An information system dedicated to pesticides users security

Gil de Sousa, Sonia Grimbuhler, Jean-Pierre Chanet, J.C. Champomier

An information system dedicated to pesticides users security

G. De Sousa[1], S. Grimbuhler[2], J.-P. Chanet[1], J.-C. Champomier[1]
[1]Cemagref, UR TSCF, 24 Avenue des Landais, BP 50085, 63172 Aubière cedex, France
[2]Cemagref, UR TSAN, Parc de Tourvoie BP 44, 92163 Antony cedex, France
Correspondent author: gil.de-sousa@cemagref.fr

Abstract

Many recent studies deal with the pesticides. According to their characteristics, good practices in their usage are then essential to avoid any kind of exposure. To answer to this problematic, the development of an application which aims at grouping into the same information system, all the most relevant existing information about this topic, has been investigated. This article presents the followed approach to reach that purpose. The first step is the establishment of a knowledge base in order to build the information system model. This base is elaborated from the different identified sources and from inquiries realized with future users. The second is the development of a module dedicated to the integration of data from several sources. A method to extract information from online databases has then been developed. The integration of other types of information sources is in progress. A configurable interface is finally associated to the information system to display the data according to the end users needs.

Keywords: Pesticides, several data sources integration

1. Introduction

The use of pesticides raises many problematic addressed in the context of various studies and research works. However, as shown by initiatives such as the ECOPHYTO 2018 plan initiated by the French government, which aims at significantly reduce the use of this kind of products before 2018, different ways are possible to try to reach this goal. In this example, the recommended methods range from developing new agricultural techniques, which need no or less pesticides, to providing more information and advices to the users. The latter method has the both advantage to reduce pesticides use and to take care of users health. It is closed to the subject of the research project that we actually investigate and that will be exposed in this article.

The first main purpose of our work is to provide good practices in pesticides usage, mainly, in the pesticide occupational exposure. Pesticide users have to respect some security protocols in order to reduce or even avoid contamination risk. However, sometimes, these protocols can be ignored or misinterpreted. It may, for example, be important to emphasize that an exposure can occur by different means either by respiratory or dermal way, this latter being the most common source of contamination (Aprea et al. 2001) (Aprea et al. 2004).

The second purpose is the design and the development of an application that could merge, into the same information system, data provided by several sources. In this case, the researched data are related to pesticides. However, the approach presented in this article can be applied to others thematics, not necessary agricultural or environmental, since we try to group data from different information sources.

To access to the data stored in this “global” and centralized information system, an adaptive interface is being developed to display the essential information at the level of the security and some others that can be required by the end user.

The first section presents the proposed approach in attempting to reach the objectives just explained. The second section describes the methodology adopted in the design of the model of the centralized information system. The third section addresses the technical problem of the data integration. A presentation of the desired adaptive interface is given in the next section. Finally, this article ends with the conclusion.

2. The proposed approach

The development of the desired application is made following different steps which are listed below before being described, in detail, in a dedicated section. There are four main steps:

- Construction of the knowledge base.
- Design of the centralized information system model.
- Integration of the different sources listed.
- Development of the adaptive interface.

The first step consists of identify all available information about risk prevention, good practices at the level of the security related to pesticides usage. These information can concern these topics directly or indirectly. For example, the concentration of the active substance of a given pesticides is an information that is important not only for environmental aspect but also for the user health.

The information system is naturally built according the information listed in the knowledge base. However, its model or its organisation depends on users needs, generally, expressed in a scope statement. In our case, this document is built from interviews realized with potential future users. The knowledge base and the information system are highly linked and are so presented in the same section.

During the construction of the knowledge base, several sources have been listed. A first study has to be conducted in relation with the type and the accessibility of each source. Indeed, the type of the source brings with it the data integrating problematic. The accessibility includes two elements. The first corresponds to the quality of the network connexion existing with the data source. The second involves the copyrights that can exist on the data. Particular use of data available on the Web requires, generally, author permission. Without this latter, data cannot be used and are considered unavailable.

The information system gathers a significant amount of data which provide to the end user, without be formatted, would annoy him. The role of the interface is to match user needs with available information. In addition, each person has their preferences on the display of information on, for example, software or a website. The knowledge degree of the user is also an important thing to take into account. Information should be given according to the users skills. These elements show the advantages of an adaptive and configurable interface in order to present information to the end user.

The final application is also intended for students in agriculture. It can be seen as an additional way to educate them on good practices in pesticide usage. Then, in the process of interface creation, not only farmers have been consulted, teachers in agricultural topics have also been.

3. Design of the information system model

Conventionally, the design of the information system model is based on classical methodology. Whatever the scope, generally, the model has to strongly take into account the end users needs. Identify, classify and group, by entities, users needs are complex operations that require a close collaboration between the users and the designer. In a first step, each user tries to list all the needed information and the business process in which it is involved. These latter are generally hard to establish. Indeed, when you make the same action or process, repeatedly, during a certain time, you develop a focused vision of your activities as opposed to a global one. In a second step, the work of the model designer is to build this global vision according to indications given by the future users.

The needs of users are obviously integrated into our approach. To define them, several persons belonging to the two groups mentioned in the previous section, farmers and teachers, have been interviewed. Then, thirty people from the education and eighty farmers, working on different type of agricultural cultures, have been questioned. Some users needs have been also deduce from existing applications not necessary directly dealing with security but, naturally, with pesticides. However, while a traditional approach would stop here and the next phase would start, in our case, users needs have to be confronted to the data available in the listed sources.

Our application doesn't aim at merging, into a central information system, data stored in existing databases. It doesn't consist of the design of an empty information system that would be filled by end users. It tries to match needs observed among potential users with information already available, regardless their format.

From this study, different diagrams of class from the modelling language UML (Unified Modeling Language) have been edited to establish the information system model. An example of these diagrams of class is presented in Fig. 1. Cardinalities and classes roles have been hidden for more clarity. Main classes “Pesticides” and “Active substance” are present.

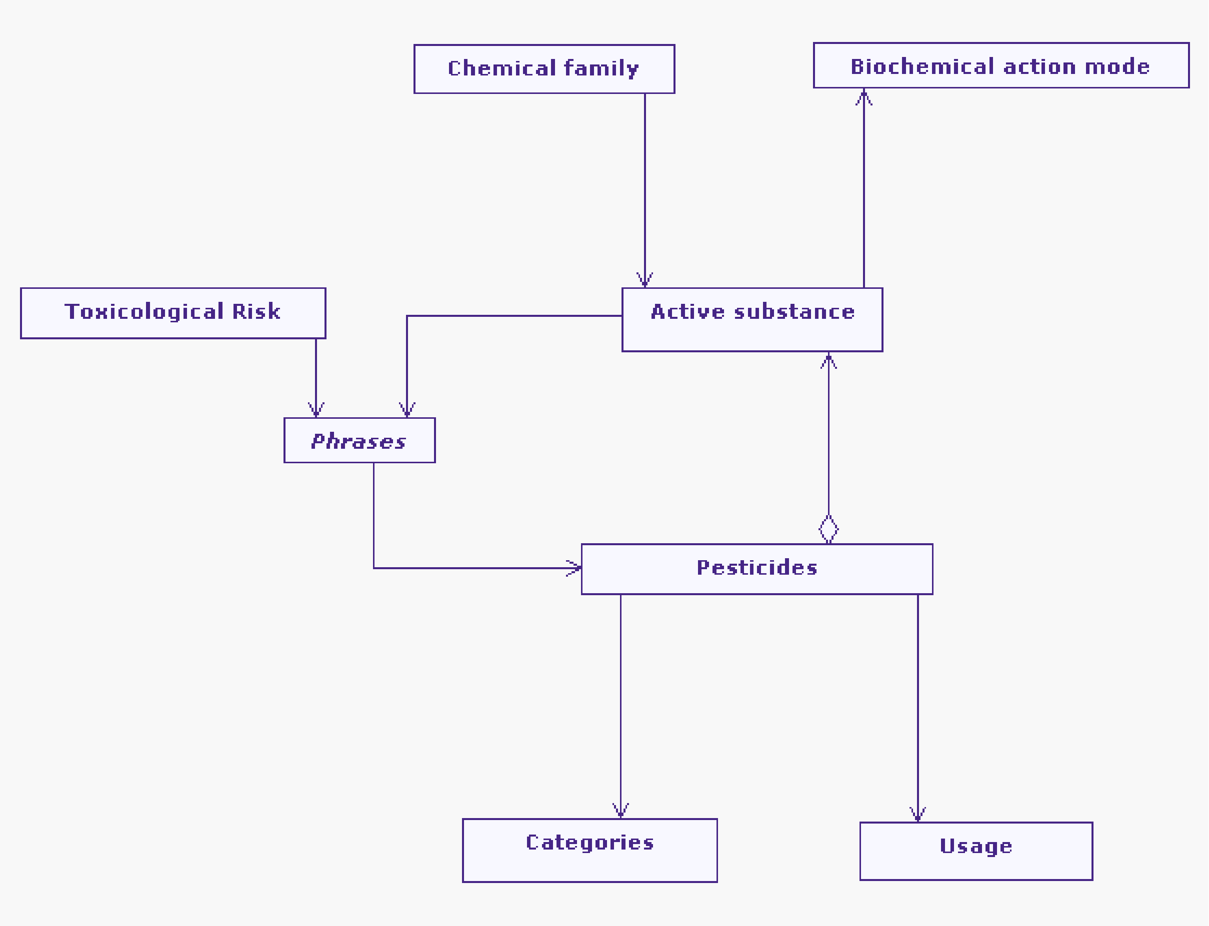

Fig 1: Part of the information system model

4. Data integration

The data warehouse research topic was initially investigated in order to provide solutions for the storage of huge quantities of data (Calvanese et al. 1998) (Schneider 2007) (Fig. 2). Such large quantities that classical database can not bear with it correctly. However, this technology not only answers to this problematic, it also provides two main functionalities:

- Data integration from several databases.
- Decision support methods.

The decision support part integrates all the mathematical and aggregation functions that can be applied to the data according to the needs of the end users. Diagrams or dashboards can, for example, be generated to offer a specific view that facilitates user decision.

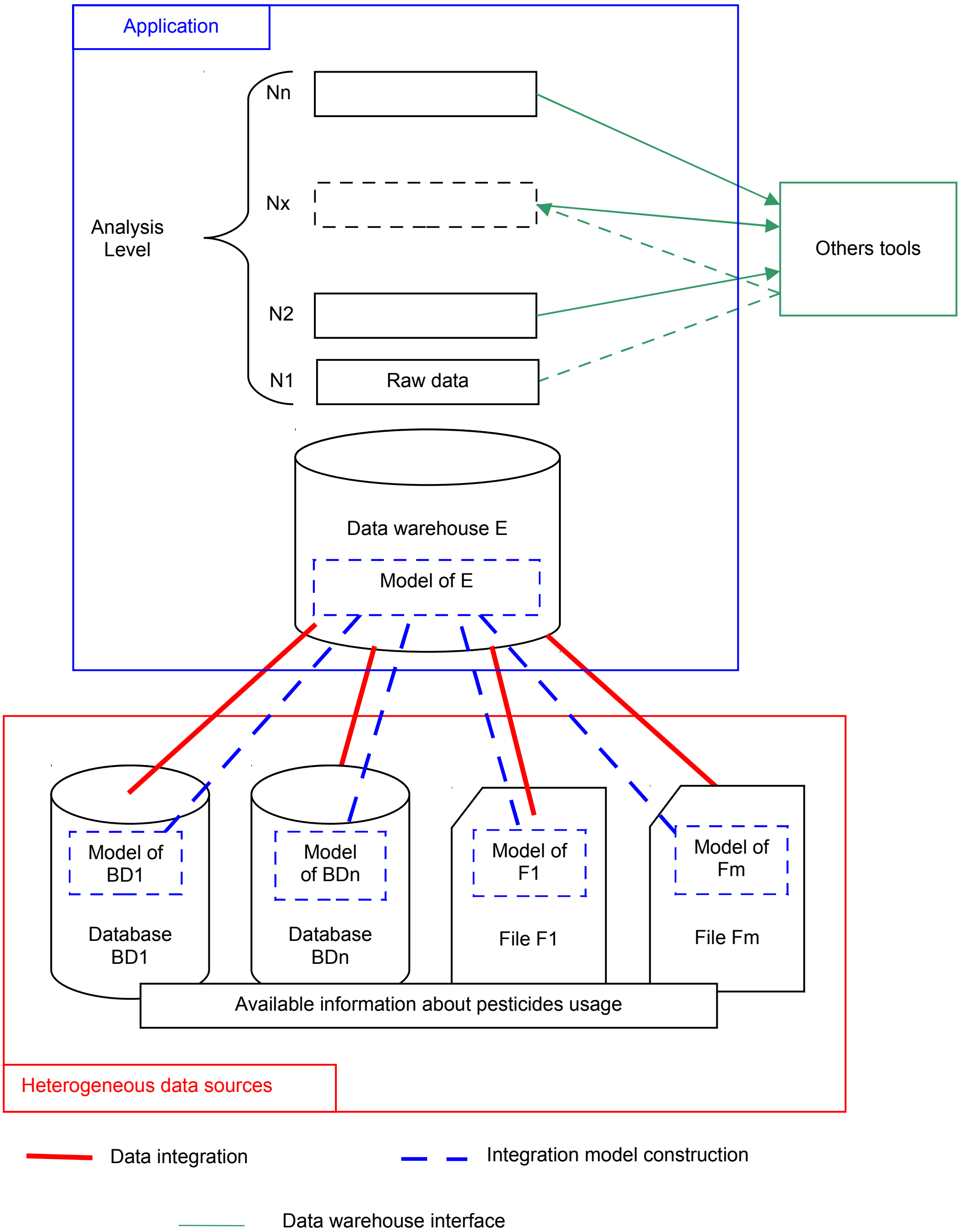

Fig 2: Data integration model

The integration of data providing by several databases into a data warehouse is an issue well-studied. It can be achieved by using an ETL (Extract Transform Load). But, in our application, data are, generally, not accessible directly from a database but from websites. So, some previous works have led to the implementation of software able to collect data from a website and to store them into a classical database. It is a first version of the software and additional developments are required to obtain a satisfactory solution.

Furthermore, data integration can be virtual or physical. In the virtual mode, data are still only stored in the source database. In the physical mode, data are copied into the information system that makes the integration. In our application, the physical mode was chosen in order not to loose some data in case of unavailability of its source.

Data integration also involves many unresolved problematic such as:

- Incomplete data.
- Conflicting data.

Incomplete or missing data can, for example, occur when an element of a given source doesn't match with any element of another source. In this case, only a part of the “tuple” resulting from the join of the two sources is filled. The problematic doesn't lie in filling missing data but the operating mode of the information system facing these incomplete data.

The second problematic is more complex. Sometimes, for the same information, two sources can provide different values. Various policies can be applied in this case. The first is to ignore the information and to consider the two values provided incorrect. The second is to store the two values advertising the end user of the ambiguity. The third is to try to find another source that could confirm or refute one of the two values. Nevertheless, we must keep in mind that this source may also give an invalid value. This latter introduces the complex problematic of the quality of the integrated data.

A future perspective for the information system is the addition of analysis axis in a data warehouse way. Then, the obtained application could be used as main source for tools that need pesticides information.

5. Presentation of the adaptive interface

As we mentioned in the first section, different elements lead us to the development of a highly configurable interface. Requests on the information system are realized by indicating a name of a pesticide or of an active substance. An authentication is required to adapt the display to the user. Furthermore, according to the profile of potential future users, three categories have been proposed:

- “Farmer”.
- “Student”.
- “Pesticides expert”.

In the first profile, user has directly access to main information about pesticides usage. Time and data relevance are the key metrics. Little or no undesired information are admitted.
The second profile is addressed to students. Then, additional information can be displayed to help the learning of some concepts. Detail about Personal Protective Equipment (PPE) can, for example, be given.
The last profile provides to the user all the available information for a given pesticides.

To increase reconfigurability and to prevent “pesticides expert” users to be submerged by information, the interface is also two-level configurable. The first level has just been presented and is associated to the profile of the user. Moreover, inside a given profile, user can choose to display or not some information. This corresponds to the second level. Whatever the level selected, crucial information, about pesticides usage, should always be displayed. Fig. 3 illustrates the functioning of the adaptive interface. User selects first its profile, in this case, the “student” one. Then, he can access to different class of data or entities and its associated attributes. Finally, he can choose the information relevant for him. In this example, “data1”, “data3” and “data4” are selected.

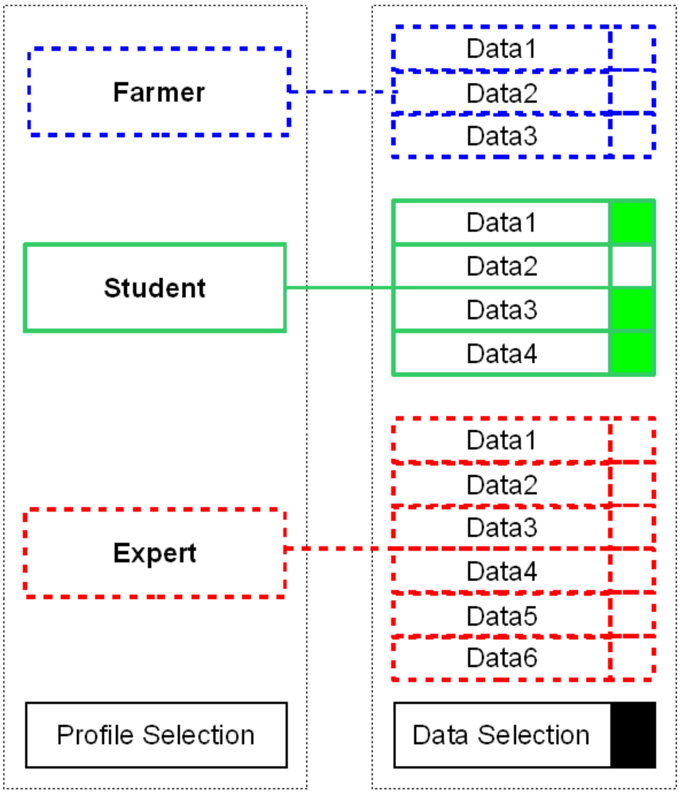

Fig 3: Adaptive Interface illustration

Conclusion

This article presents a complete approach which purpose is the development of an application dedicated to user pesticides security. The two main components of this application are an information system, fed by different heterogeneous data sources, and an adaptive interface. The role of these components is to offer to the end user all the relevant information that he can need or that can help him during pesticides usage, and allow him to preserve his health.

References

Aprea C., Sciarra G., Lunghini L., Centi L., Ceccarelli F. (2001). Evaluation of respiratory and cutaneous doses and urinary excretion of alkylphosphates by workers in greenhouses treated with omethoate, fenitrothion, and tolclofos-methyl. *AIHAJ*. 87-95.

Aprea C., Terenzoni B., De Angelis V., Sciarra G., Lunghini L., Borzacchi G., Vasconi D., Fani D., Quercia A., Salvan A., Settime L. (2004). Evaluation of skin and respiratory doses and urinary excretion of alkylphosphates in workers exposed to dimethoate during treatment of olive trees, *Archives of Environmental Contamination and Toxicology*, 127-134.

Calvanese D., De Giacomo G., Lenzerini M., Nardi D., Rosati R. (1998). Source integration in data warehousing, *9th International Workshop on Database and Expert Systems Applications DEXA-98*, 192-197.

Schneider M. (2007). A general model for the design of data warehouses. *International Journal of Production Economics, Elsevier*, 18 pages.