EXTERNAL SCIENTIFIC REPORT

Implementation of Electronic Transmission of Chemical Occurrence Data in Portugal [1]

CFP/EFSA/DATEX/2011/01/02

Francisco Ravasco, Ana Lopes, João Pereira, Sidney Tomé, Fernanda Vilarinho, Ricardo Valente,

João Pedro Nunes, Luísa Oliveira

National Health Institute Doutor Ricardo Jorge, INSA, IP

ABSTRACT

The present report describes the work done in the project Implementation of the Electronic Transmission of Chemical Occurrence Data in Portugal. The objective of the project was to implement and test an electronic system for the transmission of food contaminant data to EFSA and create a national database, according to EFSA standards. This document describes also the challenges encountered during the implementation of the standard model and makes a general analysis on its limitations and potential developments.

KEY WORDS

electronic transmission, chemical occurrence data, xml, data collection framework, standard sample description, food, feed

DISCLAIMER

---

[1] Question No EFSA-Q-2011-01273.

SUMMARY

The collection of analytical data on food and feed is an important task of the European Food Safety Authority - EFSA and an essential component in risk assessment. Following the statements of Regulation (EC) n.º 178/2002 Member States (MS) have the responsibility to report to EFSA data produced under the official control of food and feed. This information reached EFSA under various formats and supports - paper, various digital formats, among others - compromising their use due to the slowness and difficulty of compiling and comparing data. Therefore, and in order to improve the comparability of technical data received and analyzed and facilitate their transmission, EFSA created in 2010, a data model that standardizes the language and the information to report about each sample and each test result designated by "Standard Sample Description for food and feed" – SSD and the "Guidance on Data Exchange", both recommended for use for data transmissions by MS.

Portugal, through INSA, applied to EFSA's call for proposals CFP/EFSA/DATEX/2011/01   with the project "Implementation of Electronic transmission of chemical occurrence data in Portugal", (Grant agreement CFP/EFSA/DATEX/2011/01/02) which had the Food Safety Authority of Ireland (FSAI) as a mentor country.

FSAI past experience on implementing proper software and tools to support the SSD format, was a crucial starting point for executing the project and provided Portuguese authorities an important assistance with standards compliance, database structure and data submissions.

This project that started in December 2011, was implemented in close collaboration with the competent authorities and official national laboratories, and had as main objectives: i) the creation of an informatic system and a national database for the collection and transmission of electronic data from chemical contaminants in food and feed according to the data model SSD, ii) the compilation and transmission of national data produced from 2009 to 2011, and iii) the standardization and improvement of data collection in this area.

This document describes the encoding and mapping strategy for chemical occurrence data produced by the Portuguese authorities, how the national database is populated, the transformations required to support the SSD standard terminology supplied by EFSA and how these transformations will be maintained in the future. It also describes the initiatives to standardize and improve data collection for chemical occurrence in food and feed.

TABLE OF CONTENTS

BACKGROUND

In order to improve the technical comparability of the data received and analyzed, and to simplify and make faster the communication to EFSA, the creation of a data model that uses an uniform language with information as complete as possible has become of vital importance. Given these assumptions, in 2010, EFSA created the data model "Standard Sample Description for food and feed" - SSD and the "Guidance on Data Exchange" mandatory for chemical occurrence data transmission by Member States.

Portugal currently has different National Competent Authorities for operating control plans (DGAV, ASAE and IPMA) and official laboratories (LSA/ASAE, INSA, INIAV and IPMA) carrying out the collection and/or analysis of samples of the official control of food and feed, and specifically the chemical contaminants, and the information generated (sample information and results) are collected and stored in different forms. At the moment of the article 36 call for proposals CFP/EFSA/DATEX/2011/01, no national database with detailed analytical data on food chemical occurrence existed and each authority had different processes for recording samples and laboratory data (three different LIMS: ASAE-LABWAY/INIAV-NAUTILUS/INSA-SIGALIS and IPMA collected data in paper format), and therefore the vocabulary used was different and not always controlled, making it difficult to compile data. The project "Implementation of Electronic Transmission of Chemical Occurrence Data in Portugal" was an opportunity to develop a national database at INSA for the integration of data from the different mentioned systems and of EFSA catalogues, and also to improve the standardization and quality of national collection of chemical occurrence data in food and feed.

TERMS OF REFERENCE

The deliveries of the project "Implementation of Electronic Transmission of Chemical Occurrence Data in Portugal (CFP/EFSA/DATEX/2011/01/02)" have been agreed as follows:

D1. Document D1 – Data Standardization: the encoding and mapping strategy developed and used.

D2. Three Data transmissions providing country's data in electronic format according the standard during the data transmission of October 2012.

D3. Document "D3 - Final report".

This grant was awarded by EFSA to: Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA, IP)

Avenida Padre Cruz, 1649-016 Lisboa, Portugal.

Grant title: Implementation of Electronic Transmission of Chemical Occurrence Data in Portugal

Grant number: CFP/EFSA/DATEX/2011/01/02

INTRODUCTION AND OBJECTIVES

Up to 2012, no national database with detailed analytical data on food chemical occurrence existed. The implementation of this project was the first step for the creation of an automated electronic data collection system in Portugal and the gained experience and challenges identified and hopefully overcome will allow the country to improve its data management capacity. The existence of such an information resource with standardized data will provide data users (EFSA, national risk assessors and risk managers, and others) with an essential work tool.

Portugal has currently various national authorities who plan and execute different official control plans and several official laboratories that perform analyzes of samples taken on the official control of food and feed.

The official food control in Portugal is currently coordinated by the Food and Veterinary Directorate-general (DGAV) from the Ministry of Agriculture, Sea, Environment and Territorial Planning (MAMAOT). The official control is based on a Multiannual Integrated National Control Plan that comprises 36 sectoral plans including targeted and random sampling responding to European Legislation requirements and is run by several National Competent Authorities belonging either to MAMAOT or to the Ministry of Economy and Employment (MEE). The sectoral control plans that include chemical analysis are run by the:

- Food Safety and Economic Authority/Autoridade Segurança Alimentar e Económica (ASAE/MEE);

- Food and Veterinary Directorate-General/Direcção Geral de Alimentação e Veterinária (DGAV/MAMOT);

- Portuguese Institute for Ocean and Atmosphere/Instituto Português do Mar e da Atmosfera (IPMA/MAMAOT);

The chemical analysis are performed essentially by Official Laboratories, where the detailed analytical information required for data transmission to EFSA reside in sources like LIMS, Access databases, Excel files or paper reports.

The flowchart below has a summary description of the entities involved in the official control programs.
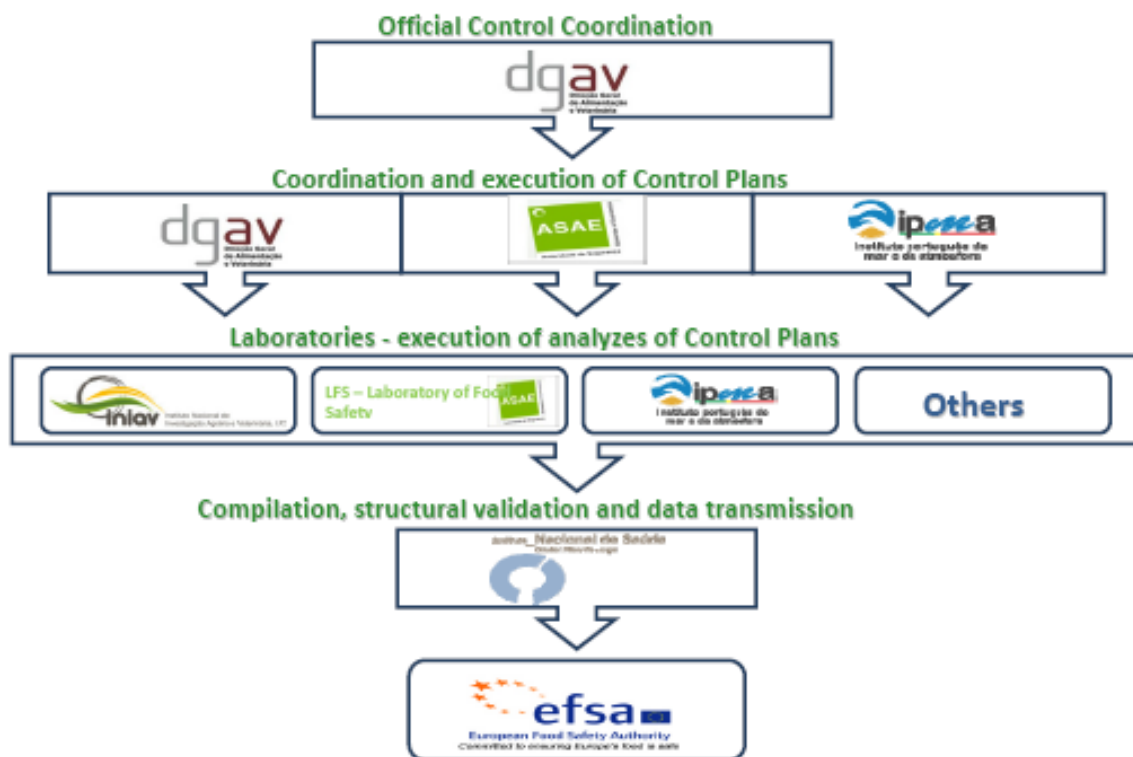
Figure 1:   Flowchart of the information circuit on official control of food chemical contaminants

The Electronic Transmission project main objective was the creation and maintenance of a national database, which compiles data concerning the plans of official control performed in Portugal for the occurrence of chemical contamination in food and feed. It aimed also to collect all the national data (official checks) existed from 2009 until 2011 and to develop the existing LIMS' at the national authorities so that in the future they include SSD data fields previously missing, hence improving the quality of national data.

For the development of the project it was necessary to create, implement and test an electronic system for the collection of data from different national sources, transforming them in the Standard Sample Description (SSD) format using also the food classification and description system FoodEx and proceed to validation and subsequent transmission to EFSA in XML format. The transmission process focuses in the first instance on chemical contaminants data, but it has to be flexible enough to support other types of data that can fit into the same Standard Model.

The specific objectives of this project were:

➢ To develop a national database according to SSD

➢ To map the different national systems to SSD format and controlled vocabularies

➢ To develop INSA's LIMS to include the missing SSD data elements for future data

➢ To develop translation tools

➢ To populate the national database with national historical data since 2009

➢ To transmit data from national database to EFSA in XML file through EFSA data collection framework (DCF) according to the EFSA Guidance on data exchange (EFSA Journal 2010; 8(11):1895)

➢ To improve the quality of future national data collection (guidelines and training to sampling officers, training on SSD, development of other national LIMS to include the missing SSD data elements where possible)

MATERIALS AND METHODS

MATERIALS

1.   Documents used for the project

To perform its work INSA used the harmonized documents established by EFSA and that are listed below:

✍ **Standard Sample Description for food and feed (SSD)** - available online: http://www.efsa.europa.eu/en/datex/datexsubmitdata.htm.

✍ **Generic Reporting Format (GRF)** - available online: http://www.efsa.europa.eu/en/datex/datexsubmitdata.htm.

✍ **The Guidance on Data Exchange** - available on: http://www.efsa.europa.eu/en/datex/datexsubmitdata.htm.

✎ **The EFSA food classification system FoodEx** - available online: http://www.efsa.europa.eu/en/datex201001/docs/cfpefsadatex201001ax1.xls.

✎ **Call for continuous collection of chemical contaminants occurrence data in food and feed** - available online: http://www.efsa.europa.eu/en/art36grants/article36/cfpefsadatex201101.htm.

2. Software tools used for the database development

For the system's database we have decided to use the Microsoft SQL Server 2008 R2. This relational database management system has an easy method of importing the data related to the SSD format. This data was present on an Excel file sent by EFSA, namely the *StandardSampleDescription* file. Each sheet inside that file has been mapped to a table through the SQL Server Import and Export Wizard, resulting on a database server that contained all the necessary information for the project and that could be easily accessed by our web platform that is being developed.

In order to develop the proposed web application, Microsoft Visual Studio 2010 was used as the preferable IDE, since it would be easily integrated with the remaining systems and tools, along with the ASP.Net technology.

To create the XML files, the tool XmlSampleGenerator was used. This tool generates an XML document based on an XML Schema (XSD) file. The *StandardSampleDescription.1.0.xsd* XSD file provided by EFSA was used to generate the XML.

METHODS

The project was executed in two areas in parallel, data collection and software development, followed by data validation and transmission to EFSA.

1. DATA STANDARDIZATION

The initial phase of the project consisted in the verification of the available information in each of the competent authority LIMS and its correspondence with the SSD system. Mapping between the controlled vocabularies used in each entity and the SSD codes also performed.

2. COLLECTION OF HISTORICAL DATA

a. Survey of existing data

The following stage of the project consisted of a survey of the national scene, among the authorities carrying out official controls for food and feed, with regard to existing data concerning chemical contamination since 2009.

b. Data collection and mapping to SSD

The collection of the historical data since 2009 was performed presentially on site during 2012 and the beginning of 2013 and the mapping was done manually on the MS Excel® Generic Reporting Format (GRF), provided by EFSA, and accordingly to the SSD controlled vocabularies. We emphasize that the information gathered refers to the data provided by the different institutions mentioned previously, so we cannot ensure that they correspond to the totally of the samples analyzed in the country concerning the official control plans.

3. SOFTWARE DEVELOPMENT

With the purpose of having a solution that could achieve the goals and the objectives previously described, it was established that a Web Platform should be developed. This platform would handle the data input from the several Competent Authorities and perform the SSD mapping defined by EFSA. It would also provide numerous amounts of tools to help every step of the mapping process.

4. DATA VALIDATION AND TRANSMISSION TO EFSA

The information collected manually on the GRF was validated by the software previously described accordingly to the EFSA validation rules and the data was transmitted to EFSA as XML through the Data Collection Framework (DCF).

5. QUALITY IMPROVEMENT OF FUTURE DATA COLLECTION

a. Sampling Data Collection Form

In order to improve the quality of national data and meet specific needs of each one of the national competent authorities and the information requested by EFSA, a draft of a harmonized sampling collection form based on FSAI's – "Food Safety Laboratory Service - Sample Request Form" was created.

b. Training of Sampling Officers

Initially it was thought to develop a training manual for the use of the Data Collection Form and the preparation of training sessions for its use by the sampling officers and data managers whenever the document was fully accepted by all official control entities and ready to be implemented in the control programs.

RESULTS

1. DATA STANDARDIZATION

A summary of the data received from the different data sources and how it can be mapped to the EFSA SSD elements is presented in table 1. This summary takes special account of the mandatory elements. In order to meet SSD requirements, we needed to add additional information which applies to all samples in each data set (e.g. fields in the table that are constant/hardcoded or derived). Table 1 also shows which data elements in the LIMS used by three entities collaborators of the project (ASAE, INIAV and INSA) have controlled vocabularies.

Table 1: Information received from data sources

EFSA			INIAV			ASAE			IPMA			INSA		
Element Code	Element Name	Mandatory	Received in Extract	Element label	Controlled terminology	Received inExtract	Element Name	Controlled terminology	Received in Extract	Element label	Controlled terminology	Received in Extract	Element Name	Controlled terminology
L.2	labAccred	Yes	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
L.3	labCountry	No	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
O.1	localOrg	No	LIMS	sample_user_u_requisitado_por	no	LIMS	Cliente	yes						
O.2	localOrgCountry	No	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
R.01	resultCode	Yes	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
R.02, R.03, R.04	analysisY analysisM analysisD	Yes No No	LIMS	sample_authorised_on	yes	LIMS	DataResultado	no	PAPER			LIMS	DTFIMANALISES	
R.05	EFSAParamCode	No												
R.06	paramCode	Yes	LIMS	test_description	yes	LIMS	Cód. Método	yes	PAPER			LIMS		
R.07	paramText	No				LIMS	Parâmetro	yes				LIMS	PARATXT	
R.08	paramType	Yes				LIMS			PAPER			LIMS		
R.09	anMethRefCode	No				LIMS	Descrição	yes	PAPER			LIMS		
R.10	anMethCode	No (Yes - for some parameters)	LIMS	test_description	yes	LIMS	Técnica	yes	PAPER			LIMS		
R.11	anMethText	No										LIMS	METODEXT	
R.12	accredProc	No				LIMS	Parâmetro Acred.	yes	PAPER			LIMS		
R.13	resUnit	Yes/No	LIMS		yes	LIMS	Unidade	yes	PAPER			LIMS		
R.14	resLOD	No	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
R.15	resLOQ	No	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
R.16	CCalpha	No												
R.17	CCbeta	No												
R.18	resVal	No				LIMS	Res.Trat.		PAPER			LIMS	RESULTADOTRATADO	
R.19	resValRec	No				LIMS								
R.20	resValRecCorr	No				LIMS						LIMS		
R.21	resValUncertSD	No												
R.21	resValUncertSD	No												
R.22	resValUncert	No												
R.23	moistPerc	No												
R.24	fatPerc	Yes/No (for some parameters)				LIMS								
R.25	exprRes	Yes	LIMS			LIMS			PAPER			LIMS		
R.26	resQualValue	No				LIMS								
R.27	resType	Yes				LIMS			PAPER			LIMS		
R.28	resLegalLimit	No										LIMS	APRADM	
R.29	resLegalLimitType	No												
R.30	resEvaluation	Yes/No				EXCEL	Parecer Técnico	yes						
R.31	actTakenCode	No												
R.32	resComm	No				EXCEL	Parecer Técnico	yes	PAPER				OBSERVACOES	

EFSA			INIAV			ASAE			IPMA			INSA		
Element Code	Element Name	Mandatory	Received in Extract	Element label	Controlled terminology	Received inExtract	Element Name	Controlled terminology	Received in Extract	Element label	Controlled terminology	Received in Extract	Element Name	Controlled terminology
L.01	labCode	No	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
S.01	labSampCode	Yes	EXCEL	sample_name	yes	LIMS	Amostra	yes	PAPER			LIMS	NRAMOSTR	
S.02	labSubSampCode	No												
S.03	lang	Yes	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
S.04	sampCountry	Yes	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
S.05	sampArea	No				PAPER/EXCEL	Local de Colheita	no	PAPER					
S.06	origCountry	Yes				PAPER/EXCEL	Origem	no						
S.07	origArea	No				PAPER/EXCEL	Local de Colheita/ Origem	no	PAPER					
S.08	origFishAreaCode	No				PAPER	Origem	no	PAPER					
S.09	origFishAreaText	No				PAPER	Origem	no	PAPER					
S.10	procCountry	No				PAPER/EXCEL	Produtor (animal/ vegetables products)	no						
S.11	procArea	No				PAPER/EXCEL	Produtor (animal/ vegetables products)	no						
S.12	**EFSAProdCode**	Yes	LIMS	sample_sample_type	yes	LIMS/EXCEL	Sub-grupos de Géneros Alimentícios	yes	PAPER			LIMS		yes
S.14	prodText	Yes	LIMS	sample_user_u_product_description	no	PAPER/EXCEL	Identificação do produto	no	PAPER			LIMS	AMOSDESC	
S.15	prodProdMeth	No							PAPER					
S.16	prodPack	No				PAPER	Forma de embalamento	no	PAPER			LIMS	ACONDICIONAMENTO	
S.17	prodTreat	Yes				PAPER/LIMS/ EXCEL	Sub-grupos de Géneros Alimentícios	yes	PAPER			LIMS		
S.18	prodBrandName	No				PAPER/EXCEL	Designação / Marca	no						
S.19	prodManuf	No				PAPER/EXCEL	Produtor (animal/ vegetables products)	no						
S.20	prodIngred	No												
S.21	prodCom	Yes/No (for some parameters)										LIMS	OBSERVACOES	
S.22	prodY	No							PAPER			LIMS		
S.23	prodM	No												
S.24	prodD	No												
S.25, S.26, S.27	expiryY	No				PAPER/EXCEL	Validade	no						
S.28, S29, S.30	sampY	Yes	LIMS	sample_sampled_on	yes	PAPER/EXCEL	Data Colheita	no	PAPER			LIMS	DTCOLHEITA	
S.31	progCode	No	LIMS	sample_user_u_teste_reason		PAPER	no	no						
S.32	progLegalRef	No				PAPER	no	no	PAPER					
S.33	progSampStrategy	Yes	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
S.34	progType	Yes	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
S.35	sampMethod	Yes	Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database			Value automatically filled in the future Database		
S.36	sampleNum	No												
S.37	lotSize	No												
S.38	lotSizeUnit	No												
S.39	sampPoint	Yes				PAPER/EXCEL	Local de Colheita	no	PAPER			LIMS		

In order to maximize the future automation of the data collecting system, mapping tables (already submitted to EFSA with the Document D1 – Data Standardization: the encoding and mapping strategy developed and used) between the controlled vocabularies used in the entities mentioned above and the controlled vocabularies used in the SSD system were made. These mapping tables will be included in the software, currently in the final development stage, and INSA will be responsible for their

maintenance. This will allow the country to continue using and updating the terminologies employed on each data source and correspond them to the EFSA controlled terminology values.

2. COLLECTION OF HISTORICAL DATA

a. Survey of existing data

The information collected from ASAE refers to food in general which is analysed for numerous chemical substances such as mycotoxins, metals, inorganic elements and derivates. This information was available in different formats. Regarding the year of 2009, part of the information was stored in paper format and other part in Excel files. From July 2011 laboratory result information is recorded in a LIMS (LABWAY) and the sample information is reported into a Database (GESTASAE) and in an Excel file. Some of the LIMS vocabularies (parameters, matrix, analytical methods, sample number, etc.) are controlled, and conversion to the vocabulary used by EFSA, with the help of the translation mapping tables previously developed, will be relatively simple. The information in GESTASAE and LABWAY can be linked through a code (Reference PNCA). Since the GESTASAE system is still under development, it is not possible to extract information regarding sample collection, and because of that all this information is recorded manually in an Excel file (without controlled vocabularies), leading to several mistakes, slowness, work duplication, which complicates automatic data compilation.

The information collected from INIAV refers to feed and food of animal origin which is analysed for numerous chemical substances such as mycotoxins, metals, inorganic elements and derivates, veterinary drugs and pesticides residues. This information was recorded in a LIMS (NAUTILUS), from 2009. Although the system has some controlled vocabularies, they seem to be very simplified (the field that represents sample description is very generic). FoodEx coding will require consideration of several data fields to ensure it is as specific as possible.

The information collected from IPMA refers to seafood in general which is analysed for various chemical substances such as marine biotoxins, metals and inorganic elements. This information was stored in paper format. Until this time, they do not have any electronic system for recording samples and laboratory analysis data, nor do they have any type of controlled vocabulary.

The information collected from INSA refers to food for infants and small children and rice and rice derived products which are analysed for different chemical substances such as mycotoxins, metals, inorganic elements and derivates. This information was recorded into a LIMS (SIGALIS), from 2009.

Although the system has some controlled vocabularies, they seem to be very simplified and not completely compatible with SSD.

b. Data collection and mapping to SSD

IPMA had all the information on paper, therefore this collection was made directly from the analysis reports and the laboratory records book.

As for ASAE, since part of the information was stored in an Excel file and other on paper or in the LIMS, the collation of all this information was very difficult and because of that it was decided that the data collection should be easier if performed from the analytical reports/processes which had also the sample collection form used on the sampling procedure. The data collection was made directly on the GRF and mapping to EFSA controlled terminologies was made at the same time, which resulted in achieving more detailed information.

With regard to INIAV the information was received in Excel format (exported from Nautilus) and the mapping was done manually directly on the GRF.

Regarding INSA, the information was received in Excel format and the mapping was done manually directly on the GRF.

A particular note has to be made regarding the information available in each data sources, which was not sufficient to complete the data requested by EFSA, although all the mandatory fields were able to be fulfilled according to EFSA's SSD.

For the period to which the project relates (2009 to 2011), information on more than 31000 analytical determinations was collected as shown in table 2.

Table 2:   Number of determinations and analytical parameters/year

Data collection groups	2009		2010		2011		2012[(*)]	
	Determ.	Param.	Determ.	Param.	Determ.	Param.	Determ.	Param.
Group 1	896	20	779	19	975	19	1260	19
Group 2	188	3	0	0	11	1	0	0
Group 3	7696	15	5638	14	4214	15	2636	8
Group 4	5405	22	2351	8	3361	21	1465	8
Total	14185	60	8768	41	8561	56	5361	35

(*) Data submitted through DCF, valid but not yet accepted by EFSA.

Even though it is possible for the systems administrators to introduce new fields/elements in the LIMS' used by some entities, there are some particular situations in which the supplier has to be involved, and that might cause some difficulties in the development of the project, given the costs associated with these changes. In INSA's LIMS, some additional new fields from the SSD were already included, in order to meet the information required by EFSA. Nevertheless it is not possible to include the totality of terms of some controlled vocabularies e.g. (Packaging – S.16 and Product treatment – S.17) due to system limitations. In these cases the most commonly used terms were chosen and included.

Some LIMS also allow the introduction of new terms in the controlled terminology, which may adversely affect the uniformity and the automatic encoding of data according to the SSD, if not used properly. Whenever the controlled terminology is updated on LIMS' the same operation should be performed on the database, this could lead to constant updating. This point will have to have special attention from the systems administrators in order to improve standardization and conformity with the SSD.

3. SOFTWARE DEVELOPMENT

In order to build an accessible and dynamic platform, different types of technologies and resources had to be used since the aim was to provide a solution on a different kind of environment. All the methods and materials used will be described in the following subsections, for a better understanding of the approach taken for this kind of problem.

3.1. System and Tools

3.1.1. Database: Microsoft SQL Server 2008 R2.

Although it may seem that the process of importing the file's data was simple, it actually required some steps before the whole information was persisted in the database. Unfortunately, it wasn't possible to instruct SQL Server to import each sheet and create a table for each one of them. SQL Server would only consider that the main information is located on the first sheet of an Excel file so, before importing, we had to take the data inside each sheet and save it to a separate Excel file. When this step of separating the sheets was finished, we then proceeded to import each one of them with the help of the wizard.

After analysing the information located in the several tables and comparing it with the original information in the Excel sheets, we noticed that there were some inconsistencies between some of the entries. Several cells were empty, forced to be filled in with the NULL value when they were supposed to have something. Some other cells had characters (symbols and letters) between their own values, which would possibly affect the application in a negative way when trying to perform operations with them. An additional step was taken and it consisted on cleaning the Excel's file sheets.

Finally, once all of the previous steps were finished, we had the complete SSD data properly stored in the database. Figure 2 tries to better illustrate what was described about the process of creating the database.
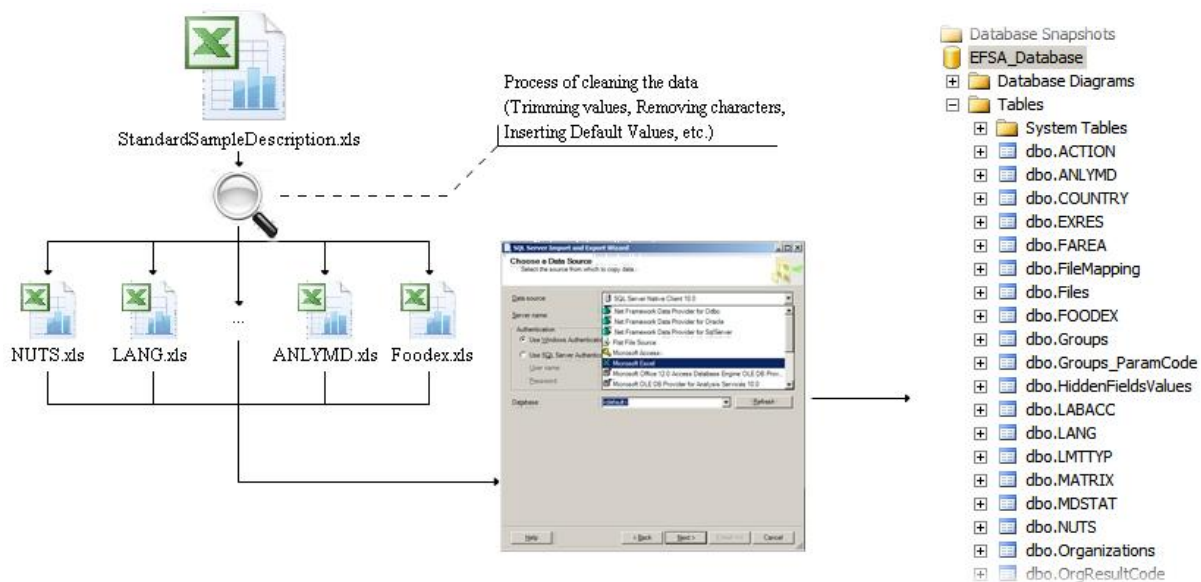
Figure 1: Diagram showing the process of storing the SSD data into the Database.

3.1.2. Development Tools: Microsoft Visual Studio 2010

As for the revision, version and source control of the code produced by each member of the development team, as well as for the overall project tracking, the Microsoft's Team Foundation Server was used.

With the purpose of testing the application behaviour as it grew, and also in the most realistic scenario possible, we installed the web server software named Internet Information Server (IIS) on a machine with the Windows Server 2008 R2 Standard operating system. Before concluding each stage of the project, the application source would be deployed on the server machine for testing, before initiating the next stage.

3.2. Implementation

3.2.1. System Design

In this chapter is presented the design of the developed system. Figure 3 shows the main processes of the system "alimentos PT.ON.DATA". The first process consists in the extraction of data from laboratories to files Microsoft® Excel®. The next process is the import of Excel files with laboratory data. After import files it is necessary to map each column of the Excel datasheet to the respective SSD fields. After performing the mapping process, the data is stored in the database and the process of validation starts. Sample records that successfully pass the validation process are stored in "SSD" table of the database. Records of samples that contain invalid or missing data are stored in the "Quarantine" table. In the following process the data stored in the "SSD" table are exported to XML files according to the XML schema (XSD). Finally the files are transmitted to EFSA.
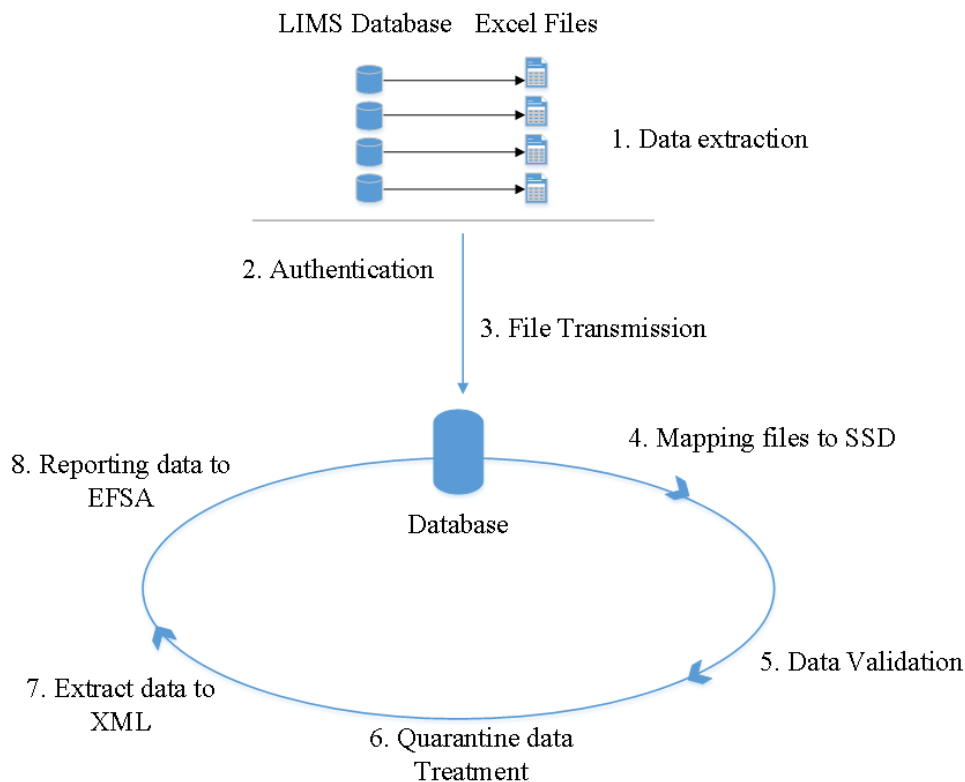
Figure 2: Graphical representation of system processes "alimentos PT.ON.DATA"

The various system services are available to users through a web platform. The web platform is a client-server architecture. The access to the platform services are conditioned on the following user profiles: Administrator, Supervisor and Accountable_CA.

Figure 4 illustrates the overall system architecture and shows the various modules that each user profile has access. In the following chapters will be approached in detail the system's modules.
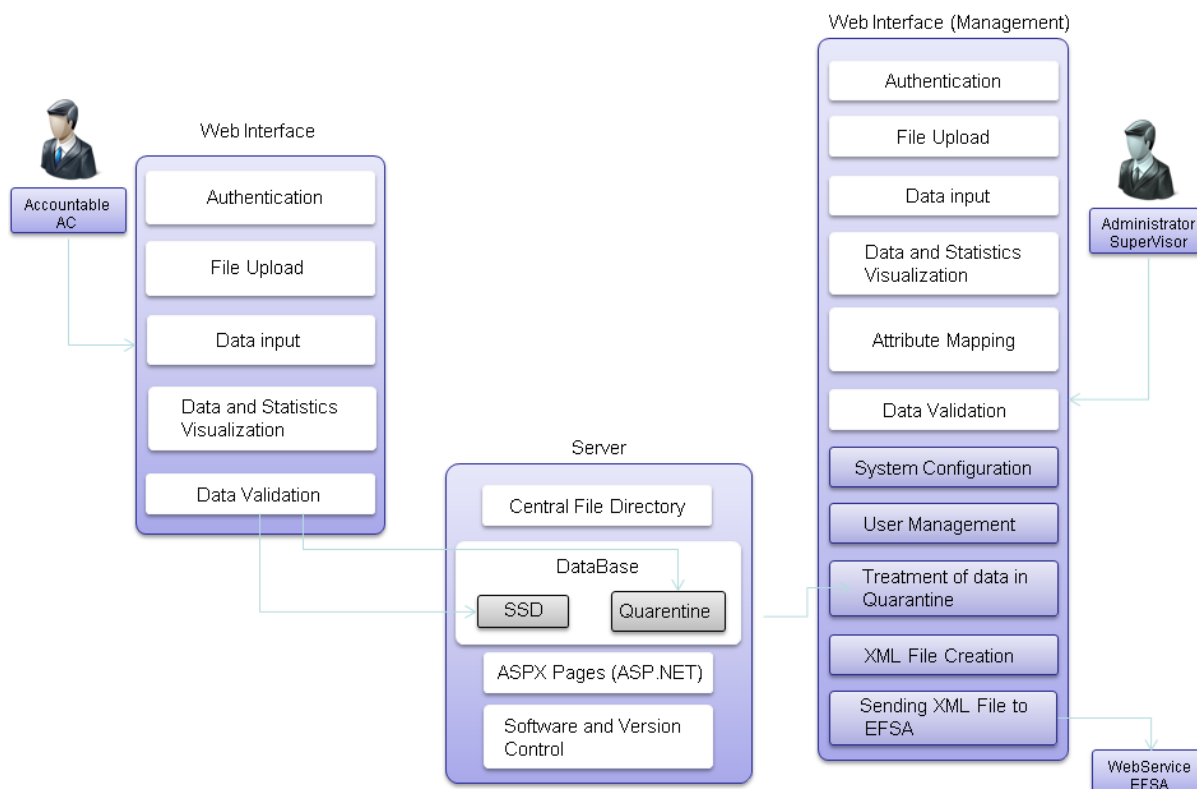
Figure 3: Overall system architecture "alimentos PT·ON·DATA"

3.2.2. User Profiles

Various types of profiles are defined so it is possible to have a greater control and safety whenever accessing the web platform. The aim is to assign different profiles to different users, since some of them have different functions and permissions, and therefore, it was necessary to adapt the platform to the respective user. This section describes the features to which each type of profile will have access.

3.2.2.1. Administrator

This user profile has access to all system features; however it has the main purpose of making global setting changes to the system, create new profiles and perform user management.

3.2.2.2. Supervisor

This user profile has the main function of overseeing and control the whole process of collecting and processing information from the several Competent Authorities (CAs), and sending data to EFSA.

- ✓ Perform the mapping of files that were delivered by the Competent Authorities;

- ✓ Processing all information that is under quarantine;

- ✓ Validation, editing or removal data;

- ✓ Creation of XML files with the information contained in the SSD table;

- ✓ Submission of XML files to EFSA.

3.2.2.3. Accountable CA

This user profile has the main role upload files containing the analysis results of the samples taken by the Competent Authority. In each institution there shall be a person responsible for this activity. After a Supervisor finish the mapping of these files, the Accountable CA is responsible for conducting a first validation of the mapped data.

Valid data will go to the SSD table and incorrect data will be put in the "quarantine" table. Any doubt that the Supervisor has in relation to any data, he will be able to clarify it with Accountable CA, who has submitted the respective data.

- ✓ Uploading files to the System;

- ✓ Validation, editing or removal of certain data.

3.2.2.4. Guest

Users with Guest profile can only view data and statistics that the Supervisor or system Administrator have made available.

3.2.3. Authentication and User Management

In order to users make use of the developed web platform, it is necessary to enter a username and password on the homepage authentication. Once identified and authenticated, and according to their access privileges, users will have at their disposal all the system tools available.

The user management tool is only accessible to system administrators. The use of the platform becomes more secure using this kind of system access. After authenticating a user with an Administrator account, can navigate to the User Management page, located under platform Tools, to create or edit data for a particular user.
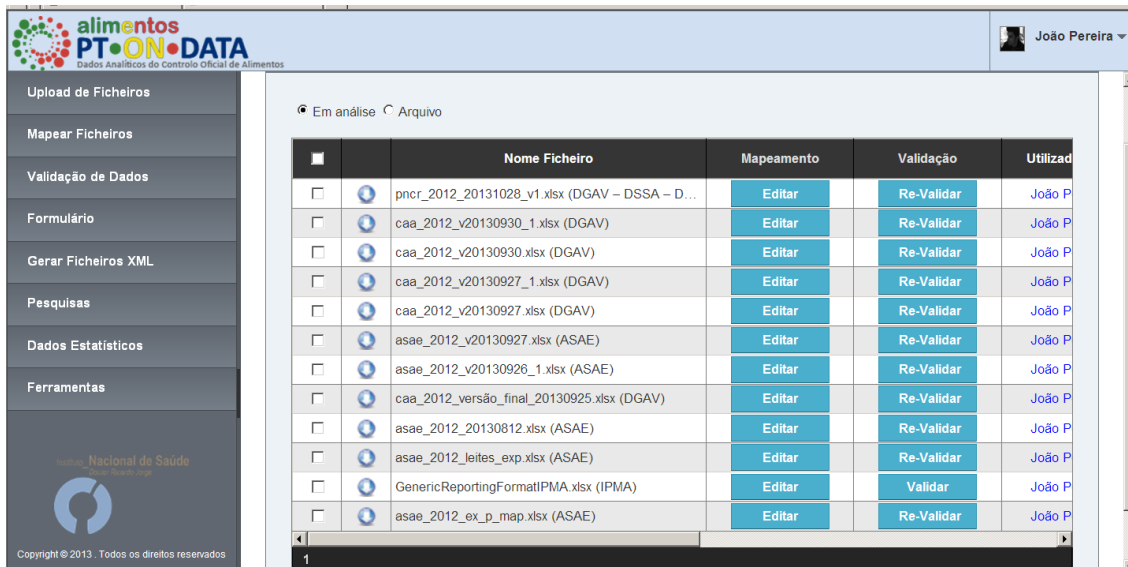
Figure 4: "Gestão de Utilizadores" Window

As shown in Figure 5, it is possible to associate a large set of information to the created user, such as a photo ID, contact details or additional notes. However, the most relevant information in addition to the username and password is the one that identifies the entity that should be associated with the user created. As the user, the entity itself (CA) can be created below this tool, and it will be important to associate a specific user to an entity, so that management of the uploaded data by each entity can be carried out and subsequently sent to EFSA.

3.2.4. File Upload

There are two ways for inserting data into the system: by uploading data files with the sample results or via the web form by manual entry. On the "Upload File" page (Figure 6) the user can upload files in Microsoft® Excel® format (extension xls or xlsx). This page is available to all registered users of the system. To read the files contents, was used the Application Programming Interface (API) OLE DB® developed by Microsoft. This API is constituted by a set of interfaces Component Object Model (COM) which allows universal access to various data sources.

Figure 5: Web Platform: page "Upload de Ficheiros"

The files must have the following requirements to be accepted by the system:

✓ Excel files (since all LIMS used by CAs can export data to this format - XLS or XLSX extension);

✓ At least one column with data;

✓ At least one row with data;

✓ One sheet named "report" (once an excel file can have multiple sheets, it is necessary to indicate which sheet containing the data to store in the database).

If a file doesn't meets all these requirements it is rejected by the system. In case the file is accepted, it is stored in the database with associated information such as the hash value. The hash value is used for example to check / prevent duplicate files from being entered into the database.

After the file has been saved in the database, the file and data are visible only to the Competent Authority that performed the upload, ensuring confidentiality between the various CAs. Only users with Supervisor or Administrator profiles will see the data from all authorities.

The next step is to map the columns of the excel file with the SSD fields. This mapping is always performed by the Supervisor who has to associate each column of the excel file with the corresponding SSD field (figure 7).

For example: "Column1" ⟶ "labSampCode"

"Column2" ⟶ "labSubSampCode"
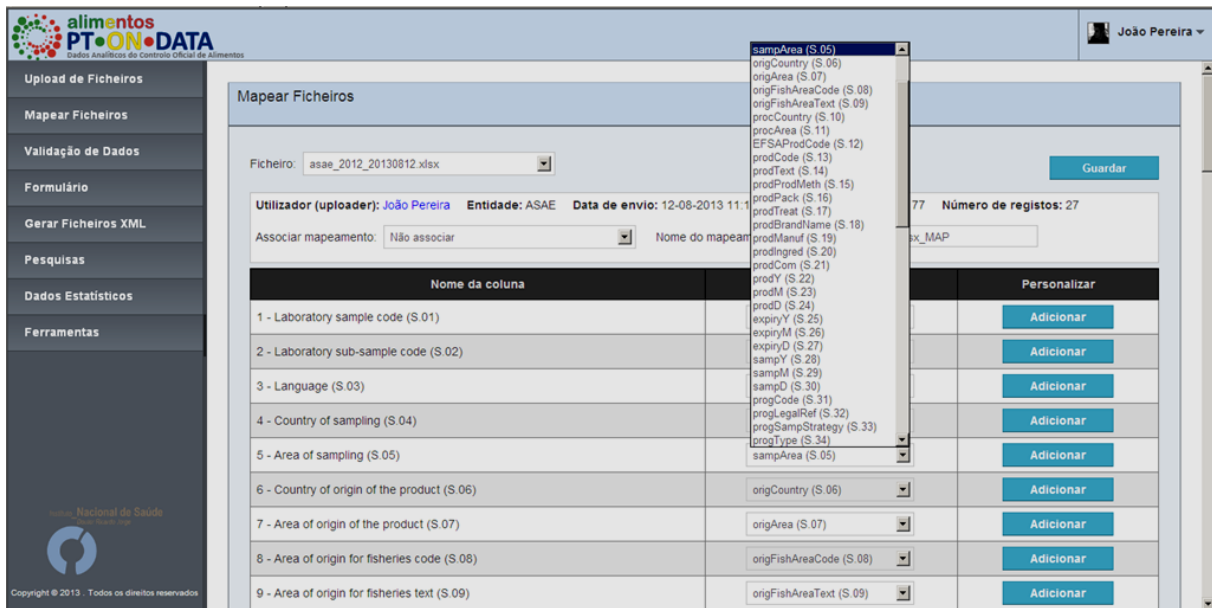
Figure 6: Screenshot of "Mapear Ficheiros" between excel columns and SSD fields

It is also possible to associate an excel column to multiple SSD fields (figure 8). This is particularly useful when the date is in a single column and it is necessary to match the day, month and year with SSD fields.
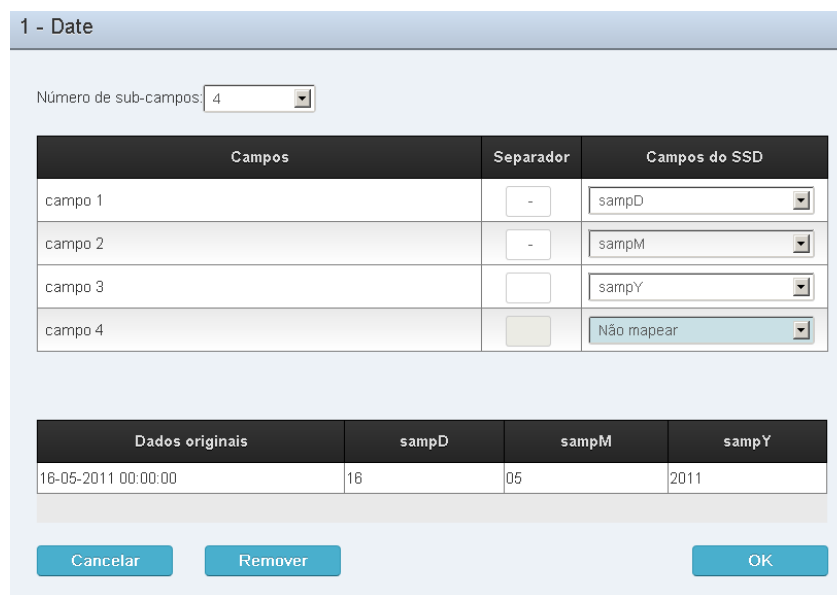
Figure 7: Screenshot of Custom Mapping

If the system recognizes that the file to map is identical to one already mapped, it suggests an automatic mapping, which is always validated by the Supervisor.

Finally the mapping information is stored in the database in XML format.

3.2.5. Web Form

The developed platform was meant to be used by any Portuguese Food Control Authority. Some of these Authorities had their own LIMS, where they could select the data they wanted to report for a certain year, export it in the Excel format, and finally import the Excel file into our platform in order to have it transformed into the SSD format. However, some of these Authorities are still managing their data using physical paper documents. For them, the whole process of sending the information through the platform would be more complicated compared to the others Authorities.

For this case, and as a main objective to reduce the human error affecting the data, we had decided to develop a web form within the platform.

This web form is divided into four different steps, where each one of these steps is intended to represent each step that a sample actually goes through, i.e. the first step in the form is the Generic Data step and every field contained in it are related to information that are common between all the samples, for example, there will be shown the field *Country of Sampling,* since we are working with Portuguese Authorities and, for every sample, the country value will be the same; the second step is the Sample Acquisition Data step where all information about process of harvesting/acquisition should be inputted; next there is the Laboratory Data step where all the information related to the procedures and results obtained when analysing the sample so be introduced; lastly there is the step, titled Overall Sample Assessment, where some of the most important information previously introduced can be reviewed.

The main focus when creating this form was to make it user friendly, because anyone that would use it would have to handle and understand a large quantity of data, due to the fact that there is a lot of variables used for the SSD standard. We managed to make it simple to use thanks to the wide range of controls that the ASP.Net framework possesses.

For every field in the form that was supposed to have some SSD controlled language, we had that same language loaded from the server database and displayed to the user. This helps the user to fill in more quickly and avoids any types of errors and mistakes. We also helped the process of creating

entries on the form by reducing the number of fields that the user had to interact with. Knowing that most of the 76 variables of the SSD have the same information in them for the same sample, only differentiating at the variables regarding to the laboratory results, we came up with a method of associating multiple results for the same sample. Figure 9 helps explaining better the problem and the solution found.
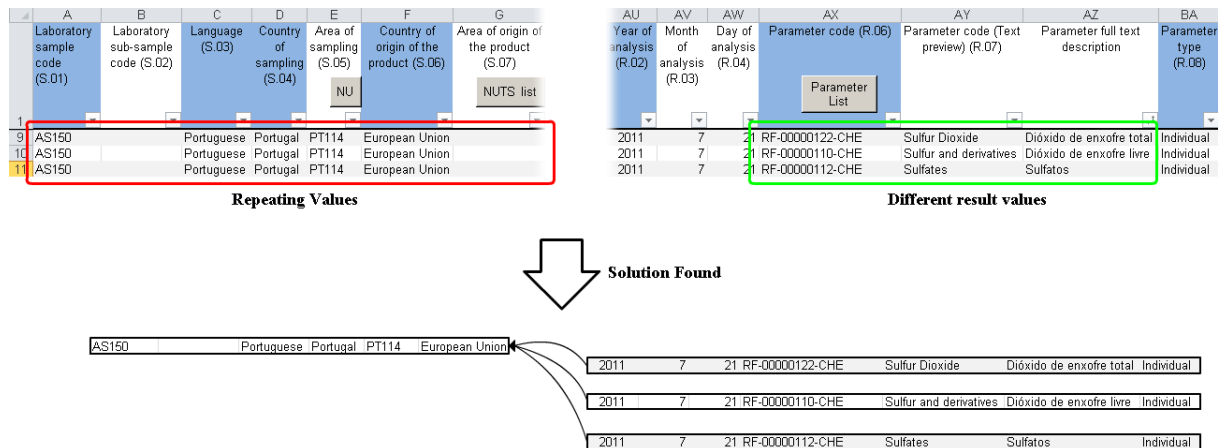
Figure 8: Problem and Solution found for the repeating values for the same Sample

To implement the above solution we made use of the GridView control and placed it on the third step of the form. Each time the user filled all the fields in that step and clicked on the *Add* button, a new result would be added to the sample information present on the previous steps, removing all the work of having to fill every individual entry.
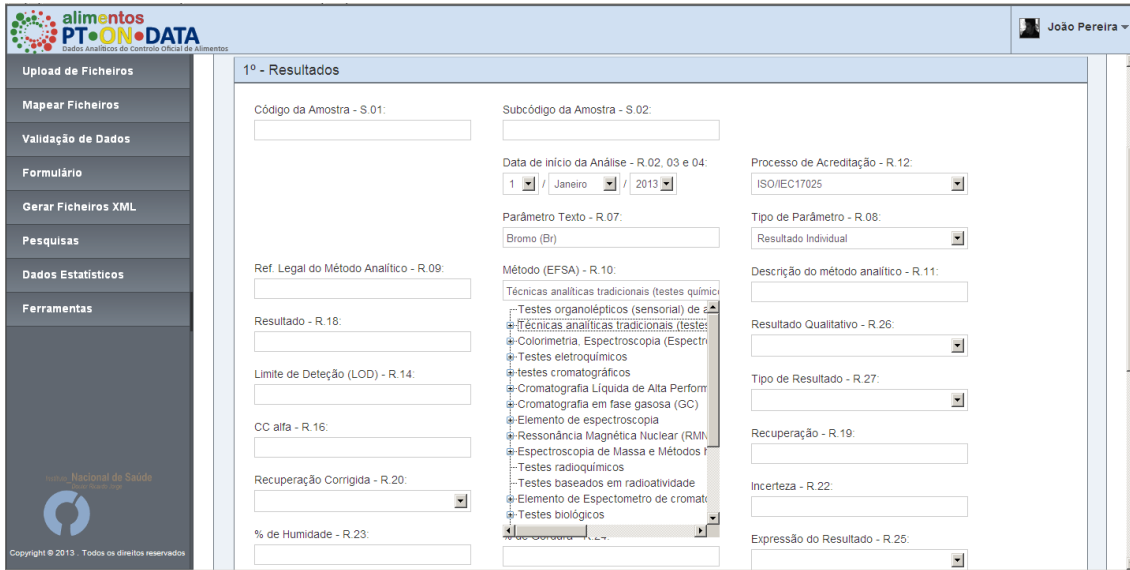
Figure 9: Screenshot showing the introduction of laboratory data on the web form

On figure 10, it is possible to see the resulting implementation of the form. One thing that is also possible to be noticed is the use of the TreeView and DropDownList controls populated with some SSD data from the server's database. In the early stages of the form development, the amount of data being sent to the client was immense, and when trying to keep the state of these controls (TreeViews, DropDownLists, TextBoxes and GridViews) between the several steps, this amount of data, which was approximately 8 Mb, was being download and upload whenever an operation occurred on the page. To solve this problem, preventing the page from taking near a minute from loading, we have partially disabled the upload of the control states and resorted to the use of Javascript techniques to prevent the huge amount of data to be uploaded to the server.

At the end of submitting the sample entries inserted through the form, the application redirects the user to the entries summary page, where all previous sample insertions can be viewed. On this page, the user can choose whether to start inserting a new entry, or use the other important feature of the web form, which is the ability to export all the form data to an Excel file. This Excel file can later be uploaded to our server using the platform's File Upload functionality.

It is also essential to highlight that all the sample data inserted via the Web Form are encrypted, if the user chooses to do so. Whenever an account is created for any Authority, an e-mail is sent to them with a randomly generated password only for the web form. The password method of protection, together with the AES encryption algorithm, makes possible to maintain the confidentiality of data that are only for the use of the authority that created them. Once encrypted, the records entered in the

---

form will be stored in the database system in a format unreadable and impossible to decrypt by brute force.

Additional developments

After first phase of implementation is finished, it was decided to focus on making improvements that could be made to the existing processes. As a result, new features were implemented in order to enrich the existing tools. These features arose from the needs identified users to make use of the platform. Following are the description of the features and changes made and considered of greater importance.
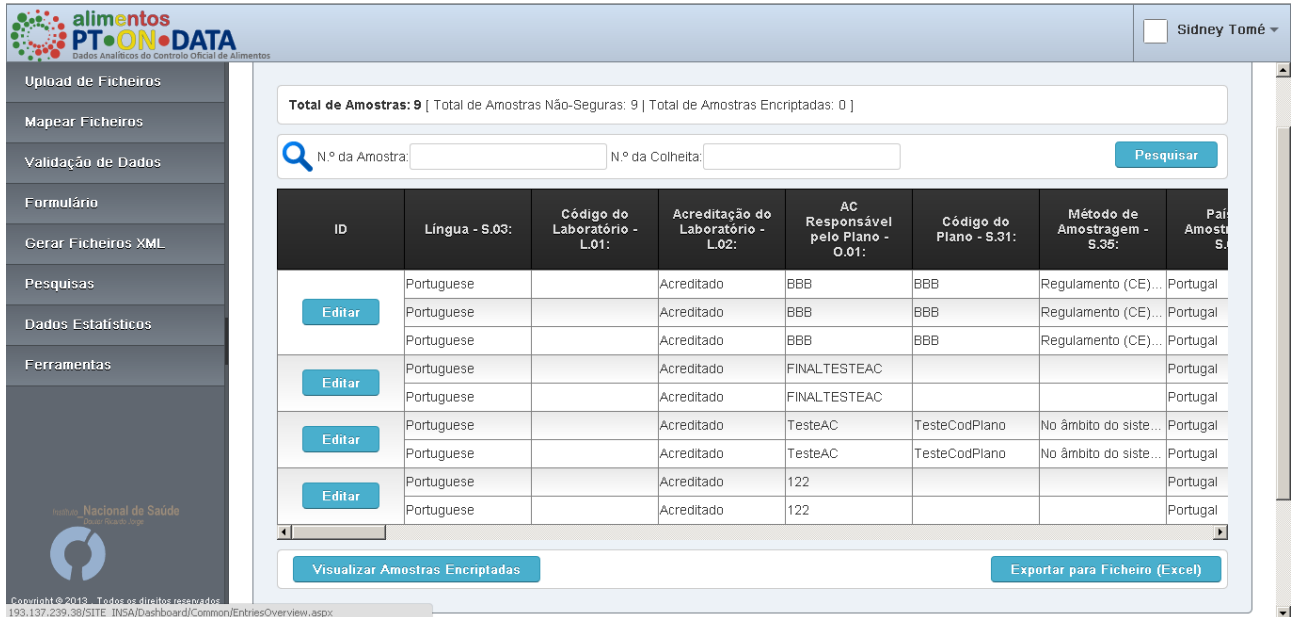
⇨ **Web Form visual appearance and functional changes**

Multiple changes were imposed on the Web form in order to facilitate the registers insertion process. Most of these changes consisted largely in defining the value that should appear by default in some fields. The dynamism with which the form was developed enabled a rapid and effective definition of the default values for any field present in it. In addition to these visual changes, certain features have also been implemented, such as the possibility to edit an entry representing a sample result. In addition it was also developed the possibility to save information from a sample result when the user has forgotten to do it, that it's done when the last step on the form is carried out.

It was also created a new feature that consists in the possibility to load the database with an English - Portuguese translation table for all the values included in SSD fields. This new functionality allows the user to have a better understanding of the values whenever he's filling the form.

⇨ **Research, editing and removing Web Form entries**

This feature has been developed for the overview page of the entries made in the form, and allows a user to manage these entries. Figure 11 shows the different possibilities that can be applied to entries.

Figure 10: Web Form records management

⇨ **Web Form records classification**

Once it was decided to make use of some of the validation rules imposed by EFSA in Web Form itself, such as the verification of the fulfilment of the required fields, the form is now able to sort the entries as complete, if all mandatory fields have been filled or as incomplete otherwise. This information can also be exported by the user to an Excel file where entries are classified as complete or incomplete

3.2.6. Data Validation
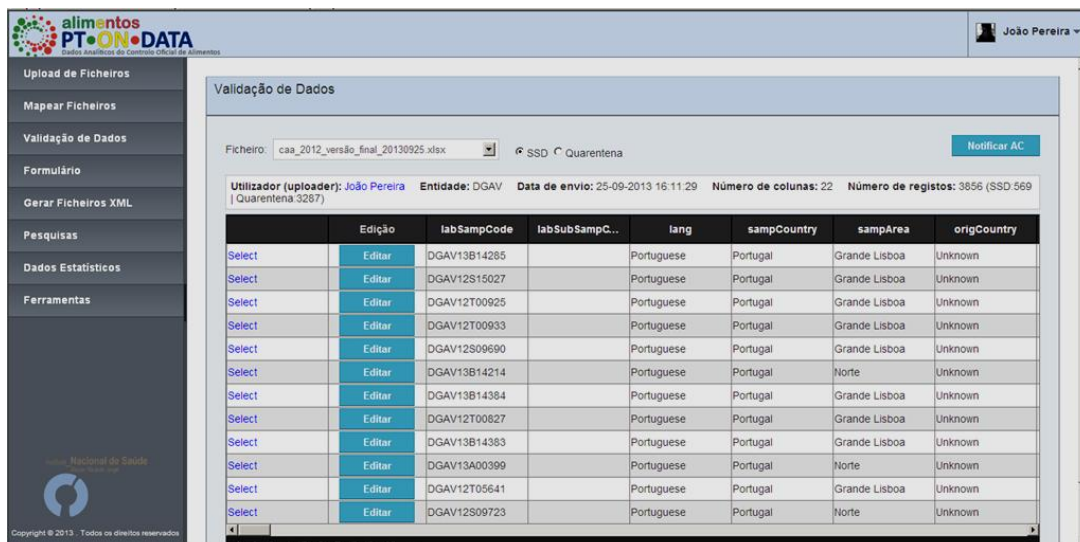
After data entry, was necessary to implement mechanisms for validation according to the SSD rules and correction of such data. Since at national level doesn't exists a set of terms common to all authorities to describe the samples and their outcomes, it was necessary to create repositories where it was possible to map uncontrolled data terms to SSD.

This validation is made in two steps:

1. The system makes a first validation on each value of each data element individually. Whether a data element fails in some validation, these data are automatically moved to the quarantine table and it is added a note about the validation rule that failed.

   - Checks that all required SSD fields are filled;

- Verifies that the controlled data are in accordance with the list of SSD field values. If the values are written by name, the system tries to associate the corresponding SSD code (e.g. sampCountry: "Portugal" → "PT");

- Checks if the data type for each field is in accordance with the SSD standards.

2. The system only performs this second step to the data elements that have been validated in previous step (data that has not been moved to the quarantine). The system validates the data as a whole, applying the SSD rules present on a PDF file "Guidance of EFSA - Guidance on Data exchange.pdf" submitted by EFSA. A tool that allows the system to read and apply the rules directly as they are written in the PDF file was developed. Thus, if it is necessary to add new rules to the SSD, this is a process that can be done easily by any user. If the data are considered valid by this second step, then that data is ready to be sent to EFSA, otherwise the data is sent to quarantine table.

In figure 13 it is shown the flowchart of the whole process of data validation. In an initial stage, only the user with the role "Supervisor" has access to the process of data validation. If desired, the supervisor can provide access to the entity who owns this data to help in the validation process. This action is performed by activating the control "Notificar AC". It is recommended that the Supervisor always requests a check to the CA, even if the data is consistent with the SSD rules. If the data has been corrected by the entity, they are moved to SSD table, otherwise they will remain in quarantine.

The "Data Validation" (Figure 12) allows us to visualize all data of each file, the data can be in two different steps: SSD and Quarantine.

Figure 11: Web platform: page "Validação de Dados"

Once the Supervisor approves the data, the state of these data elements changes to "approved" and they can be exported to an XML file.
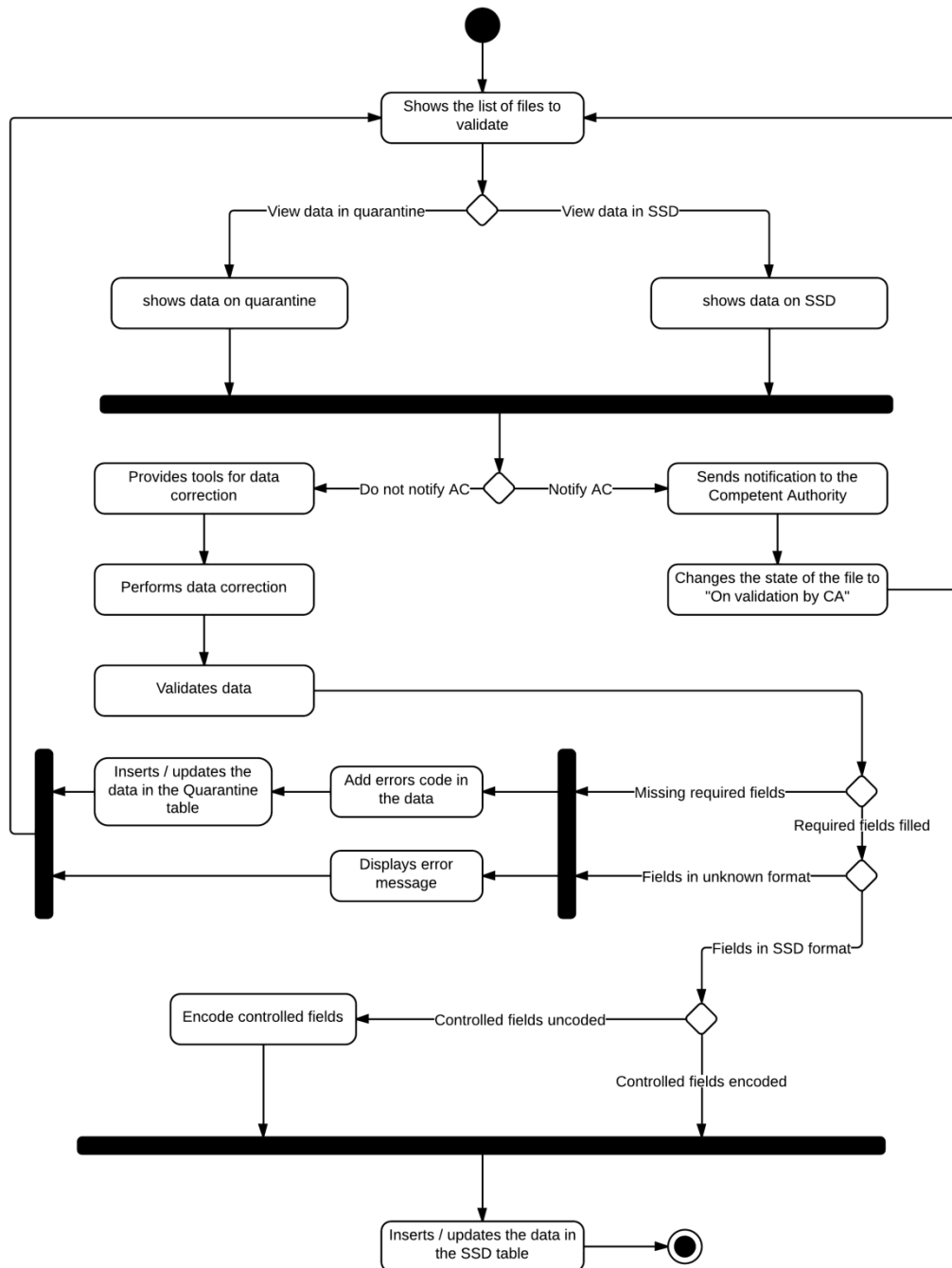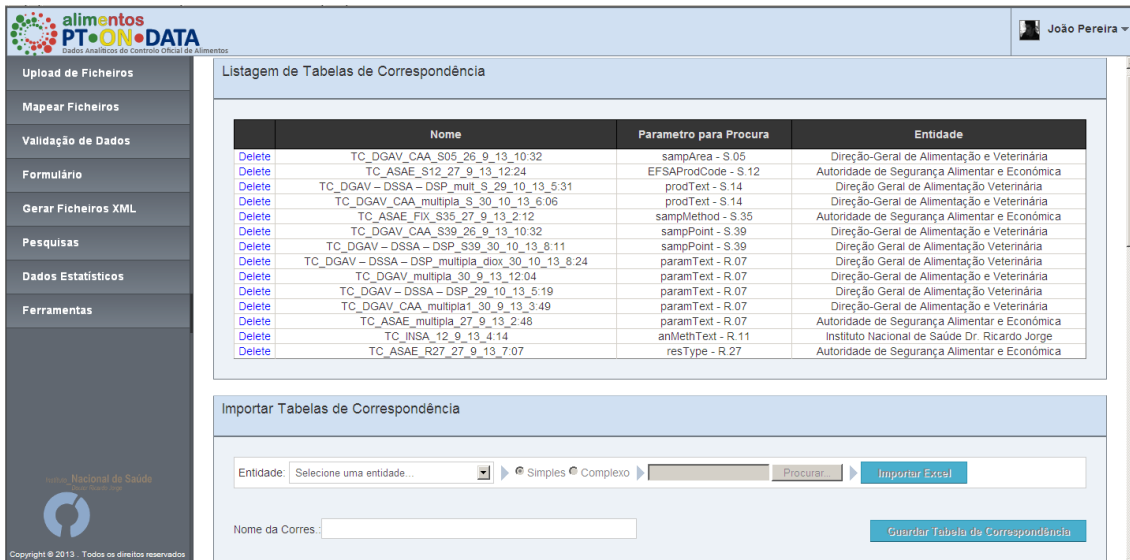
Figure 12: Data Validation flowchart

3.2.7. Creation and use of correspondence tables

After data entry, mapping and validating, it can be performed an additional step in order to correct wrong records. This step involves the application of correlation tables, which basically will look for certain values and replace them with another value considered correct, thus making the entries valid. These correlation tables are created through the tool illustrated in Figure 14, of single access by system supervisors. From this tool, the supervisor can choose to create matching tables in two different ways, either by importing Excel files with the contents of the tables or through manual creation using the values already present in the database. The user who will apply the correlation tables will always have the possibility to select which tables can be applied to records mapped and validated. After selection and subsequent application of the tables it is done, data is again validated with the rules imposed by EFSA to ensure that they are actually valid or still invalid.

Figure 13: Correspondence tables management tool

3.2.8. Generating and Transmitting XML to EFSA

The last process of the system development consists in generating XML files containing the data of analyses of samples previously validated by the Supervisor. In figure 15 is a screenshot of the window to generate XML files. Only users with the role "Supervisor" have access to this page.

There are two filters which may be used in order to filter the data to be included in the XML file:

- Parameter groups - it is possible to filter the data elements by Parameter Code values (check *Data Collection Groups Management Tool* in chapter 3.2.9 - platform tools).

- Report year - it is possible to export data that belong to a specific year. The field SampY (year of sampling) is used to filter the data elements.

Figure 14: Screenshot of Web platform page "Gerar Ficheiros XML"

After the data has been exported to an XML file, the status of each data element is changed to "Reported". It is also possible to access the history of all submitted files to EFSA.

In order to generate XML documents it was used Microsoft®XmlSampleGenerator22, which allows generating an XML structure from an XSD document. To generate the XML structure it was used the StandardSampleDescription.1.0.xsd scheme provided by EFSA. After adding data to report in the XML document, it is performed a final validation of the document and if it's according to the XSD schema, the document is placed to perform download in the web platform.

3.2.9. Platform Tools

While developing the platform, we noticed that it was becoming more and more complex and hard to manage all the data around it. So, in order to make the platform handling and maintenance an easier task, the team built a set of tools (System Tools) that could only be used depending on the profile associated to the logged user, meaning that some tools would only be able to be accessed by Admins, while some others were only accessible by regular users.

Below is a list of all the tools developed for the platform:

- **Database Management Tool:** This lets the Administrators avoid the effort of having to use the SQL Server whenever there is a need to alter some value in a certain table, either by a request of a user or just by necessity. This tool allows managing all tables and data from the Database via web interface.

- **User/Authority Management Tool:** This tool allows the creation and editing of users and authorities for the platform. An authority is created to be linked to a user.

- **Data Collection Groups Management Tool:** This tool is used when data has to be submitted to EFSA in the XML format, but separated by group. Each group is divided according to its *Parameter Code* (R.06). Basically with this tool it is possible define which Parameter Code values should be exported into a different group, resulting on separate XML files.

- **Form Settings:** This tool allows regular users to define additional fields that they would like to appear in the form when inserting a new entry, thus responding to a specific need of the organization. This tool is very helpful because some of the authorities require more information about the sample that would useful for them but isn't present among the SSD variables.
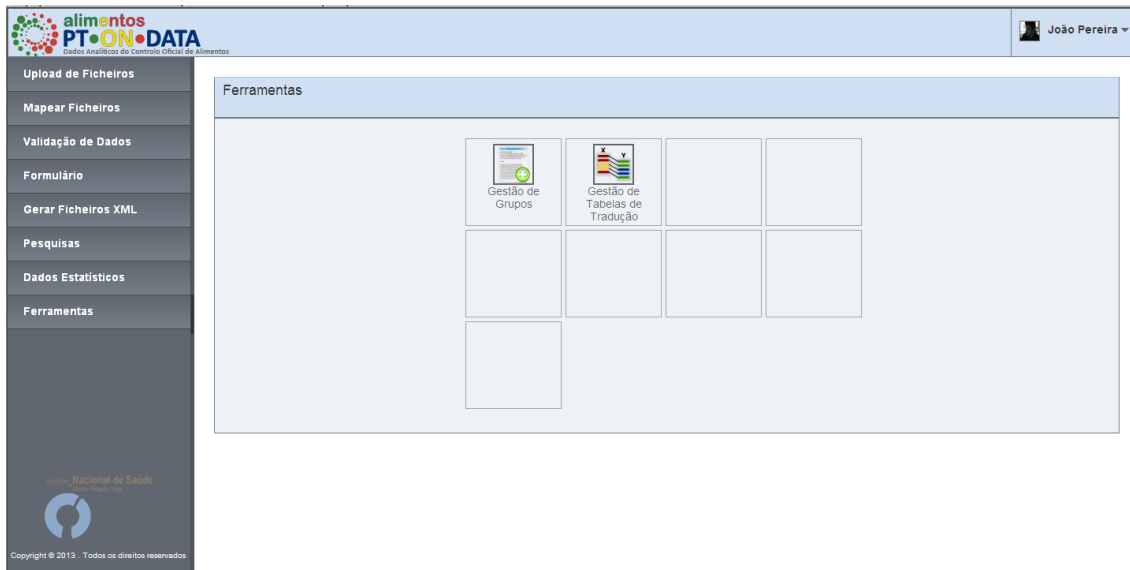
Figure 15: Web platform page "Ferramentas"

3.2.10. Research

The search tool allows searching for data in the SSD table using specific filters. After the search you can export data to a Microsoft® Excel®. This tool is available on the web platform page "Pesquisas" (Figure 17).
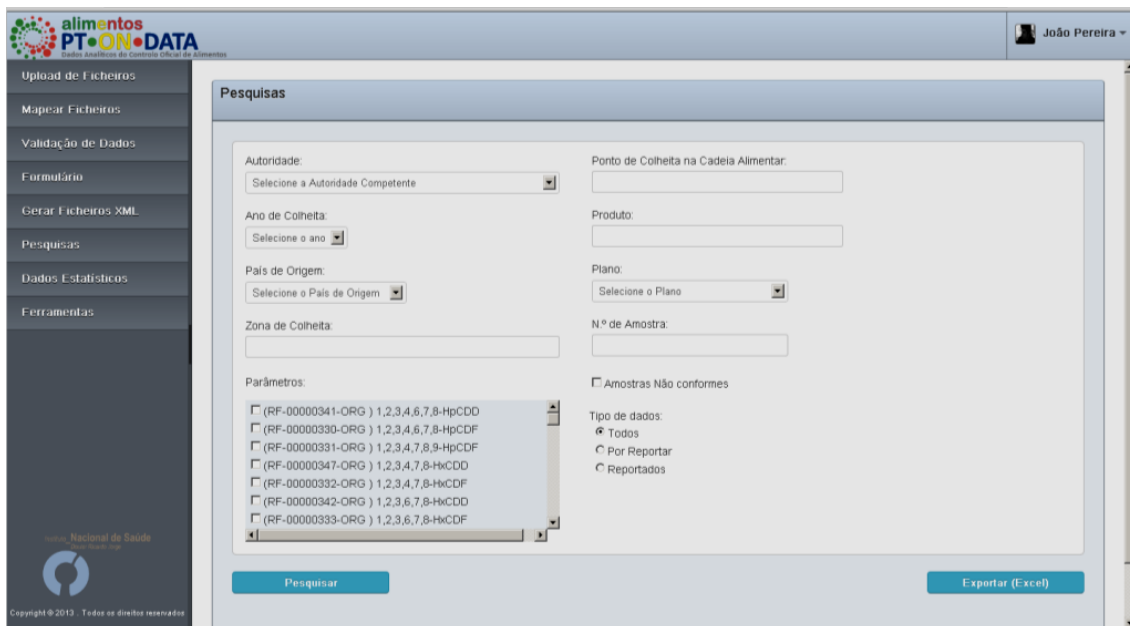
Figure 16: Web platform page "Pesquisas"

3.2.11. Statistics

On the "Dados Estatísticos" page (Figure 18) it is possible to visualize statistics about the data entered into the system. Statistics can be global or grouped by years.
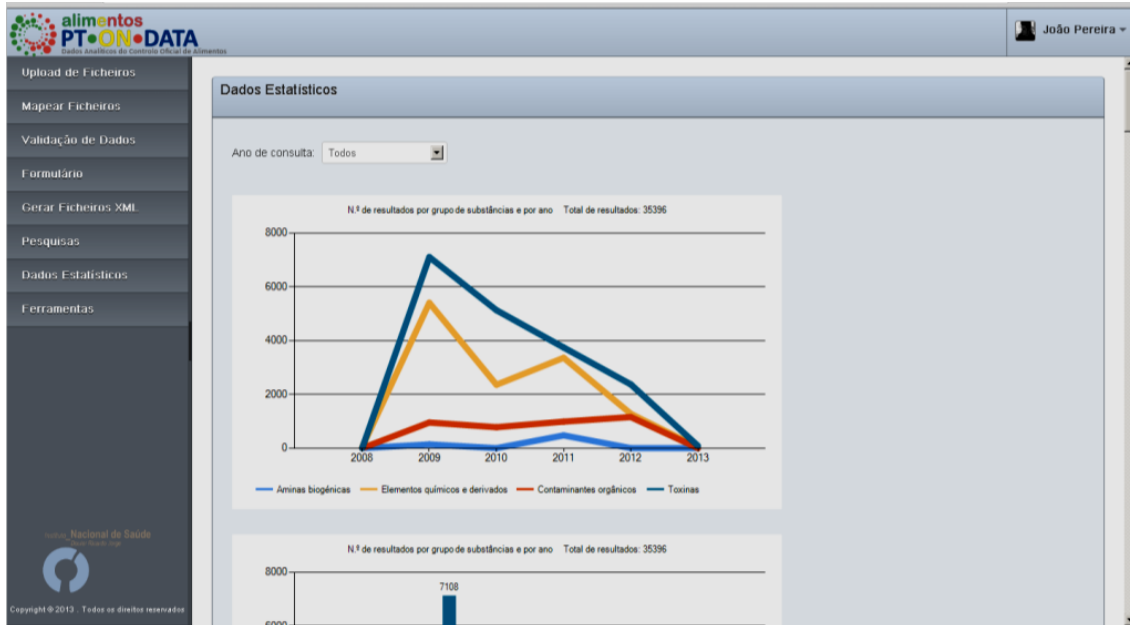
Figure 17: Web platform: page "Dados Estatísticos"

4. DATA VALIDATION AND TRANSMISSION TO EFSA

The 2009 and 2010 data (some of 2010 data were not reported during the 2011 annual call for data), were compiled in the GRF and were transmitted to EFSA in XML format through the Data Collection Framework (DCF) web interface during the 2012 annual call. The 2011 data was validated by the software validation process referred in 3.2.6, and successfully reported to EFSA through the Data Collection Framework (DCF) web interface in a XML format by the end of 2012 and considered accepted. The transmission process was made according to the process described in the EFSA document "Guidance on Data Exchange". The 2012 data was also validated and sent during the annual call according to the procedures mentioned above and were considered valid but are still under analysis by EFSA for acceptance.

Table 3 represents a summary of the reported data between 2009 and 2012.

Table 3: Summary table of the reported data between 2009 and 2012

Group 1 Parameter Code	2009 Frequency	2010 Frequency	2011 Frequency	2012 Frequency
1,2,3,4,6,7,8-HpCDF	48	41	51	50
1,2,3,4,7,8,9-HpCDF	48	41	51	50
1,2,3,4,7,8-HxCDF	48	41	51	50
1,2,3,6,7,8-HxCDF	48	41	51	50
1,2,3,7,8,9-HxCDF	48	41	51	50
1,2,3,7,8-PeCDF	48	41	51	50
2,3,4,6,7,8-HxCDF	48	41	51	50
2,3,4,7,8-PeCDF	48	41	51	50
2,3,7,8-TCDF	48	41	51	50
OCDF	48	41	51	50
1,2,3,4,6,7,8-HpCDD	48	41	51	50
1,2,3,6,7,8-HxCDD	48	41	51	50
1,2,3,7,8,9-HxCDD	48	41	51	50
1,2,3,7,8-PeCDD	48	41	51	50
2,3,7,8-TCDD	48	41	51	50
OCDD	48	41	51	50
1,2,3,4,7,8-HxCDD	48	41	51	50
TEQ dioxins (PCDD and PCDF) UB	48	41	51	50
Melamine	20	41	57	0
Polychlorinated Biphenyls	12	0	0	360
Total	**896**	**779**	**975**	**1260**

Group 2 Parameter Code	2009 Frequency	2010 Frequency	2011 Frequency	2012 Frequency
Hydroxymethylfurfural	29	0	11	0
Benzo-a-pireno	27	0	0	0
Polycyclic aromatic hydrocarbons	132	0	0	0
Total	**188**	**0**	**11**	**0**

Group 3 Parameter Code	2009 Frequency	2010 Frequency	2011 Frequency	2012 Frequency
Fumonisin B1	271	137	23	0
Fumonisin B2	271	137	23	0
Zearalenone	242	226	91	10
Patulina	242	226	2	10
Ochratoxins	483	212	118	76
Histamine	139	0	474	0
Aflatoxins B1	1106	628	425	626
Aflatoxins B2	1103	578	384	627
Aflatoxins G1	1103	576	384	627
Aflatoxins G2	1103	580	384	626
Aflatoxins M1	78	13	21	34
Deoxynivalenol/Vomitoxina	229	251	55	0
Domoic acid group toxins	463	682	634	0
Okadaic acid group toxins	413	739	565	0
Saxitoxin-group toxins	450	653	631	0
Total	**7696**	**5638**	**4214**	**2636**

Group 4 Parameter Code	2009 Frequency	2010 Frequency	2011 Frequency	2012 Frequency
Cadmium (Cd)	1042	631	586	419
Iron (Fe)	39	0	7	0
Copper (Cu)	645	336	112	132
Total mercury	532	231	481	313
Lead and derivates	1196	768	572	448
Manganese (Mn)	25	0	2	0
Selenium (Se)	3	0	120	13
Zink (Zn)	643	337	106	131
Sodium (Na)	38	0	14	0
Phosphorus (P)	5	0	2	0
Iodine	18	0	12	0
Potassium (K)	26	0	8	0
Magnesium (Mg)	20	0	6	0
Nitrate	186	0	307	3
Nitrites	126	0	45	0
Sulfates	250	0	235	0
Sulfur Dioxide	328	0	296	0
Sulfur and derivatives	238	0	269	0
Arsenic and derivatives	3	18	104	0
Chromium and derivatives	11	15	39	6
Chlorates	21	0	0	0
Nickel and derivatives	10	15	38	0
Total	**5405**	**2351**	**3361**	**1465**

	2009	2010	2011	2012
Nacional total/year	14185	8768	8561	5361

5. QUALITY IMPROVEMENT OF FUTURE DATA COLLECTION

a. Sampling data collection form

An harmonized "Sampling data collection form" was initially created as a result of the congregation of the information used by each entity involved on the project and the model used by our mentor FSAI (Food Safety Authority of Ireland). Given the different and specific characteristics of each of the entities involved and the numerous information required by EFSA, this document was discussed during a workshop held in February 2013 subordinated to this topic. Due the complexity and the different scope of the document, it was suggested then that it should be discussed individually with each entity in the coming months. At this point, and as a result of the close work performed with each of the official entities, the documents used in the past by each organization were subjected to improvement and several changes (restructuring of the visual information available and inclusion of new fields) in order to respond not only to their particular requests, but also to answer EFSAs demands, and are fully implemented and in use by each competent authority, with the exception of IPMA, which is still evaluating the document and its applicability on its official control program (figures 19, 20, 21 and 22).

These harmonized documents, are going to facilitate the pooling of information, by adopting a uniform structure and language. This will improve the appropriate use of information available and increase the range of possibilities for its use on the food safety area.

Figure 18: First page of the sample collecting form used by ASAE

Figure 19: First page of the sample collecting form used by DGAV_CAA

Figure 20: First page of the sample collecting form used by DGAV_PNPR

Instituto **Nacional de Saúde**
Doutor Ricardo Jorge

Instituto Português do Mar e da Atmosfera

Auto de Colheita de Amostras
Géneros Alimentícios

Laboratório de destino ☐FQ ☐LMB ☐MIC ☐SEN ☐PAT ☐Outro: _____

1. Colheita

Data: ___/___/___ Hora: ___:___ Amostra nº_____

1.1. Motivo

☐PNMMB ☐P.I.F. _____ ☐Prestação de serviços _____ ☐Outra: _____

1.2 País de Origem: ☐ Portugal ☐ Comunitário Qual _____ ☐ Países 3ᵒˢ Qual _____

2. Amostra

2.1. Tipo	☐Água ☐Peixe ☐Fígado ☐Músculo ☐Moluscos Bivalves ☐Gastrópodes
	☐Cefalópodes ☐Crustáceos ☐Outro _____

Descrição da amostra: _____

Observações: _____ Selo nº _____

O produto encontrava-se à temperatura de _____ °C tendo sido colocado em:

☐Refrigeração _____°C | Conservado/Transportado | ☐mala térmica Com _____ termoacumuladores
☐arca frigorifica

☐Congelação/Ultracongelação -_____°C | Conservado/Transportado | ☐viatura frigorifica | ☐Congelação ☐Refrigeração Termostato regulado para _____°C
☐Outro Qual _____

2.2. Amostragem

2.2.1. Método de produção: ☐Mar ☐Água doce ☐Selvagem ☐Aquicultura Zona (FAO) _____

2.2.2. Zonas costeiras

Norte	☐L1 ☐L2	Centro	☐L3 ☐L4	Lisboa e Vale do Tejo	☐L5	Alentejo	☐L6	Algarve	☐L7 ☐L8 ☐L9

2.2.3. Zonas estuarino-lagunares

Norte	☐ELM ☐EMI		Lisboa e Vale do Tejo	☐ETJ ☐LOB		Algarve	☐VRSA1-TAV1 ☐TAV2 ☐FUZ1 ☐OLH1 ☐OLH2	☐OLH3 ☐OLH4 ☐OLH5 ☐FAR1 ☐FAR2	☐POR1 ☐POR2 ☐LAG
Centro	☐RIAV1 ☐RIAV2 ☐RIAV3	☐RIAV4 ☐EMN1 ☐EMN2	Alentejo	☐ESD1 ☐ESD2 ☐LAL					

3. Tipo de processamento

☐Sem processamento ☐Congelação
☐Conserva ☐Refrigeração
☐Salga/Secagem ☐Secagem
☐Outro _____

4. Tipo de Acondicionamento Acondicionamento original ☐Sim ☐Não

☐Frasco vidro ☐Saco plástico ☐Metal ☐Papel/Cartão
☐Tubo plástico ☐Saco rede ☐Granel ☐Outro _____

5. Ensaios pretendidos

5.1. FQ	☐Cádmio	☐Chumbo	☐Mercúrio	☐Histamina	☐Outro _____
5.2. LBM	☐ASP	☐PSP	☐DSP (AO+DTXs)	☐YTXs	☐PTXs ☐AZAs
5.3. MIC	☐Listeria	☐Salmonella	☐Clostridium	☐E. coli	☐S. aureus ☐Outro _____
5.4. SEN	☐Cheiro	☐Sabor	☐Cor	☐Textura	☐Frescura ☐Outro _____
5.5. PAT	☐????	☐????	☐????	☐????	☐???? ☐Outro _____

O Gestor de amostras _____

É EXPRESSAMENTE PROÍBIDA A ALTERAÇÃO OU REPRODUÇÃO PARCIAL DESTE DOCUMENTO

Figure 21: First page of the sample collecting form used by IPMA

b. Training of Sampling Officers

Whenever asked by the competent authorities, it was prepared an instruction sheet for completion of the sample collection forms and upon request, it is foreseen close collaboration in the planning and in the training of staff assigned to these functions of sampling collection.

6. LAUNCH OF THE PORTUGUESE SYSTEM "alimentos PT·ON·DATA"

The official launch of the Portuguese system "alimentos PT·ON·DATA" was held on the 24th of October which was attended by all competent and official authorities involved in the project. We have also had the presence of Eileen O'Dea from FSAI – project mentor – and Stefano Càppe representing EFSA. About 50 people attended this event and the topics discussed were the ones referred in the following program (figure 23).

Figure 22: "alimentos PT·ON·DATA" launch program

CONCLUSIONS AND RECOMMENDATIONS

In future it is expected that sample information and results will be stored in the LIMS of the national authorities and will be periodically extracted and sent to INSA in CSV or other agreed format.

This will be achieved by the development of an extraction interface form the national data base developed which will be able to extract sample data from the laboratories LIMS.

TECHNICAL ISSUES

We have encountered some difficulties during the data collection and the classification of the lists with controlled vocabularies from the national data sources, some of them were already reported on the Document D1 – Data Standardisation and are outlined again:

✧ We have many Portuguese traditional/regional and other specific products that do not have any correspondence to the FoodEx list.

✧ Many of the marine species consumed in Portugal are not listed in FoodEx list.

✧ We have many refrigerated products (0 – 6 °C) and there is no specific product treatment to be chosen to classify these types of products.

✧ Many of EFSA required information were not available or were difficult to obtain.

✧ Most of the historical data information was on paper format and that required the presence of INSA technicians in the various entities, often for extended periods of time, which delayed the data collection.

✧ During the use of GRF we have experienced some difficulties that we believe were due to Office updates that disable the DLL files which prevent the proper use of the macros in this file.

✧ The XML schema validation rules don't allow the use of comma as a decimal separator, although this can be used on the GRF whenever numerical data is introduced. Since it is not allowed the use of national language this rule should be revised in order to accepted the numerical rules from each MS.

✤ As already have been mentioned before we had many SSD fields with no controlled vocabularies and because of that we received a large amount of data that had to be subjected to technical screening to correspond to the SSD system.

The collection of national historical data and the development of the mapping tables were tasks that involved qualified personnel and took a long period of time, but allowed us to have a strong notion of the national scenery on this area and to point us some future directions of work.

CONCLUSIONS

Near the official deadline of this project, and due to several constraints, INSA has requested an extension of this delivery date. Some of the constraints were already mentioned on this report, such as the absence of a national database on chemical occurrence data, different data sources and formats for the data available and no national normalization for detailed raw data as requested by SSD guidance. Another thing that contributed to the delay of the full implementation of the project had to do with national legal difficulties related to the economic crises that resulted on the late hiring of technical and IT grantees to develop the IT system. Also we have implemented the project in parallel action lines namely software development and manual data harmonization and entering. During 2012 submission Portuguese Authorities/Data sources sent data either in GRF format when they have no electronic system in place (LIMS) or in different excel files exported from their LIMS and sample collection registering system. In this last case, data was inserted manually in GRF by our technical grantees and subsequently uploaded in our system by our IT grantees and then validated by the system and submitted to EFSA. All this involved a huge amount of technical time just to enter and transform the data to SSD. It was foreseen this way because we had to prepare the data collection while at the same time we were developing the IT system.

The 2013 data collection was thought to be as much automatic as possible, in other words, national authorities uploaded files as they have them from their systems into our newly developed national system. The upload of real life files, the translation tables and the functionality for mapping fields to SSD was only tested at this time, and was only then that we have found some improvement needs like a functionality for multiple upload when an authority sends separate files for sample collection data and for sample results data, and further harmonization actions to improve data completeness and quality at sample collection and laboratory level. For that reason, to refine and optimize the system to its full potential and also to improve data quality at data sources level, the extension of the original

deadline and the maintenance of the grantees until the 2013 continuous call is finished were absolutely necessary.

The project was of considerable interest to Portugal since there was no national database in which the official control plans information was recorded, nor was the information gathered harmonized among the entities that coordinate and perform these controls.

The dispersion and heterogeneity of data hinder their use for various purposes, particularly for risk assessment. The joint work of the various entities that coordinate and perform the official control of food and feed on the same project, certainly improved the systematization and quality of existing and will improve future data, as well as will facilitate their use that will certainly bring an improvement in the use of national resources. The existence of interfaces that allow the exchange of information between different official organizations and a national database capable of storing all the information available and compatible with other databases nationally and/or internationally, facilitates the exchange of data and helps to improve food security globally.

The use of the national database – currently in the final stages of implementation - will allow the collection, mapping and archiving all data from the official controls, preferably automatically, which after validation will be reported to the central database of EFSA in a XML format.

This project allows promoting standardization and data quality, optimizing the use of resources at national level, processing of data and improving its availability and use for academics. Also, this project enabled the creation of sustainable and durable structures that can maintain data viability and hence the optimization of food security indicators to quantify the impact of risk management decisions.

By participating on this project, Portugal had an opportunity to create from scratch a national database where all data can be recorded, mapped and automatically reported to EFSA. This project made it possible to develop harmonized procedures, and alert the Portuguese official control authorities to the importance of quality and availability of analytical controls data.

Our mentor, FSAI, during the initial phase of the project provided an indispensable help in understanding the importance of this project either at national or European level. FSAI participated in a Workshop at Lisbon with all the competent authorities and official laboratories involved in the project. Through its experience in implementing a similar project in which to collect information that

also came from different sources, they provided us assistance in structuring and organizing work to develop with official authorities involved in order to facilitate data collection, fill appropriately the GRF files, in preparing the report D1-Data Standardization and in finding the best solution to adopt to built, develop and implement the Portuguese database. FSAI also provided us an example of the form used in the sampling collection that served as a guide for the preparation of the draft document of the future Portuguese harmonized collecting form. During the course of the project FSAI have always been available and in constant contact for any needed clarification or assistance, largely facilitating further work, for which we thank our mentor, the enormous availability shown.

RECOMMENDATIONS

The most important recommendation for EFSA is to be aware of the efforts for the MS while modifying the SSD structure. In order to avoid mistakes and the use of a wrong document version EFSA should clearly notify about any changes implemented in the SSD catalogues. The updates should have clear information about changes made, so that MS can adapt their mapping tables. Just providing updated information, without explaining the changes, could cause an enormous deal of work for the MS, because sometimes requires a manual comparison of both versions.

New versions of SSD catalogues should keep the current structure and only add/erase data elements and respective controlled vocabularies. Changes in controlled vocabularies should not include changes of codes and should be clearly notified and identified. A modification to the schema of the SSD will have a huge impact in the currently developed system and will require the need of additional resources by the MS for this implementation.

REFERENCES

⇨ Standard Sample Description for food and feed (SSD) - available online: http://www.efsa.europa.eu/en/datex/datexsubmitdata.htm.

⇨ Generic Reporting Format (GRF) - available online: http://www.efsa.europa.eu/en/datex/datexsubmitdata.htm.

⇨ The Guidance on Data Exchange - available on: http://www.efsa.europa.eu/en/datex/datexsubmitdata.htm.

⇨ The EFSA food classification system FoodEx - available online: http://www.efsa.europa.eu/en/datex201001/docs/cfpefsadatex201001ax1.xls.

⇨ Call for continuous collection of chemical contaminants occurrence data in food and feed - available online: http://www.efsa.europa.eu/en/art36grants/article36/cfpefsadatex201101.htm.

⇨ Regulation (EC) n.º 178/2002 of the European Parliament and of the Council, of 28 January - Laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety. OJ L 31, 1.2.2002, p. 1–24. (http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2002R0178:20090807:EN:PDF)

GLOSSARY/ABBREVIATIONS

AC/CA – Autoridade Competente/Competent Authority

ASAE – Autoridade de Segurança Alimentar e Económica/Food Safety and Economic Authority

DCF - Data Collection Framework

DGAV – Direção Geral de Alimentação e Veterinária/Food and Veterinary Directorate-General/ Direção Geral de Alimentação e Veterinária

EFSA – European Food Safety Authority

FSAI – Food Safety Authority of Ireland

INIAV- Instituto Nacional de Investigação Agrária e Veterinária/ National Institute of Agricultural and Veterinary Research

INSA- Instituto Nacional de Saúde Doutor Ricardo Jorge/ The National Health Institute Doutor Ricardo Jorge

IPMA - Portuguese Institute for Ocean and Atmosphere/Instituto Português do Mar e da Atmosfera

MS – Member State

SSD - Standard Sample Description

XML - Extensible Markup Language