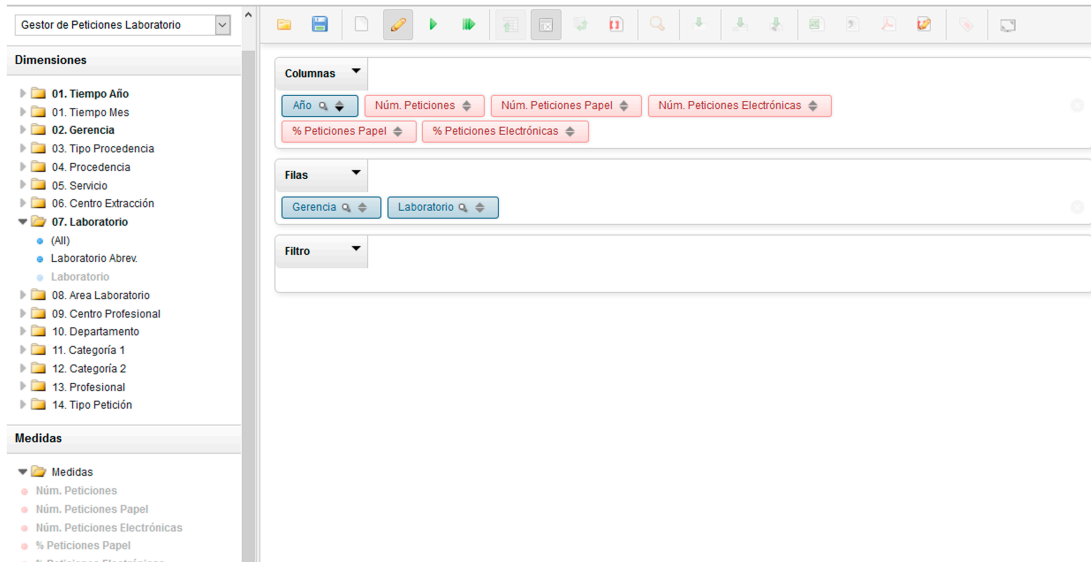
Una vegada accedit a l'apartat d'Anàlisi Lliure veurem els cubs disponibles i prement el botó *obrir* accedirem a *Saiku*.

Una vegada hagi carregat, seleccionarem el cub que volem consultar (veure *il·lustració 39*).

Apareix una secció amb tres apartats:

- Columnes
- Files
- Filtres

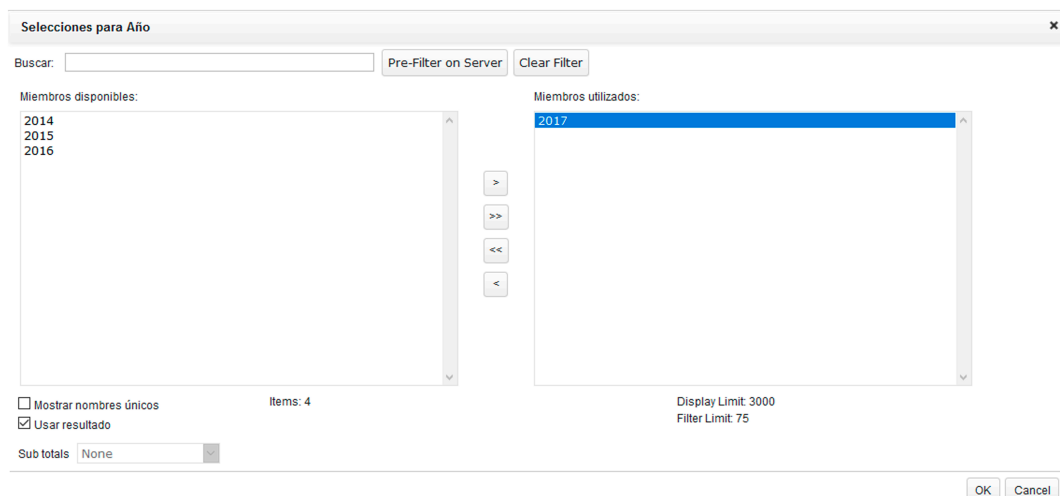
Aquests apartats marcaran l'estructuració final de la informació elegida, si en columnes o en files, i els filtres a aplicar sobre les dades perquè compleixin uns requisits específics.



Il·lustració 39: Selecció de les dades a mostrar del cub.

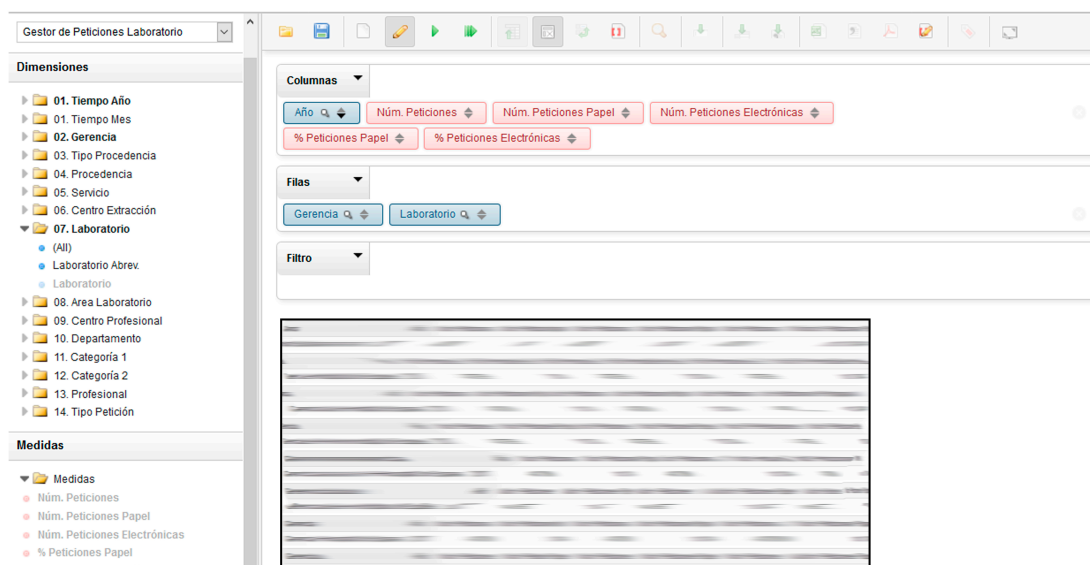
Es trien els valors que es volen representar i de la forma en que es volen estructurar. En aquest cas volem que el cub imprimeixi per columnes la informació referent al total de peticions i al tipus de petició (electrònica o en paper) i els percentatges de cada un dels tipus. Com a files volem que ens mostri la informació anterior per hospitals (gerències) i per laboratoris (veure [il·lustració 39](#)).

Si volem que únicament ens mostri les dades, descrites al paràgraf anterior, per a l'any 2017 podem definir la següent condició:



Il·lustració 40: Filtrar la informació del cub per any.

Si premem el botó perquè ens mostri la informació, aquesta apareixerà a la part inferior mostrant tots els valors que compleixen les restriccions elegides:



Il·lustració 41: Resultat de executar la consulta sobre el cub.





## 7 Conclusions

---

Actualment qualsevol acció que es realitza tant al *mon digital* com en el dia a dia genera informació i un conjunt de metadades associades.

El context de la societat de la informació ha propiciat la necessitat de tenir millors, més ràpids i més eficients mètodes per extreure i transformar les dades d'una organització en informació i distribuir-la al llarg de la cadena de valor.

Tenint en compte la tendència a l'alça de l'ús de sistemes o dispositius electrònics tant a nivell laboral com a nivell personal i la gran quantitat de dades generades en qualsevol d'aquests àmbits, esdevé una tendència cada vegada més comú interessar-se per aquests grans volums de dades.

Aquesta pròpia informació pot generar activitat comercial entre diferents organitzacions per benefici comú o simplement per benefici econòmic per una de les parts.

Per aquest motiu, es de gran importància plantejar-se noves formes de poder fer ús de la informació que es genera, emmagatzemar-la de manera correcta i estructurada per poder treure un rendiment o tenir la capacitat de prendre decisions en funció dels resultats.

La gran majoria d'organitzacions necessiten actualment sistemes de la informació per prendre millors decisions i ser més competitives i han identificat el *Business Intelligence* com una de les principals necessitats. Per aquest motiu, és necessari poder construir solucions.

Als darrers anys, el mercat *Business Intelligence* s'ha vist marcat per una clara evolució que el destaca com un mercat madur ja que s'ha produït una consolidació mitjançant la compra d'empreses petites per part dels principals agents del mercat (*SAP, IBM o Microsoft*). També s'han enriquit amb solucions *Open Source* que cobreixen l'espectre de necessitats d'una organització per l'explotació de la informació. Inclús hi ha una tendència a l'aparició de noves empreses innovadores que centren els seus esforços en el mercat de la intel·ligència de negoci (*visualització, l'anàlisi predictiu, les virtual appliances i el real-time Business Intelligence*).

Concretament, les solucions *OLAP* permeten oferir i estructurar un gran nombre d'informació de forma resumida, donant la possibilitat de consultar aquestes dades, independentment de coneixements tècnics, ja que no és necessari conèixer sofisticades tècniques d'extracció de dades ni tampoc tenir experiència en el maneig de bases de dades ni en llenguatges *SQL*.

Aquest treball m'ha ajudat a aprofundir sobre metodologies que actualment estan en ús i veure com la sinergia entre el factor humà i el factor tecnològic pot oferir solucions de gran utilitat donant un alt valor a la informació.



# Referències

---

[1] *Pentaho Data Integration*:

<http://community.pentaho.com/projects/data-integration/>

[2] *Saiku Analytics*:

<http://meteorite.bi/products/saiku>

[3] *Álamo Consulting*:

<http://alamoconsulting.com/>

[4] Ministeri de Sanitat, Serveis Socials i Igualtat:

[http://www.ingesa.msssi.gob.es/estadEstudios/documPublica/internet/pdf/Catalogo\\_prueb1.pdf](http://www.ingesa.msssi.gob.es/estadEstudios/documPublica/internet/pdf/Catalogo_prueb1.pdf)

[4] Conesa, J; Curto, J; *Introducción al Business Intelligence*. (Editorial UOC) nombre 163; Barcelona, 2010:

<https://www.amazon.es/Introducción-al-Business-Intelligence-Manuales/dp/8497888863>

[5] Manual de salud electrónica para directivos de servicios y sistemas de salud:

[http://repositorio.cepal.org/bitstream/handle/11362/3023/S2012060\\_es.pdf](http://repositorio.cepal.org/bitstream/handle/11362/3023/S2012060_es.pdf)

[6] Llei Orgànica 15/1999, de 13 de desembre, de Protecció de Dades de Caràcter Personal:

<https://www.boe.es/buscar/doc.php?id=BOE-A-1999-23750>

[7] Pleg de Prescripcions Tècniques *FIC*:

<http://www.plataformadecontractacio.caib.es/DocumentoAdjuntoView?idLicitacion=39966&idTipoDocumento=129>

[8] *Pentaho Big Data – Business Intelligence* España:

<http://www.stratebi.com/pentaho>

[9] Documentació propietat d'Álamo Consulting (informació no pública).

- Totes les referències han estat consultades entre el dia 26 de juny de 2017 i el dia 07 de juliol -



## Agraïments

En primer lloc, agrair a la meva família tot el suport.

A l'equip d'*Álamo Consulting* per donar-me l'oportunitat de treballar com a consultor en pràctiques i per tot el coneixement que he adquirit durant tot aquest temps. Destacar el gran grup humà que forma l'equip d'*Álamo Consulting*.

Al meu tutor de *TFG*, Gabriel Fontanet Nadal, per tots els consells, l'ajuda que m'ha ofert i per la seva predisposició en tot moment perquè aquest treball arribés a bon port.

