Our experience developing a terminology server

"Consumer driven innovation"

26-10-2012

Javier Fernández

ITServer Project manager

j.fernandez@indizen.com
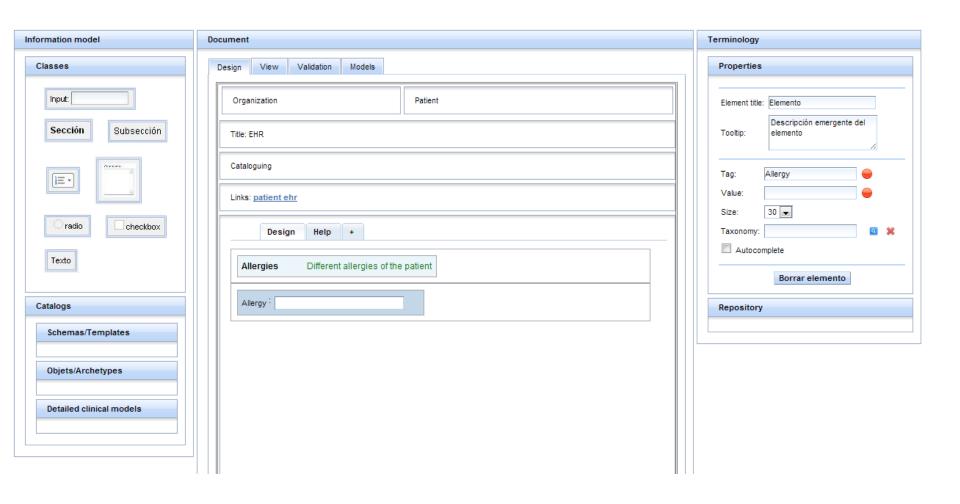
The beginning

Joining together **domain knowledge with technical knowledge**

indizen
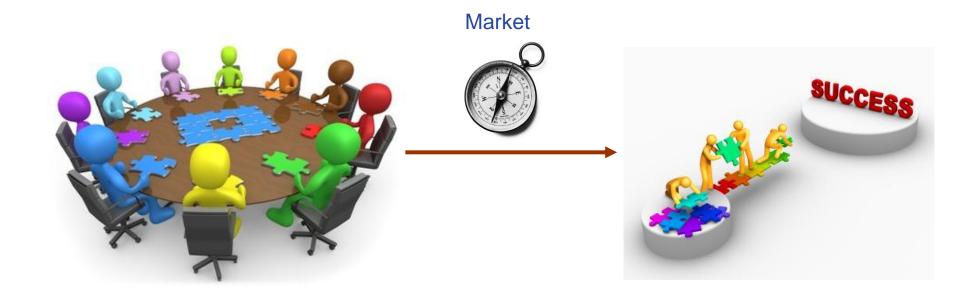
Our initial goal

Our initial goal

Market

indizen

The beginning

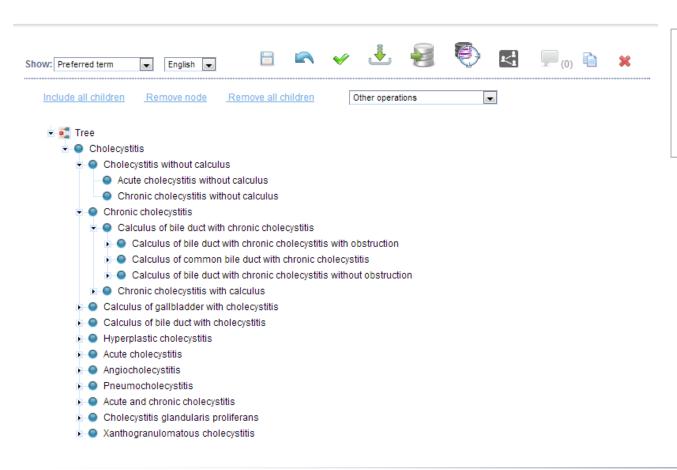1st step → Snomed CT Browser

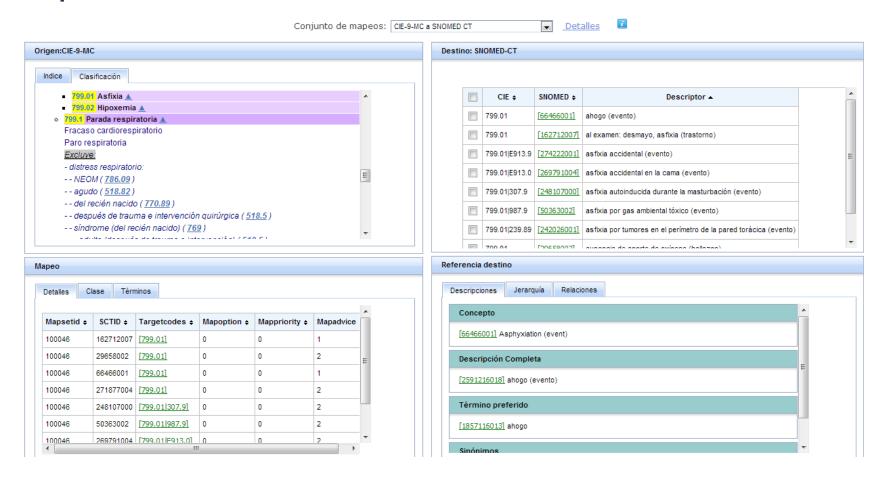# SNOMED CT is too big.Let's create subsets

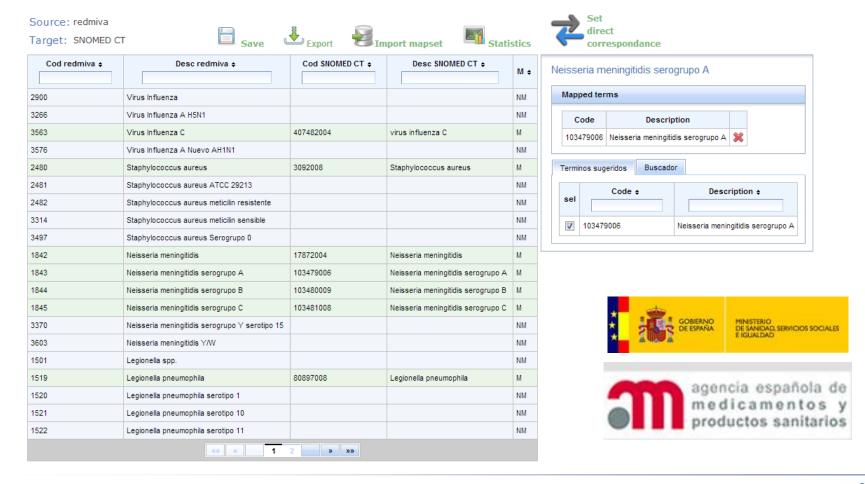# Well…but I want something interoperable

7

Mapset browser

Mapsets. Why not edit it?

9

NLP Toolkit

Language dependent

Techniques:
- Tokenization
- Lemmatization
- Spelling error detection
- Chunking
- Acronym expansion
- …

We need human revision:
- Negation
- Certainty
- Subject of information
- Planning stage

What if I want to create a subset with 50K expressions in natural language?
Solution: Batch processes

Tasks

Active tasks:

Task name:setDescriptionsInCSV   **Begin date:**06 Mar 2012 15:55.04

Finished tasks:

Task name:setSnomedElementsInCSV

Begin date: 09 Oct 2012 15:17.37
End date: 09 Oct 2012 15:17.37

Download file:

Create subset:

Task name:setSnomedElementsInCSV

Begin date: 09 Oct 2012 15:14.04
End date: 09 Oct 2012 15:14.04

Download file:

Create subset:

Services

suggestSNOMEDElements
It's one of the most popular services. It receives a list of descriptions in natural language and returns a list of suggestions in SNOMED CT® for each description ordered by similarity range.

generateRelations
It's a service that receives a sctid list as an input and returns a table in csv format with their definition relationships

generateDescriptions
It's a service that receives a sctid list as an input and returns a table in csv format with the following columns: ConceptId, DescriptionId, FSN, DescriptionType and Term.

generateCrossmapCIEToSNOMED
It's a service that receives an ICD code list as an input and returns the mapping of these codes with SNOMED CT® codes.

generateCrossmapSNOMEDToCIE
It's a service that receives a sctid list as an input and returns the map of these codes with ICD codes.

generateDescendants
It's a service that receives one sctid and returns all the descendants of the code using the SNOMED CT® hierarchy.

How can they use the content that has been created?

-REST services

GetSctDescriptionById: Service that returns all the attributes of a particular description

Parameters:
- descriptionid: Description SctId

xml: rest/core/getSctDescriptionById/1132518011
json: rest/core/getSctDescriptionById/1132518011
jsonp: rest/core/jsonp/getSctDescriptionById/1132518011

GetSctDescriptionsByConceptId: Service that receives a conceptId as an input and return all the attributes of the associated description

Parameters:
- conceptid: Concept SctId

xml: rest/core/getSctDescriptionsByConceptId/127496003
json: rest/core/getSctDescriptionsByConceptId/127496003
jsonp: rest/core/jsonp/getSctDescriptionsByConceptId/127496003

GetSctConceptById: Service that returns all the attributes of a particular concept

Parameters:
- conceptid: Concept SctId

xml: rest/core/getSctConceptById/127496003
json: rest/core/getSctConceptById/127496003
jsonp: rest/core/jsonp/getSctConceptById/127496003

GetSctConceptByDescriptionId: Service that receives a sctid as an input and returns all the attributes of a concept

Parameters:
- descriptionid: Description SctId

xml: rest/core/getSctConceptByDescriptionId/1132518011
json: rest/core/getSctConceptByDescriptionId/1132518011
jsonp: rest/core/jsonp/getSctConceptByDescriptionId/1132518011

We need standards: CTS2

indizen

Current architecture

ITS-WI

| Browser | Subsets | Mapsets | REST Services | Community |
| Extensions | Translation | Statistics | Tools | Administration |

ITS-CORE

ITS Services

ITS DAO

ICD
LOINC
SNOMED CT
Mesh
…

Subsets
Mapsets
Local terminologies

Administration and security

ITS-API

ITS-WS

indizen

Our sandbox- ClinicalDocs

ClinicalDocs - Hospital de Huascarán

Dr Manuel Robles

Q Buscar otro paciente

Nombre: Ignacio Suarez Molina
NHC: 11234567800
Fecha nacimiento: 29/12/1985
Sexo: Hombre

DOCUMENTOS
- Episodio clínico 1234
- Fractura de tobillo

PRUEBAS
Microbiología
Laboratorio
Radiología

Datos Generales del Paciente

| Identificación | Episodios | Alertas 2 | Factores de riesgo 1 | Tratamiento 1 | Pendiente 1 | Más... |

| | Id | Buscar | [+] Nuevo episodio clínico |

Id	SubId	Descripción	Estado	Fecha	Servicio	Operaciones
1234		**Episodio clínico 1234**	**Cerrado**			Operaciones
	2011-1234567	Episodio de atención 2012-1234567	Cerrado	20/02/1999	ORL	Operaciones
	2011-1234568	Otitis media	Cerrado	20/02/1999	ORL	Operaciones
1235		**Fractura de tobillo**	**Abierto**			Operaciones
	2012-1234569	Fractura de tobillo	Cerrado	20/02/1999	Urgencias	Operaciones
	2012-1234570	Fractura de tobillo	Abierto	20/02/2000	Traumatología	Operaciones
	2012-1234571	Fractura de tobillo	Abierto	20/02/2000	Radiología	Operaciones

martes, 16 de octub

indizen

Roadmap

CTS2

New Browsers

RF2

ITServer

Extensions
Module

NLP Toolkit
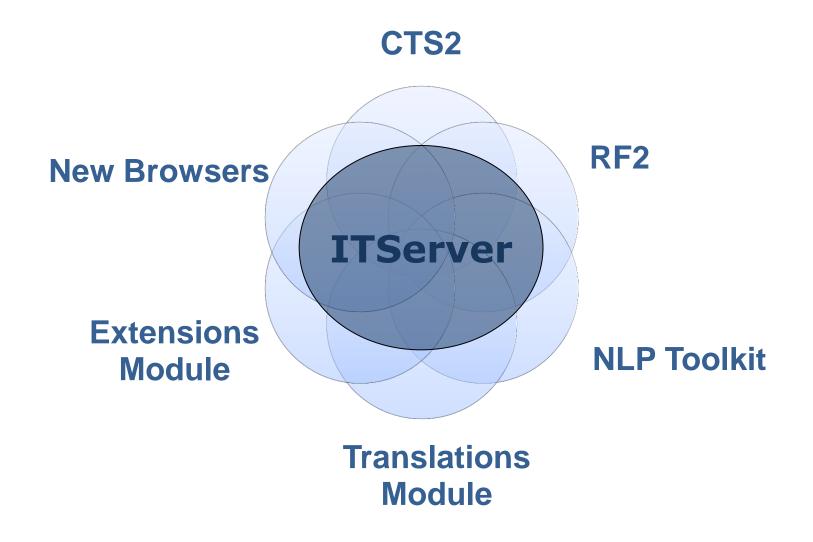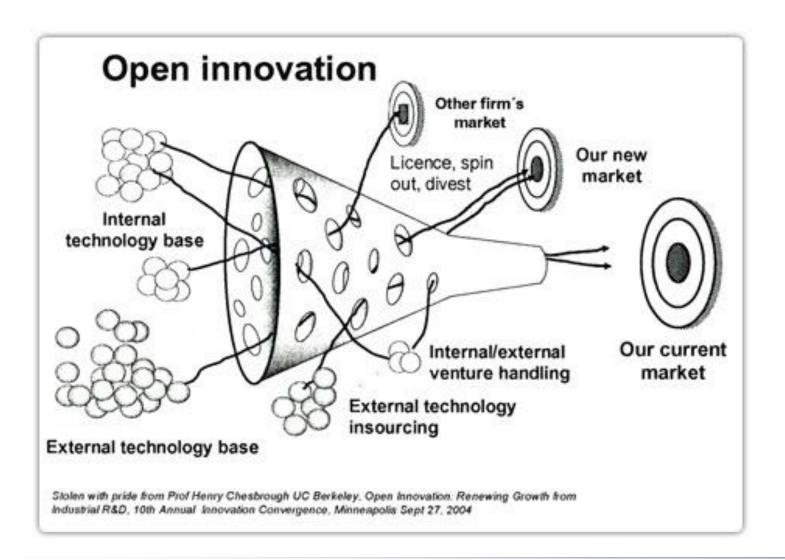
Translations
Module

indizen

Open Innovation

Conclusions

- We have to provide advanced **tools to connect local terminologies and standard terminologies**

- **NLP tools are the key for a smooth transition** between free text and the use of standard terminologies

- The most important part of a terminology server is the services that it provides in order to **guarantee an easy integration for the external systems**

- We need a **scalable environment** that allow us to have fast terminology services (Cloud)

- To create a good terminology server **we need the collaboration of universities and other organizations**

indizen

Questions

indizen

j.fernandez@indizen.com

C/ Tarragona, 30
28045 Madrid
Spain
Tel. +34 91 535 85 68
www.indizen.com

www.itserver.es

indizen