# PETALE: Case Study of a Knowledge Reengineering Project

Olivier Glassey[1, 2]

[1] Swiss Graduate School of Public Administration, Route de la Maladière 21,
1022 Lausanne, Switzerland
olivier.glassey@idheap.unil.ch
[2] Fraunhofer FOKUS, Kaiserin-Augusta-Allee 31,
10598 Berlin, Germany
olivier.glassey@fokus.fraunhofer.de

**Abstract.** In 2003 the author followed a knowledge reengineering project and workflow project in an administrative unit of the Canton of Vaud in Switzerland. The goal of this paper is to briefly present the case and to show what lessons were learned. We will show how this organization went from a mainly paper-based mode of operation to an integrated electronic workflow system and how we built a domain model using a combination of UML diagrams and RDF schemas.

## 1 Introduction

Between January 2003 and September 2003 we followed the development of an information system in the Service of Economy and Tourism (SET) of the Canton of Vaud in Switzerland. We focused on the knowledge engineering aspects of the project and the goal of this paper is to present the case study and to show what lessons were learned. We believe it is of interest because the SET is a typically knowledge-based administrative office with very few structured processes. Furthermore this organization went successfully from a paper-based mode of operation (although they used office suites) to an integrated electronic workflow system in less than a year. This integration was not only realized internally, but also with different external users.

The SET was created in 1998, out of a large-scale reorganization project, largely based on New Public Management ideas. It is a relatively small unit, employing 26 persons that belong to five main organizational sectors (Tourism Promotion, Local and Regional Development, Trade Police, Economic Promotion, New Technologies). Each of these sectors relies on designated legal texts to accomplish its tasks, for a total of more than twenty laws and decrees, both at the Federal and Cantonal levels. To give a brief overview of the variety of the SET's tasks, let us mention that in 2001 this service treated about 200 enterprises files in the domain of economic promotion and that it delivered more than 12'000 authorizations and business licenses for restaurants, shops, movie theatres, hawking, lotteries, concerts, ski teachers, etc. In

simple cases, the SET can take decisions without any form of notice from other public administrations or administrative units, but in more complex ones it might have to consult up to ten of them, at the communal, regional, cantonal and federal levels. For example, the SET has to check whether a company respects environmental and labor laws before considering tax exoneration. It also has to verify whether a building is compliant in terms of safety, hygiene and insurances before it delivers an authorization to operate a restaurants or a bars. All these decision processes are at the best semi-structured, although most of them are rather unique negotiation-based cases. For a detailed typology of different types of administrative services such as routine processes or individual case solving, see [1].

Before the introduction of the new system, SET employees processed all the incoming files "by hand": they had to check whether a file was complete and valid before they could register it. After a decision was taken, they also had to track the file and the following procedures: inspections, payment, renewal, etc. Finally they had to provide statistical data to the relevant administrative department. In order to accomplish these tasks the only software tools they had were a Lotus Notes contact base, three disconnected Access databases and several consolidated Excel sheets. In some extreme cases, software tools and paper-based files were overlapping five times without any type of "communication" between them, meaning that data had to be updated manually. Furthermore, the tracking index of each file consisted of a separate Word document, with the consequence that there was no automated way to see all files related to a designated company.

## 2 PETALE Project

The SET has five main strategic missions to fulfill and the PETALE system was designed to support most of the tasks realized within these five domains. However it is out of the scope of this paper to go into the details of these tasks. To illustrate the knowledge reengineering that the SET went through and the daily operations of this administrative service, we will concentrate on one example: delivering authorizations to operate a restaurant or a bar. Furthermore we decided to study only one area of knowledge analysis that the CommonKADS approach proposes [2]. This knowledge-modeling suite is based on three groups of models that answer three essential questions:

- *Why* is a knowledge system a potential solution? Which costs, benefits and organizational impacts does it have?
- *What* is the nature and structure of the knowledge involved?
- *How* must the knowledge be implemented in as system? What software architecture is suitable?

The SET developed a complete IT strategic plan before starting the PETALE project and we consider that it answers the first question quite thoroughly. Indeed this strategy provided amongst other things an inventory of existing procedures and a new organizational architecture. Furthermore this IT strategy pointed out that a new legal basis was needed in order to share data amongst the administrative services using the

PETALE system. A private IT development company was hired to implement the system and did a "state-of-the-art" job, with several mock-ups of the system and corresponding validation rounds with users representatives, hence we believe that the last question was cared for. Thus we focused on the nature and the structure of knowledge and in the following section we will show what was done during the project and we will propose a few additions in terms of knowledge models. The main processes of the SET were described in the IT strategy. However these descriptions were text-based and we wanted to somewhat formalize them. We followed the approach described in [3], based on use cases and scenarios, with addition of business rules. For more on use cases requirements we suggest [4] and for a complete presentation of business rules we recommend [5]. We will not explain that work here, as it is not directly related to the knowledge reengineering case we want to present.

However we will show one example of a UML collaboration diagram (Fig. 1): obtaining an authorization for restaurants, bars, nightclubs, sport clubs, street vendors or for special events such as parties. To get the full picture on this graphical notation language, we advise reading a reference book written by the creators of UML [6].

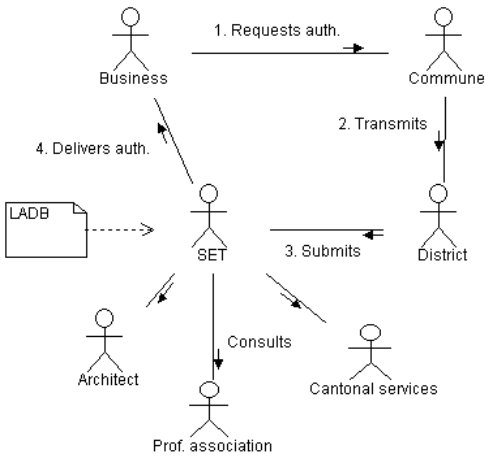


**Fig. 1.** A business requests a license from the commune where it is located. The commune then has to transmit this authorization request to the cantonal level through the prefecture (public administration at the district level), this procedure being the official channel of communication between a commune and the Canton. At the cantonal level, the SET relies on the "Loi sur les Auberges et Débits de Boissons", which is the legal basis treating all the cases related to selling drinks. In complex cases, the SET has to consult other cantonal services, professionals or business associations in order to deliver an authorization.

Most of this information can be found with full details in the IT strategy of SET [7] and in the PETALE users manual [8]. In the following section, we will present the domain model that we developed as a side-project: they were not used by the SET in order to develop the system.

## 3 Domain Models for PETALE

We were able to attend project meetings where software engineers, managers and user representatives gathered in order to define the scope and the requirements of PETALE. Furthermore lawyers were there in order to verify that the system respected the new regulation on data sharing and data privacy. Although terms such as "knowledge components" or "ontology" were never outspoken, the meetings were often revolving around these concepts. Before the PETALE project and in a paper-based mode of operation, these knowledge components were not explicitly identified: the SET employees simply knew that a "contact" could be a person, a business or a commune, that a "file" contained all similar authorizations related to one contact, that they had to look up by hand the different files with different authorizations related to a single contact, that a bar can have three "owners" (one for the building, one for the goodwill and one for the mandatory professional license "attached" to the bar), etc. These are just a few cases amongst many ambiguities that an experienced worker can handle but that need to be modeled in order for a computer system to work properly. Indeed, many hours of discussions were necessary in order to define concepts that were satisfactory to all the attendees. For example the concept of "signboard" was introduced at the request of the communes, because that is how a bar or a restaurant is identified at their administrative level: a signboard is linked to a physical address and a goodwill, moreover with the address one can find the owner of a building from land registers.

The software engineers developed a large and complex database architecture based on entity-relationship diagrams. A computer (or a skilled database developer) understands these models, but it is difficult for the average user to use them. In order to obtain a more explicit domain model, we created simple models based on Resource Description Framework (RDF) schemas. This W3C recommendation is an emerging standard that enables the definition of metadata for encoding machine-readable semantics [9]. They are shown here under a basic form of labeled directed graph, but let us quickly mention that they can be developed into full-scale knowledge models. With the use of tools such as Protégé2000, it is possible to generate the corresponding XML code or RDF statements under the form of triples that can be used in relational database [9]. Examples of the mathematical equivalence of these RDF statements (graphs, XML and triples) as well as further references are given in [10]. We believe that these RDF schemas provide the link between domain models that are understandable by human beings and domain models that are machine-readable.

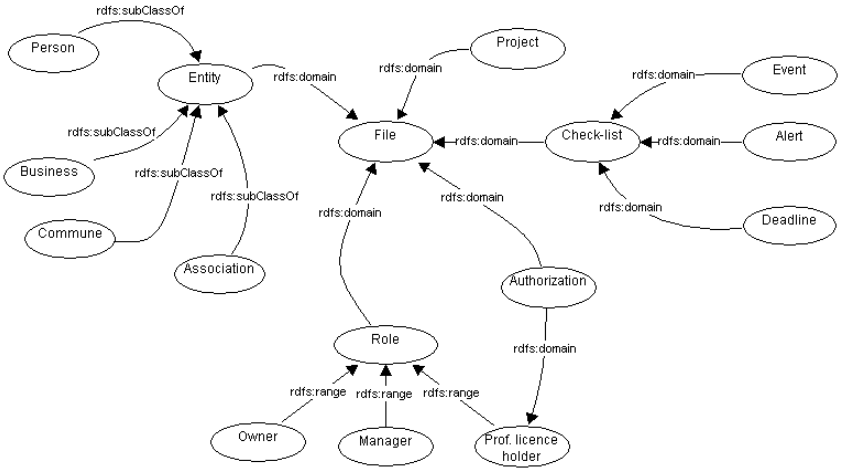


**Fig. 2.** A file is made of one or several authorizations concerning one entity and is completed by a check-list of temporal milestones containing data about special events that might take place at a site selling drinks (e.g. authorization to open later than usual for a particular occasion), deadline (e.g. annual renewal of the authorization) and the alerts that are sent to the PETALE user in charge of that file (e.g. check payment for renewal of authorization). The concept of project allows different files to be grouped thematically. An entity can be either a person, a business, a commune or an association (non-profit organization). Although this is not shown on the graph for readability reasons, owner, manager and professional license holder are all sub-classes of entity. They are linked to a file through a concept of "role". These three roles can be endorsed by a single person or by any combination of persons and companies. The only constraint is that the professional license holder must be an individual and the authorization is legally bound to the license holder.

The framework we propose is rather simple but we think it can be used efficiently for acquiring, representing and sharing domain knowledge, with both dynamic and static representation and a strong focus on the nature and the structure of knowledge, as we mentioned in the introduction.

## 4 Conclusions

The first authorization was printed on October 1st 2003, the same day the system became operational. The PETALE project was thus successful, for several general reasons: thorough reorganization prior to the project, sound IT strategy, strong support from the head of the SET service, high technical skills of the IT company, etc. We also believe that knowledge engineering was a key factor:

- The organizational aspects of the project were taken into account very carefully and detailed task analysis was made.
- All the actors concerned by PETALE were involved in the identification and refinement of knowledge components during the full cycle of the development,

from the first draft of the graphical interface to the different mock-up versions of the system.
- A functional prototype of the system was validated by a sample of users within the SET, in other public units of the Canton of Vaud and from four "test" communes.

As we already said in the text, most of the participants of the PETALE project probably would not speak of knowledge engineering or ontology development, although we judge that it is exactly what they did. However we feel that the use a formal methodology such as CommonKADS would have made their job easier, or at least the use of a graphical notation language such as UML for the description of information and control flows.

## Acknowledgement

This case study was realized thank to the support of M. Philippe Sordet (Head of the Service of Economy and Tourism of the Canton of Vaud) and his team.

## References

1. Lenk, K. & Traunmueller, R.: Perspektiven einer radikalen Neugestaltung der oeffentlichen Verwaltung mit Informationstechnik. Oeffentliche Verwaltung und Informationstechnik, Schriftenreihe Verwaltungsinformatik, Decker's Verlag (1999)
2. Schreiber, G., Akkermans, H., Anjewierden, A., de Hoog, R., Shadbolt, N., Van de Velde, W. & Wielinga, B.: Knowledge Engineering and Management: The CommonKADS Methodology, MIT Press (2000)
3. Glassey, O.: A One-Stop Government Prototype based on Use Cases and Scenarios. In Traunmüller R. & Lenk, K. (eds.): Electronic Government. 1st International Conference EGOV 2002, Aix-en-Provence, France. LNCS 2456, Springer Verlag (2002)
4. Kulak, D. & Guiney, E.: Use Cases: Requirements in Contexts. Addison-Wesley (2000)
5. Ross, R.: The Business Rule Book: Classifying, Defining and Modeling Rules, Version 4.0. Business Rules Solutions, Inc. (1997)
6. Booch, G., Rumbaugh, J. & Jacobson, I.: The Unified Modeling Language User Guide, Addison-Wesley (1999)
7. Minfray, M.N.: Schéma directeur informatique: Système d'information du SET/PPC. Document interne de l'Etat de Vaud (2002)
8. Elca Informatique, SA.: PETALE: Manuel d'utilisation DEC-SET. Document interne de l'Etat de Vaud (2003)
9. Noy, F.N., Fergerson, R.W. & Musen, M.A.: The Knowledge Model of Protégé-2000: Combining interoperability and flexibility (2000). Available at: http://www.smi.stanford.edu/projects/protege/. Accessed November 28, 2003.
10. Glassey, O.: Developing a One-Stop Government Data Model. In Bertot, J.C. (ed.): Government Information Quarterly, Volume 21, Issue 2, Elsevier Science (2004)