XVIIth World Congress of the International Commission of Agricultural and Biosystems Engineering (CIGR)

Hosted by the Canadian Society for Bioengineering (CSBE/SCGAB)
Québec City, Canada June 13-17, 2010

CONCEPTUAL MODELING FOR THE DESIGN OF A COMPUTER-BASED TRACEABILITY SYSTEM

GIOVANNI CASCONE[1], CLAUDIA ARCIDIACONO*[1], SIMONA M. C. PORTO[1], UMBERTO ANGUZZA[1]

[1]University of Catania (Italy) - Department of Agricultural Engineering - via S. Sofia 100, 95123, Catania.
* corresponding author <carcidi@unict.it>

CSBE101471 – Presented at 8th World Congress on Computers in Agriculture (WCCA) Symposium

ABSTRACT The European Standard UNI EN ISO 22005:2007, which defines the principles and objectives of traceability, mandates each organization belonging to a food-and-feed supply chain to choose appropriate tools for the implementation of a traceability system. Literature points out that the most suitable tool for the implementation of a traceability system is a computer-based information system (IS). In the context of the life cycle of a IS, the phase known as 'conceptual modeling' makes it possible to define models describing information content and functions of the IS, without taking into account the particulars of coding it. In this paper, the Unified Modeling Language (UML) was used in the phase of 'conceptual modeling' to model the information content of traceability IS. This modeling language was proposed as an alternative to the entity-relation model (E-R) since it provides the conceptual modeling phase with the same language for both the information content modeling and the function modeling. An application of UML was carried out to a case study concerning nursery-chain traceability of certified citrus plants. The importance of 'tracking' and 'tracing' citrus-plant production arises from the need to limit the diffusion of plant diseases, e.g. the *Closterovirus Citrus Tristeza Virus* which causes one of the most damaging diseases in citrus orchards and is the most important pathogen associated with this crop and has detrimental financial effects.

Keywords: plant propagating material, voluntary certification.

INTRODUCTION Several plant diseases are caused by unhealthy plant propagating materials (seeds, cuttings, scions, plants) which nowadays are exchanged faster and wider than in the past, due to trade globalization. Consequently, the prevention of plant-disease diffusion is expected to be performed at the plant nursery chain level by systems able to 'track' and 'trace' its production. In this regard, the European UNI EN ISO 22005:2007 standard (International Organization for Standardization, 2007) defines principles and objectives of traceability and specifies the steps of traceability-system design. Furthermore, according to the standard organizations belonging to food-and-feed supply chains are expected to implement such a system by choosing the most suitable tools to collect, record, and communicate product information. From a literature survey it has come out that the most appropriate tool is a computer-based information system (IS)

(Opara, 2003; Bertocco, 2004; Information Society Techonologies, 2005; Regattieri *et al.*, 2007; Niederhauser *et al.*, 2008).

In a previous study (Cascone *et al.*, 2009) the phases of the life cycle of a computer-based IS for traceability procedures implementation were identified and the steps needed to develop the requirements analysis phase, which may represent a weak point within a computer-based IS design (Cherry & Macredie, 1999), were highlighted. This methodology was applied to the case study of citrus-plant nursery chain with the aim to develop the requirements analysis of a computer-based IS capable to 'track' and 'trace' plant propagating materials. In particular, in that study the citrus-plant production certified under the Italian Ministry Decree of May 4, 2006 (Ministry of Agriculture and Forestry, 2006) were taken into account. This type of production would receive a significant support from 'tracking' and 'tracing' procedures insofar as they could contribute to limit the diffusion of plant diseases. Among them, *Closterovirus Citrus Tristeza Virus* (CTV) causes one of the most damaging diseases in citrus orchards (Bar-Joseph & Lee, 1989) and is the most important pathogen of this crop from the viewpoint of economic harm (Lee & Bar-Joseph, 2000).

On this basis, this paper proposes a methodology to perform the phase of 'conceptual modeling' within the life cycle of a computer-based traceability IS. In particular, to develop this phase the Unified Modeling Language (UML) was used as an alternative to the entity-relation model (E-R) since it provides the same language for information content and function modeling (Booch *et al.*, 1999; OMG, 2002).

1. MATERIALS AND METHODS The life cycle of a computer-based IS can be subdivided into phases, each involving well-defined entities (activities, human resources, and material resources) and producing a result (Batini *et al.*, 2001; Atzeni *et al.*, 2002). In the methodology proposed in a previous study (Cascone *et al.*, 2009), the life cycle of a computer-based traceability system (Fig. 1) was characterized by the possibility of iterating over the phases, moving backwards and forwards, or round and round, as needed (Mike O'Docherty, 2005). Phases 1 to 4 (stage 1) of the life cycle constituted the basis to design a computer-based IS, whereas phases 5 to 9 (stage 2) were more strictly linked to the implementation and maintenance of the system. Phases 2 to 4, in particular, have been object of investigation in several research studies (Dunn *et al.*, 2005; Nierderhauser *et al.*, 2008). In this regards, in a previous study (Cascone *et al.*, 2009), each step of phase 2 of a computer-based traceability system was made explicit by considering the requirements of the UNI EN ISO 22005:2007.

In the requirements analysis phase, information on procedures (step g) and on data to be managed by the traceability system (step f) constituted the input of the next phase, called 'conceptual modeling', which in this study was developed by using UML language. This language allows the building of conceptual schemes describing the information content and the functions of the traceability system. In particular, UML makes it possible:

- to model functions taking into account also users who will interact with the system (actors). In detail, from the requirements analysis phase (Cascone *et al.*, 2009), the functions to be modeled are: product track, product trace, nonconformity management, statistics, data management, information backup, and disaster recovery.

- to model system information content taking into account the relation among data and without considering the particulars of coding them. In detail, this information regards: products and their ingredients, supply-chain actors, and supply-chain processes.

Among the views provided by UML (i.e., use-case view, design view, implementation view, component view and deployment view), in this work use-case view was applied

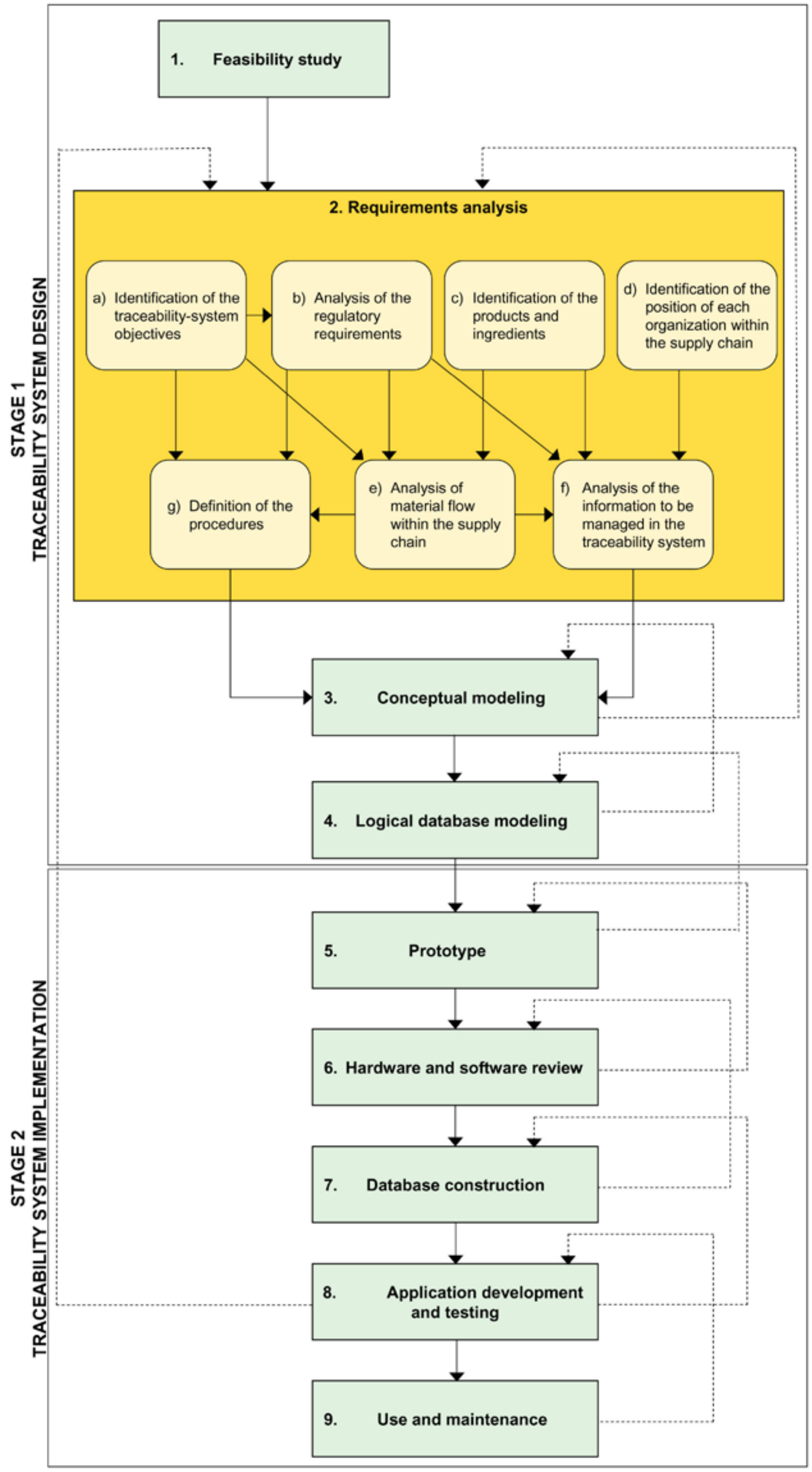


Figure 1. Flow chart of the iterative development methodology of the life cycle of a computer-based traceability system.

and developed to carry out the conceptual modeling phase of a traceability system.

To build the use-case view the following steps must be carried out:

1. identification of the actors of the system. An actor is defined as an identity (e.g., human operator, an organization, a physical device, another information system) which is external to the system and interacts with it. In a use-case view an actor is represented as a stick figure.

2. description of the system functionalities by using a set of use cases. In a use-case diagram, a horizontal ellipse is used to describe a 'use case', i.e., an action that provides something to an actor. Use cases are contained in a solid rectangle describing the system. The associations between actors and use cases are depicted by solid lines, whereas broken lines indicate an extended or included action.

3. characterization of the defined use cases by using activity diagrams and/or interaction diagrams (i.e., sequence diagrams and collaboration diagrams). An activity diagram is a simple way to represent the activity workflows involved in a single use case or only in a part of a use case, or even in many use cases, and how those activities depend on one another. A sequence diagram shows how and when the actors interact with the system to reach the goal of the use case, including information exchanged, decisions taken, and work products created. A collaboration diagram conveys the same information such as sequence diagrams, but a collaboration diagram focuses on object roles instead of the time sequence in which messages are sent.

4. production of a class diagram by using the data which the system handles for use cases. In a class diagram, each resource (concept that describes a class of objects having an independent existence) is represented by a class in terms of its structure and behavior. Classes are often associated with, or related to, other classes. Associations are modeled as lines connecting two classes. In a class diagram, each class is modeled with a rectangle having three sections: the top section for class name, the middle section for class attributes (characteristics of the class which may have a value), and the bottom section for class methods (functions operating on the class).

1.1 The case study of the citrus plant nursery chain. Citrus production in 2005 reached about 105 million tons worldwide and a cultivated surface of about eight million hectares. In Europe, the production was about ten million tons with a surface of about 0.5 million hectares. In 2005, about four million tons have been produced in Italy, the second nation after Spain in citrus production (ISMEA, 2006). Production is threatened by diseases affecting citrus orchards. Among them the most widespread is the CTV disease, which has been disseminated to almost all citrus-growing countries by the movement of infected plant material or by the aphid vector *Toxoptera Citricida* (Bar-Joseph & Lee, 1989). The fight against CTV is carried out, among other methods (Gmitter *et al*., 2007), by using healthy propagating material and through the eradication of outbreaks. To this end, in Italy a regulation on voluntary certification was introduced by the Italian Ministry Decree of May 4, 2006 (Ministry of the Agriculture and Forestry, 2006). Specifically, this regulation defines the requirements for the actors of the nursery chain to comply with the National Service of Voluntary Certification (NSVC) established by the Italian Ministry Decree of July 24, 2003 (Ministry of Agriculture and Forestry, 2003). Components of

NSVC are described in Tab 1. Outside the NSVC there are the breeders who produce the primary sources and farmers who will buy the citrus plants from the accredited nurseries (Fig. 2). To reduce the risk of diffusion of several plant diseases and of CTV, a computer-based IS could be a strategic tool as it could provide citrus plants and propagating materials with chain traceability.

NSVC COMPONENTS	DESCRIPTION
CCP	Conservation and pre-multiplication centers for mother plants breeding to produce "pre-base" materials (e.g. seed, grafts).
CP	Pre-multiplication centers to produce mother plants and "base" materials.
CM	Multiplication centers for mother plants breeding and production of certified propagating materials.
Nursery	Accredited nurseries for production of plants to be certified.

Tab. 1 - NSVC's components and related products.

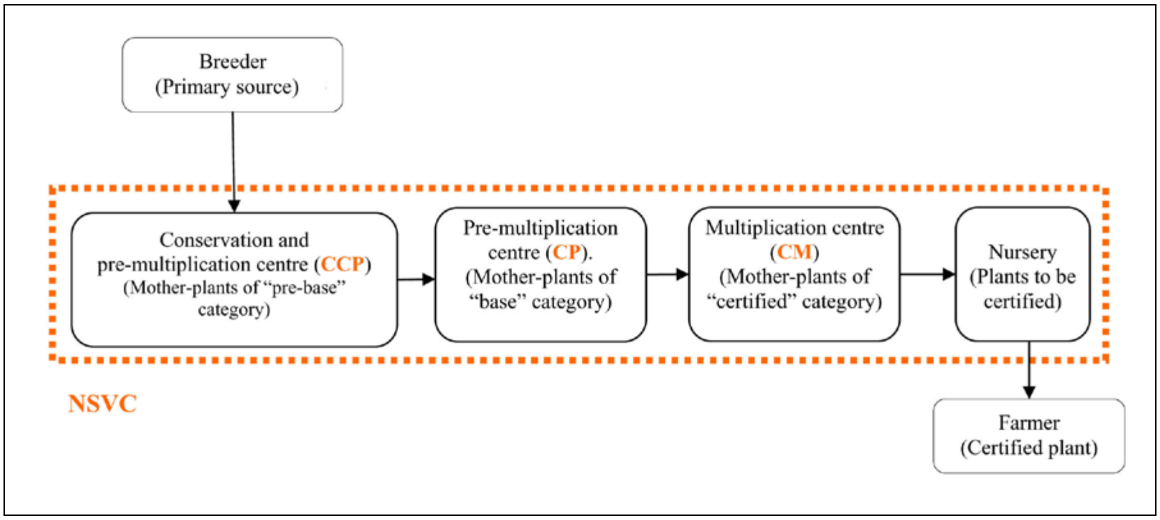


Fig. 2 - Material flows within the citrus-plant nursery chain.

2. RESULTS To perform the requirements analysis phase, data acquisition was performed by interviewing the computer-based system end-users and by collecting existing materials concerning the functions that the computer-based traceability IS should provide, for instance, regulations in force (e.g., laws, standards, and rules), farm rules and procedures. The development of this phase had generated a document in which the requirements and goals were formally defined. In this section the results of step f) and step g) of that phase (Fig. 1) are summarized since they constitute the input of the conceptual modeling phase. In detail, Tab. 2 reports a description of traceability system information, whereas Tab. 3 lists and describes the functions to be implemented in the system.

UML use-case view was used to carry out conceptual modeling phase of the life cycle of the computer-based traceability system. Firstly, actors were identified and assigned to a user profile. In detail, the actors belonging to the Nation Service of Voluntary

Certification, i.e., CCP, CP, CM, and Supplier, were associated to the *NSVC* profile, whereas Breeder, Farmer, and Phytosanitary Service (PS) were associated to the *external_NSVC* profile. Afterwards, use-case diagram was elaborated (Fig. 3) by using the information contained in Tab. 3.

Each user can activate a set of system functionality associated to a well-defined actor role by means of a personal login. To characterize the use cases, sequence diagrams were used. To this aim, each use case was decomposed into simple interactions between object (i.e. actors, classes, system components). Fig. 4 shows the sequence diagram of two use cases: 'information request about a nonconformity' and 'trace'.

INFORMATION	DESCRIPTION
NSVC's centers	*CCPs, CPs, CMs, increase blocks, and nurseries*: incoming plant and propagating material verifications, visual and genetic verifications, phytosanitary control, and transfer of materials to the next center. *CCPs and CPs*: screen-house characteristics, thermotherapy treatments, characteristics of grafts and rootstocks. *CMs*: cultivation practices. *Nurseries*: cultivation practices, grafts, and rootstocks.
Primary sources	Information gathered from the National Register of Accessions of cultivars, clones, and selections which have already been certified. This register is kept and updated by the Operating Secretary of the NSVC within the Ministry for Agricultural and Forestry Policies, as instituted by the decree of July 24, 2003.
Plants and propagating materials	Pomologic characterization and protocols of any trials carried out to verify the sanitary state of plant material.
Farmers	Information provided to farmers (certification card): plant code, plant botanical name, category, species, variety, family, accession code, health status, rootstock variety, nursery code, producer name, and data related to supervising or certifying authorities.

Tab. 2 - Information to be managed by the computer-based traceability system.

FUNCTIONS	DESCRIPTION
Traceability	"Track" function recognizes all the citrus plants derived from grafts and rootstocks identified by a code number. "Trace" function, by using the lot number or other identification datum of a citrus plant, provides data on the processes carried out in the nursery or in the centers (CCPs, CPs, CMs, and increase blocks) to obtain the plant or the propagating materials.
Nonconformity management	This function identifies the origin of the propagating materials used to produce the citrus plant(s) when a nonconformity has been identified. By using the code number of the noncompliant propagating material, the function recognizes the nurseries where other noncompliant plants are present, as well as the final customers.
Statistics	This set of functions regard activities undertaken by various organizations in the chain (marketing, cultivation practices, etc.).
Data management	These functions regard data acquisition, storage, production, processing, and communication.
Backup and disaster recovery	A set of technological measures and organizational processes to backup data on different media and recover the computer-based IS (data and infrastructures) in serious emergencies.

Tab. 3 - Functions to be implemented in the computer-based traceability system.

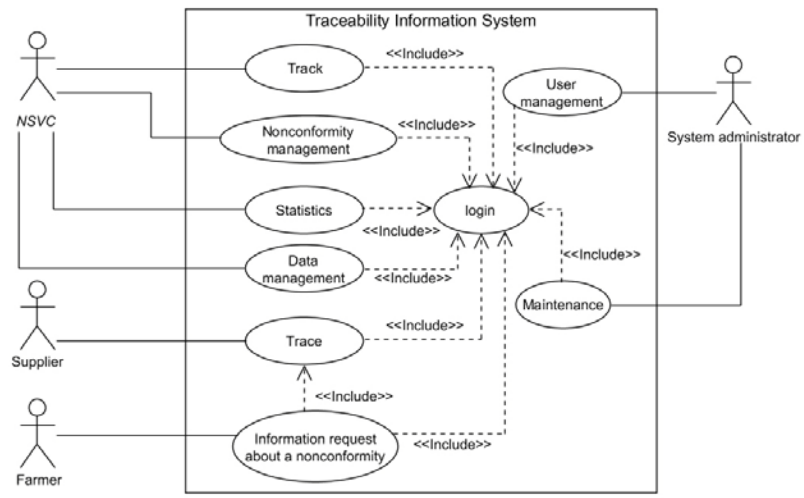


Fig. 3 - Use-case diagram of the computer-based traceability system.

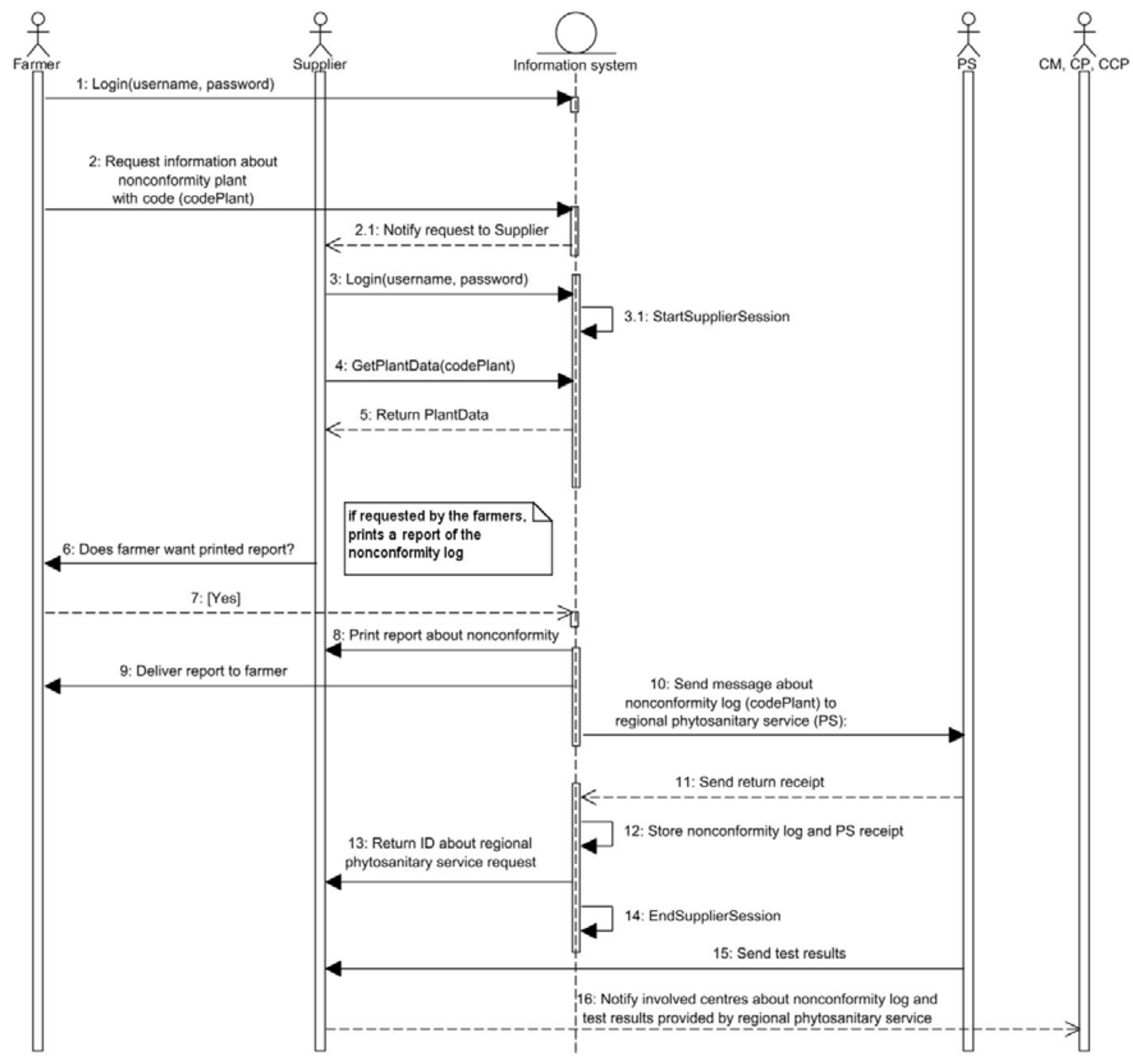


Fig. 4 - Sequence diagram of 'information request about a nonconformity' and 'trace' use cases.

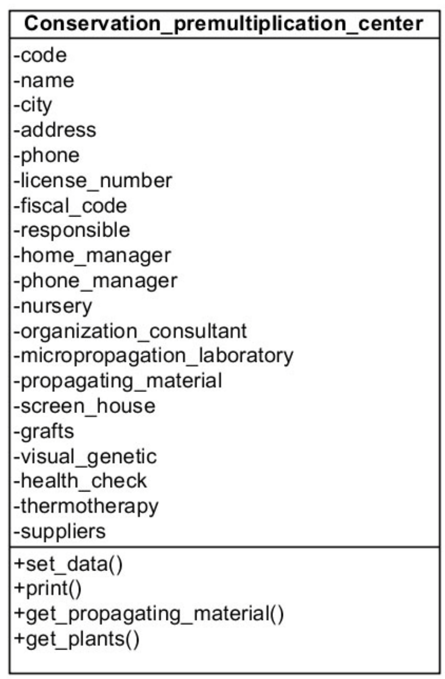
The interaction objects are reported at the top of the diagram. For every object a vertical line (life line) represents the object’s life during the interaction. Vertical rectangles, called activation bars, indicate when an object is executing an operation. An interaction between two objects is performed as a message represented by an arrow between their life lines. In particular, a dashed arrow indicates the return value of an object request.

Class diagram was used to describe data which the system handles in the use cases. A set of classes corresponds to the identified actors (Fig.5a), other classes were derived from the information contained in Tab. 2. Each class is characterized by a name, a set of attributes and a list of functions (methods) operating on attributes belonging to the same class or to a associated class. For instance, Fig. 5b shows the details of the *Conservation_premultiplication_center class*. In particular, methods *set_data* and *print* allow to store and print data of the same class, whereas *get_plants* and *get_propagation_material* operate on attributes of *plant* and *propagating material* associated class.

3. CONCLUSION The methodology proposed in this study makes it possible to implement, following the UNI EN ISO 22005:2007 standard, a set of traceability procedures by designing a computer-based traceability system. The methodology was applied to a case study to define the phases of conceptual modeling for the nursery chain of citrus plants certified under the Italian Ministry Decree of May 4, 2006. The conceptual modeling phase produced, as its final result, a set of UML schemes describing information and functions of the computer-based traceability system. These schemes allow to develop the next phase of the life cycle of the computer-based traceability system. This phase provides the production of the system logical model where data

Classes derived from actors	*Breeder* *Conservation_premultiplication_center* *Premultiplication_center* *Multiplication_center* *Nursery* *Farmer* *Phytosanitary_service*
Other classes	*Increase_block* *Propagating_material* *Plant* *Pomological_characteristic* *National_register_accession* *Genetic_phytosanitary_test* *Certification_card*

(a)



(b)

Fig. 5 - (a) Classes used in the class diagram. (b) Details of one of the classes.

structures can be specified in detail together with the most suitable type of Data Base Management System (DBMS) to contain them (e.g., relational, hierarchical, reticular, object-oriented). As regards traceability functions, they can be implemented, by using an appropriate programming language, in the phase called 'Application development' within the stage of traceability system implementation.

REFERENCE

Atzeni, P., Ceri, S., Paraboschi, S., & Torlone, R. (2002). *Basi di dati. Modelli e linguaggi di interrogazione* (*Databases. Query models and languages*), McGraw-Hill, Milano.

Bar-Joseph, M., & Lee, R., F. (1989). *"Citrus tristeza virus"*. AAB Descriptions of Plant Viruses n. 353. AAB, Wellesbourne (GB).

Batini, C., Pernici, B., & Cantucci, G. (edited by). (2001). *Sistemi Informativi* (*Information Systems*): vol. I, II, III, V - Franco Angeli Editore, Milano.

Bertocco, M. (2004). Più facile la tracciabilità con un sistema informatico (Easier is traceability by using an information system). *Colture protette*, 12: 41-47.

Booch, G., Rumbaugh, J., & Jacobs, I. (1999). *The Unified Modeling Languages User Guide*. Addison-Wesley.

Cascone, G., Arcidiacono, C., Porto, S. M. C., Anguzza, U. (2009). Progettazione di un sistema informativo per la tracciabilità delle piante di agrumi: il caso studio della filiera vivaistica (Information system design citrus-plan traceability. The case study of the nursery supply chain). IX Convegno Nazionale dell'Associazione Italiana di Ingegneria Agraria (IX National Conference of the Italian Association of Agricultural Engineering), Ischia Porto (Italy), 12-16 september, 2009, http://www.aiia2009.it/CD/pdf/5-8.pdf.

Cherry, C., & Macredie, R. D. (1999). The importance of context in Information System design: an assessment of participatory design. *Requirements engineering*, 4: 103-114.

Dunn, C., L., Gerard, G., J., & Grabski S., V. (2005). Critical evaluation of conceptual data models. *International Journal of Accounting Information Systems*, 6: 83-106.

Gmitter, Jr., F. G., Deng, Z., & Chen, C. (2007). Cloning and characterization of disease resistance genes in Citrus. In: *Citrus Genetics, Breeding and Biotechnology* (pp. 287-305). Kahn, I. H. (Ed.). CAB International, Oxfordshire.

Information Society Technologies. (2005). GeoTraceAgri final project report (GTA). URL: http://www.geotraceagri.net/ Accessed May 16, 2006.

ISMEA - Istituto di Servizi per il Mercato Agricolo ed Alimentare (Service Institute for agricultural and food market). (2006). *Filiera agrumi in Calabria* (*Citrus supply chain in Calabria*). ISMEA, march 2006.

Lee, R., F., & Bar-Joseph, M. (2000). *"Tristeza"*. *Compendium of Citrus Diseases*. 2nd edition. Ed. Timmer W., Garnsey S.M. & Graham J.H., APS Press, St Paul (US): 61-63.

Ministry of the Agriculture and Forestry. (2003) Decree of July 24, 2003 - Organizzazione del servizio nazionale di certificazione volontaria del materiale di propagazione vegetale delle piante da frutto (Organization of the National Service of Voluntary Certification of fruit plant propagating materials). Official Gazette of the Italian Republic n. 240 of October 15, 2003, Rome.

Ministry of the Agriculture and Forestry. (2006). Decree of May 4, 2006 - Disposizioni generali per la produzione di materiale di moltiplicazione delle specie arbustive ed arboree da frutto, nonché delle specie erbacee a moltiplicazione agamica (General principles for the production of shrubbery and fruit tree propagating materials as well as of herbaceous species characterized by agamic propagation). Official Gazette of the Italian Republic n. 168 of July 21, 2006, Rome.

Niederhauser, N., Oberthür, T., Kattnig, S., & Cock, J. (2008). Information and its management for differentiation of agricultural products: The example of specialty coffee. *Computers and electronics in agriculture*, 6 (1), 241–253.

OMG (Object Management Group). (2002). The object management group site. Unified modelling language specification. Ver. 1.4 ftp://ftp.omg.org/pub/docs/formal/01-09-67.pdf.

O'Docherty, M., (2005). Object-Oriented Analysis and Design: Understanding System Development with UML 2.0. John Wiley & Sons Ltd, West Sussex, England.

Opara, L., U. (2003). Traceability in agriculture and food supply chain: a review of basic concepts, technological implications, and future prospects. *Food, Agriculture & Environment*, 1, 101-106.

Regattieri, A., Gamberi, M., & Manzini, R. (2007). Traceability of food products: General framework and experimental evidence. *Journal of Food Engineering*, 81: 347-356.