Towards Geographic Requirements Engineering

Mounir Touzani, Anne Laurent, Thérèse Libourel Rouge, Joël Quinqueton

Towards Geographic Requirements Engineering

M. Touzani[1], A. Laurent[1, 3], T. Libourel[2, 3], J. Quinqueton[1, 4]

1. LIRMM (UM, CNRS), Montpellier
2. Espace-Dev (UM, UAG, UR, IRD), Montpellier
3. University of Montpellier (UM)
4. University Paul Valery Montpellier (UPVM)

mtouzani64@gmail.com
laurent@lirmm.fr
libourel@lirmm.fr
jq@lirmm.fr

ABSTRACT. It is now well recognized that information management systems often deal with geographic information even if they often do not take advantage of it. In these systems, requirements engineering (RE) is becoming more and more important in helping stakeholders to model their requirements regarding the quality of the data.
However, to date, few RE projects take geographic information and geographic constraints into account. We thus propose to study in this paper the question of managing such geographic constraints in the framework of RE. We propose an extension of the RE model and its implementation within the KAOS framework, and the use of geographic ontologies.

KEYWORDS: Information management systems, Requirements Engineering, Geographic Information, Geographic Ontologies.

Submission to KMIKS, March 25, 2015.

1. Introduction

Experts estimate that most of the information handled in information management systems is geographic. Employees and suppliers addresses are often known, however, these data are not used with their geographic specificities. For many years, the specific management of geographic information was indeed solely performed by experts in so-called geographic information systems. Geographic data are now becoming very popular, easy to manage and massively used in many systems, from management information systems [Ser 06] to Web and social networks (e.g., GoogleMaps, Open-StreetMaps).

Many organizations now exploit such geographic information, thus requiring high quality. Bad quality can impact the whole system and result in considerable costs and losses. The quality of geographic data is crucial and sensitive. Bad quality can lead to tragic accidents and/or loss of money and time.

The context of our research lies in the design and development of managerial IS (e.g., universities), which requires the use of RE techniques.

More specifically, we focus on the evolution of such a system by considering the following two challenges :

– Taking spatial dimension and its impact in a multiple-localization framework into account (the geographic localization of administrative departments, curriculum, research laboratories are of crucial importance);

– Taking restructuration requirements into account in the RE framework.

Nearly thirty years ago, Brooks [BRO 87] stated, “The inability to produce complete, correct, and unambiguous software requirements is still considered the major cause of software failure today”. This statement is still valid.

As Brooks, we believe that it is essential to take into account the specificities of spatial information in the RE cycle. We illustrate this point through the KAOS method.

This paper is organized as follows : Section 2 presents the state of the art for spatial data specificities. In section 3 we give some definitions of requirements engineering (RE). Section 4 introduces our proposition while Section 5 illustrates it with a case study. Section 6 concludes and proposes some perspectives for our work.

2. Geographic Information

Geographic information (GI) describes objects, phenomena or actions from real world [BEC 90]. It provides information on the name, type, shape, color of the object, together with its geographic location or information from nearby objects. The localization of objects is defined by Cartesian coordinate system (rarely used) or geographic coordinate system (latitude, longitude) and system of projection for planar maps (3D to 2D). These projections are either **(a)** equivalent to conservation of surface reports or **(b)** comply with conservation angle measurements observed in the field for navigation (in France: Lambert 93 Projection).

Indeed, the perception of geographic information can be defined in two views : either discretized or continuous [Ser 06].

This perception is currently translated via ontologies. From the definition of Thomas R. Gruber, an ontology is defined as the representation of a conceptualization [Gru93].

The usual two kinds of ontology are[1] :

– An object-based ontology describes the world as a space that is filled with discrete, identifiable units (i.e., objects), usually in the form of geographic coordinates (houses, factories, roads, rivers, lakes, or pollution plumes).

– A field-based ontology, describes the world as a collection of spatial distributions of phenomena (e.g., elevation or temperature) in the particular space. In fact, we look at all of the locations we are interested in, and we determine how much of that attribute or what category of that attribute is linked.

In this world-view, you might think of location as being an independent variable, and the attribute of interest as being a dependent variable (Worboys, 1995). As GI-Scientists, we are not usually really worried about whether some phenomenon (e.g., a mountain or lake) exists, but in how best to describe that phenomenon using numbers in a computer. A vector data structure is a computer implementation of an object based ontology, while a raster data structure is a field-based implementation.

In terms of modeling of geographic information must include the spatial relationships between objects that are as important as the entities themselves [PAP 94].

Many directions have been taken to define spatial relations and specify three classes of spatial relations i.e. topological, projective and metric.

– Metric relations are of distances or angles. They can be defined by measurable methods (e.g., the town is located 5km away from the beach), cognitive methods (e.g., forest is near river), or fuzzy methods.

– Topological relations are about connections between objects. These relationships are generally defined by measurable methods (e.g., via the DE9IM matrix [EGE 89]), but can also be expressed by terminologically cognitive methods (e.g., next to, touches, within).

– Projective relations are described by space projections such as cardinal relationship (e.g. east of, north of) or orientation relations of the objects against each other (e.g. left, down, front).

The quality of such data has always been a challenge [DEV 05]. Geographic data is specific and must be distinguished from other data as they :

[1] https://www.e-education.psu.edu/geog486/l1_p8.html

– require specific skills;

– contain or have subjective descriptions that can lead to various interpretations.

They correspond to several potential perceptions (multi-percetion); for instance a road can be seen as an object in a graph if the aim is to manage the traffic; as a surface if the aim is to manage the asphalt, as a volume if the aim is to manage the gas and water networks;

– made of imprecise descriptions (e.g. “this place is close to this other one”);

– require their quality to be dealt with [DEV 05];

– may require visualization tools;

– require meta-data, especially for relying to standards and master data.

Correctly setting data quality does not only require information on the data being used but also on the needs of users [DEV 05].

Quality is defined as internal and external quality concepts [Dav 97] :

– the internal quality is measured as the difference between the data that should be generated and the data that were actually produced. It concerns the errors, measured using the data producer;

– the external quality is measured by the difference between the desired data by the user and the actual data produced. It is linked to user requirements and varies with each user. This is the adequacy of the specification requirements of the user (fitness for use).

Two types of conflicts can thus emerge : producer/producer and producer/user.

Assessing the quality is also making a description of the data or metadata. The current standard for metadata geographic information is the ISO 19115. This standard takes into account mainly the view of the data producer. The metadata are used to exchange data and information between different users on the quality of the data by filling the information related to the data produced.

Metadata is “data about data”, and is defined as the data providing information about one or more acpects of the data.

Five quality criteria have been specified by [nat87] : genealogy, position accuracy (geometric), attribute accuracy (semantics), completeness and logical consistency.

Others were added: news (temporal accuracy) [GUP 95], textual fidelity and semantic coherence [SAL 95].

A misunderstanding of data (inaccuracy) or an imprecise qualification of a system characteristic of natural language [SME 96] can also lead to spatial data quality imprecision.

Below are two examples of inaccuracies :

– “The monument is higher than 100 meters”. In this case the possible values of the height may be between 80 and 120 meters;

– “In the amphitheater there are about fifty students”. In this case the inaccuracy in modeling is to formalize the term about.

As we have mentioned in the abstract, we propose to use the existing geographic ontologies.

3. Requirements Engineering

Requirements engineering (RE) is a discipline capable of guiding us toward the modeling, collection and formalization of the requirements linked to geographic information.

RE can specify concise, accurate and complete requirements reflecting an expressed need [LAM 09].

A requirement is a condition or capability that the user needs in order to solve a problem or achieve a goal [IEE 98].

A requirement is a condition or capability that must be met by a system or system component to satisfy a contract, a standard, specification, or other formally imposed document [IEE90].

According to AFIS[2], Requirements Engineering aims at establishing and maintaining a single repository through methods, rules and processes. It is completed during the development and is maintained throughout the life of the system. The basic elements of this discipline are the needs and requirements. Contrary to the requirements that represent a vision of the perceptions and expectations of the end user, the requirements are the designer’s vision of these needs.

Specifically, Requirements Engineering process is composed of four steps [SOM 98] : elicitation, analysis, specification and verification.

Among the methods used in requirements engineering, the KAOS[3] method is well recognized [VAN 01].

After comparison of various RE methods, at first level, KAOS seems to be the most relevant to manage geographic information, regarding to many criteria (simplicity, easy understanding and communication around, reuse), reading diagrams allow to catch up the information and detect an emerging ideas.

---

[2] http://www.afis.fr

[3] KAOS : Knowledge Acquisition in autOmated Specification

Even after approval by stakeholders, requirements may continue to evolve over time due to errors or context change requests. This requires to effectively manage these requirements during a project, or even after the completion of a project.

KAOS has the advantage of managing traceability, prioritization and changes in the requirements.

At a second level, KAOS consists in a goal-oriented approach that provides a graphical specification language for prioritizing objectives and sub-objectives, conflicting elements, constraints and obstacles. It is also possible to assign requirements to officials of each operational objective, and how and when to meet this goal.

A KAOS model contains the goals and requirements on the computer system, the expectations on the environment of the system, the conflicts between objectives, the obstacles, the entities, the software agents, the operations to implement, the refinements of goals into sub-goals and the purpose of obstructions by obstacles. Goals must be refined until the software operations to realize them. The information provided by the purposes to detect and resolve conflicts resulting from multiple points of view and propose alternative combinations for the specified constraints can be met [LAM 09].

The KAOS metamodel consists of four models [LAM 09] :

1) Goals model is to define the goal that the system must be achieved by the cooperation of agents;

2) Responsibilities model means describes the responsabilities of the agents in the implementation of the requirements that are assigned to them;

3) Operations model describes how agents need to cooperate to make the system work;

4) Objects model represents and brings up all the elements constituting an objective in view of the objects, which must be categorized entities, relationships, events and agents.

4. Proposition

In our research context, we propose to analyze and define geographic requirements, using KAOS as a method of requirements engineering. A case study illustrating this proposal is proposed in the section 5.

In performing this study, we will benefit from knowledge drawn from existing geographic ontologies.

We thus consider building a repository of quality requirements, according to the criteria and themes related to the field of activity.

The proposed approach is to summarize the development process [WIE 13] and adapt to geographic information using the KAOS method :

Figure 1 illustrates this approach :

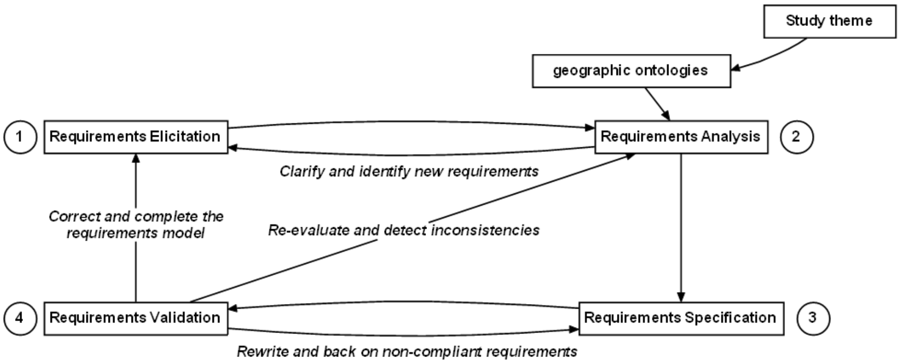

Figure 1. Development process of geographic requirements

1) Requirements Elicitation. This step allows stakeholders to express their need and to define the spatial requirements. At this stage, we have a requirement in natural language (text) and high-level goals to answer the question WHY?

The goals of KAOS models and responsibilities are used to perform the following actions :

– Identifying stakeholders;

– Organizing meetings;

– Carrying out interviews;

– Using questionnaires;

– Identifying high level objectives or goals;

– Revealing the rules and geographic constraints;

– Considering geographic standards;

– Identifying responsibilities (actors).

2) Requirements Analysis. At this level, gathered goals are refined and structured into sub-goals to answer the question HOW? This is the step where we need to take advantage of existing ontologies and define business rules and spatial constraints. Requirements can be identified. Missing information can be completed iteratively. This step comes along with any KAOS models (goal, responsibility, operation and object modeling). Among other actions to perform, we consider :

– Refining and enriching the goals;

– Restructuring the collected goals;

– Identifying the studied theme;

– Considering existing geographic ontologies according to the study theme;

– Identifying incompleteness;

– Detecting conflicts and suggesting workarounds;

– Detecting obstacles and propose resolutions or alternatives.

3) Requirements Specification. We aim at providing detailed documentation to stakeholders. A KAOS report can be generated by the "**Objectiver**" software.

The general geographic ontologies can be used according to theme designed for business requirements.

According to [4], a geographic ontology example set by Robert Laurini (Figure 2) :

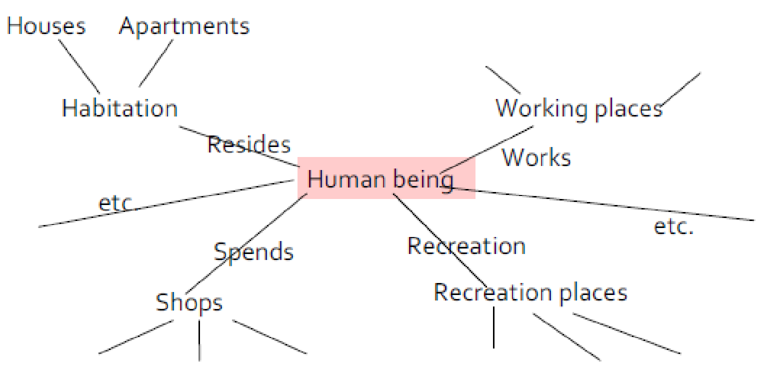

Figure 2. Beginning of an urban ontology

4) Requirements Validation. This activity aims to ensure that the product design meets the needs of users expressed as requirements [BOE 88].

It is a communication tool that will allow stakeholders to check the compliance with the requested requirements. The aim is to review in detail the points that are not clear enough before committing.

With KAOS, alternative solutions can be modeled in order to achieve the objectives. The documents provided by the requirements specification phase will serve as a bargaining tool with stakeholders to effectively build the model goals.

Validation generally used to detect errors in requirements, keep account of conflicts and obstacles to better resolve them.

---

[4] http://liris.cnrs.fr/Documents/Liris-5671.pdf

In our article, we propose to use a systematic method for developing and managing geographic requirements. So we make the choice of a goal-oriented method (KAOS) and understandable by the stakeholders. We will show through a case study it is possible to extend the use of KAOS operating on geographic objects (spatial type).

It is also possible to take advantage of existing geographic ontologies to reflect all geographic constraints tailored to each theme (tourism, education, ecology...).

5. Case Study

5.1. Description

We consider a problem that may result from the merging of two academic institutions being geographically distant.

Before the merging, we consider that every academic institution was taking some pieces of spatial information in a crude manner and were organized with their own specificities regarding human resource management. After the merging, spatiality becomes dramatically important as multiple localizations of some services are now considered. The merging raises new problems, as for instance the tradeoff to be made between proximity to private residence and the proximity to the spatial localization of the service. Every employee is offered to accept or reject the workplace he is been proposed to work in.

To take the spaciality into account, it is necessary to refer to ontologies (e.g., INSEE ontology[5] or IGN Data[6] or geonames[7] that provides geographic information for many places in the world), as well as the reference specifications issued by international standardization organizations: ISO [ISO 94], OGC [OGC 99].

The description of spatial objects in data models as illustrated by Figure 2 result from these specifications.

Employees of different services complete a form arguing their preferences. For this purpose, the application allows them to evaluate the duration of their itineraries.

From this analysis, an arbitration committee collects all information to make a decision and answers the empoyee through the managerial hierarchical system.

In general, one of the parameters to be taken into account by the business expert to evaluate and give their point of view could be to locate the ZIP code corresponding to the home address of the employee (this data may be derived from the Human Resource database).

[5] RDF Format - http://rdf.insee.fr/geo/index.htm
[6] http://data.ign.fr/
[7] http://www.geonames.org/

The quality aspect, the geographic, standards and existing ontologies will also have to be taken into account.

All these parameters are important to make a good assessment and identify improvement actions, while avoiding the conflicts that may result from the integration of geographic information within an Information System (IS).

Many benefits can come from such use. As part of a process of sustainable development, the transport sector is the perfect example to provide travel plans shared internally by the institution or by the urban communities. This would reduce the use of private vehicles and promote public transportation or biking.

Thus, the employee could benefit from the optimization of transportation infrastructure.

5.2. Requirements Engineering with KAOS

After identifying the stakeholders who participated in the meetings and interviews to obtain the first requirements, we made the assessment of goals that can be used as starting points for building the model goals.

However, geographic information remains at the heart of our research. That is why we focus on the objectives given set by manager of the building and logistics departement.

In this case study, we will use the "Objectiver" software to implement the KAOS method and its models (Figure 3) :

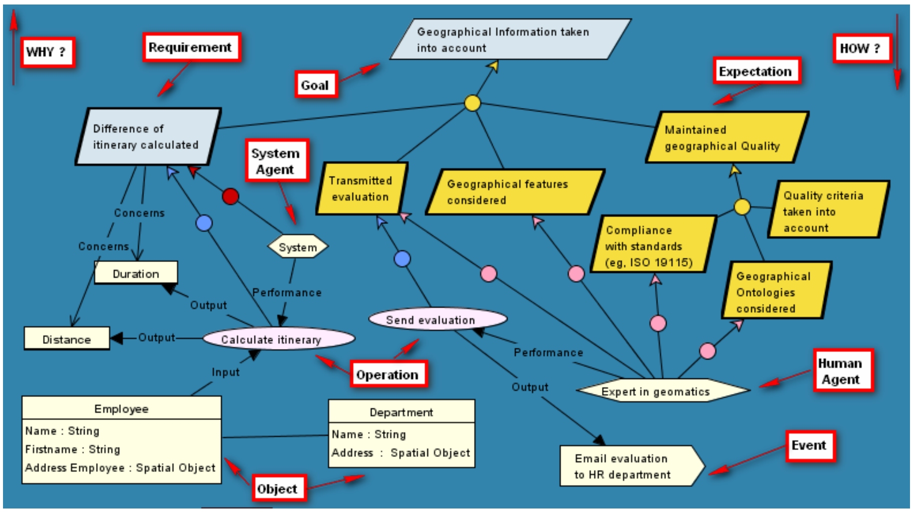

Figure 3. Case Study with KAOS Modeling

For the purpose of this modeling, the goal model is in the form of a tree goals, requirements, and expectations of their systems and/or human agents; the object model includes the UML objects and operations model represents various actions that will be performed by agents on objects to achieve a goal.

After requirements elicitation step, requirements analysis identifies the operations performed, the relevant objects assigned to these operations and goals, chosen requirements and expectations.

The modeling diagram of the case study is not complete. It helps to show how to apply the KAOS requirements engineering method for geographic information.

Obstacles and/or conflicts have not been addressed in this study, but may be detected during the validation activity.

6. Conclusion and Further Work

This paper introduces an extension of requirements engineering systems for managing geographic information. Such information are now becoming massively used and requires high quality management. These extension will be implemented in the KAOS method.

The perspectives opened by this work are various. The KAOS method could be enriched regarding to geographic information context, another perspective is to improve the quality of geographic information by the use of the needed referential, building master data management (MDM) system and developing business rules management system (BRMS) that could manage geographic information system.

Acknowledgements

The authors would like to thank the Région Languedoc-Roussillon for its financial support of this work through the HPC-Data@LR project.

7. References

[BEC 90] BECKER R., “Network Visualization”, 4th International Symposium on Spatial Data Handling, Zurich, Switzerland, , 23-27 juillet 1990.

[BOE 88] BOEHM B. W., “A Spiral Model of Software Development and Enhancement”, Computer, vol. 21, num. 5, 1988, p. 61–72, IEEE Computer Society Press.

[BRO 87] BROOKS F., “No Silver BulletâEssence and Accidents of Software Engineering”, Computer, vol. 20, num. 4, 1987.

[Dav 97] DAVID B., FASQUEL P., Bulletin d’information de l’IGN - "Qualité d’une base de données géographique : concepts et Terminologie ", N. 67, IGN France, 1997.

[DEV 05] DEVILLERS R., Qualité de l’information géographique, IGAT.: Série Géomatique, Hermès Science Publications, 2005.

[EGE 89] EGENHOFER M., “A Formal Definition of Binary Topological Relationships”, In Proc. FODO 1989, vol. 367 of LNCS, Springer, 1989, Page 457â472.

[Gru93] Toward Principles for the Design of Ontologies Used for Knowledge Sharing, Stanford University. Computer Science Department. Knowledge Systems Laboratory., 1993.

[GUP 95] GUPTILL S., MORRISON J., Elements of Spatial Data Quality, The International Cartographic Association Series, Elsevier Science Limited, 1995.

[IEE 98] IEEE, “IEEE Guide for Developing System Requirements Specifications”, IEEE Std 1233, 1998 Edition, , 1998.

[IEE90] IEEE 610.12- 1990 : Standard glossary of software engineering terminology, IEEE, 1990.

[ISO 94] ISO/TC211, “ISO geographic information/geomatics technical committees 211”, 1994.

[LAM 09] VAN LAMSWEERDE A., Requirements Engineering: From System Goals to UML Models to Software Specifications, Wiley, 2009.

[nat87] “A Draft Proposed Standard for Digital Cartographic Data”, 1987. [OGC 99] OGC/TC, “Opengis abstract specification”, 1999.

[PAP 94] PAPADIAS D., KAVOURAS M., “Acquiring, Representing and Processing Spatial Relations”, Presented at Sixth International Symposium on Spatial Data Handling, Taylor Francis, 1994.

[SAL 95] SALGÉ F., “{CHAPTER} {SEVEN} - Semantic accuracy”, GUPTILL S. C., MORRISON J. L., Eds., Elements of Spatial Data Quality, International Cartographic Association, p. 139 - 151, Pergamon, Amsterdam, 1995.

[Ser 06] SERVIGNE S., LIBOUREL T., “Fondements des bases de données spatiales”, Oct. 2006, Hermes Sciences-Lavoisier. Traité IGAT : Information Géographique et Aménagement du Territoire. 237 p. ISBN 2-7462-1378-8.

[SME 96] SMETS P., “Imperfect Information: Imprecision and Uncertainty.”, Uncertainty Management in Information Systems, p. 225-254, 1996.

[SOM 98] SOMMERVILLE I., KOTONYA G., Requirements engineering: processes and techniques, John Wiley & Sons, Inc., 1998.

[VAN 01] VAN LAMSWEERDE A., “Goal-Oriented Requirements Engineering: A Guided Tour”, Proceedings of the Fifth IEEE International Symposium on Requirements Engineering, RE ’01, Washington, DC, USA, 2001, IEEE Computer Society, p. 249–.

[WIE 13] WIEGERS K., BEATTY J., Software Requirements, Developer Best Practices, Pearson Education, 2013.