Jacques Teller
Chris Tweed
Giovanni Rabino (Eds.)

Conceptual Models for Urban Practitioners

ESCULAPIO

ESF provides the COST Office through an EC contact

COST is supported by the EU RTD Framework programme

Jacques Teller
Chris Tweed
Giovanni Rabino (Eds.)

Conceptual Models for Urban Practitioners

Proceedings (Reviewed Papers) of the
2nd Workshop of the COST Action C21 - Towntology
Urban Ontologies for an Improved Communication in Urban Civil Engineering Projects
Castello del Valentino, Turin, Italy

Jacques TELLER
LEMA - Universite de Liege
1 chemin des Chevreuils B52
4000 Liege
Belgium
Jacques.Teller@ulg.ac.be

Giovanni RABINO
Politecnico di Milano
Dipartimento di Architettura e Pianificazione
Piazza Leonardo da Vinci 32
20133 Milano
Italy
giovanni.rabino@polimi.it

Chris TWEED
BRE Trust Chair in Sustainable Design of the Built Environment
Welsh School of Architecture
Cardiff University
Bute Building,
King Edward VII Avenue
CF10 3NB Cardiff
United Kingdom
tweedac@cf.ac.uk

ISBN 00-0000-000-0

SOCIETA' EDITRICE
ESCULAPIO

December 2008 – Società Editrice Esculapio s.r.l.
40131 Bologna – Via U. Terracini 30 – Tel. +39.050.63.40.343 – Fax +39.051.63.41.136

www.editrice-esculapio.it

Foreword

COST- the acronym for European COoperation in Science and Technology - is the oldest and widest European intergovernmental network for cooperation in research. Established by the Ministerial Conference in November 1971, COST is presently used by the scientific communities of 35 European countries to cooperate in common research projects supported by national funds.
The funds provided by COST - less than 1% of the total value of the projects - support the COST cooperation networks (COST Actions) through which, with EUR 30 million per year, more than 30 000 European scientists are involved in research having a total value which exceeds EUR 2 billion per year. This is the financial worth of the European added value which COST achieves.
A "bottom up approach" (the initiative of launching a COST Action comes from the European scientists themselves), "à la carte participation" (only countries interested in the Action participate), "equality of access" (participation is open also to the scientific communities of countries not belonging to the European Union) and "flexible structure" (easy implementation and light management of the research initiatives) are the main characteristics of COST.
As precursor of advanced multidisciplinary research COST has a very important role for the realisation of the European Research Area (ERA) anticipating and complementing the activities of the Framework Programmes, constituting a "bridge" towards the scientific communities of emerging countries, increasing the mobility of researchers across Europe and fostering the establishment of "Networks of Excellence" in many key scientific domains such as: Biomedicine and Molecular Biosciences; Food and Agriculture; Forests, their Products and Services; Materials, Physical and Nanosciences; Chemistry and Molecular Sciences and Technologies; Earth System Science and Environmental Management; Information and Communication Technologies; Transport and Urban Development; Individuals, Societies, Cultures and Health. It covers basic and more applied research and also addresses issues of pre-normative nature or of societal importance.

Website: http://www.cost.esf.org

This publication is supported by COST

Table of Contents

Introduction

The domain of Transport and Urban Development (TUD) of the European programme for Cooperation in the field of Scientific and Technical Research (COST) aims at fostering international research networking activities of scientists and experts dealing with transport systems and infrastructures, urban land use and development, architecture and design, and civil engineering issues. The focus is on interdisciplinary approaches and the intention is to cover both basic and applied research activities including technical and technological developments and their changeovers that are relevant to policy and decision making processes. A significant concern is devoted to activities exploring new research needs and developments.

This domain is by definition cross-sectoral and multidisciplinary, encompassing a wide range of scientific expertises within the transport and land use planning, design, and management activities with a special emphasis on the strong interrelationships among the relevant policy fields as well on all aspects related to sustainable development. The domain activities should be innovative and complementary to other European programmes in the relevant fields.

The main objective of the Action C21, in the TUD domain, is to increase the knowledge and promote the use of ontologies in the domain of Urban Civil Engineering (UCE) projects, with a view to facilitating communications between information systems, stakeholders and UCE specialists at a European level.

Secondary objectives of the Action are:

- producing a taxonomy of ontologies in the UCE field, contrasting existing design methodologies, techniques, glossaries and production standards;
- developing an urban civil engineering ontology both in textual and visual (graph) presentation and a visual editor to integrate and update concepts, definition, photos, etc., into the ontology (software tool);
- developing a set of guidelines for the construction of multi-lingual UCE ontologies, based on practical examples (cases);
- analysing the role of ontologies as a tool to foster an improved communication between UCE stakeholders.

The Action C21, which is known informally as "Towntology", brings together a large and heterogeneous grouping from across Europe, whose interests range from construction to urban tourism from transport infrastructure to resource visualisation. On 17-18th October 2007, in Turin, Italy, the Action convened a successful workshop to address emerging issues in the field. This volume presents the contributions to that workshop, in order to capture the essence of the meeting.

The 2nd Workshop of the Action C21, which took place in Turin, was titled "Ontologies for Urban Development: Conceptual Models for Practitioners", and the emphasis was on the developing a deeper understanding of how ontologies work in practice with a view to informing the development of future ontologies and conceptual tools that will make communication between different urban development disciplines easier.

The theme of the Workshop was set by Working Group 3 (WG3) of the Action. WG3 is interested in the socio-technical issues that emerge during the development and use of ontologies in organisations. The development of this Group within an otherwise technically focused topic signals the importance of investigating the social and cultural context in which ontologies are expected to operate. It recognises that the introduction (and revision) of ontologies will impact on working practices within organisations. In most previous research, ontologies are conceived of as highly technical and formal tools for systematising and organising knowledge and information about different topics in the transport and urban development domain, without much thought about who will use them and how they will use them.

WG3 seeks to examine the use of ontological frameworks and systems in practice and to find out how actual use differs from designers' intentions and what we can learn from these anomalies to design better ontologies in future. The Call for Papers, therefore, emphasised the need to document experiences of using ontologies in practice and the focus of the Workshop was on ontologies in practice and the impact they have on the organisations in which they are used.

Despite this emphasis, the contributions to the Workshop remain largely technical and in some cases target quite detailed formal and even philosophical problems in ontology development. And yet there are hopeful signs in many of the papers that research in this area is at least keeping one eye on the eventual applications of ontologies and their take-up by end-users. Ontology development is not immune to developments in other areas and a several of the papers refer to the potential inherent in social networking as a means of building and maintaining ontologies or, in some cases, 'folksonomies' that are largely abstracted from large sets of tagging information provided by many users. The use of folksonomies is still in its infancy and there are few examples in the transport and urban development domain.

The first paper is provided by one of the keynote speakers, professor and philosopher Maurizio Ferraris, who works in the Laboratory for Ontology at the University of Turin. His contribution, titled "*Social Ontology*", is theoretically deep and it helps to elaborate and explain the complex matter of ontologies. This paper is complemented by the second keynote speaker's contribution, which is a good example of urban ontologies applied in practice. It is written by Philip Carlisle, from English Heritage, National Monuments Record in the UK, and it concerns "*Practical approaches to standardizing vocabularies: the Cultural Heritage experience*".

The first group of contributions considers the relationships between ontologies and environmental data, such as air quality and hydrology. This group includes: "*Ontologies for the Integration of Air Quality Models and 3D City Models*" (Metral, et al.), and "*Using a hybrid approach for the development of an ontology in the hydrographical domain*" (López-Pellicer, et al.).

Another important aspect treated in the Workshop is different web-based tools, as useful instruments to connect the users to data, through ontologies. These papers include: "*Web-based Interactive Visualization of Uncertain Concepts in Urban Ontologies*" (Ban, et al.), "*Generating an urban domain ontology through the merging of cross-domain lexical ontologies*" (Lacasta, et al.), "*An Ontology-Based Intelligent Information System for Urbanism and Civil Engineering Data*" (Trausan-Matu, et al.), "*Semantic search engine for geographic data*" (Pegoraro, et al.), "*Geo Semantic Web Communities for Rational Use of Landscape Resources*" (Marcheggiani, et al.), and "*Using an Ontology-based Model for Knowledge Representation in Rural Landscape*" (Cataldo, et al.).

Another group of contributions takes into account the policies, urban planning and urban development aspects. In particular, this refers to: "*Ontologies for urban regeneration: opportunities and weakness for their development in cohesion policies for cities*" (Rotondo), "*Colour plan for urban design*" (Ceconello, et al.), and "Generating Urban Forms from Ontologies" (Caneparo, et al.).

Finally, this volume includes contributions about landscape characterization and interpretation of design processes: "*Exploring Ontologies of Historic Landscape Characterisation: Towards an approach for recognising the impact of incremental change to historic legibility in urban areas*" (Dobson), "*Elaboration and application of ontology in a process of architectural project*" (Pellegrino), and "*Expressing Urban Development Concerns within a Domain Ontology*" (Athanassopoulou).

The absence of large numbers of papers addressing particular problems of using ontologies in practice could be taken as a good sign. Perhaps there is little to say, because ontology developers are getting it right and users have nothing to complain about. Maybe ontologies are not as problematic in practice as we might imagine. More likely, however, is that the very idea of ontologies in practice is so novel that many people are unaware of the terminology and so fail to recognise ontological issues when they arise. If that is so, then we are ahead of the game with this Workshop and have, in some small way, prepared the ground for research in this area in years to come.

Social Ontology and Documentality[1]

Maurizio Ferraris
University of Torino
maurizio.ferraris@labont.it
COST ACTION C21 - TOWNTOLOGY
Urban Ontologies for an improved communication
in urban civil engineering projects

1. Introduction

Encyclopedia entries, bets, gains and losses, research projects, books, lessons, relationships, votes, credits, exams certificates, exams, records, academic degrees, students, professors, art works and consumerist literature, cathedrae, aulas, application forms, hiring, revolutions, work-shops, conferences, dismissals, unions, parliaments, stock societies, laws, restaurants, money, property, governments, marriages, elections, games, cocktail parties, tribunals, lawyers, wars, humanitarian missions, voting, promises, buying and selling, prosecutors, physicians, perpetrators, taxes, vacation, medieval soldiers, presidents. What are all those objects made of? And, first of all, are they objects? Some philosophers would say they are not, since—according to them—only physical objects exist. Other philosophers would dare say that even physical objects are socially constructed, since they are the result of our theories. For real, thus, the world would be Prospero's world: *We are such stuff / As dreams are made on and our little life / Is rounded with a sleep*. That is not the case, though: social objects exist indeed, the proof being the difference between thinking to promise something, and actually promising something: once you give your word, the promise keeps on existing, even in case you forget about it, or—as more frequently happens—you change your mind.

The first aim of this article is to expand on the nature of social objects, as contrasted with physical and ideal objects, and to spell out the steps that lead to their discovery. Secondly, I will illustrate and criticize the major contemporary theory on social objects, John Searle's theory, and compare it with another theory, according to which social objects are a kind of inscription. Lastly, I want show how, from this standpoint, a social ontology evolves naturally into a theory of documents, which I propose to name "documentality".

2. Physical, Ideal and Social Objects

For a long time philosophers have underestimated the dimension of social objects, focusing exclusively on physical and ideal objects. This fact, probably is a consequence of an ambiguity concerning the nature of social objects, which is apparent as soon as we confront them with two other classes of objects into which reality can be divided. Physical objects, such as tables, lakes, occupy a place in space and in time, and exist even if we do not think about them; ideal objects, such as numbers, relations or theorems, differently from physical objects, do not occupy any place in space and in

[1] I would like to thank Pier Carlo Rossi and Giuliano Torrengo for important comments on the theses contained in this article.

time, but, as much as physical objects, exist even if we do not think about them. Social objects, on the other hand, such as marriages and graduations, occupy a modest amount of space (more on this later: it is, roughly, the amount of space a document occupies) and a more or less extended portion of time—they cannot be eternal though (differently from ideal objects, social objects seem to tend towards their own end: the theorem of Pythagoras is meaningful exactly because it is eternal, a promissory note for the opposite reason, i.e. because it will expire sooner or later; although there may be social objects, such as the Roman Empire or the Egyptian Dynasties, that last longer than the life of an individual). Thus, social objects looks like being somewhere in between the materiality of physical objects and the immateriality of ideal objects. I explain in details this point later. What I would like to underline in the first place, both in order to explicate why philosophers, and common people with them, have discovered social objects so late, and to draw attention to the most peculiar aspect of social objects, is the following: *differently from physical and ideal objects, social objects exist only in so far as there are men thinking that they exist*. Without men, mountains would remain what they are, and numbers would have the same properties they actually have—but it would be complete nonsense to talk about offences and loans, Nobel prizes and years in jail, art works or pornography. This feature has been misconstrued, and this fact has lead to the spreading out, in various ways, of a conceptual ambiguity. That is, the idea to the effect that social objects are utterly relative, or that they are nothing over and above a manifestation of the will. What is denied to social objects here is their object-like nature: they are reduced either to something indefinitely interpretable or to a bare psychological act.

We can find out how little true this reduction is thanks to a simple experience. I can decide to go to the cinema; if, eventually, I change my mind, this decision does not constraint me in any way. It is really just an expression of the will, which, since it has no outward manifestation, has a purely psychological dimension. Things are different if I propose to someone to come along with me to the cinema; if I change my mind, I have to tell her or him and provide a justification. What I have constructed, then, is an object that is not nullified by a bare change of my will. Let us further assume that my invitation had the form of a promise; for instance, I have told my son: “I promise you that, if you keep on being a good boy, I take you to the movie tonight”. Now, if I had told him only “I promise that”, I would have not promised; a promise has a beginning only when there is an object of reference, and a time limit, if only a vague one (“I promise you that sooner or later I quit smoking”). If, on the other hand, social objects were utterly relative constructions, they would not carry within them any necessity, and it should be possible for “I promise” to be a promise, but “I promise” is just the first singular person of the indicative present tense form of the verb “to promise”.

3. The discovery of social objects

This phenomenology of social objects should enable us to detect the features of social objects[2] that motivated their discovery. We are not dealing here, strictly speaking, with a historical progression (none of the following authors is likely to have ever read any of the others), but rather with a theoretical progression.

---

2 Di Lucia, ed., 2003; Ferraris 2003b; Gilbert 1989 and 1993; Johansson 1989; Kim-Sosa, ed., 1999; Moore 2002; Smith 1998, 1999, 2002; Tuomela 2002.

The first stage of the history consists in recognizing the specificity of social objects, and the first one who did this is the Italian Giambattista Vico (1668-1744)[3], who—quarreling with Cartesian rationalism and naturalism—defended the original character of the sphere of human interaction. In order to individuate this sphere, which identifies the passage from animal to man, and from nature to culture (the latter, thus, is essentially meant to be social progress), Vico points towards marriages, tribunals and burials. Those are social acts, that do not describe anything nor add anything new to the physical or ideal world, nonetheless they trace the change from the animal to the man, and from nature to culture.

The second stage of the history concerns the Scottish philosopher Thomas Reid (1710-1796)[4], who underlines the autonomy of social objects, distinguishing them form mere psychological construction or manifestation of the will. Reid claims that the premise for the constitution of a social object is an act concerning at least two people. As in the previous example, thinking about going to the cinema is not, whereas proposing someone to go to the cinema is, a social act.

The third stage, in the middle of the Nineteen hundred, amounts to the theory of linguistic acts, by the English philosopher John L. Austin (1911-1960)[5]. Linguistic acts are somehow an explication of the specific character of social acts. Social acts, insofar as they require to be expressed, are linguistic acts (we will see how this conclusion is partly misleading); and since they are not just a description of something (think to the "yes" said at a wedding cerimony, a paradigmatic example), but they produce something, they possess an original feature with respect to other parts of language. While saying "this is a chair" does not amount to acting upon the chair in any way, saying "the meeting is open" or "I hereby declare you doctor in philosophy" produces an object that was not there before.

The fourth stage of our history, relatively eccentric if compared with the previous ones, is provided by the German philosopher and law theorist Adolf Reinach (1883-1917)[6], who proposes a typology of social objects, described as a priori derivable (namely as endowed with a logical form, roughly what I was underlining when I let you notice that "I promise" is not a promise), and insists upon the fact that what is produced by social acts is not a self-contained *praxis*, but it is a *poiesis*, the construction of an enduring object (a graduation or a marriage, with respect to other social events, such as a party or a ruffle in which nobody gets seriously injured, have consequences that reach farther than the event).

3 Vico 1744.

4 Reid 1785.

5 Austin 1962.

6 Reinach 1913; see Mulligan, ed., 1987.

4. X counts as Y in C

In the contemporary debate, the standard theory of social objects has been proposed in the nineties by the American philosopher John R. Searle (n. 1932)[7]. The building up of this ontology can be described as a strategy in four steps.

The first step is set at Oxford, during the fifties, at the school of—among the others—John Austin, and continues at Berkeley during the sixties and seventies. Here Searle's activity focused on linguistic acts, an especially subtle part of language. When I say "yes" at a wedding ceremony, I am not describing anything that is already there, I am constructing something that is born in that very instant.

The rhapsodic analyses of Austin get a systematic dimension in Searle's work. Searle offers a complete classification of them[8], but this is not the only thing he is doing. On the one hand (and the upshots in social ontology derive from here), Searle does not limit himself to a classification of linguistic acts, but he acknowledges also the existence of objects that may be borne to life by performative acts—a particular kind of linguistic acts. A marriage and a conviction, for instance, understood as rites, may last just few minutes in their culminating moment. The corresponding social objects may last years though, and it is the philosopher's task to account for those objects' existence. By doing this, however, a philosopher should deliver a theory of mind too[9], since the peculiar feature of objects such as marriages or penal convictions, differently form cows or mountains, is that they exist only if there are minds believing that they exist.

And here it is Searle's second step, set in Berkeley, during the Eighties. Austin was exclusively concerned with language (and perception), Searle, I have just suggested, was seeking for a theory of mind. Could a computer passing the Turing test get married? Is a computer used in a betting shop really betting? Can it christen a boat? Can it bequest something to another computer? The obvious answer to all these questions is 'no'. And this depends on the fact that human mind has something that computers do not have: intentionality. Intentionality is the capacity to refer to things in the world, by using the representations we have at our disposal here, roughly, under the hair, beyond the eyes, and between the ears—namely in our minds. Intentionality, however, is not a ghost, a feeble mist descending on the world as postmodernists uphold, when they claim that Being is reducible to Language. Not at all, it is something as real as photosynthesis or digestion is. We should not misunderstand this point, because one thing is maintaining that human mind is not a computer, and quite another maintaining that Darwin was wrong. This is a very delicate knot, because claiming that the *individual I* is in many cases the result of a collective intentionality does not mean that reality is constituted in an inter-subjective way. Not at all, there are pieces of reality perfectly capable of staying by themselves, and they do not depend on language or conscience. Other pieces, surely, do depend on them. Still, we should not mix up those two cases, if we do not want every honest philosophy coming to an end.

---

7 Koepsell – Moss, ed., 2003; Smith 2003a.

8 Searle 1969 and 1975.

9 Searle 1980, 1983 and 1992.

The third step is set here and there around the world between the seventies and the eighties[10]. Searle assists—not inert, but astonished—to postmodernists spreading though the departments of comparative literature, under the risk that sooner or later they would get to philosophy departments too. Here is the crowd: who says that the Being that can be understood is language, and who says that nothing exists outside texts, and you would eventually meet a fun-loving person maintaining that there are no facts, but only interpretations. In conclusion, the moral seem to be that—quite paradoxically—words, but not things, exist; concepts exist, but not the objects they refer to.

We would be mistaken in seeing in this reaction to the postmodern idealism simply a polemical phase, since it is in this framework that Searle elaborates the theory of reality as 'background'[11]. Reality is something that does not require to be demonstrated, because it is at the ground of our demonstrations. Reality constitutes the basic element of Searle's general ontology, it delivers us the deep sense of his realism—and, at the same time, the deep sense of the non-realism of the postmodern speaker who, with their laptop, on the plane, polishes the talk s/he will give at a conference in an American University, on the topic of the non-existence of the external world.

And here the last step comes[12], which is set in Paris. During the nineties Searle enters in a Café, and pronounces a French sentence "*Un demi, Munich, à pression, s'il vous plaît*". Searle makes us notice that this very simple sentence triggers a huge invisible ontology: the social exchange between he and the waiter, a lattice of norms, prices, fares, rules, passports and nationalities, an universe of such a complexity that would have had Kant shivering, if only Kant had thought about it. We are the postmodernists' antipodes. If postmodernists dissolve tables and seats by reducing them to interpretations, Searle's ontology asserts that also things such as promises and bets, shares and debts, medieval knights and Californian Professors, tenures and symphonies possess their own specific reality. They are neither ghosts, nor movements of the consciousness or of the will (given that promises exist when we sleep, and in case we change our mind too, and contracts can bind institutions independently from the people who run them), they are higher order objects with respect to physical objects, in accordance with the rule '*X* counts as *Y* in C'—meaning that the physical object *X*, for instance a colored piece of paper, count as *Y*, a 10 euro banknote, in C, the Europe of the year 2006.

It is not hard to see how here we are approaching the closure of a system. The philosopher of language who has studied linguistic acts comes across performative acts, and notices that with them we can construct social objects; the philosopher of mind, who has studied intentionality, understands intentionality's role in the construction of social reality; the anti-postmodern polemist, in turn, elaborates a realist ontology that enable us to understand that—for some reasons and our intentions and hopes notwithstanding—it is useless to try avoiding to pay the beer by saying that reality (social and maybe physical reality as well) is socially constructed. The last move was left for the social ontologist: discovering this new reign of objects that—please notice—cannot be defined as 'mental' just because they need human minds.

---

10 Searle 1993a, 1993b, 1998.

11 Searle 1999.

12 Searle 1995.

5. No (social) thing exists outside texts

Very well. Now, we know that this theory (and Searle knows it very well too) has counterexamples, together with the general difficulty of clarifying the key notion of "collective intentionality"[13]. Even if we limit ourselves to considerations concerning the object, the problem is two-fold: it is not obvious at all how, from the physical object, we manage to get to the social object; and it is not clear at all how, given the social object, we should individuate a corresponding physical object.

In order to explain the shift from the physical to the social, Searle makes the example of the transformation of a wall into a boundary. The idea is the following: firstly, there is a physical object, a wall that divides the inside from the outside, and defend a community. Then, step by step, the wall deteriorates, and only a line of stones is left—unhelpful as a physical shelter—to identify a social object, namely a boundary: the very same that, later on, will be the yellow line that in the post-offices and airports allocates an insuperable threshold. Now, we can understand as a wall, by falling apart slowly, can, in certain circumstances, turn into a boundary. But it is not at al evident how, on the ground of this simple analogy—a lucky chance that who knows how few times occurred—the yellow line or the center line of the road are born. The question is further complicated by the following consideration: if really a physical object can constitute the origin of a social object, the *every* physical object would turn into a social object, every wall would signify a prohibition. But clearly this is not the case, as everyone who decides to tear down a wall in their house can verify, provided that the demolition would not contradict certain norms—which not necessarily concern the physical solidity of the wall. And, lastly, we should not forget that one of the most famous wall in contemporary history, the Berlin Wall, was begot by a boundary—the opposite of what should have happened according to Searle's explanation.

As for the second aspect of the problem—the aspect concerning the reversibility from the social to the physical sphere—it is rather intuitive to assert that a banknote is also a piece of paper, or that a president is also a person. As much as it is true that when Searle is alone in a hotel room there is only a physical object, but many social objects (an husband, a employee of the state of California, an American citizen, a driving license holder…). In this case, the passage back from Y the social) to X (the physical) goes smoothly. However, things change in different, although not very peculiar, situations. How should we deal with vague[14] or vast entities, such as a State, a battle, a university? And how about negative entities, such as debts?

The English philosopher Barry Smith (n. 1952),[15] has rightly pointed out that in many cases we have to acknowledge the existence of Y independent entities, namely entities that do not ontologically coincide with any part of physical reality. Here, according to Smith, we are dealing with "representations". In order to better define the

---

13 I have criticized at length this aspenct of Searle's theory in Ferraris 2005. Collective intentionality has been originally elaborated by the Finnish philosopher Raimo Tomela (1995), whereas the psychologist M.E. Bratman (1992) was concerned with "shared intentionality".

14 Williamson 1994 and 1998.

15 Smith 2003b.

notion of "representation", Smith qualifies it as a "quasi-abstract entity", providing as an example, a chess match played at random. The idea is that chess may be played independently from any physical support. You can play in Internet, where the chessboard is not "present" as a physical chessboard is (for instance, it has two localizations, corresponding to the two computers). Moreover, two experts can play by heart, without there being even a chessboard represented on a screen, but rather through two barely thought chessboards. Smith expands the model to the paradigm of money. Also in this case, from a certain moment on (and more and more as technology develops), we loose the physical counterparts, substituted by traces on the computer. Also in this case there is a social object to which it does not correspond any physical object, but rather a representation.

This is all fine, but really the computer *blips* are not physical at all? Are they really a *res cogitans* utterly detached from a *res extensa*? It takes only visiting a technological cemetery (a huge Chinese landfill, or the corridor of a Department where out of order computers have been stored) to realize how much plastic and silicon is necessary for magnetic traces to exist. And, unless we want to say that computers have souls, separated from their bodies, the *blips* will be material things as well. Indeed, it is difficult—nay, impossible—to uphold that, in the case of money that is transformed into traces on a computer, there are only representations, and not a physical thing sustaining them, although something endowed with a rather light physicality. But let as suppose that this is indeed the case, that representations do not need anything physical. Then, there would be no way to answer the question: *how should we distinguish in principle 100 real thalers from 100 ideal thalers?* How are we to distinguish the representation of 100 thalers from 100 merely imagined, or dreamed of, thalers ?[16]

The difficulties emerging both out of Searle's theory and Smith's correction help us to spot the way to the solution of the problem of social objects, which I propose to develop after the theory of the French philosopher Jacques Derrida (1930-2004)[17]. Derrida has elaborated a philosophy of writing that finds its most correct application in the social sphere. What is more interesting is that Searle knew this theory, but the alliance was rendered impossible by a reciprocal misunderstanding. Actually, Derrida dedicated an essay to Austin's linguistic acts[18]. Those acts, Derrida observed, are mostly inscribed acts, since without records of some sort the preformatives would not produce social objects such as conferences, marriages, graduation ceremonies, or constitutions. The point is simple, if we imagine a graduation or a wedding ceremony in which there are no registers and testimonies, it is difficult to maintain that a husband, a wife, a

---

16 If one maintains that it is false that a social object depends from a particular physical substrate, but it is true that every social object generically depends on some physical substrate (namely an inscription of same sort), one can keep on criticizing Searle's position (that concerns rather the fact that Searle points to the "wrong" physical substrate, somehow), and at the same time avoid Smith's "representational" conclusions. The chess match does not depend on a particular chessboard, neither depends it on two particular computers, or some particulars neurons. Still, if a match is there, then some physical substrate is also there, and therefore the match generically depends on some physical substrate. On the distinction between particular and generic dependence see Simon 1987: 296-7.

17 Derrida 1967; see Ferraris 2003a, 2006. On the social role of writing see Ong 1982.

18 Derrida 1971.

graduated person has been produced. This amounts to saying that social objects turn out to be (as much as the ideal ones) closely linked to the forms of their inscription and recording. That article irritated Searle, who few years after replied[19] (the reply was followed by a exceedingly long response by Derrida[20]) against what to him was nothing but a misunderstanding of Austin. Thus, the meeting seemed not to bring anything to the point. Still, we can see in it the solution of Searle's puzzle.

Indeed, the problem in Searle's social ontology depends on non having investigated the hypothesis that the physical counterpart of a social object is a trace, namely exactly what Derrida has brought attention to during all his career—be it a trace on the paper, or a trace in the brain, an inscription in the memory that remind us a promise, a debt, a duty or a fault. Derrida, having at hand the evidence to the effect that money has turned into inscribed paper, although not yet the (more striking) evidence to the effect that it would have turned into computer bytes, provided as soon as 1967, through his hypothesis on the nature of writing, the ground of an extremely powerful ontology. However, Derrida was wrong in claiming that "nothing exists outside texts" (and Searle was entitled to reproach him on this). Actually, as we have seen, physical and ideal objects exist independently from every recording, as much as independently from there being an humanity. This is not the case for social objects, which depends tightly on records and the existence of humanity. It is in this sense that, by weakening Derrida's thesis, I propose to develop a social ontology starting form the intuition that no *social* thing exists outside texts.

6. Object = Inscribed Act

Keeping this in mind, my thesis[21] is that, contrary to Searle's idea, the constituting rule of a social object is not *X counts as Y in C* (social objects are higher order objects with respect to the underlying physical objects), but *Object = Inscribed Act*: social objects are social acts (concerning at least two people) characterized by the fact of being inscribed, in a document, in a computer file, or simply in people's head. With respect to Searle, we solve all the problems of the shift from the physical to the social; with respect to Smith, we have a way to distinguish an actual social object from a purely thought one; with respect to Derrida, we acknowledge a specific sphere of social objects, separated by physical and ideal objects. The essential lines of this theory are the following: in the world there are subjects and there are objects. The subjects refer to objects (the former represent, think about, somehow deal, with the latter), namely they possess intentionality; objects do not refer to subjects.

Objects come in three kinds: (1) physical objects (mountains, rivers, human bodies, and animals) that exist in space and in time, and are independent from subjects knowing them, even though they may have built them, as for artifacts (chairs, screwdrivers); (2) ideal objects (numbers, theorems, relations) that exist outside of space and time, and are independent from the subjects knowing them, but which, after having been discovered, can be socialized (for instance, a theorem can be published: still, it is

---

19 Searle 1977.

20 Derrida 1988.

21 Ferraris 2005.

the publication, not the theorem, that has a beginning in time); (3) social objects, that do not exist *as such* in space, since their physical presence is limited to the inscription (money is such because of what is written on the coin, on the banknote, on the memory of the credit card), but last in time, and whose existence depends on the subjects who know, or at least can use, them and who, in certain cases, have constituted them. This latter circumstance display us the fact that social objects, for which construction is necessary, depends on social *acts*, whose *inscription* constitutes the *object*.

As I have indicated through the law Object = Inscribed Act, social objects consist in the recording of acts that encompass at least two people, and are characterized by being inscribed, on a physical substrate what so ever, from marble to neurons, passing through paper and computers. I do not consider obnoxious the idea that also bran processes are to be described in terms of a sort of writing, since they manifest to us exactly in those terms, as it is revealed also by the fact that the mind has always been described as a *tabula rasa*, i.e. a writing table.

From this standpoint, and weakening Derrida's axiom, one can state that "no social thing exits outside texts". Physical as much as ideal object exist independently form inscription and records, but this is not the case for social objects. Without some sort of recording is impossible to conceive any kind of society, and—even more so—any social object. However, the recording is a necessary but not sufficient condition for the existence of social objects: without recordings there are no social objects, but not necessarily a recording (for instance a recollection of mine) constitutes a social object.

Social objects are constituted by inscribed acts, but not every inscription is a social object. Fingerprints become social object when they are registered by the police and used as evidence in a trial, and in this case they are actually part and parcel of an inquiring procedure. And when fingerprints are taken on a passport, they become part of a document, which is endowed with an even more evident social character, since it incorporates this social character—so to speak. From this standpoint, the document has to be conceived, rather than as something which is done once for all, and constituting a class of stable objects, as a teleological end of a theory of social objects. Not all inscriptions are documents, but there is no inscription that, in certain conditions and once it has acquired a certain social power, cannot become such.

7. Documentality

If all this is true, then a theory of social objects develops naturally into a theory of the document, understood as an inquire centered on the definition of what I call "documentality", namely the properties that constitute, in each case, the necessary and sufficient conditions (starting from two very general conditions: being an inscription and being a document or a "documental" thing) to be a social object. At last, there is no society if there are no documents, and documents are records with a particular social value. On this ground, a theory of documentality can develop along three directions. The ontological dimension, answering the question: what is a document? The technological dimension, concerning the means through which documentality can be spread in a complex society. The pragmatic (and forensic) dimension, which concerns the care of documents in a society characterized by the explosion of writing, and in world dominated by information technology[22].

1. As for the first question—what is a document?—we need to articulate the law Object = Inscribed Act. Documentality comprises a sphere encompassing so different things as memories, notes (a memo can, although not necessarily must, acquire social value), and international treaties; all such things can be realized through the most different media (paper writing, electronic writing, pictures …); they can refer to the most different activities (borrowing a book, getting married, being named, declaring war…). In the vast majority of those realization it is possible to spot the structure of documentality: first of all, a physical substrate; then, an inscription, which obviously is smaller than the substrate and which defines its social value; finally, an idiomatic thing, typically a sign (and its variation, such as the electronic signature, the debit card's or mobile's PIN, …), which guarantees its authenticity.

It is important here to single out a point. Sounds, signs, and thoughts are not physical objects as hefty as States or persons. They possess far less molecules. Still, they are not completely void of physical bulk: a sound needs vibrations, a thought requires electric activity in the brain, and this, obviously holds for signs over a piece of paper too (even more manifestly so). This last circumstance, if one think over it a while, is more relevant than it is generally believed, since paradigmatic social object such as banknotes are "signs over a piece of paper": banknotes qualify as social, and not only physical, objects because of few molecules, those in the inscription and possibly in the thread-mark too. Actually, the really important aspect of a banknote, what turns it from a physical object—a drawing, let us say—into a social object, are the few molecules of the inscription that declare its value, along with the ones of the signature of the governor who states its validity—and not the bunch of molecules constituting its form and matter––proof being that a very big banknote may have lesser value than a much smaller one. Those few molecules, moreover, are not very different from the *blips* in the computer of a bank: they are objects of the same sort, showing similar characteristics; and this holds also for what Searle calls "status indicators", such as, for instance, passports and driving licenses. At the same time, those molecules are *something*, although there is only a little amount of them, and not just a representation. Those few molecules account for Smith's expression "quasi-abstract entity": the entity must be recorded somewhere in space. At the ontological level, my proposal is to spot in documentality five ascending degrees (from physical to social): traces, recordings, inscriptions, documents, idioms.

I call a "trace" what, endowed with a rather small number of molecules, serves as the physical substrate of a record. Only for social objects the trace has a constitutive value. In the world of physical objects, there are traces only for minds that are able to recognize them. In the world of ideal objects, traces operate only in the socialization of an entity that does not depend on an inscription. Things are different for the constitution of a social object, since the trace openly indicates a beginning in time, and moreover it motivates the chronology of the object also beyond the intentions of the people involved in its constitution, and the length of their life.

A trace, in a mind or for a mind, become a record; this record can, in certain circumstance, acquire social value, for instance when the agents of the scientific police

---

[22] Koepsell 2000 and 2003.

transform a piece of DNA attached to a cigarette butt into a proof. But, indeed, the mere recording is not, as such, a social thing.

An "inscription" is a record with a social value. Within a society, spoken or written words, as much as hand-shakings, can be relevant things. The inscription possesses the following laws of essence: it is the necessary but not sufficient condition of the social object; it is smaller than its substrate; its size has no bearing on the size of the corresponding social object; the inscription is true if it is idiomatic.

Inscriptions that can acquire a legal value are documents, which are, along with Smith's perspective[23], acts fixing[24]. And here it is the last ingredient of our ascending hierarchy: the idiom. With "idiom" I mean a specific way of presentation of an inscription that links a particular instricption to an individual. Its more evident model is the signature (on a document, a check, a banknote: an element that is almost everywhere in social reality, although it is often unobserved), but it can also be a specific way in which some expresses themselves, for instance their normal tone of voice. Its aim is the object's individuation, and exactly in so far as it individuates an object it can play the role in the validation of social objects, which, thanks to the signature, emerge as the expression of the intentionality of someone.

2. As to the second question—how documentality can be spread in a complex society?—if it is true that in our developed society the demand for documentality is growing at a fast pace, it is also true that our society is provided growing resources from the electronic supports, which enhance and multiply the law Object = Inscribed Act.

This element is apparent in financial transactions, and in all that can be done through them. On a financial level, and already in an economy based on paper supports, documents are what fix the values, compose different values within a single system, stir resources and energies, relate people, protect transactions[25]. On this ground, a shift form the paper support to the electronic support de-locates the operations by extending the capacity of writing. It becomes, thus, possible to accomplish differently natured operations: paying taxes, fines, bills (also those that, differently from power and gas

---

23 Smith 2006.

24 According to Smith (2006), thus, it is possibile to develope a theory of what he calls "document acts", i.e. a theory "1. of the different types of document, ranging from free-text memos to standardized forms and templates, and from single documents as self-contained collections of information to bodies of documents incorporating various sorts of riders, codicils, protocols, addenda, amendments, endorsements and other attachments, including maps, photographs, diagrams, signatures and other marks, 2. of the different types of physical medium or bearer for a document's content (most important here is the distinction between paper and electronic documents), 3. of the different sorts of things we can do to documents (fill in, sign, countersign, stamp, copy, notarize, transfer, invalidate, destroy), 4. of the different sorts of things we can do (achieve, effect) with documents (establish collateral, create organizations, record the deliberations of a committee, initiate legal or military actions), and of the different ways in which, in performing such acts, we may succeed or fail to achieve the corresponding ends, 5. of the institutional systems to which documents belong (marriage, property, law, commerce, trade, credentialing, identification, movement of goods and people), and of the different positional roles within such systems which are occupied by those involved in the performance of the corresponding acts, 6. of the provenance of documents (of what distinguishes an original, authentic document from a mere copy or forgery)."

25 De Soto 2000.

bills, cannot be domiciled in a bank, such as the garbage tax) and union fees (for every kind of employee); booking medical visit, lawyers, public office; obtaining certificates (family state, identity documents, house certificates); doing bank transactions; obtaining postal services (at a virtual counter you can send a registered letter, a telegram, and, in general, a letter that will be delivered in a paper form); and purchasing on line (in this case, physical commodities—we are delivered our shopping on line—, as much as events or social objects: plane tickets, museum tickets, concert tickets).

The problems concerning identification are tougher. Electronic documents are not localized, or, at least, they are far less localized than paper documents. If I fill on line an application form in a public administration site, the form will be the same whether I fill it at a computer in Italy or in Mexico, but, then, where exactly is the form? Moreover, who answers me is not a person (someone will read the form only later, if anyone), but a program. And a program cannot talk, unless it has been enabled to do so through another program, that is a written thing. The document is not any longer the transcription of a voice localized in a physical person, is a written thing de-localized in each computer through which we can access it. In this framework, we find the issue of the digital signature, which constitutes a remedy to the impersonality and de-location of the digital, and which sum up within itself exactly two fundamental features of the document: individual reference (idiomaticity) and deontic power [26].

3. Finally, let us face the third question. How are we to take care of documents in a world characterized by the explosion of writing? The growing problems of privacy in the advanced societies are usually read from the stand point of a Big Brother, namely a big watching eye, in accordance with the model of Bentham's Panopticon, but this image is partly misleading. Actually, it is true that there are more and more cameras observing (also with an infra-red eye) our everyday life, in banks, stations, supermarkets, private buildings, and satellites. But the strength of this eye would be nothing, were it not accompanied with the capacity of recording,—which is exactly what turns an act of vision into a document. In this case too, the debates on the interceptations are just the tip of the iceberg: democracy requires to be investigated through the central questions rising within the category of "documentality".

All this suggests two complementary, although contrasting, considerations. On the one hand, the growing role of documentality shows undoubtedly why it is so bad being "sans papier"; it is exactly the lack of those paper—which are more and more turning into computer blips—the starting point of the process leading to the bare life, namely the offended life, a life liable to anyone's offence. In this sense, then, documentality looks like a safeguard. On the other hand, obviously enough, documentality deprives us of the right to a secret and private sphere, it creates a sort of

---

26 See the characterization of the digital signature to be found in the Italian Legislation (Art. 24, March 5th 2005, n. 82, Digital Administration Code). The digital signature ought to univocally refer to one and only subject and to the document or set thereof to which it is affixed or associated. The affixing of a digital signature integrates and substitutes the affixing of seals, stamps, and marks of whatever kind, and used to every aim to which the current normative applies. For the generation of the digital signature, a qualified certificate has to be used, whose validity, at the time of the subscription, is not expired, revoked, or suspended. The validity of the certificate, along with the identifying elements of the titular, of the certificating officeholder, and possibly the constraints on its use has to be established through the qualified certificate itself, according to the technical rules established by article 71.

universal control. Therefore, what ha been called the *habeas data*, namely the acknowledgement of the privacy of the records concerning us[27], turns out to be not less important than the acknowledgement of the *habeas corpus* that was ratified eight hundred years ago.

Conclusions

I maintain to have demonstrated that the critical category for social ontology is the category of "documentality", in accordance with the constitution law Object = Inscribed Act. Through this category it is possible to develop a unified theory of social objects, going over the difficulties found in the previous theories.

Bibliography

AUSTIN, J. L., (1962), *How to do Things with Words*, Oxford, Oxford University Press.

BRATMAN, M. E., (1992), "Shared cooperative activity", *The Philosophical Review*, 101, pp. 327-41.

DERRIDA, J., (1967), *De la grammatologie*, Paris, Ed. de Minuti.

—, (1971), "Signature, événement, contexte", poi in Id., *Marges de la philosophie*, Paris, ed. de Minuit 1972.

—, (1977), *Limited inc.: Abc*., Baltimore, Johns Hopkins University Press.

DE SOTO, H., (2000), *The Mistery of Capital. Why Capitalism Triumphs in the West and Fails Every Where Else*, New York, Basic Books.

DI LUCIA, P., (2003), ed., *Ontologia Sociale. Potere deontico e regole costitutive,* Macerata, Quodlibet.

FERRARIS, M., (2003a), *Introduzione a Derrida*, Roma-Bari, Laterza.

—, (2003b), "Oggetti sociali", *Sistemi intelligenti*, XV, 3, pp. 441-466.

—, (2005), *Dove sei? Ontologia del telefonino*, Milano, Bompiani

—, (2006), *Jackie Derrida*, Torino, Bollati Boringhieri

GILBERT, M., (1989), *On Social Facts*, New York, Routledge.

—, (1993), "Group Membership and Political Obligation", *The Monist,* 76, pp. 119-131.

JOHANSSON, I., (1989), *Ontological Investigations. An Inquiry into the Categories of Nature, Man and Society*, London, Routledge; 2a ed. Frankfurt/M. Ontos-Verlag 2004.

KIM, J., – SOSA, E., (1999), eds., *Metaphysics: An Anthology*, Oxford, Blackwell.

KOEPSELL, D. R., (2000), *The Ontology of Cyberspace*, La Salle, Open Court.

—, (2003), "Libri e altre macchine: artificio ed espressione", in R. Casati, ed., (2003), pp. 429-439.

KOEPSELL, D. R. - MOSS, L. S. , (2003), eds. , *John Searle's Ideas About Social Reality*, monography, *American Journal of Economics and Sociology*, 62.

MOORE, M.S., (2002), "Legal reality: a naturalist approach to Legal Ontology", *Law and Philosophy*, 21, pp. 619-705.

MULLIGAN, K., (1987), ed., *Speech Act and Sachverhalt. Reinach and the Foundations of Realist Phenomenology*, The Hague, Nijhoff.

ONG, W. J., (1982), *Orality and Literacy. The Technologizing of the Word*, London and New York, Methuen.

REID, TH., (1785), *Essays on the Active Powers of the Human Mind*, in Id.,

27 Rodotà 2006.

Philosophical Works, Hildesheim, Olms 1967.
REINACH, A., (1913), “Die apriorischen Grundlagen des bürgerlichen Recht”, *Jahrbuch für Philosophie und philosophische Forschung*, 1, pp. 685-847.
RODOTÀ, S., (2006), *La vita e le regole*, Milano, Feltrinelli
SEARLE, J. R., (1969), *Speech Acts*, Cambridge, Cambridge University Press.
—, (1975), *A Taxonomy of Illocutionary Acts*, Cambridge, Cambridge University Press.
—, (1977), “Reiterating the Differences: A Reply to Derrida”, *Glyph*, I, pp.172-208.
—, (1980), “Minds, brains and programs”, *Behavioral and Brain Sciences*, 3, pp. 417-58.
—, (1983), *Intentionality. An Essay in the Philosophy of Mind*, New York-Cambridge, Cambridge University Press.
—, (1992), *The Rediscovery of the Mind*, Montgomery (Vt.), Bradford Books.
—, (1993a), “Rationality and Realism, What is at Stake ?”, *Daedalus*, pp. 55-83.
—, (1993b), “The World Turned Upside Down”, and “Reply to Mackey” in G. B. Madison, *Working through Derrida*, Evanston, Northwestern University Press, pp. 170-188 and pp. 184-188.
—, (1995), The Construction of Social Reality, *New York, Free Press.*
—, (1998), *“Postmodernism and Truth”, TWP BE (a journal of ideas), 13, pp. 85-87.*
—, (1999), Mind, Language and Society. Philosophy in the Real World, *New York, Basic Books.*
SMITH, B., (1998), “Ontologie des Mesokosmos: Soziale Objekte und Umwelten”, *Zeitschrift für philosophische Forschung*, 52, pp. 521-540.
—, (1999), “Les objets sociaux”, *Philosophiques*, 26, pp. 315-347, http://www.erudit.org/erudit/philoso/v26n02/smith2/smith2.htm English version: “Social Objects” http://wings.buffalo.edu/philosophy/ontology/socobj.htm.
—, (2002), The Ontology of Social Reality, http://ontology.buffalo.edu/smith//articles/searle.PDF, 2002
—, (2003a), “John Searle: From speech acts to social reality”, in *John Searle*, ed. Id., Cambridge, Cambridge University Press.
—, (2003b), “Un’aporia nella costruzione della realtà sociale. Naturalismo e realismo in John R. Searle”, in P. Di Lucia, ed., 2003, pp. 137-152.
—, (2006), [“Document Acts”. http://ontology.buffalo.edu/document_ontology/]
SIMONS, P., (1987), *Parts. A Study in Ontology.* Oxford: Clarendon Press
TOMELA, R., (1995), *The Importance of Us*, Stanford (CA), Stanford University Press.
—, (2002), *The Philosophy of Social Practices*, Cambridge, Cambridge University Press.
VICO, G. B., (1744), *La scienza nuova*, in Id., ed., N. Abbagnano, *La scienza nuova e altri scritti*, Torino, Utet 1952, pp. 247-748.
WILLIAMSON, T., (1994), *Vagueness*, London, Routledge.
—, (1998), a c. di, *Vagueness*, fascicolo monografico, *The Monist*, 81, pp. 193–348

Practical approaches to standardizing vocabularies: the Cultural Heritage experience

Philip Carlisle
English Heritage, National Monuments Record
Kemble Drive, Swindon SN2 2GZ
Wiltshire, United Kingdom
philip.carlisle@english-heritage.org.uk

Abstract: The European Heritage Network was established as part of an EU funded project (HEREIN) which ran from 1998-2003 with the aim of creating a pan-governmental network for the Cultural Heritage Sector. One of the key deliverables of the project was a multilingual thesaurus of Cultural Heritage Policy Terminology (the HEREIN thesaurus) to allow the indexing of the National Heritage Policy reports of the EU member states in the two official languages of the Council of Europe (English and French) and Spanish. This paper will outline the practical approaches taken to develop the thesaurus and its extension from the original three languages using the HEREIN thesaurus as a case study.

Keywords: Cultural Heritage; thesauri; indexing; international standards; multilingual; European cooperation; mapping.

1. Introduction

The European Heritage Network brings together government departments and agencies responsible for cultural heritage under the umbrella of the Council of Europe. It was established in 1999 as part of the HEREIN project following the 4th European Conference of Ministers responsible for the Cultural Heritage held in Helsinki in 1996 which recommended that the Council of Europe "consider setting up a permanent information system for the benefit of national authorities, professionals, researches and training specialists in touch with heritage developments in other countries" [01]. The network has since become a reference point for government bodies, professionals, research workers and non-governmental organisations active in this field and currently includes representatives from 41 countries.

The website of the network [02] provides a database of the national heritage policies of the EU member states as well as a multilingual thesaurus allowing users to access the policies of other countries using their own language.

This paper will outline the development of the thesaurus and the methodology and approaches used to provide a practical, multilingual tool using the Thesaurus of Cultural Heritage Policies (HEREIN thesaurus) as a case study.

2. Standardizing terminology

With the increasing availability and use of the internet the need to standardize terminologies to facilitate searching has become a major issue for information providers. This is particularly important where the same information is available in multiple languages and needs to be conveyed to users from across a wide geo-political area.

Without standardized vocabularies

2.1. Wordlists

Most people are aware of the existence of controlled vocabularies. The simplest form is the wordlist; an alphabetical list of keywords to be used to control the entry to a database field.

Simple wordlists are limited by their inability to express any relationships between the words contained in the list. In the following list there are four types of building:

Bungalow
Castle
Fort
House

This simple wordlist, although useful, could be given greater meaning by grouping the terms together into hierarchies related by function:

- Dwellings
  - Bungalow
  - House
- Defensive Buildings
  - Fort
  - Castle

This more complex wordlist is already more useful and allows the user to search at the broader level using the “parent” (eg. Dwellings) to retrieve all records indexed with any of the “child” terms or simply to search on the individual child terms.

2.2. Thesauri

The structure of a thesaurus provides even greater flexibility than a complex wordlist. It allows three basic relationships to be established between terms:

- hierarchical - groups terms together according to certain criteria for example, function or type.
- equivalent - allows the user to define which of two synonyms should be used as the “preferred term” when indexing.
- associative – allows terms to be related to one another where no explicit hierarchical or equivalent relationship may exist

In addition to creating relationships a thesaurus allows user to define the terms using "scope notes". These can simply be a dictionary-style definition of the term but they can also include guidance on how the term should be used.

The construction of a thesaurus should conform to the standard as defined in the relevant ISO document. There are currently two standards, one governing monolingual thesauri (ISO2788) [03] and one governing multilingual thesauri (ISO 5964) [04].

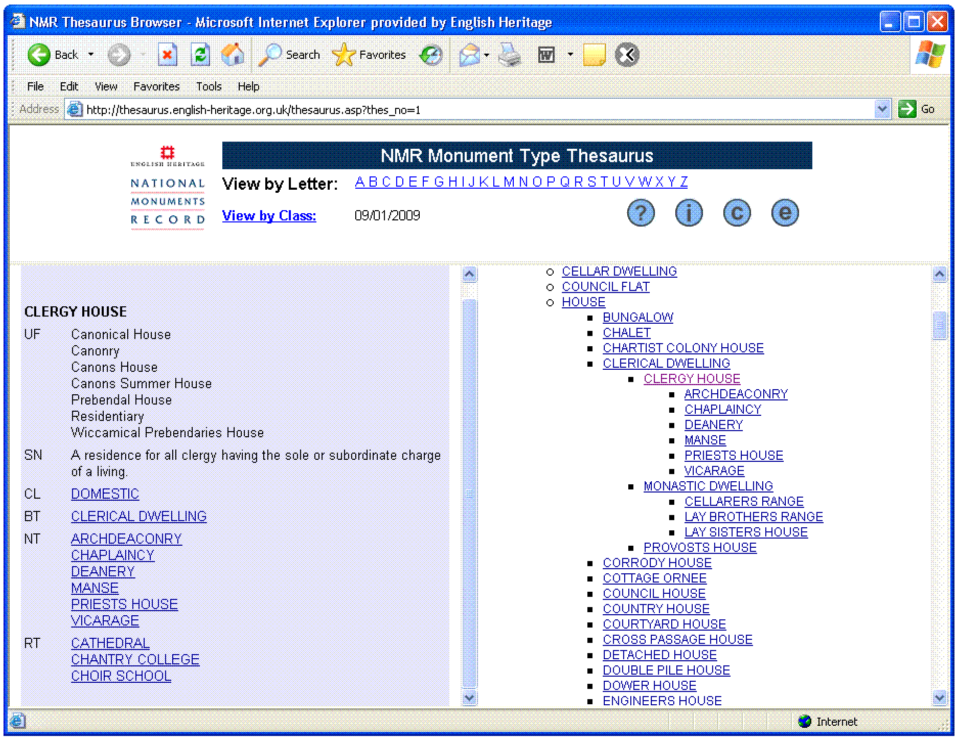


Fig.1 Screenshot of the term CLERGY HOUSE, showing hierarchical and alphabetical displays. Copyright, English Heritage, 1999.

These standards have been in use since the late 1980s and are currently being revised but the basic principles are still valid.

In a display, whether printed or digital, the following conventions are used to show the relationships:

- USE – indicates the "preferred term" to be used where synonyms exists
- UF- indicates the "non preferred term" or "guide term"
- SN – indicates the scopenote or definition
- BT – indicates the Broader Term or "parent"
- NT- indicates the Narrower Term or "child"
- RT – indicates the Related Term

3. Developing the HEREIN Thesaurus

It was recognized early on, that a network such as that proposed by the Conference of Ministers, would need a database which would allow searching across multiple languages, as the main function of the network was to provide access to the national heritage policies of the EU member states. As such, some form of multilingual vocabulary would be required and a working group was established to develop the HEREIN thesaurus.

As the project was being run under the auspices of the Council of Europe it was necessary for each policy document to be produced not only in the native language of the country, but also in either English or French[1] and so those two languages would form the basis of the thesaurus with Spanish[2] being added as the third language.

3.1 Methodology

Before the thesaurus could be developed it was necessary for the thesaurus working group to identify the terms which would be required for the thesaurus. In addition to the documents from France, Spain and the UK the national policies of Norway, Eire and Hungary were chosen to be 'mined' for terms. These terms were then augmented with others derived from various legislation and specialized documentation such as legal dictionaries. This initial phase resulted in approximately 1200 terms in each language.

Many of the terms identified were deemed to be too specific to the particular legislation of one country and so it was decided to limit the initial number of preferred terms to 500 with the more specific terms being denigrated to non-preferred status. Not only would this make the thesaurus easier to manage but it would also ease usability by focussing on more generic terminology common to the 3 languages.

The terms were then grouped into 9 broad categories and a hierarchy was constructed in each of the 3 languages.

1 *Agents (organisations and people)*
General terms for people and organisations involved in heritage, for example local authority.

2 *Heritage Category*
Specific categories and objects connected with heritage, for example protected sites.

3 *Documentation*
The tools and references or standards used in the creation of documentation, for example computerized database.

4 *Legal systems*

---

[1] The two official languages of the Council of Europe

[2] Spanish was chosen as the third language as the website was also being translated into Spanish and it was felt that 3 languages would provide a stronger foundation for future expansion of the thesaurus

Terms used to represent specific legal concepts, for example listed building. It contains policies. The act of carrying out a policy would be in Interventions

5 *Interventions*
Actions carried out related to heritage, for example archaeological excavations It would include the act of carrying out a policy or strategy but not the policy itself which would be in Legal systems

6 *Professional training, skills and qualifications*
Terms for specific skills or professions and training connected to heritage, for example art restorer.

7 *Access and Interpretation*
Terms covering the provision of access to heritage and the interpretation of heritage, for example school trips.

8 *Economic and Financial systems,*
Terms covering finance and finance related activities, for example public grants.

9 *Broad concepts*
General concepts related to heritage, for example archaeology.

As no single language was being used as a source language it was necessary to ensure that each of the hierarchies was exactly equivalent to its two linguistic counterparts. As the structures of the French, Spanish and UK legal and governmental organization differed it was often necessary to create new terms in the other two languages to ensure that the hierarchies matched. This was particularly true for the first group where the loan terms of ‘autonomous communities’ and ‘communautés autonomes’ were created in English and French to allow the inclusion of an essential Spanish term ‘comunidades autónomas’.

Once the hierarchies had been completed each term was then defined in each of the three languages and then translated. These definitions were then compared to define the degrees of equivalence. Five degrees of equivalence as defined in ISO 5964 were used and each language was compared to the other two languages, ie. En to Es and Fr, Fr to En and Es, Es to En and Fr to ensure that the equivalences were true for all three languages:

1 Exact equivalence -where the meaning and scope of the terms are the same in both languages being mapped and both terms are capable of being preferred terms.

2 Inexact equivalence - here a term in the target language expresses the same general concept as the source language but the meanings are not precisely identical

3 Partial equivalence – where the term in the source language has no exact match but a near translation can be made by choosing a term in the target language which has a slightly broader or narrower meaning

4 Single to Multiple – where the term has no exact match but can be expressed by a combination of two or more preferred terms in the target language

5 Non-equivalence – where the target language contains no equivalent concept

Where terms were found to have no equivalents and the term was deemed to be essential, it was agreed that either the source term would be taken as a loan term (ie. in its original form) or translated.

3.2 Mapping and extending the thesaurus

Once the initial tri-lingual thesaurus had been established on paper it was decided to extend it to four languages. Hungarian was chosen as the fourth language as part of the HEREIN 2 extension project.

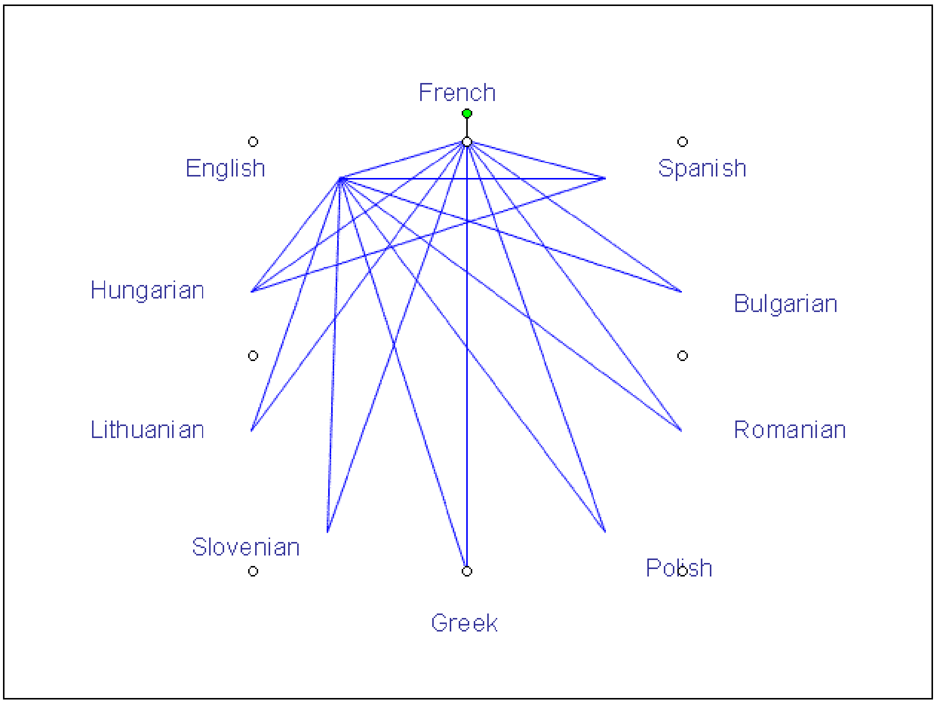


Fig.2. The working thesaurus showing the mappings

Having successfully integrated Hungarian it was then decided to extend to include 6 other languages. These were Bulgarian, Greek (Cypriot), Lithuanian, Polish, Romanian and Slovenian.

To facilitate the extension, software was procured to enable online editing. Unfortunately this was not successful and so new open source software was developed whilst the thesaurus continued to be developed using spreadsheets.

Each new country was also given the option of choosing just 2 of the original languages to map to (see Fig. 2). Inevitably this meant that the new languages chose to map to English and French. This meant that the mapping process would be quicker but didn't accurately reflect the degrees of equivalence between all languages. In fact this form of mapping actually created more problems than it solved particularly where a term in one of the new languages claimed an exact equivalence with term A in English but a single to multiple with term B in French. If term A and B had been defined as begin exactly equivalent then the mapping was not logical.

It was decided in 2005 to go back to basics and to ensure that the software was capable of dealing with any problematic logic loops caused by this and to simplify the mapping process.

The German language was introduced as the new fourth language and the degrees of equivalence were painstakingly reviewed by the thesaurus working group.

Once the software had been loaded with the new core thesaurus, the other languages were reinstalled one by one. The software now checks that the mappings do not create any logic loops and produces a report detailing any errors encountered.

4. Conclusion and future work

Developing a thesaurus from scratch is a resource-intensive, time-consuming process and should not be entered into lightly. Particularly one which involves politics and more than one language! The European Heritage Network is now entering its second decade and work still continues on the thesaurus. It has taken a long time to find a solution which is both practical, financially viable and, with the number of languages now approaching 14, increasingly essential but with the software which is now in place the thesaurus working group are confident that any new countries wishing to join the thesaurus will find the work quicker and easier than those who were there at the start. As well as looking to expand with more languages the thesaurus working group is currently looking at how to handle the problem of synonyms not only within a language but across borders. The Francophone team are trying to solve the problem to ensure that the same concept can be expressed using different terms in French, Belgian, Swiss and Luxembourgeoise whilst still maintaining their national preference.

References

[01] IVth European Conference of Ministers responsible for the Cultural Heritage, Helsinki (1996) *Report by the Secretary General submitted in pursuance of paragraph 8 of Resolution (71) 44 of the Committee of Ministers* https://wcd.coe.int/com.instranet.InstraServlet?command=com.instranet.CmdBlobGet&InstranetImage=259999&SecMode=1&DocId=546330&Usage=2

[02] Council of Europe: http://www.european-heritage.net/sdx/herein/index.xsp (2008)

[03] International Organization for Standardization (1986): *ISO 2788: Guidelines for the establishment and development of monolingual thesauri*, 2nd Ed. Geneva: ISO, 1986

[04] International Organization for Standardization (1985a): *ISO 5964: Guidelines for the establishment and development of multilingual thesauri*, Geneva: ISO, 1985

Ontologies for the Integration of Air Quality Models and 3D City Models

Claudine Metral
Institut d'architecture - University of Geneva
Site de Battelle - 7, route de Drize - CH 1227 Carouge/Geneva - Switzerland
claudine.metral@unige.ch

Gilles Falquet
Department of Information Systems - University of Geneva
Site de Battelle - 7, route de Drize - CH 1227 Carouge/Geneva - Switzerland
gilles.falquet@cui.unige.ch

Kostas Karatzas,
Informatics Applications and Systems Group, Dept. of Mechanical Engineering
Aristotle University, GR-54124 Thessaloniki, Greece
kkara@eng.auth.gr

Abstract. In the perspective of a sustainable urban planning, it is necessary to investigate cities in a holistic way and to accept surprises in the response of urban environments to a particular set of strategies. For example, the process of inner-city densification may limit air pollution, carbon emissions, and energy use through reduced transportation; on the other hand, the resulting street canyons could lead to local levels of pollution that could be higher than in a low-density urban setting.

The holistic approach to sustainable urban planning implies using different models in an integrated way that is capable of simulating the urban system. As the interconnection of such models is not a trivial task, one of the key elements that may be applied is the description of the urban geometric properties in an "interoperable" way. Focusing on air quality as one of the most pronounced urban problems, the geometric aspects of a city may be described by objects such as those defined in CityGML, so that an appropriate air quality model can be applied for estimating the quality of the urban air on the basis of atmospheric flow and chemistry equations.

It is generally admitted that an ontology-based approach can provide a generic and robust way to interconnect different models. However, a direct approach, that consists in establishing correspondences between concepts, is not sufficient in the present situation. One has to take into account, among other things, the computations involved in the correspondences between concepts.

In this paper we first present theoretical background and motivations for the interconnection of 3D city models and other models related to sustainable development and urban planning. Then we present a practical experiment based on the interconnection of CityGML with an air quality model. Our approach is based on the creation of an ontology of air quality models and on the extension of an ontology of urban planning process (OUPP) that acts as an ontology mediator.

Keywords: Ontology, 3D city model, air quality model, street canyon model, sustainable urban planning

1. Introduction

Sustainable development and urban planning

One of the most often cited definitions of sustainability is the one created by the Brundtland Commission in its report (Brundtland, 1987) which defined sustainable development as development that "meets the needs of the present without compromising the ability of future generations to meet their own needs." The report highlighted three fundamental components to sustainable development: environmental protection, economic growth and social equity.

Regardless of whether environmental issues are not the only aspects of sustainable planning and decision processes, they are impossible to circumvent. In urban areas, one of the most important environmental problems is air pollution, mostly induced by vehicle traffic. This pollution can cause severe damages on health and is particularly crucial because of the high density of population in cities. The improvement of air quality is therefore imperative and must be taken into account in a sustainable urban planning process. In fact a deep understanding and an accurate prediction of urban flow and pollutant dispersion in cities is required for a more proactive urban planning. In addition, the existing EU legal framework (Dir 96/62 and the associated “daughter” directives), dictate the use of air quality models as one of the ways to be applied for atmospheric quality assessment and related policy making.

Air quality models

Air quality models are important tools to study, understand and predict air pollution levels. Among the various categories, 3D air pollution models are those that aim at “reconstructing” the environment, its properties and governing physical laws. Current air quality models and the associated simulation systems apply a number of (simplifying) assumptions about the environment, in particular its geometry and its physical properties, according to the scale of interest. Thus, models that are used for the study of air quality at a local-to-regional scale make use of a grid-base spatial resolution, where each grid cell has dimensions in the order of magnitude of 1 x 1 km. Coming down to the urban scale, simulation models try to reconstruct the building shapes and basic geometrical characteristics and dimensions, in order to improve simulation results.

3D city models

At the same time, 3D city models are rapidly developing (a number of cities have already been modeled in CityGML). These models could provide the air quality models with (some of) the missing information they require to be more precise and effective. On the other hand, urban planners and stakeholders need information about the impact of their development plans on air quality evolution (current air quality and future scenario projections). Thus their models must be complemented with data originating from air quality models. However, these data do not fit directly into the urban planning models, they must be adapted (aggregated, computed, etc.) to these models.

2. CityGML

What is CityGML?

CityGML is an open information model for the representation and exchange of virtual 3D City Models on an international level (OGC, 2006) (Kolbe et al, 2005) (Kolbe et al, 2006). CityGML defines the classes and relations for the most relevant topographic objects in cities and regional models with respect to their geometrical, topological, semantical, and appearance properties. It has been defined in this way because purely graphical or geometrical 3D models, if appropriate for visualization purposes, are not sufficient for applications such as urban and landscape planning, architectural design, touristic and leisure activities, 3D cadastres, environmental simulations, mobile telecommunications, disaster management, homeland security, vehicle and pedestrian navigation, training simulators or mobile robotics.

CityGML differentiates five consecutive Levels of Detail (LOD), from terrain alone to architectural models (outside and interior). With LOD increasing, objects become more detailed, both at the geometry and the thematic level. Besides their attributes, objects in CityGML can have external references to corresponding objects in other databases or data sets. For example, a building in CityGML may be related to a cadastral database.

With the Application Domain Extensions (ADE), CityGML provides an extension mechanism so as to enrich the data with identifiable features for specific domain areas, such as environmental noise mapping.

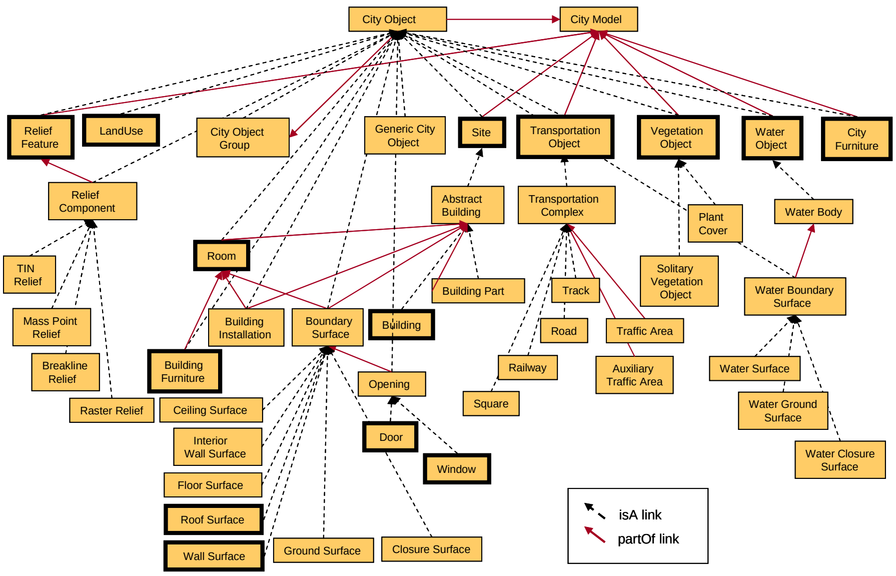
CityGML is implemented as an XML application schema of the Geography Markup Language 3 (GML3), the extendible international standard for spatial data exchange and encoding issued by the Open Geospatial Consortium (OGC) and the ISO TC211.

Underlying concepts of CityGML

The main thematic fields are:

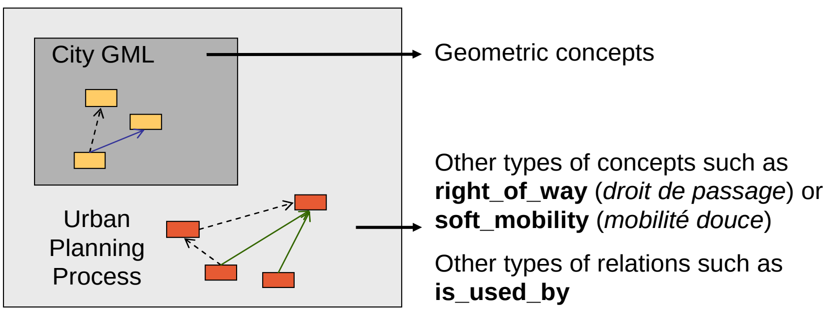
- the terrain (named as Relief Feature),
- the coverage by land use objects (named as Land Use),
- transportation (both graph structures and 3D surface data),
- vegetation (solitary objects, areas and volumes, with vegetation classification),
- water objects (volumes and surfaces),
- sites, in particular buildings (bridge, tunnel, excavation or embankment in the future),
- City Furniture (for fixed object such as traffic lights, traffic signs, benches or bus stops).

So the main underlying concepts of CityGML (see figure below) are :
Relief Feature, Land Use, Site, Building (as subconcept of Site), Transportation Object, Vegetation Object, Water Object, City Furniture.
Buildings in CityGML can be represented with roofs, walls, windows, doors, rooms and even furniture (depending on the LOD) associated to the concepts Roof Surface, Wall Surface, Window, Door, Room and Building Furniture (see figure below).



3. OUPP: an ontology of urban planning

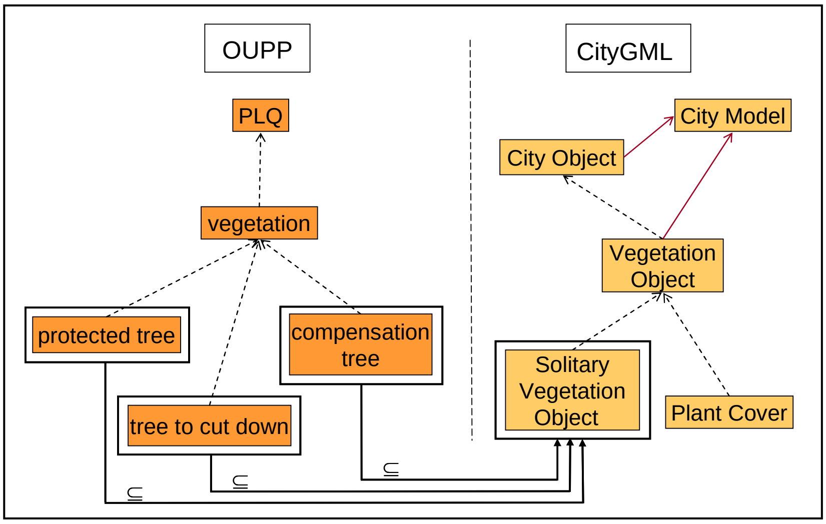
CityGML, with its urban and geometric concepts and their semantic relationships, is useful for representing an urban planning project, but not sufficient. Indeed, some non geometric concepts such as right of way or more abstract concepts such as soft mobility don't exist in CityGML. Similarly, relationships such as is used by or participates to are missing (see figure below). This is why we decided to define an ontology of urban planning process (OUPP).



OUPP has been defined and is still under development in the framework of the "Integrating Urban Knowledge into 3D City Models" research project funded by the Swiss Confederation and part of the European COST Action "Towntology". The main objective of this research is to contribute to a better communication between the various actors involved in an urban planning process. The main sources for OUPP are master and local plans (such as PLQ), as well as GIS data of Geneva.

An urban planning project (that uses urban data and documents such as spatial data from GIS, 3D city models, texts, maps, plans, pictures) is represented by an ontology with concepts and instances related to the data and documents. With this representation, a user can not only view in an integrated way the project but also accede only to the knowledge that fits its profile and centers of interest (Metral et al, 2006).

In our research, use of CityGML is done from OUPP. More precisely, OUPP refers to the concepts that exist in CityGML. An ontology alignment has been realized between OUPP and CityGML, with 1 to 1, 1 to many and many to 1 alignment relations (an example of a n to 1 alignment relation is given in figure below).



4. Air quality models

The aims of air quality models

In urban areas, traffic-related pollution is one of the more environmental crucial problems (EEA, 2006). As air pollution entails damages on humans, vegetation and ecosystems, reducing pollutant emissions to the atmosphere is imperative. Models of air pollutant dispersion, transport and transformation have been defined. They are used for supporting such reductions and, more generally, for supporting environmental decision making

The diversity of air quality models

According to (Moussiopoulos et al 2006), models describing the dispersion and transport of air pollutants in the atmosphere can be distinguished in many ways such as:

- the spatial scale (global; regional-to-continental; local-to-regional; local),
- the temporal scale (episodic models, (statistical) long-term models),
- the treatment of the transport equations (Eulerian, Lagrangian models),
- the treatment of various processes (chemistry, wet and dry deposition)
- the complexity of the approach.

Urban scale models

At this scale, in general, air flow is very complex, as it depends not only on the meteorological conditions but also on the characteristics of urban obstacles such as buildings (form, orientation with regard to the wind direction, etc.). The air flow patterns that develop around buildings govern the dispersion mechanisms and dictate concentration levels in a built environment.

Air quality may become seriously poor in densified urban areas. Street canyons, that are relatively narrow streets between buildings lining up continuously along both sides, are more and more frequent in such areas. Unfortunately they can induce serious pollution problems because pollutants emitted in a street canyon (such as vehicle exhaust gases) tend to disperse less than those emitted in an open area, while the combination of pollutants emitted and local meteorological conditions may trigger pollution production mechanisms that result in accumulated concentration levels and bad air quality

Street canyon models

Various types of numerical models have been employed to simulate flow and dispersion of pollutants in urban street canyons. While most are two-dimensional models such as (Baik & Kim, 1999) (Huang et al, 2000), there exists some three-dimensional models such as (Kim and Baik, 2004) (Santiago et al, 2007).

The flow and pollutant dispersion in urban street canyons are mainly controlled by meteorological conditions (such as the prevailing wind direction and speed) and canyon geometry, in particular the street aspect ratio (ratio of the building height to the width between buildings). When wind blows over buildings with different shapes and sizes, it is disturbed by the buildings, and turbulent eddies, that can stream into the street canyon, are produced. The turbulence intensity, which depends on the geometric configuration of the surrounding buildings, has an effect on the pollutant dispersion in the street canyon (Kim & Baik, 2003). Furthermore, when the street bottom or the building wall is heated by solar radiation, thermally induced flow is formed that combines with mechanically induced flow formed in the canyon (Kim & Baik, 2001) (Huang et al, 2002) (Liu et al, 2003). Lastly, pollutants emitted from automobiles, for example NO and NO2, are chemically reactive (Baik et al, 2007).

Air pollution modeling within street canyons has to deal with problems at the so called neighborhood scale, i.e. 100 m to 1 km. As the flow around a building is influenced by the geometry and the scale of the obstacle (Fig 1), the modeling attempts applied so far focus on the representation of both aspects of the urban web via detailed 3d representations. For this reason, mesh-based 3D modeling is applied, trying to represent as much as possible the actual geometry, scale, and physical characteristics of the area of interest (Fig 2).

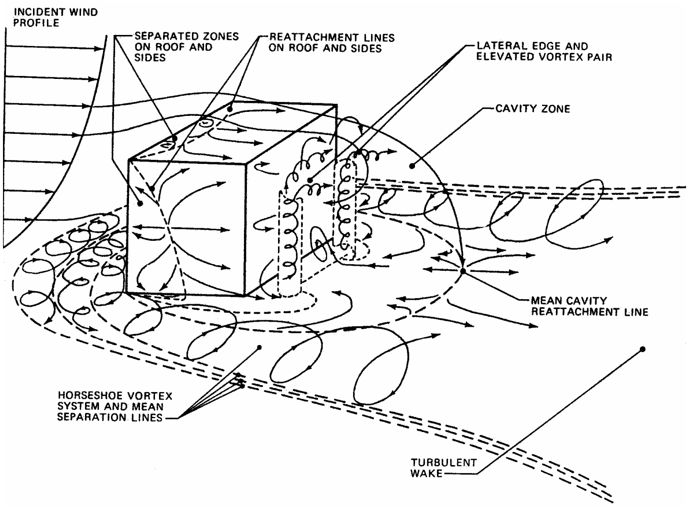


Figure 1: Flow around a building (cube)

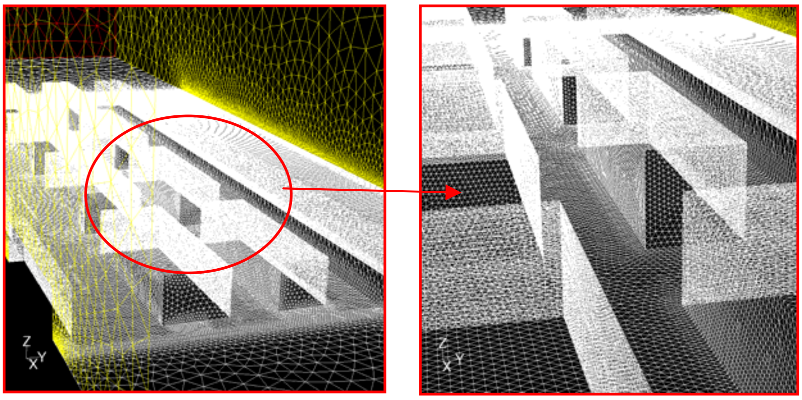


Figure 2: 3-d modeling within a city. Figure taken from E. Solazzo: Advanced Tools for Rational Energy Use towards Sustainability with emphasis on microclimatic issues in urban applications. Presentation made duting the ATREUS project meeting in Thessaloniki,Greece (http://aix.meng.auth.gr/atreus/thessaloniki_meeting.html)

The basic problem resulting from this approach is the fact that it is very demanding when it comes to computational resources. In addition to that, the mesh-based simulation needs to be supplemented by the proper parameters, representing the thermal characteristics of the surfaces, their relative position in accordance to the sun, mechanically generated turbulence (from traffic). These types of problems could be automatically solved, if urban modeling were possible via the use of semantically rich 3D objects (such as CityGML objects). For instance, the thermal characteristics of a surface could be derived from data such as surface geometry and orientation, material, color, etc. that exist in 3D city models. In addition, some of the most important parameters like the roughness length in urban airflow modeling, also result "automatically" from 3D data about obstacle shape, dimensions and relative position, thus making the use of a 3D urban model extremely advantageous in comparison with the way that it is currently being done.

Studies realized with street canyon models revealed that favorable meteorological conditions and appropriate urban geometry can reduce pollutant concentration levels (Kovar-Panskus et al., 2002). Therefore, results generated by these models are crucial for a sustainable and proactive urban planning process. On the other hand, street canyon models take as input parameters relative to the geometry, the topology and the semantics of streets and buildings. Such data are available in 3D city models expressed in CityGML, either directly, either in another form and possibly through some computations. Thus, the interconnection of air quality models with CityGML and OUPP is a first step towards taking into account the city in a more holistic way.

5. Towards an ontology of air quality models

The first phase of our ontology of air quality models focuses on street canyon models.

Underlying concepts and properties of street canyon models

Street canyon models are based on equations taking as input:

- the **pollutant source** characteristics (source location, emitted product, etc.)
- the **meteorological conditions**, mainly the ambient wind conditions (speed, direction relative to the street canyon, etc.) but also, to some extent, the thermal conditions (solar heating)
- the **street canyon** geometry**,** in particular its aspect ratios (height-to-width ratio for 2D and 3D models, height-to-length ratio for 3D models) or its orientation relative to the ambient wind

and producing as output:

- a **flow** characterized mainly by its **vortices** (intensity, rotation direction, location, etc.)
- a **pollutant dispersion distribution**.

Thus, the main underlying concepts are: Street Canyon, Meteorological Conditions, Pollutant Source, Flow , Vortex (as part of Flow) and Pollutant Dispersion Distribution. The concepts Pollutant Source and Pollutant Dispersion Distribution refer to each other. The Ambient Wind Conditions and Thermal Conditions are also main concepts, defined as part of Meteorological Conditions. As a street canyon is the space between buildings that line up continuously along both sides of a relatively narrow street (Liu et al, 2003), the concept Street Canyon has to be defined from the concepts Street and Building. A flow and its vortices are both vector fields. Thus the concept Vector Field has to be added as superconcept of Flow and of Vortex. The pollutant concentration distribution is a scalar field that we want to associate to isosurfaces for visualization of the values. Thus we also need the concepts Scalar Field and Isosurface (see figure below).

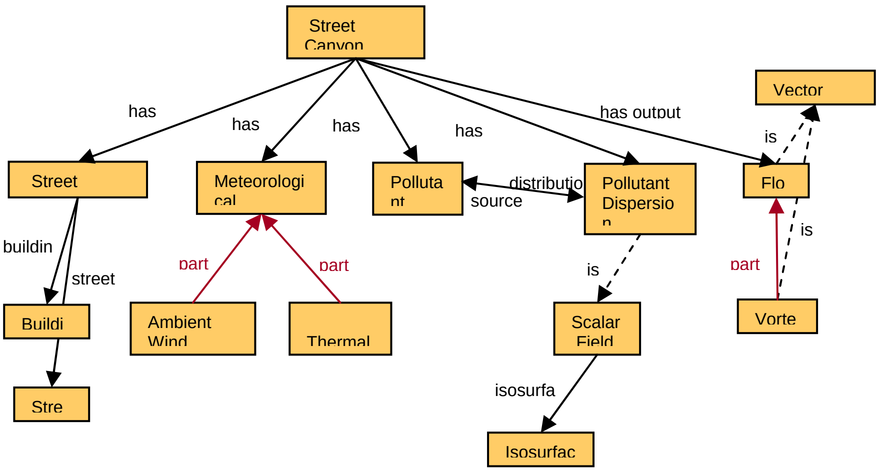


Figure 3: The Street Canyon part of the air quality ontology

The concepts with their properties are shown below:

Street Canyon

shape	wide / square / narrow
height-to-width ratio	ratio of the Building height to the Street width
height-to-height ratio	ratio of the Building height at windward side to the Building height at leeward side
orientation	(absolute)
leeward side location	(depends on the current wind conditions)
windward side location	
level	street-level / building-level / roof-level

Street

location	(absolute)
width	
orientation	(absolute)
albedo	(average reflection coefficient)

Building

location	(absolute)
height	

roof slope	
albedo	(average reflection coefficient)

Meteorological Conditions

Ambient Wind Conditions (as part of Meteorological Conditions)

speed	
direction	(absolute)
turbulence intensity	

Thermal Conditions (as part of Meteorological Conditions)

sunshine duration
temperature of air
temperature at street bottom
temperature on surface of building at leeward side
temperature on surface of building at windward side

The direction of the ambient wind to the street canyon (required in the equations of the model) can be computed from the absolute direction of the ambient wind and the absolute orientation of the street

Pollutant Source

emitted product	
origin	traffic-related / chemical toxics / biological toxics
reactivity	reactive / non-reactive
source location	street-level / advected
emission rate	continuous / non-continuous
dispersion distribution	which is a Pollutant Dispersion Distribution

Flow (which is a Vector Field)

regime	isolated roughness flow / wake interference flow / skimming flow

Vortex is a Vector Field

part of	a Flow
intensity	weak / strong
rotation direction	clockwise / counter-clockwise
location	
origin	mechanically induced / thermally induced
shape	portal vortex / roll-type vortex / horseshoe vortex

Isosurface

value range	
geometry	a closed surface

Vector Field

Scalar Field

isosurface	property which refers to Isosurface concept

Pollutant Dispersion Distribution (which is a Scalar Field)

source	a Pollutant Source

6. Interconnecting CityGML and air quality models

In this section we will present different integration patterns that occur when interconnecting CityGML (CG) and an air quality model (AQ) and how to represent them in the OUPP ontology. Our goal is to provide a formal (logical) description of the interconnections at the conceptual level that can easily be turned into an automated (computerized) process. The main idea is to create "associative" concepts and axioms in OUPP that refer to the interconnected concepts in AQ and CG. The aim is

1. to represent potentially complex integration patterns
2. to put all the integration related entities in OUPP and avoid modifying AQ or CG
3. to provide a kind of centralized mediator for integrating other models (water quality, noise reduction, etc.)

The interconnection technique we propose here is inspired by the work of Mitra, Wiederhold and Kersten (2000) on articulation ontologies. In their interconnection architecture the semantic relationships between two source ontologies are described by articulation rules (implications between concepts of the two ontologies), which are then translated to yield concepts in an articulation ontology and semantic implication edges between the articulation ontology and the source ontologies, They also propose functional rules that are intended to normalize values expressed in different systems of measurement. In the following, we extend this approach by defining interconnection concepts that may have different types of semantic links with the source (CG and AQ) ontologies. Another aspect that distinguishes our approach from most of the other approaches (see (Kalfoglou and Schorlemmer, 2003) or (Shvaiko and Eurenat, 2005) for surveys) is the role of the OUPP ontology. In our context, the OUPP is not simply an articulation or a semantic bridge, it is a ontology with its own concepts that may not exist in the other ones.

Interconnection patterns

0. Property ranges

The simples interconnection pattern occurs when a property of a concept in one ontology has its range in the other ontology. For instance, the geometry property of an isosurface in AQ is a surface, which is a concept defined in CG. In this situation it is not necessary to define an interconnection concept.

1. Concepts with different viewpoints

The properties of the two concepts are not the same in both ontologies and some properties can be computed from those in the other ontology. For instance, street in AQ has a width property while street in CG has a geometry. In addition, the width in AQ can be derived from the geometry in CG.

We introduce the following interconnection concept (concepts with the same name are distinguished by prefixing them with the ontology name)

OUPP:street

properties

in_AQ	a AQ:street
in_CG	a CG:street

Each instance of OUPP:street refers to the same street in AQ and CG.

The following axiom expresses the dependency between the width and geometry properties

for all s **in** OUPP:street : s.in_AQ.width = width_computation(s.in_CG.geometry)

where width_computation represents the (possibly complex) geometric function that computes the width of the polygon that forms the street geometry in CG.

Another similar examples is AQ:building, with its roof slope computed from the roof surface of CG:building and its albedo obtained from the surfaces, material, and color of CG:building.

2. Instances derived from several instances in the other ontology

This is a more complex integration pattern where a concept instance in one ontology corresponds to a set of concept instances in the other one. For example, a street canyon (in AQ) exists only if there is some street (in CG) bordered by buildings in a particular configuration. The interconnection concept has the following definition:

OUPP:street_canyon

properties

in_AQ	a AQ:street_canyon
street	a CG:street
buildings_1	a set of CG:building
buildings_2	a set of CG:building

where buildings_1 and buildings_2 refer to the sets of buildings that border the street on both sides. The following axiom defines the notion of street canyon

for all c **in** OUPP:street_canyon :
for all x **in** c.buildings_1 : borders(x, s)
and
for all y **in** c.buildings_2 : borders(y, s)
and
continuously_aligned(c.buildings_1)
and
continuously_aligned(c.buildings_2)

where borders and continuously_aligned are geometric predicates. In addition, the properties of the street canyon in AQ depend on its components (in CG) properties (as in case 1) and possibly on other factors. For example

for all c **in** OUPP:street_canyon : c.height-to-height_ratio =
average_height(c.buildings_1) / average_height(c.buildings_2)

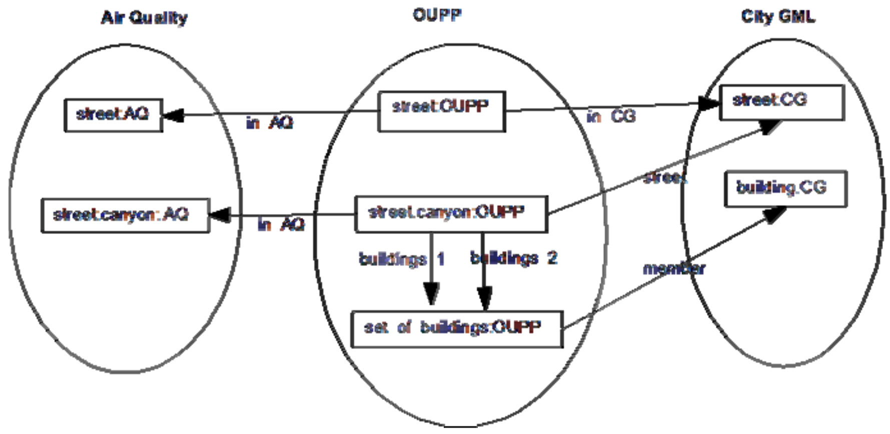


Figure 4: Concepts in OUPP as mediators between the Air Quality and CityGML ontologies

This interconnection technique can be used to “import” theoretical results from one domain into another. For instance, consider the following result:

An experiment led by Baik and Kim (2002) showed that, with buildings of similar height on both sides of the canyon and a street aspect ratio H/W of 2 (building height of 80m and street width of 40m for example), the pollutant concentration in the lower region of the street canyon is higher near the downwind building than near the upwind building. On the other hand, in the upper region of the canyon, the concentration is higher near the upwind building than near the downwind building.

As the lower part of the street canyon corresponds to the area most people walk or cycle, it is important to take into account (import) this result in sustainable urban planning. For example, if most of the time the wind blows in the same direction, it is preferable to position a cycle path on the upwind side of the canyon as it is the side with the lower level of pollutant.

This can be formally represented by defining a new concept “street canyon with favourable upwind side” (SCFUS for short) in OUPP. This concept would be associated with the following axiom:

for all s **in** OUPP:SCFUS : (0.9 < x.in_AQ.height-to-height_ratio < 1.1)
and (1.9 < height-to-width ratio < 2.1)

Similarly, we could define the concept of potentially polluted street as streets that belong to a street canyon where air flow simulation produces a flow with one or more vortices (thus keeping pollutants in the canyon)

From a software engineering point of view, the integration patterns we have shown here specify the computational processes that are necessary to transfer and translate data from one model into the other one.

7. Conclusion and future work

In this paper we have argued that the integration of air quality models with 3D city models can greatly improve air quality modeling and sustainable urban planning. Air quality modeling can take advantage of the semantically rich representations of 3D city models to facilitate air flow simulation. The simulation results and other theoretical results from air quality modeling can then enrich 3D city models to help urban planners deal with air pollution issues.

We have proposed a theoretical and practical approach to air quality and 3D city model integration that is based on the definition of an air quality ontology and the use of an urban planning process ontology that acts as a mediator. Through this approach it is possible to support sophisticated interconnection patterns and to formally specify them. In addition, the OUPP mediation ontology can be re-used to integrate several other models that are of interest for sustainable urban planning.

In the near future we plan to enrich this integration example with other results and concepts from air quality modeling. Then we will apply this same approach to integrate other models, in particular water quality models. We will also develop computational tools to effectively re-use and transfer data between models.

Bibliography

1. Baik, Jong-Jin & Kim, Jae-Jin (1999). A Numerical Study of Flow and Pollutant Dispersion Characteristics in Urban Street Canyons. In *Journal of Applied Meteorology*, 38(11), 1576-1589.
2. Baik, Jong-Jin; Kang, Yoon-So & Kim, Jae-Jin (2007). Modeling reactive pollutant dispersion in an urban street canyon. In *Atmospheric Environment*, 41(5), 934–949.
3. Brundtland Commission (1987), *Our Common Future*. Report from the UN World Commission on Environment and Development (WCED). Oxford University Press, Oxford.
4. EEA (2006). *Air pollution at street level in European cities*. European Environment Agency (EEA) Technical report, No 1/2006, Copenhagen.
5. Huang, H.; Akutsu, Y.; Arai, M. & Tamura, M. (2000). A two-dimensional air quality model in an urban street canyon: evaluation and sensitivity analysis. In *Atmospheric Environment*, 34(5), 689-698(10).
6. Huang, Zhen; Wang, Jiasong; Wu, Zhijun & Xie, Zheng (2002). Study of Impact of Street Layout and Thermal Radiation on Air Quality by Numerical Simulation. In *Proceedings of Better Air Quality in Asian and Pacific Rim Cities (BAQ 2002)*, 16-18 Dec 2002, Hong Kong SAR.
7. Kalfoglu, Y., Schorlemmer, M., Ontology mapping: the state of the art. Knowledge *Engineering Review*, Vol 18:1, 1-31, 2003.
8. Kim, Jae-Jin & Baik, Jong-Jin (2001). Urban street-canyon flows with bottom heating. In *Atmospheric Environment*, 35(20), 3395–3404.
9. Kim, Jae-Jin & Baik, Jong-Jin (2003). Effects of inflow turbulence intensity on flow and pollutant dispersion in an urban street canyon. In *Journal of Wind Engineering and Industrial Aerodynamics*, 91, 309–329.
10. Kim, Jae-Jin & Baik, Jong-Jin (2004). A numerical study of the effects of ambient wind direction on flow and dispersion in urban street canyons using the RNG k–e turbulence model. In *Atmospheric Environment*, 38(19), 3039–3048.
11. Kolbe, T. H.; Gröger, G. & Plümer, L. (2005). CityGML - Interoperable Access to 3D City Models. In *Geo-information for Disaster Management*. *Proceedings of the First International Symposium on Geo-information for Disaster Management*, van Oosterom, Peter; Zlatanova, Sisi & Fendel, Elfriede (eds.), Delft, Netherlands.
12. Kolbe, T. H.; Gröger, G. & Plümer, L. (2006). CityGML - 3D City Models for Emergency Response. In *Geoinformation-Technology for Emergency Response*, ISPRS book series, Taylor & Francis.
13. Kovar-Panskus A., Louka P, Sini J.-F., Savory E., Czech M., Abdelqari A., Mestayer P.G. and Toy N. (2002). Influence of Geometry on the Mean Flow within Urban Street Canyons – A Comparison of Wind Tunnel Experiments and Numerical Simulations. In *Water, Air and Soil Pollution: Focus,* 2(5-6), 365-380(16).
14. Liu Huizhi, Liang Bin, Zhu Fengrong, Zhang Boyin & Sang Jianguo (2003). A Laboratory Model for the Flow in Urban Street Canyons Induced by Bottom Heating. In *Advances in Atmospheric Sciences*, 20(4), 554–564.
15. Metral C., Falquet G. and Vonlanthen M. An Ontology-based System for Urban Planning Communication. In *Ontologies for Urban Development: Interfacing Urban Information Systems*, Geneva, 7-8 November 2006 (forthcoming by Springer).
16. Mitra, P., Wiederhold , G., Kersten, M. A Graph-Oriented Model for Articulation of Ontology Interdependencies. In: Proceedings Conference on Extending Database Technology 2000 (EDBT'2000), Konstanz, Germany, 2000.
17. Moussiopoulos, Nicolas; Berge, Erik; Bøhler, Trond; de Leeuw, Frank; Grønskei, Knut-Erik; Mylona, Sofia & Tombrou, Maria (2006). *Ambient air quality, pollutant dispersion and transport models*. Report of the European Topic Centre on Air Quality, prepared

under the supervision of G. Kielland, Project Manager, European Environment Agency, Copenhagen.

18. OGC (2006). *Candidate OpenGIS CityGML Implementation Specification (City Geography Markup Language)*. Approved discussion paper of the Open Geospatial Consortium, Inc. related to CityGML specification, 06-057r1, available on http://www.citygml.org, consulted on 8 June 2007.
19. Santiago Jose Luis, Martilli Alberto & Martin Fernando (2007). CFD simulation of airflow over a regular array of cubes. Part I: Three-dimensional simulation of the flow and validation with wind-tunnel measurements. In *Boundary-Layer Meteorology*, 122(3), 609-634(26).
20. Shvaiko, P., Euzenat, J. A survey of schema-based matching approaches. *Journal on data semantics*, Vol. 4, 146-171, Springer Verlag, 2005.

Using a hybrid approach for the development of an ontology in the hydrographical domain

F. J. López-Pellicer[a], L. M. Vilches-Blázquez[b], J. Nogueras-Iso[a],
O. Corcho[c], M. A. Bernabé[d], A. F. Rodríguez[b]

[a] Computer Science and Systems Engineering Dept., University of Zaragoza, Zaragoza (Spain)
{fjlopez, jnog}@unizar.es

[b] National Geographic Institute, Madrid (Spain)
{lmvilches, afrodriguez}@fomento.es

[c] School of Computer Science, University of Manchester, Manchester(UK)
ocorcho@cs.man.ac.uk

[d] Technical University of Madrid, Madrid (Spain)
ma.bernabe@upm.es

Abstract: This work presents a hybrid approach for domain ontology development, which merges top-down and bottom-up techniques. In the top-down approach the concepts in the ontology are derived from an analysis and study of relevant information sources about the domain (e.g., hydrographic features). In the bottom-up approach the concepts in the ontology are the result of applying formal methods on a analysis of the data instances on the repositories (e.g., repositories containing hydrographical features)

Key words: Hydrography, Urban Ontologies, Ontological Engineering.

1. Introduction

This work presents a hybrid approach for domain ontology development, which merges top-down and bottom-up techniques. In the top-down approach the concepts in the ontology are derived from an analysis and study of relevant information sources about the domain (e.g., hydrographic features). In the bottom-up approach there is an analysis of application domain repositories (e.g., repositories containing hydrographical features). The results of this analysis are applied to generate dynamically the ontology.

The purpose of applying this hybrid approach is to provide a pragmatic aspect which might help to verify the appropriateness and feasibility of the theoretical domain ontology proposed in top-down approaches with the application ontology obtained in the bottom-up approach. Additionally, the merging of top-down and bottom-up approaches facilitates the mapping between the domain ontology and a particular repository, a task which is usually required for projects related to data harmonization of heterogeneous repositories. This hybrid approach represents a novel way of developing ontologies, which has not been usually applied in the literature of ontological engineering until now. However, we think that it can provide important benefits in contexts that require the harmonization and conversion of heterogeneous data repositories.

Additionally, this work describes as a use case the applicability of this methodology in the context of the Hydrography and Urban Civil Engineering domains. Hydrography and related phenomena represent an essential part of reality in our cities as a consequence of the water supply needs they all have. This is going to characterize some aspects of city planning owing to the presence of water infrastructures and to the addition of certain hydrographic features in urban landscapes (Vilches-Blázquez et al., 2007). Even natural features such as rivers, when crossing urban environments, have their boundaries shaped by people and can be considered as artificial objects (Fonseca et al., 2000).

The Spanish National Geographic Institute (IGN), the organizational body leading the development of the Spanish Spatial Data Infrastructure (IDEE), is defining a hydrographic domain ontology to establish mappings between the IGN feature catalogues and others managed at local, national, regional, and European level. IGN has begun to build a domain ontology of hydrographic features, which is called “hydrOntology", whose purpose is to serve as a harmonization framework among Spanish cartographic producers. For the development of “hydrOntology” we have followed the proposed hybrid approach.

The rest of this paper is organized as follows. The next section describes the hybrid approach methodology, describing the activities involved in this methodology and the techniques applied for each activity. Then section 3 shows the applicability of this approach for the development of this methodology in the hydrography domain. Finally, this paper ends with some concluding remarks and proposals for further research.

2. Hybrid approach methodology for the development of a domain ontology

The methodology for the hybrid approach proposed consists of the following activities:

- Application of the top-down approach: The objective of this activity is the development of a draft version of the ontology following a top-down approach.
- Application of the bottom-up approach: The objective of this activity is the development of a draft version of the ontology following a bottom-up approach.
- Comparison of ontologies: The objective of this activity is to find a consensus between top-down and bottom-up approaches.

Figure 1 shows the process proposed for this hybrid approach methodology. The following subsections describe in more detail the top-down and bottom-up approaches.

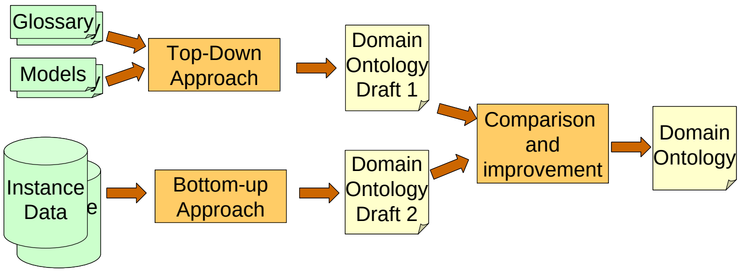


Figure 1: Hybrid approach methodology

2.1. Top-down ontology

For the top-down approach we propose the use of METHONTOLOGY, a widely-used methodology for building ontologies. METHONTOLOGY emphasizes the reuse of existing domain and upper-level ontologies and proposes to use, for formalization purposes, a set of intermediate representations that can be later transformed automatically into different formal languages. Therefore this methodology is suitable for developing ontologies at the knowledge level. Moreover, it takes into account the main activities identified by the IEEE software development process (IEEE, 1996) and other knowledge engineering methodologies.

METHONTOLOGY has been used by different groups to build ontologies in different knowledge domains, such as Chemistry, Science, Knowledge Management, e-Commerce, etc. A detailed description of the methodology of this ontology building can be found in (Gómez-Pérez et al., 2003). Figure 2 shows the ontology building tasks suggested in the METHONTOLOGY framework (Corcho et al., 2005).

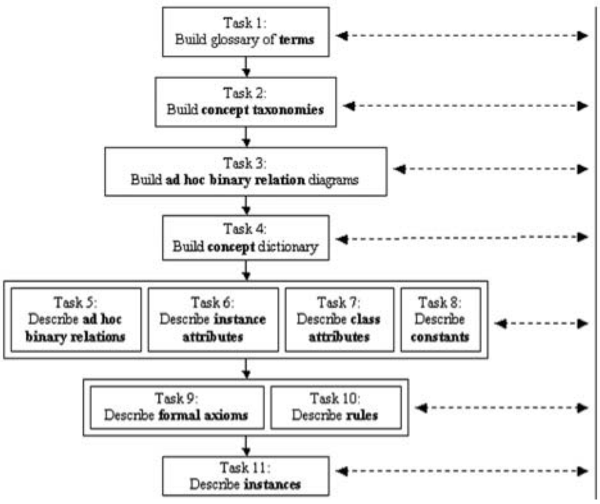


Figure 2: Tasks of the conceptualization activity according to METHONTOLOGY (Gómez-Pérez et al., 2003)

The figure 2 emphasizes the ontology components (concepts, attributes, relations, constants, formal axioms, rules and instances) built inside each task. Also, this figure illustrates the steps this methodology proposes for creating such components during the conceptualization activity. This is not a sequential modelling process, though some order must be followed to ensure the consistency and completeness of the represented knowledge(Corcho et al., 2005).

METHONTOLOGY proposes a set of tasks for capturing a knowledge domain(Gómez-Pérez et al., 2003). These tasks can be divided into three groups.

The first group would be steering to enclosure and structure the domain by means of tasks 1 to 4 (see figure 2).

- Task 1: To build the glossary of terms that identifies the set of terms to be included on the ontology, their natural language definition, and their synonyms and acronyms.
- Task 2: To build concept taxonomies to classify concepts. The output of this task could be one or more taxonomies where concepts are classified.
- Task 3: To build ad doc binary relations diagrams to identify ad hoc relationships between concepts of the ontology and with concepts of other ontologies.
- Task 4: To build the concept dictionary, which mainly includes the concept instances for each concept, their instance and class attributes, and their ad hoc relations.

The second group of tasks, from 5 to 7, would help to document the acquired knowledge from the previous tasks.

- Task 5: To describe in detail each ad hoc binary relation that appears on the ad hoc binary relation diagram and on the concept dictionary. The result of this task is the ad-hoc binary relation table.
- Task 6: To describe in detail each instance attribute that appears on the concept dictionary. The result of this task is the table where instance attributes are described.
- Task 7: To describe in detail each class attribute that appears on the concept dictionary. The result of this task is the table where class attributes are described.

Finally, METHONTOLOGY proposes others tasks, from 8 to 11, to complete a domain knowledge.

- Task 8: To describe in detail each constant and to produce a constant table. Constants specify information related to the domain of knowledge; they always take the same value, and are normally used in formulas.
- Once concepts, taxonomies, attributes and relations have been defined, METHONTOLOGY proposes to describe formal axioms (task 9) and rules (task 10) that are used for constraint checking and for inferring values for attributes. Optionally, information about ontologies should be introduced (task 11).

It is important to mention that different domain ontologies may have different knowledge representation needs, so this methodology suggests that the previous set of tasks should be reduced or extended as needed.

2.2. Bottom-up ontology

For the development of a domain ontology following a bottom-up approach we propose the applicability of (FCA) techniques (Ganter and Wille, 1999; Stumme and Maedche, 2001) to output a hierarchy of concepts from the feature instances contained in the repositories used as data sources (See figure 3).

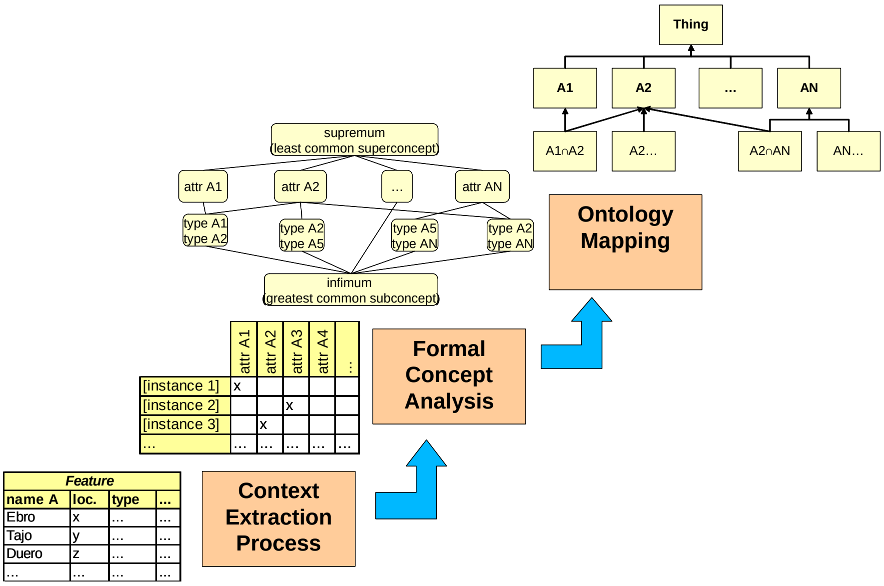


Figure 3: Bottom-up process

The basis of FCA is the definition of a *formal context* (), which consists in a triple where G is a set of *objects* and M is a set of *attributes*. I (*incidence matrix*) represents the binary relation between “objects" and “attributes" with only two possible values, *present* or *absent*.

There are two *closure operators* that link G and M within a *formal context* **K**:

$$A \subseteq G,\ A' = \{m \in M \mid \forall g \in A,\ (g,m) \in I\} \quad (1)$$

$$B \subseteq M,\ B' = \{g \in G \mid \forall m \in B,\ (g,m) \in I\} \quad (2)$$

A' can be understood as the maximum set of attributes common to the objects in A and B' as the maximum set of objects which have in common the attributes in B. Given these definitions, the pair (A,B) is called a *formal concept* if and only if:

$$A \subseteq G,\ B \subseteq M,\ A' = B \wedge A = B' \quad (3)$$

In other words, (A,B) is called a *formal concept* if and only if the maximum set of attributes shared by the objects in A is B and, on the other hand, A is the maximum set of objects which share the attributes in B. A is called the concept *extent* and B the concept *intent*. The set of all the *formal concepts* of the *formal context* is partially ordered by the order induced by the set inclusion:

$$(A_1,B_1) \leq (A_2,B_2) \Leftrightarrow A_1 \subseteq A_2 \; (\Leftrightarrow B_2 \subseteq B_1) \quad (4)$$

Where the formal concept (A_1,B_1) is called *subconcept* of the formal concept (A_2,B_2), and (A_2,B_2) is called *superconcept* of the formal concept (A_1,B_1). Furthermore, the induced partial order is a complete *lattice*, known in this context as *concept lattice*

Comparing FCA with respect to Object Orientation, *formal concepts* are equivalent to *classes*, and *superconcepts* and *subconcepts* relationship between concepts are equivalent to the *generalization* and *specialization* relationships.

FCA techniques have a direct application on relational database repositories consisting of one single table. In this case, each row is mapped to an object and each column to a set of attributes. The incidence relation I can be derived derives from the contents of the table: for each row the presence or absence of a value or range of values in a column determines the presence or absence of one or several attributes. However, the mapping between rows, columns and data values from the repository and objects and attributes of I is a non trivial task if the relational schema is not normalized.

Therefore, previous to the application of FCA techniques, the main issue is how to obtain from different repositories a unified and harmonized view of the data in terms of *objects* and *attributes*, i.e. the formal context required by FCA. As our purpose is to create an ontology draft, the selected data should contain thematic attributes. The data that best fit to this requirement is *hydrologic gazetteer data*. Among other thematic attributes, each gazetteer feature is described as belonging to a feature type and its name may contain valuable thematic data in the *generic name* (identification of feature type that may be extracted as a substring from the name).

Our approach is as follows:

1. Select the gazetteer entries related to hydrography. Also prepare a set of common hydrographic names with their variants.

   $GAZ \leftarrow Hydrographic\ gazetteer$

   $GEN \leftarrow Hydrographic\ names$

2. Set initially G as the set of features contained in the gazetteer.

   $G \leftarrow \{g|g \in GAZ \cdot isFeature(g)\}$

3. Set M as the set of feature types used in the gazetteer that belong to the hydrographic domain along with those generic hydrographic names which appear in the selected features.

   $M \leftarrow \{(g,m)|g \in GAZ\cdot\ isFeatureType(t)\} \cup \{n|n \in GEN \cdot \exists\ g \in G, contain(g \rightarrow name, n)\}$

4. Define I initially as the incidence relationship between features and generic names.

   $I \leftarrow \{(g,m)|g \in G \cdot \forall m \in M \cdot isGeneric(m) \wedge contain(g \rightarrow name,m)\}$

5. Remove from *G* features whose name does not contain a generic name.

   $G \leftarrow G \setminus \{g | g \in G \cdot \nexists m \in M \cdot (g, m) \in I\}$

6. Complete *I* with the incidence relationship between the remainder features and their feature types.

   $I \leftarrow I \cup \{(g,m_1). | m_1 \in M \cdot \exists (g,m_2) \in I \cdot isFeatureType(m_1) \wedge g \rightarrow type = m_1\}$

The work with generic names is not an easy task. Gazetteers can contain *multilingual generic names* and *synonyms*. Another issue is the existence of *slight differences* between the generic name of a feature name and those in *M*. Finally there exists the possibility that the generic name and the feature type of a feature have *different semantics*. The multilingual generic names problem is solved with the use of a dictionary. The most promising approach to the matching problem is the use of robust string matching libraries, e.g. *SecondString*(Cohen et al., 2003). And the occasional different semantics problem is solved by counting duplicate rows in *I* and removing them, and hence the correspondent *g*, if their number is below a threshold.

Once the incidence matrix has been obtained, the concept lattice is generated using one of the several algorithms available, in our case *next closed set* (Ganter, 1987). This generated lattice identifies:

1. Relevant feature types from their *extent*.
2. New feature types derived from *formal concepts* that contain a generic as attribute.
3. Feature types that are candidate to a disjoint-decomposition.

Thanks to the FCA technique and some minor adjustments, the original feature type taxonomy can be enriched in a way that helps the ontology designer to understand better the domain.

3. Experiment: Applying the hybrid approach methodology to the hydrography domain

As mentioned in the introduction, IGN is defining a hydrographic domain ontology to establish mappings between their own feature catalogues and others managed at local, national and European level. This domain ontology is called “hydrOntology” and it has been developed following the hybrid 1approach described in the previous section.

The following subsections describe the applicability of the hybrid approach methodology to the development of “hydrOntology".

3.1. Top-down ontology

In order to develop our ontology following the top-down approach, we have taken into account different knowledge models (the feature catalogues of the National Geographic Institute of Spain, the European Water Framework Directive, the Alexandria Digital Library, the UNESCO Thesaurus and other resources), some integration problems of geographic information and several structuring criteria (Vilches-Blázquez et al., 2007). We have tried to cover most of existing GI sources in order to build a full domain ontology. For that reason, this ontology contains more than a hundred relevant concepts related to hydrography (e.g. river, reservoir, lake, channel, pipe, water tank, siphon and so on).

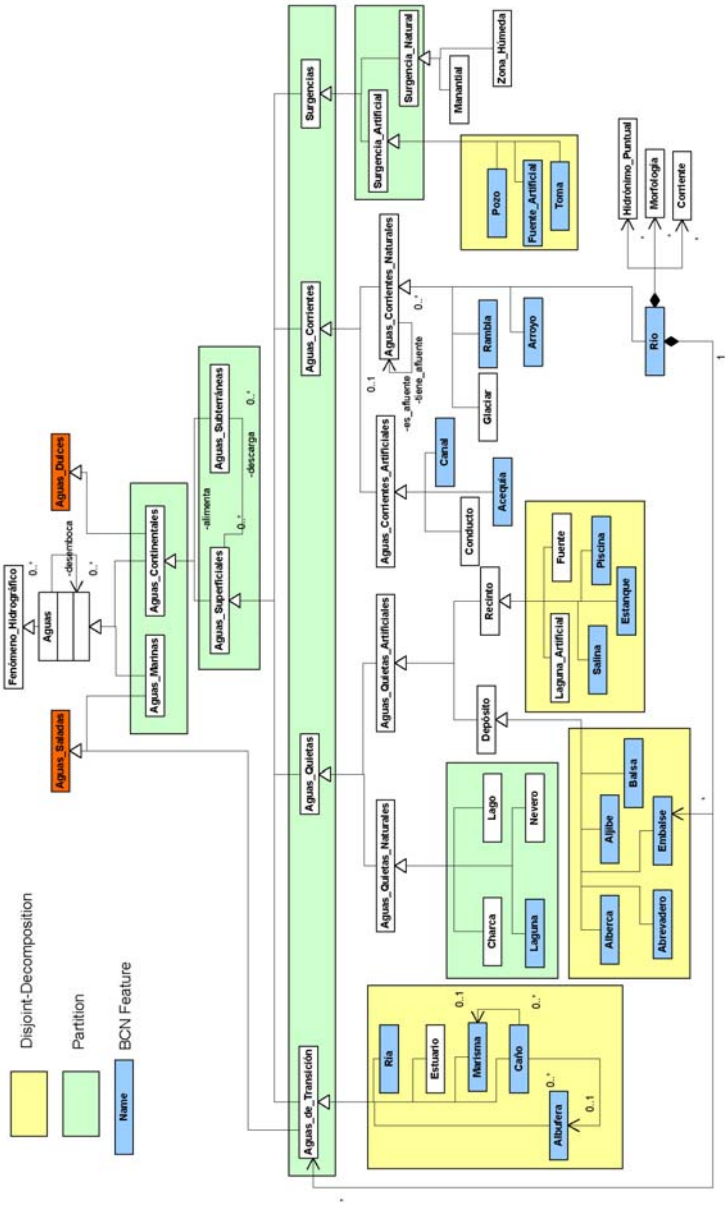


Figure 4: Top-down ontology

Figure 4 shows a “hydrOntology” model overview. It is divided into two levels; the upper level represents the most abstract features in the ontology and the lower level describes a set of well-known hydrographic features. The upper level contains the “Hydrographical Feature” concept, and other specialised concepts like “Inland Waters” and “Sea Waters”. There is a different degree of specialisation in each of these concepts, since the current focus of this ontology is on “Inland Waters”. According to the Water Framework Directive (European Parliament, 2000), these concepts are divided into “Superficial Waters” (“Transitional waters”, “Stand Waters”, “Flowing Waters” and “Sources” are subclasses of “Superficial Waters”) and “Groundwaters”. For each of these classes we have identified concepts in the lower level, where a detailed set of hydrographic features is provided.

Furthermore, in the “hydrOntology” development we have taken into account some concepts about feature capture that depend exclusively on different Spanish geographic regions. Among these features one could mention “ibón”, “lavajo”, “chortal”, “bodón” and “lucio”. These concepts are designated by their local name and they are synonymous to the feature “Charca” [1].

Moreover, this figure shows some examples of the four taxonomic relations defined in the Frame Ontology (Farquhar et al., 1997) and the OKBC Ontology (Chaudhri et al., 1998), both used by METHONTOLOGY (Vilches-Blázquez et al., 2007).

A concept *C1* is a *Subclass-Of* another concept *C2* if and only if every instance of *C1* is also an instance of *C2* (Corcho et al., 2005).

A *Disjoint-Decomposition* of a concept *C* is a set of subclasses of *C* that do not have common instances and do not cover *C*, that is, there can be instances of the concept *C* that are not instances of any of the concepts in the decomposition (Corcho et al., 2005). Some examples of this type of relationship are shown in figure 4.

An *Exhaustive-Decomposition* of a concept *C* is a set of subclasses of *C* that cover *C* and may have common instances and subclasses, that is, there cannot be instances of the concept *C* that are not instances of at least one of the concepts in the decomposition (Corcho et al., 2005). Figure 4 shows an example of this type of relationship.

A *Partition* of a concept *C* is a set of subclasses of *C* that do not share common instances and that cover *C*, that is, there are not instances of *C* that are not instances of one of the concepts in the partition (Corcho et al., 2005). Some examples of a partition are shown in figure 4.

At the moment we are working on providing mappings of this ontology with other databases at several levels (from local to national level). Furthermore, we are planning to provide multilingual support for “hydrOntology” (English, French, Portuguese, Catalan, Basque, Galician languages) and to merge this ontology with other domain ontologies (e.g. Urban Civil Engineering).

3.2. Bottom-up ontology

For the bottom-up approach we have analyzed the repositories that have been used to build a gazetteer at the Spanish National Geographic Institute. In particular, we have focused on the part of the repositories used as source for the generation of the hydrographic names.

---

[1]“Charca” is a small pond of shallow water. The above mentioned terms are Spanish local names.

Figure 5 shows how the process has been applied to the feature repositories. As it can be observed, the *Thematic Analysis* module determines the feature type and the generic name of each feature. Both the feature type and the generic name are the thematic signature of a feature. Then, the *Signature Filter* selects the distinct signatures that represent a significant number of features; and the *Formal Context Builder* creates an incidence matrix whose rows are these signatures. Finally, the *Lattice Builder* applies FCA and the *Ontology Generator* transforms the formal concept lattice into OWL (Web Ontology Language) (Bechhofer et al., 2004), an RDF-based language to express ontologies.

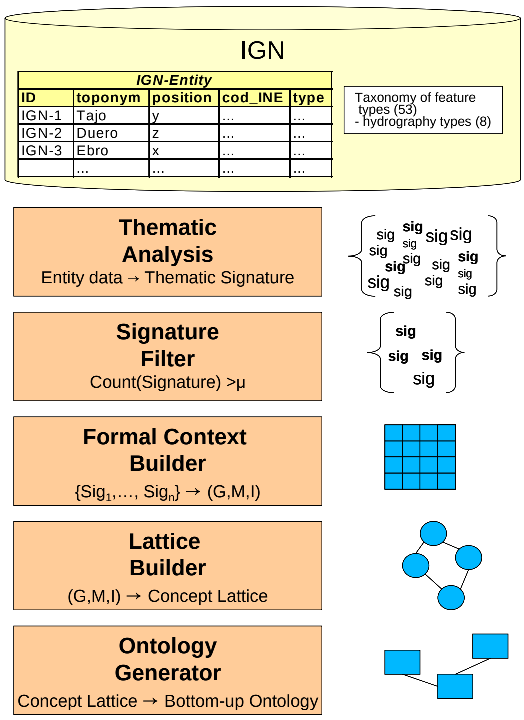


Figure 5: Building from the repository

Figure 6 shows part of the generated ontology. This ontology contains 51 concepts. They can be classified from their source as follows: the original *IGN Feature Types*, which are denoted with the suffix “IGN"; the types extracted from the name of the features, i.e. *generic names*, which are denoted with the suffix “GEN”; and the concepts derived from the combination (map) of concepts in the previous groups, which are denoted with the prefix “MAP". The most common concepts by far are “Corriente fluvial IGN" (stream of water) (71 % of instances) and its subclass “MAP Corriente fluvial arroyo" (52 % of instances), the map between “Corriente fluvial IGN" (stream of water) and “Arroyo GEN" (creek) .

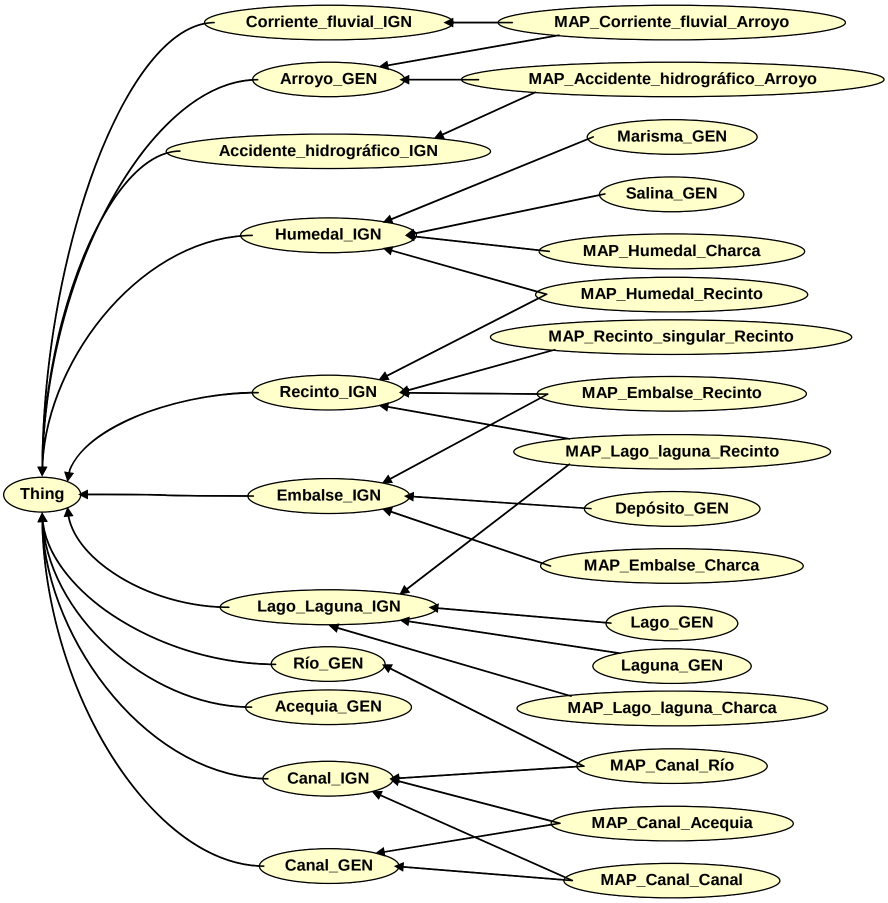


Figure 6: Bottom Up Ontology (part)

3.3. Comparison of results for the improvement of "hydrOntology"

The comparison between the results of the Formal Concept Analysis and "hydrOntology" has given the opportunity to obtain the necessary feedback to evaluate the feasibility of "hydrOntology" and enrich it, if necessary. The most outstanding results of this comparison have been the following:

Detection of concept equivalence: There are equivalent concepts in both ontologies. Some IGN feature types have the same name as concepts described in "hydrOntology" and share the same semantics, e.g. "embalses" (dams) or "corriente fluvial" (streams). This helps to reinforce the existence of concepts in "hydrOntology".

Detection of semantic heterogeneity: The comparison has detected cases of semantic heterogeneity as well. According to (Bishr, 1998), semantic heterogeneity is defined as the consequence of different conceptualizations and database representations of a real world fact. There are two types of semantic heterogeneity that can be distinguished in this comparison process:

Naming heterogeneity: It occurs when the same concept is named differently. There are different examples of this type of heterogeneity that can derive possible advices for the improvement of “hydrOntology”:

1. “hydrOntology” distinguishes locally bound concepts that share similar characteristics and are mapped to one single IGN feature type. For instance, one can find different concepts for small ponds such as “Bodón" or “Lavajo". Another example is the concept “Ría", which is a transitional water feature type only found in the north of Spain. The recommendation for the improvement of “hydrOntology” could be that these concepts should be either merged or better characterized:
   - -. Locally bound concepts should be maintained if only if they are relevant in size or number in an area. And if they were maintained, their characterization should be improved with a description of the geographic region where they exclusively occur.
   - -. If they were merged, each concept should have not only multilingual support but also the dialectal and local variants in each language.
2. There are cases where a single IGN feature type corresponds to several concepts in “hydrOntology” because all these concepts share non-thematic attributes such as position, shape or size. The best example is “Accidente hidrográfico" (hydrographic feature) that contains features described in “hydrOntology" as thermal features (“Terma”), springs (“Manantial”) and parts of rivers (e.g., “Meandro”, or “Poza”) whose only shared characteristic is their representation as a point. The detection of this *one-to-many* mapping may help to identify constants and axioms in “hydrOntology”.

Cognitive heterogeneity: It occurs when the same term is used for different concepts. We have found two separate cases of this problem:

1. Two concepts share the same term, but one of them has broader semantics. An example of this case occurs between the IGN Feature type “canal” and the concept “canal” from “hydrOntology”. The semantic meaning of “canal” in IGN is that of an artificial stream instead of the expected narrower meaning of an irrigation channel. In contrast to IGN feature types, “hydrOntology” has correctly identified both the broader meaning (“Aguas Corrientes Arficiales”, artificial stream) and the narrower meaning (“Canal”, irrigation channel). The detection of this heterogeneity will remind the ontology designer to verify the inheritance hierarchy in both ontologies to identify possible errors.
2. There are IGN concepts named with compound terms that may correspond to two (or more) concepts of “hydrOntology”, both of them named with part of this compound term. An example of this problem occurs with concepts related to wetlands. The FCA concept lattice contains several concepts derived from the IGN Feature type “Humedal” (wetland) and the *generic names* identifying the origin of this wetland: “MAP Humedal Charca” (wetland originated by a small pond), “MAP Humedal Marisma” (wetland originated by a marsh), and so on. This initial problem may help to identify two issues. On the one hand, one may identify a missing

relation between the concepts in “hydrOntology”. For instance, a new relation could be established between the wetland concepts (concept “Zona Humeda” and its subclasses) and the concepts representing the origin of these wetlands. On the other hand, the ontology may reconsider the appropriateness of defining partitions of concepts, or even defining separate concepts. For instance, “hydrOntology” defines two separate concepts related to small ponds: “Charca” (the concept originating the wetland) and “Zona Encharcable” (the wetland originated by the small pond). The question is whether in real world one can find repositories distinguishing the instances of both concepts.

4. Conclusions

This work has presented a hybrid approach for domain ontology development, which merges top-down and bottom-up techniques. Each technique produces ontologies which differ in their respective point of view. The top-down ontology draws the required/expected semantic of the data held in the repositories. The bottom-up ontology reveals the effective/possible semantics of the data held in the repositories. The comparison of both ontologies provides useful information and feedback.

As regards the experiments in the hydrography domain, we can conclude that the ontology derived from FCA has provided insight on possible missing attributes and relationships in “hydrOntology" and advice on how to improve the multilingual support or to treat locally bound feature types. Future work will be oriented to find more automatic mechanisms for the comparison and merging of top-down and bottom-up approaches. For instance, we could merge both ontologies using tools such as PROMPT (Noy and Musen, 2000).

Acknowledgements

The work of F.J. López-Pellicer and J. Nogueras-Iso has been partially supported by the Spanish Ministry of Education and Science through the project PET2006_0245 from “the National Plan for Scientific Research, Development and Technology Innovation”'.

References

Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D. L., Patel-Schneider, P. F., Stein, L. A., February 2004. OWL Web Ontology Language Reference. W3C, W3C Recommendation 10 February 2004, http://www.w3.org/TR/2004/REC-owl-ref-20040210/.

Bishr, Y., 1998. Overcoming the semantic and other barriers to GIS interoperability. International Journal of Geographical Information Science 12 (4), 299–314.

Chaudhri, V. K., Farquhar, A., Fikes, R., Karp, P. D., Rice, J. P., January 1998. Open Knowledge Base Connectivity 2.0.3. Technical Report KSL-98-06, Knowledge Systems Laboratory, Stanford, CA, http://www.ai.sri.com/ okbc/okbc-2-0-3.pdf.

Cohen, W. W., Ravikumar, P., Fienberg, S. E., 2003. A Comparison of String Distance Metrics for Name-Matching Tasks. In: Proc. IIWeb 2003 (IJCAI 2003 Workshop). pp. 73–78.

Corcho, O., Fernández-López, M., Gómez-Pérez, A., López-Cima, A., 2005. Law and the Semantic Web. Legal Ontologies, Methodologies, Legal Information Retrieval, and Applications. Springer-Verlag, Ch. Building legal ontologies with METHONTOLOGY and WebODE, pp. 142–157.

European Parliament, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. The EU Water Framework Directive - integrated river basin management for Europe. L 327, 22/12/2000 pp. 0001-0073, European Parliament.

Farquhar, A., Fikes, R., Rice, J., 1997. The Ontolingua Server: A Tool for Collaborative Ontology Construction. International Journal of Human Computer Studies 46 (6), 707–727.

Fonseca, F. T., Egenhofer, M. J., Davis Jr., C. A., Borges, K. A. V., 2000. Ontologies and knowledge sharing in urban GIS. Computers, Environment and Urban Systems 24 (3), 251–271.

Ganter, B., 1987. Algorithmen zur formalen begriffsanalyse. In: Ganter, B., Wille, R., Wolff, K. E. (Eds.), Beiträge zur Begriffsanalyse. B.I.–Wissenschaftsverlag, Mannheim, pp. 241–254.

Ganter, B., Wille, R., 1999. Formal Concept Analysis: Mathematical Foundations. Springer, Berlin-Heidelberg.

Gómez-Pérez, A., Fernández-López, M., Corcho, O., 2003. Ontological Engineering. Springer-Verlag, London (United Kingdom).

IEEE, 1996. IEEE Standard for Developing Software Life Cycle Processes. IEEE Std 1074-1995. IEEE Computer Society, New York.

Noy, N., Musen, M., 2000. Prompt: Algorithm and tool for automated ontology merging and alignment. In: Proceedings of the Seventeenth National Conference on Artificial Intelligence (AAAI-2000). Austin, Texas, pp. 450–455. http://smi-web.stanford.edu/people/noy/publications.htm%l

Stumme, G., Maedche, A., 2001. FCA-MERGE: Bottom-up merging of ontologies. In: Proc. 17th IJCAI. Seattle (WA US), pp. 225–230.

Vilches-Blázquez, L. M., Ángel Bernabé-Poveda, M., Suárez-Figueroa, M. C., Gómez-Pérez, A., Rodríguez-Pascual, A. F., 2007. Ontologies for Urban Development: Interfacing Urban Information Systems. Vol. 61 of Studies in Computational Intelligence. Springer, Ch. Towntology & hydrOntology: Relationship between Urban and Hydrographic Features in the Geographic Information Domain, pp. 73–84.

Web-based Interactive Visualization of Uncertain Concepts in Urban Ontologies

Hyowon Ban [a,*] and Ola Ahlqvist [a]
[a] Department of Geography, The Ohio State University
1036 Derby Hall, 154 N Oval Mall
Columbus, OH 43210, U.S.A.
[*] ban.11@osu.edu

Abstract: In urban ontology studies, there exist different semantic definitions of exurbanization. Uncertainty from these conceptual definitions can be represented, compared, and negotiated by using graphical representation in software interface, also with the corresponding empirical spatial data can be geovisualized interactively. We develop web-based software to aid visualization and negotiation between uncertain concept definitions and real-time graphical illustration. It aims to support communications between stakeholders in virtual environment (VE). We illustrate its use on a case study of existing multiple definitions of exurbanization. The results show that the software has sufficient flexibility for negotiation between different users in diverse applications.

Key words: Sprawl ; uncertainty ; negotiation ; interactivity ; visualization ; virtuality.

1. Introduction

In Urban Development research, different groups use different concepts about the urban ontologies [08]. This may cause ambiguity due to subjective interpretation when they make a decision [27]. Many European countries try to establish consensus in urban ontologies to deal with multiple languages [28]. Ontology research in GIScience has focused on similar problems between conceptualizations [38]. One example is research on different definitions of the same phenomena with stability of language across culture and geography [06].

The objective of this paper is to develop a web-based interactive program to aid visualizing and negotiating between uncertain concept definitions for urban planning using a real-time user interface. A few papers have suggested graphic representation of the uncertainty of concepts and these concepts are represented as triangular, quadratic, trapezoidal, Gaussian, or log-exponential shape [15]. However, there has been limited graphical interactivity of the representation as an interface.

In this paper, we present a software application ‘pinu’ - Program for Identification and Negotiation of Uncertain concept definitions - that supports user interaction to update the graphical representation and the data. It is developed with Macromedia Flash™ and Virtools™ which make the application accessible through a web browser.

Flash is a format of vector-based interactive movie proposed by Macromedia, Inc. [33], which can be embedded in web pages and delivered over the Web. Flash has several unique features including small size for fast delivery, easy composition, rich semantics, and powerful interactivity [43]. Virtools is developed by Virtools Behavior Company for designing interactive 3D objects in web environment. Each object has a model with associated behaviors in terms of movement and interaction with other objects [34].

Both Flash and Virtools can communicate with each other on the web through JavaScript. The demonstrated software pinu reads definition information of concepts related to exurbanization in XML file format from the web, visualizes the uncertainty between the definitions graphically, adjusts the degree of each definition's effect by user input from the slider bars, restores the updated data into a local XML file that can be posted again to the web as an updated definition, and update the empirical spatial data in the virtual environment (VE) interactively using the new definitions.

In the results we illustrate its use on a case study of existing multiple definitions of exurbanization [13];[35]. This is followed by a conclusion where we argue that the pinu application can be useful for stakeholders, Urban Development specialists, or organizations to improve communication in urban land-use.

2. Theoretical background

2.1. Uncertain concept definitions of exurbanization

Recently, the "exurbs" have become important since it is rapidly growing and this has major implications for land-use [22]. There are various and separate concept definitions of exurbanization and uncertainty in determining the empirical exurban boundaries based on these definitions [08]. In some studies, the exurban area is defined using population and distance from a city, although exact limits are different between researchers. For instance, according to Daniels [13], an exurban area is 10 to 50 miles away from a major urban center of at least 500,000 people, or 5 to 30 miles from a city of at least 50,000 people. It should be generally within 25-minute commuting distance and the population density is less than 500 people per square mile. However, Nelson [35] argues that exurban counties are within 50 miles of the boundary of a central city of a Metropolitan Statistical Area (MSA) with a population of between 500,000 and less than 2 million, or within 70 miles of the boundary of the central city of an MSA with more than 2-million of population. Both Daniels [13] and Nelson [35] point out the uncertainty of determining exurbanization since the boundary between rural lands and suburban areas is ambiguous.

In this study, we use the different definitions of exurbanization in Daniels [13] and Nelson [35] as competing concept definitions. In studies of complexity of social systems, Holling [20] argues that views of complexity require competing models, hypotheses, and alternative perspectives. The multiple explanations define what is known, what is uncertain, and what is unknown so best judgments, not certainties are left. Similarly, in studies of natural resource management, a decision is seldom the result of a single hypothetical decision-maker, it is rather a matter of interactions between several stakeholders [41]. According to Barreteau et al. [09], "decisions" emerge from these interactions via a small number of stakeholders. In their study, a tool for group decision support is suggested to enable the decision process. By introducing the tool, they raise issues of stakeholder participation in decision making through better interactions and negotiations.

All these interactions are based on "mediating objects" [09];[24]. The mediating object induces stakeholders to agree on a common representation of their interests and facilitates communication in relation to it. In negotiating exurban boundary-issue, the common representation of exurban area in computational way might be also useful to mediate some problems in the issue. The representation could be done with a software tool as in Teulier-Bourgine [39]: problems encountered by and known to every stakeholder are brought to common knowledge, thereby preventing certain misrepresentations. Especially, a simulation tool leads each stakeholder to specify and modify his/her representations, leading to a better knowledge of the representations of others. The shared representations contribute to stakeholders' cognitive capacities by enhancing the decision process [39];[17];[09]. In this study, we develop a web-based software tool that enables a common representation of exurban boundary issues to be negotiated between policy makers, stakeholders, or researchers.

Negotiation and conflict resolution using computational systems have been well-studied in economics, decision theory, multi-criteria optimization, social-emotional, and artificial intelligence perspectives [12]. However, it is only recently that frameworks for designing and developing computer-aided decision support for negotiation/conflict resolution have emerged [23];[12]. In other studies of collaborative decision-making, interactive computer-based systems facilitate solutions for ill-structured problems between a set of decision makers working together as a team [25];[36]. Barreteau et al. [09] found a combination of multi-agent systems and role-playing games (RPG) to be an effective discussion support tool to derive renewable resource-management. In their work, a promising potential of RPG and other VEs for research, training, and negotiation support is provided.

In our developed software, we let values of attributes in a definition be modified by users to support negotiation of semantics between the definitions of Daniels [13] and Nelson [35]. The uncertainty approach with the full negotiation of concepts could be useful to reveal differences between the competing explanations, models, or hypothesis in urban ontologies and finally negotiate them to arrive at better judgments.

3. Visualization and negotiation of the uncertainty in exurbanization

Since we perceive the space and add cognitive interpretations to it [40], one of fundamental issues in understanding space is how we conceptualize, measure, represent [05], and share the individual and not necessarily identical knowledge with others. Voudouris et al. [40] recently demonstrated the potential benefits of sharing knowledge, constructing knowledge, and exploring data collaboratively using visual methods. They argue that little attention has been given to recording uncertainties of such knowledge, however it may be necessary when different people in a collaborative environment deal with different conceptualizations of the same phenomenon. For example, exurbanization could be conceptualized and interpreted individually, and the different conceptualizations could be integrated with each other.

Therefore, in urban ontology providing a method that enables people to share the data and knowledge each other could be important to land-use policy making. To measure individual uncertainty, Lowell [29] estimates the uncertainty of multiple interpretations about spatial phenomena from individuals by using a distance measurement. During the estimation, how different people address the phenomena is interactively recorded and visualized, suggesting a shift from data uncertainty to individual knowledge uncertainty. In terms of visualization approach, Voudouris et al. [40] develop a framework of "sketching", a mechanism to visualize uncertain boundaries using rough drawings or sketches, which might support reasoning about an existence of a region. In their approach, they use the "degree" of uncertainty for each location instead the finite boundary.

Deitrick and Edsall [14] argue that research into the effectiveness of uncertainty visualization for decision supporting is beginning to emerge. They present the influence of visualizing the uncertainty in the decision making process using interactive exploratory tools and static paper maps. In earlier studies, MacEachren [30] provided general methods for depicting uncertainty by using data exploration tools to allow users to display its uncertainty. Gershon [18] suggested general categories for representation of intrinsic and extrinsic uncertainty. However, these approaches are limited by using static or paper maps which hardly provide interactive representation of spatial data.

Uncertainty visualization may affect decisions, but the degree of influence is affected by how the uncertainty is expressed [14]. Therefore, visualization of uncertainty using fuzzy boundary than crisp boundary could be useful in expressing uncertain urban ontology such as the concept of exurbanization [08]. Preliminary study from stand-alone 3D geovisualization of uncertainty of exurbanization [07] presents fuzzy representation of similarity and difference between the two definitions that is followed in this study. The relevance to explore the semantic relationships between concepts and to negotiate different definitions of exurbanization visually was demonstrated in the existing study [07].

With the aid of recent technological development in visualization, 3-dimensional representation in VE could be helpful in dealing with spatial data. For the information visualization that could be changed or negotiated by a user, interactivity of the user interface becomes important. Heim [19] argues that the interactivity from the user-engagement brings "a reciprocal relationship" between user and producer since high-speed computer transmission makes the separating line between viewer and producer "blur". The user can have a better sense of space by navigating "in" it with the aid of wearable equipments such as head-mounted display, data glove, 3D goggles, and so on. VE's web-accessibility could be useful for shared-environment for negotiation and communication between stakeholders or residents in a neighborhood, for example.

Cash et al. [11] argue that there is often a gap between a highly generalized understanding produced by formal science and the experiential and practice-based understanding. This gap can also exist within the lack of interaction between semantics of urban ontologies and empirical representation of spatial data. In our proposed approach, the formal semantics of exurban definitions and the practice-based spatial visualization interact within one software-interface.

Often, negotiations show a large variance in semantics, and some negotiations in the domain of electronic business markets are made by using interactive multi-attribute [37]. For example in the Daniels's [13] definition of exurban area, a negotiation can compromise not only certain 'values'—i.e. the "10 miles"—of certain 'attributes'—i.e. the "distance from a major urban center" —but also the attributes—i.e. "distance" and/or “population density”—themselves. Many researches argue that uncertainty information is a vital component in the use of spatial data for decision support [21];[01]. Researchers have emphasized the need for empirical research to test the effectiveness of visual variables and their usefulness in depicting uncertainty [16];[31];[26];[32];[14].

4. Result: development of a web-based interactive software

The developed software ‘pinu’ (http://accad.osu.edu/~hban/exurban) supports comparing and updating the semantics of exurban definitions using interactive graphical-representation of both concept data and empirical spatial-data. Users can recognize existing definitions and change them into a new definition of their own by using sliders in the interface. Then the spatial data is updated according to the user-input values of concept definitions. pinu is designed to allow the feedback from concept-definition data to the empirical spatial-data interactively in a VE. This immediate feedback-process could be useful for researchers or stakeholders to promote communication, discussion, and knowledge sharing about socially-sensitive characteristics such as exurban land-use issue.

pinu consists of two parts developed with different software of Flash and Virtools. The upper part is for concept definition developed with Flash and the lower part is for interactive geovisualization developed with Virtools. In the Flash program, the user can choose one original concept definition between Daniels [13] and Nelson [35]’s definitions for each of left and right side. Then the values of attributes in the original definitions are represented as numbers, texts, and rectangular shapes. The location and width of the rectangular shapes graphically represent the similarity, distance, and overlap in conceptual spaces [02] of the two definitions in an interactive way. Since we are dealing with only ‘the degree’ of uncertainty as the fuzzy characteristic of each definition, a uniform rectangular shape is used among other types of the shape.

Fig. 1 shows the interface of pinu developed to represent, redefine, and restore the uncertain concepts of exurbanization, and to update and visualize spatial data based on the redefined concepts in an interactive 3D-VE. For example of the exurban definitions of Daniels [13] and Nelson [35], any values of distance and population density can be modified by moving the sliders. If the user clicks the right-sided round button after the modification, the updated semantic definitions are converted to an XML document and XML codes appear in a text box in the interface.

As mentioned previously with the example of the ‘values’ and the ‘attributes’ of distance in Daniels’s [13] definition, not only the negotiation of the values of the attributes, but also the negotiation of the attributes in the definition could be performed interactively. To form a semantic concept multiple of values and attributes could be used [37]. There are a few existing approaches of negotiation, however the questions of re-defining existing definitions and restoring of the modified definitions are not necessarily addressed between them. Only a few works investigate into the general aspects of specifying frameworks for electronic negotiations [42] and this is similar in urban ontology domain.

We consider a weight to each attribute of the definitions to reflect importance of the definition which is made by the user. Using the weight for each attribute, pinu might link the process of managing negotiation semantics and the negotiation process with some limitations. The present version of pinu let a user only add or delete the attribute in the initial set of the original definitions of Daniels [13] and Nelson [35]. For example, there is no attribute or value about population density in Nelson's [35] original definition. However, if a user wants to "add" an attribute of population density to the Nelson's [35] definition, the user can move the weights slider (vertical one) from 0 to the preferred amount (person/mi^2). To "re-define" the distance attribute in Daniels's [13] definition, the user can move the left and right sliders (horizontal ones). Also, "deletion" of the population density attribute in Daniels's [13] could be done by moving the weights slider to 0. In these ways, the negotiation of definitions can be made based on the needs of the user. However, these addition and deletion of attributes are not sufficient for full negotiation of attributes since current version of pinu does not support addition of new attributes outside of the existing attributes. The extended version of pinu could be developed to freely add or delete attributes from user's manipulation.

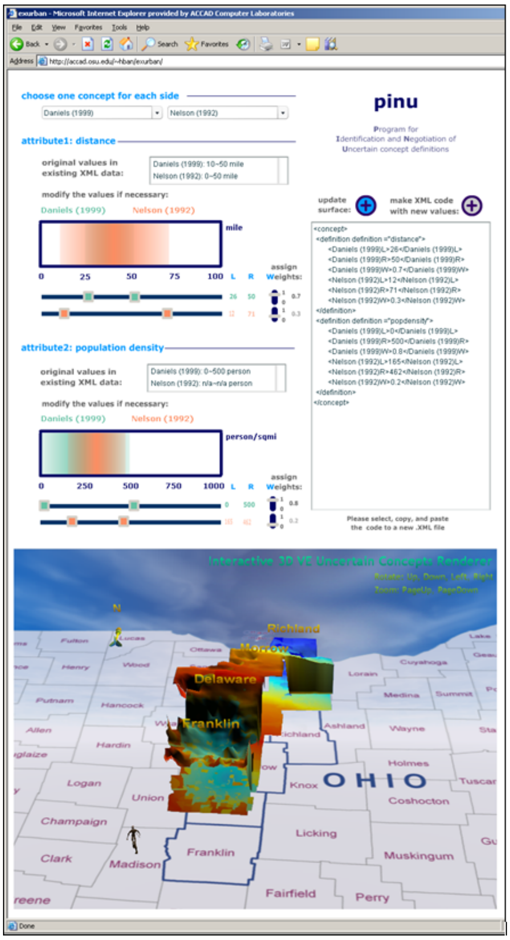


Fig. 1 User interface of the developed software pinu (Program for Identification and Negotiation of Uncertain concept definitions: prototype)

In addition to the existing studies [08];[07], we find that using a global standard XML format is able to support web-based access, representation, and restore of fuzzy concepts of exurbanization. In the interface of pinu, we convert objects of the values, attributes, and weights of exurban definition in the Flash movie into an XML document. In this way, the data becomes more convenient to interoperate with other applications since XML facilitates communication with other components or services that are either local or web-based [37].

Table 1. Concepts defined in the XML data

```
<concept>
        <definition definition ="distance">
                <Daniels (1999)L>26</Daniels (1999)L>
                <Daniels (1999)R>50</Daniels (1999)R>
                <Daniels (1999)W>0.7</Daniels (1999)W>
                <Nelson (1992)L>12</Nelson (1992)L>
                <Nelson (1992)R>71</Nelson (1992)R>
                <Nelson (1992)W>0.3</Nelson (1992)W>
        </definition>
        <definition definition ="popdensity">
                <Daniels (1999)L>0</Daniels (1999)L>
                <Daniels (1999)R>500</Daniels (1999)R>
                <Daniels (1999)W>0.8</Daniels (1999)W>
                <Nelson (1992)L>0</Nelson (1992)L>
                <Nelson (1992)R>1000</Nelson (1992)R>
                <Nelson (1992)W>0.2</Nelson (1992)W>
        </definition>
  </concept>
```

Table 1 shows the XML document consists of codes and data about the definitions of Daniels [13] and Nelson [35]. An XML document must have a root element. In the XML document, the <concept> tag describes the root element meaning which attribute is defined in it. In an XML document, the root element can have more than one child-element. In the document in Table 1, the <definition> tag is the child-element of the root element <concept>. The child elements describe attributes of “distance” and/or “population density” which are used in the definitions of Nelson [35] and Daniels [13]. In an XML document, any elements can have its sub-elements. The XML documents in Table 1, the <Daniels (1999)L>, <Daniels (1999)R>, <Daniels (1999)W>, <Nelson (1992)L>, < Nelson (1992)R>, and < Nelson (1992)W> tags describe the sub elements meaning the left (L) and right (R) values of the numeric ranges defined from the four sliders, and weight (W) values for their parent-element <definition>. For example, in the XML document, the range of population density values for Nelson’s [35] definition is set as ‘0’ and ‘1000’ since the definition does not count for the population density.

We design the visualization process of the spatial data in our application using Virtools which is based on an emerging 3D-visualization concept, VE. In pinu, numeric values defining the two definitions within Flash program are sent to JavaScript. Then JavaScript sends the values to Virtools program so that they could be used to update empirical spatial-data in VE. The Virtools program includes raster images representing exurbanization in study areas—Franklin, Delaware, Morrow, and Richland Counties in the State of Ohio, U.S.A. The raster images are generated by raster datasets consist of original fuzzy membership-functions (MFs) [08] of being exurban area: distance MF of the original Daniels's [13] definition, distance MF of the original Nelson's [35] definition, population density MF of the original Daniels [13] definition, and population density MF of the original Nelson's [35] definition. Once the new definitions are received to Virtools program, the original MFs and corresponding raster datasets are updated interactively. For example, if the weights of distance in Daniels's [13] definition are set as '0.8' from the Flash program, the values within raster image in Virtools program is updated proportionally with the weights. After each raster image is updated with the new definitions, they are synthesized into one raster image to generate a negotiated image. Then each pixel in the negotiated image has new membership values of being exurban area. Lastly, the image is blended with a color-texture and visualized as a surface in a 3D-VE. The new surface shows the MF values of being exurban area at each location as its height (Fig. 1).

pinu interactively shows initial results from the users' changes of definitions without disrupting the original data in the definitions of Daniels [13] and Nelson [35]. It provides quick ways of looking at negotiated situations of exurbanization at once, also enabling stakeholders to construct an understanding of the spatial change to the change of semantics.

In this study, pinu updates exurban surface using only weights of distance and population-density attributes in the two definitions. In a future extension version, it could also use definition values of attributes to fully update the surface. For example, if a user changes the ranges of distance in Daniels' definition using the sliders, pinu could recalculate the raster dataset using the new MF. In this way, both updated values and attributes of a concept definition could be reflected in the spatial data.

5. Concluding discussion and future extension

Boin and Hunter [10] argue that there has been a gap in the communication between the spatial data producer and the spatial data user. When they communicate with each other, they speak in their own language consisting of different vocabularies. Such a gap exists even when people use the same vocabulary but with different definitions such as 'exurbanization'. In this paper, our proposed approach might be useulf to make a bridge between different players in exurban land-use.

The adopted definitions of exurbanization in this study is limited by using the same concepts and semantic such as distance and population density, and they only differ about the absolute values. The study could be extended by including other categorical measurement semantics [05] than the absolute values. This could be done by defining or adopting different semantics for the same term, for example "distance in space" and/or "distance in time" for the concept of exurbia using the extended version of pinu mentioned in the previous section.

We have built an application prototype to support interactive visualization of negotiated exurbanization. We demonstrate the advantages of interactive web-based representation of urban sprawl in overcoming problems generated from uncertain concepts. The developed software interface supports the exchange of knowledge, and helps communication and collaboration between stakeholders.

First experiences with the prototype interface imply that the specifications suggested in this paper supply sufficient flexibility for the use in diverse application situations. It could become part of an instrument for capturing expert knowledge about land-use concepts to enable translation between agency or administration-specific terminologies as well as novel methods for landscape-change assessment with heterogeneous data sources [04]. The completion of the software interface might support an evaluation of effectiveness of the developed approach.

Ahlqvist and Ban [05] argue that a personal narrative such as impressions about scenic qualities of a landscape can be "loosely" measured. Considering the uncertainty in defining exurbanization, subjective perception of exurban landscape could be measured and also represented with empirical spatial-data in the future extension of developed software. A flexible VE interface can support immediate feedback from the user by generating personalized landscapes. Visualizing the subjective perspectives in a VE together with empirical spatial data might be an important step towards a contemporary convergence of art and GIScience.

References

01. Aerts, J.C.J.H., Clarke, K. C., and Keuper, A. D.: Testing popular visualization techniques for representing model uncertainty, Cartography and Geographic Information Science, 30(3), 249-261 (2003)
02. Ahlqvist, O.: A parameterized representation of uncertain conceptual spaces, Transactions in GIS, 8(4), 493-514 (2004)
03. Ahlqvist, O.: Using uncertain conceptual spaces to translate between land cover categories, International Journal of Geographical Information Science, 19(7), 831-857 (2005a)
04. Ahlqvist, O.: Using semantic similarity metrics to uncover category and land cover change, Lecture Notes in Computer Science, 3799, 107-119 (2005b)
05. Ahlqvist, O. and Ban, H.: Categorical measurement semantics: A new second space for geography, Geography Compass, 1(3), 536-555 (2007)
06. Baehr, P.: The ''iron cage'' and the ''shell as hard as steel'': Parsons, Weber, and the Stahlhartes Gehause metaphor in the protestant ethic and the spirit of capitalism, History and Theory, 40, 153-169 (2001)
07. Ban, H.: 3D visualization of the semantic uncertainty of exurban boundaries, In M. Raubal, H. Miller, A.U. Frank., and M. Goodchild (Eds.), Fourth International Conference on Geographic Information Science (GIScience 2006), Münster, Germany, September 2006, IfGIprints 28, Münster: Institut für Geoinformatik, 223-228 (2006)
08. Ban, H. and Ahlqvist, O.: Visualizing the uncertainty of urban ontology terms, Studies in Computational Intelligence: Ontologies for Urban Development, 61, 85-94 (2007)
09. Barreteau, O., Bousquet, F., and Attonaty, J.-M.: Role-playing games for opening the black box of multi-agent systems: method and lessons of its application to Senegal River Valley irrigated systems, Journal of Artificial Societies and Social Simulation, 4(2), http://www.soc.surrey.ac.uk/JASSS/4/2/5.html (2001)
10. Boin, A. T. and Hunter, G. J.: Do spatial data consumers really understand data quality information? In M. Caetano, and M. Painho (Eds.), 7th International symposium on spatial accuracy assessment in natural resources and environmental sciences, 215-224, http://spatial-accuracy.org/2006/PDF/Boin2006accuracy.pdf (2006)

11. Cash, D. W., Adger, N., Berkes, F., Garden, P., Lebel, L., Olsson, P., Pritchard, L., and Young, O.: Scale and cross-scale dynamics: Governance and information in a multilevel world, Ecology and Society, 11(2), 8, http://www.ecologyandsociety.org/vol11/iss2/art8/ (2006)
12. Chang, A. M. and Han, T. D.: Design of an argumentation-based negotiation support system, Proceedings of the 28th Annual Hawaii international conference on system sciences, 242-251 (1995)
13. Daniels, T., When city and country collide, Washington, D.C.: Island Press, Preface (1999)
14. Deitrick, S, and Edsall, R.: The influence of uncertainty visualization on decision making: An empirical evaluation, Progress in spatial data handling, 12, 719-738, http://www.springerlink.com/content/w622458100542684/ (2006)
15. Duch, W.: Uncertainty of data, fuzzy membership functions, and multilayer perceptions, IEEE Transactions on Neural Networks, 16(1), 10-23 (2005)
16. Evans, B. J.: Dynamic display of spatial data reliability: Does it benefit the map user? Computers and Geosciences, 23, 409-422 (1997)
17. Funtowicz, S., Martinez-Alier, J., Mundo, G., and Ravetz, J. R.: Information tools for environmental policy under conditions of complexity, European Environment Agency, Environmental Issues Series, 9, Contract N. 302/SER/9600130/3G21, http://reports.fr.eea.europa.eu/ISSUE09/en/envissue09.pdf (1999)
18. Gershon, N. D.: Visualization of an imperfect world, IEEE Computer Graphics and Applications, 18(4), 43-45 (1998)
19. Heim, M.: Virtual realism. New York, Oxford University Press (1998)
20. Holling, C. S.: Understanding the complexity of economic, ecological, and social systems, Ecosystems, 4, 390-405 (2001)
21. Hunter, G. J. and Goodchild, M. F.: Dealing with error in spatial databases: A simple case study, Photogrammetric Engineering and Remote Sensing, 61(5), 529-37 (1995)
22. Irwin, E. G., Cho, H. J., and Bockstael, N. E.: Measuring the amount and pattern of land development in non-urban areas, Prepared for the American agricultural economics association sessions at the Allied social sciences association meeting, Chicago, IL. January 5-7, 2007 (2007)
23. Jelassi, T. and Foroughi, A.: Negotiation support systems: an overview of design issues and existing software, Decision Support Systems, 5(2), 167-181 (1989)
24. King, J. L. and Star, S. L.: Conceptual foundations for the development of organizational decision support systems, Proceedings of the 23rd Annual Hawaii international conference on system sciences, 3, 143-151 (1990)
25. Kraemer, K. L. and King, J. L.: Computer-based systems for cooperative work and group decision making, ACM Computing Surveys (CSUR), 20(2), 115-146 (1998)
26. Leitner, M. and Buttnefield, B. P.: Guidelines for the display of attribute certainty, Cartography and Geographic Information Science, 27(1), 3-14 (2000)
27. Leyk, S., Boesch, R., and Weibel, R.: A conceptual framework for uncertainty investigation in map-based land cover change modeling, Transactions in GIS, 9(3), 291-322 (2005)
28. Lima, C., El-Diraby, T., and Stephens, J.: Ontology-based optimization of knowledge management in e-construction, Journal of IT in Construction, 10, 305-327 (2005)
29. Lowell, K.: Outside-in, inside-out: two methods of generating spatial certainty maps, In R.T. Pascoe (Ed.), Proceedings of GeoComputation '97 and SIRC, University of Otago, New Zealand, 26-29 August 1997, 15-25 (1997)
30. MacEachren, A. M.: How maps work, New York: Guilford Press, Chapter 6 (1995)
31. MacEachren, A. M., Brewer, C. A., and Pickle, L.: Visualizing georeferenced data: representing reliability of health statistics, Environment and Planning: A, 30, 1547-1561 (1998)
32. MacEachren, A. M., Robinson, A. C., Hopper, S., Gardner, S., Murray, R., and Gahegan, M.: Visualizing geospatial uncertainty: What we know and what we need to know, Cartography and Geographic Information Science, 32, 139-60 (2005)

33. Macromedia, Inc.: http://www.macromedia.com (2007)
34. Maher, M. L. and Gero, J. S.: Agent models of 3D virtual worlds, In G. Proctor, (Ed.), ACADIA 2002, 127-138, Pamona, CA, Thresholds (2002)
35. Nelson, A. C.: Characterizing exurbia, Journal of Planning Literature, 6(4), 350-368 (1992)
36. Power, D. J.: A brief history of decision support systems, DSSResources.COM, http://DSSResources.COM/history/dsshistory.html, version 4.0, March 10, 2007 (2007)
37. Rebstock, M. and Thun, P.: Interactive multi-attribute electronic negotiations in the supply chain: Design issues and an application prototype, Proceedings of the 36th Hawaii international conference on system sciences, 79-88 (2003)
38. Schuurman, N.: Formalization matters: Critical GIS and ontology research, Annals of the Association of American Geographers, 96(4), 726-739 (2006)
39. Teulier-Bourginer, R.. Les représentations: Médiation de l'action stratégique, La stratégie chemin-faisant: Economica, 95-135 (1997)
40. Voudouris, V., Fisher, P. F., and Wood, J.: Capturing and representing conceptualization uncertainty interactively using object-fields, In W. Kainz, A. Riedl and G. A. Elmes (Eds.), Progress in spatial data handling, 12th International symposium on spatial data handling, 12, 755-770 (2006)
41. Weber, J.: Gestion des ressources renouvelables: Fondements théoriques, Montpellier (1995)
42. Wedekind, H.: On specifying contract negotiations, In H.-R. Hansen et al. (Eds.), Proceedings of the 8th European conference on information systems, 1, 23-30, Vienna, Vienna University of Economics and Business Administration (2000)
43. Yang, J., Li, Q., Wenyin, L., and Zhuang, Y.: Search for Flash™ movies on the web, Proceedings of the 3rd International conference on web information systems engineering, workshop on Mining for Enhanced Web Search, 257-266, Singapore, Dec, 2002, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1177870 (2002)

Generating an urban domain ontology through the merging of cross-domain lexical ontologies

J. Lacasta, J. Nogueras-Iso, F.J. Zarazaga-Soria, P.R. Muro-Medrano
Computer Science and Systems Engineering Department, University of Zaragoza,
María de Luna 1, E-50018 Zaragoza, Spain
{jlacasta, jnog, javy, prmuro}@unizar.es

Abstract: In order to classify resources, digital libraries have traditionally used different types of lexical ontologies, which describe the terminology used in an area of knowledge. This paper analyzes how lexical ontologies covering different areas of knowledge can be merged to generate an enriched urban terminology. This work proposes a method to combine these different perspectives into a single network of urban related concepts. The objective of this network is to facilitate a draft for a more formal (non lexical) urban domain ontology.

Key words: Urbanism, Lexical Ontologies, Ontology Mapping

1. Introduction

Urbanism is usually defined as the study of cities including their economic, political, social and cultural environment. As it can be observed from this definition, this discipline could be considered as an intersection of different domain areas such as economics, politics culture or civil engineering. One way to represent the knowledge behind urbanism is by means of the use of ontologies. The term ontology is used in information systems and in knowledge representation systems to denote a knowledge model, which represents a particular domain of interest. According to (Gruber, 1993) an ontology is "an explicit formal specification of a shared conceptualization". Therefore, given the multidisciplinary character of urbanism, the development of an urban domain ontology requires a revision of all the aforementioned cross-domain areas, capturing the concepts directly involved with the built environment of urbanism.

The purpose of this paper is to reproduce this exercise of revising and merging the knowledge from different domains in order to obtain a better definition of the urban domain. In particular, this work proposes a method for the definition of an urban domain ontology through the merging of thesauri representing the knowledge behind different domains. A thesaurus is a lexical ontology that defines a set of terms describing the vocabulary of a controlled indexing language, formally organized so that the a priori relationships between concepts (e.g., synonymous terms, broader terms, or narrower terms) are made explicit (ISO, 1986). The applicability of thesauri for search and retrieval in digital libraries has promoted the creation and diffusion of well-established thesauri in many different domains. Thus, a thesaurus can facilitate an important source of information when trying to analyze a specific knowledge area.

The proposed merging method takes as input a set of different multilingual lexical ontologies and obtains as a result a more consistent and formalized ontology. The main step of the merging process is the detection of intersections between concepts in the different lexical ontologies. This is performed using lexical similarity techniques that take advantage of the multilingual support given by the input lexical ontologies. Additionally, it is worth noting that the general concepts of the input lexical ontologies are pruned thanks to the use of a thesaurus specialized in urbanism. The output ontology can serve two important objectives. On the one hand, it can be used as a first draft for a more formalized urban domain ontology. On the other hand, the output ontology can be used to analyze to which extent urbanism is represented in the input thesauri.

The rest of the paper is organized as follows. Section 2 summarizes the state of the art in the comparison of lexical ontologies, the basis for the alignment and merging of ontologies. Section 3 describes the merging method for the generation of an urban domain ontology. Then, section 4 tests the feasibility of the method using EUROVOC (EUPO, 2005), GEMET (EEA, 2004), AGROVOC (Lauser et al., 2006) and UNESCO (UNESCO, 1995) as input lexical ontologies. All these resources have been established by well-known organizations and provide a shared conceptualization in the areas of economics, politics, culture and environment. The last section concludes and introduces some ideas on future work.

2. State of the art in comparison of lexical ontologies

In order to extract the urban related concepts and their relations from the analyzed thesauri, it is needed to be able to determine that two concepts of different thesauri are equivalent. This problem of finding relationships (e.g., equivalence or subsumption) between entities of different models is known as ontology alignment. In this area, many different alignment techniques that automatically identify similarities between concepts have been developed (Kalfoglou and Schorlemmer, 2003), but most of the used similarity measures are not adequate for lexical ontologies. The main mapping procedures are based in the following types of analysis:

- Similarity analysis of classes. This includes from simple string comparisons between the class names (Noy and Musen, 1999) to more sophisticated analysis that take into account lexical variants, demorphing or synonymy (Fernández-Breis and Martínez-Béjar, 2002).
- Similarity analysis of properties. This includes comparison in the number, names and types of the properties of the classes (Compatangelo and Meisel, 2002).
- Similarity analysis of relations. Analyze the similarity of names, types and structure of the relations between the classes (Kalfoglou and Schorlemmer, 2002).
- Similarity analysis between instances. Analyze the similarity of the property values of the class instances (Stumme and Maedche, 2001; Doan et al., 2002).

These different types of analysis are based in only two different procedures: the analysis of the linguistic similarity between labels and the analysis of the relation between classes. The differences between the available mapping procedures are the techniques used to identify the similarities and the types of analysis considered.

Lexical ontologies have some particularities in their structure with respect to other types of ontologies. They consist of a set of lexical concepts that share a reduced set of property types and relation types. For example, the thesauri structure described in ISO-2788 (ISO, 1986) and ISO-5964 (ISO, 1985) standards is reduced to alternative labels for a lexical term in different languages and a reduced set of possible relation types (*narrower-broader* and a general *related* type that provides little semantic).

As commented in (W3C, 2005), representing each concept of a lexical ontology as a different class produce several problems for its use in resource classification. Therefore, the most usual way to model lexical ontologies is represent each concept as an instance of a general "Concept" class, which define the available types of properties and relations. In this model, instances of a class can be directly used as values of properties in the description of a resource, and therefore, it has been used to create many different lexical ontology formats (Miles et al., 2005; Lauser et al., 2006; Miller, 1990). Its main drawback is that the provided ontological structure is very poor (only one class). The generalized use of this model to represent thesauri makes unnecessary the use of mappings techniques to relate their structure (they are equivalent). Therefore, all the mapping work has been focused in the analysis of similarity between instances.

An additional problem to compare lexical ontologies is the format used to represent them. For decades, the evolution of digital libraries has encouraged the use of lexical ontologies describing the terminology of an area of knowledge in the form of taxonomies, classification schemes or thesauri, promoting in that way the creation and diffusion of well-established collection in different domains. However, the lack of standardization has produced a huge variety of incompatible formats that increase the complexity of the comparison process.

3. Method for the generation of an urban domain ontology

Urbanism can be considered as an intersection of different domain areas such as economics, politics culture or civil engineering. In this context, the process to develop an urban domain ontology, providing explicit and formal specification of the knowledge behind the urbanism discipline, makes necessary to revise all these cross-domain areas and capture all the relevant concepts.

This section describes the process to capture the structure of relations between urban concepts through the analysis and comparison of cross-domain lexical ontologies with a thesaurus structure. The result obtained is a network of related urban concepts that shows the relevance of concept relations. Figure 1 remarks the different steps of the process, showing the inputs and the produced results. Four different tasks can be highlighted and are described in detail in the following subsections: the harmonization of the interchange format used for thesauri, the mapping of concepts, the generation of the network of urban concepts and the visualization of the generated network.

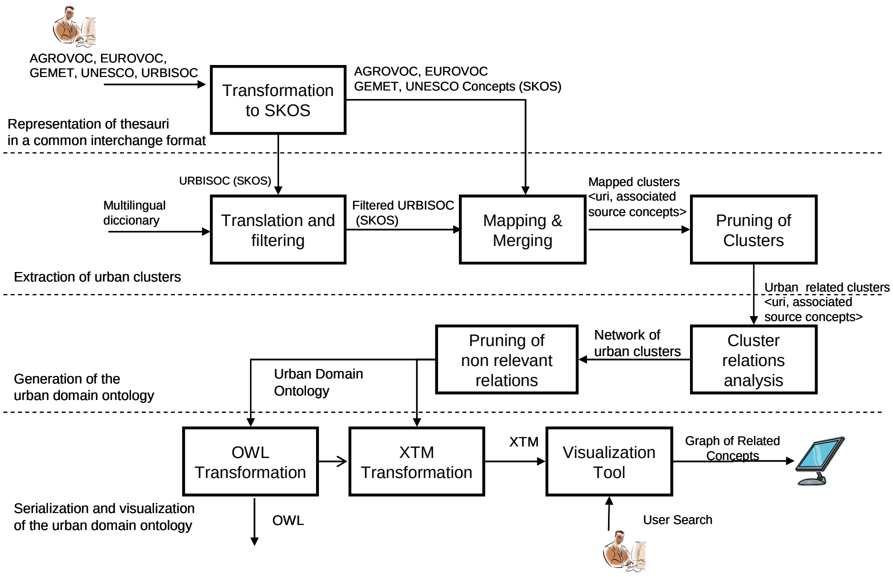


Figure 1: Work flow for the generation of an urban domain ontology

3.1. Representation of thesauri in a common representation format

The lack of a standardized representation format for thesauri has produced the spread of a great variety of incompatible formats. The system described here takes as input a set of thesauri of different knowledge areas. Therefore, to avoid format related issues in the process, thesauri must be provided in a single common format.

The British standards BS-5723 (BS, 1987), BS-6723 (BS, 1985) and their international equivalent (ISO-2788 and ISO-5964) propose models to manage lexical ontologies but lack a suitable representation format. Currently, the British Standards Institute IDT/2/2 Working Group is developing the BS-8723 standard that will be promoted to ISO when finished and whose 5th part will describe the exchange formats and protocols for interoperability of lexical ontologies. To establish the interchange format for lexical ontologies, the IDT/2/2 Working Group is involved in the SKOS project (Miles et al., 2005), a W3C initiative for the representation of simple knowledge organization systems such as thesauri, classification schemes and other types of controlled vocabularies (see figure 2). The involvement of the IDT/2/2 Working Group in the SKOS project will probably produce that the approved representation format will be SKOS related. Currently, several important thesauri of different areas are being transformed to this format. Having into account all those reasons, it has been the format selected as input of our analysis system.

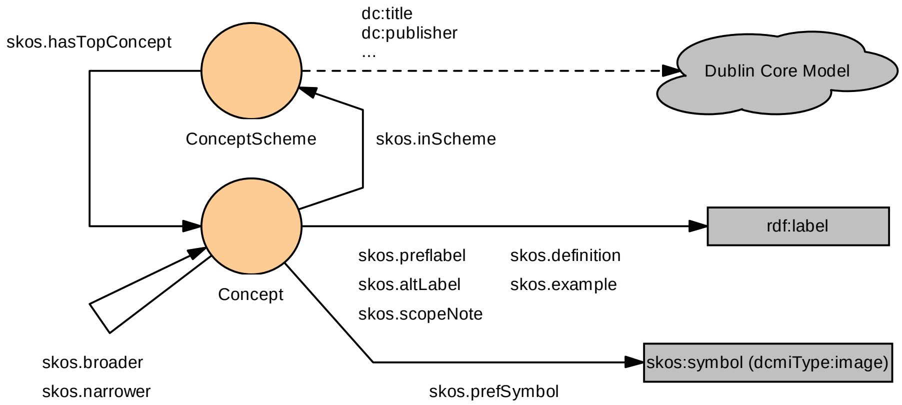


Figure 2: SKOS-Core Model

In order to facilitate the transformation of the different analyzed thesauri into SKOS, we have developed a customizable tool (developed in Java) that converts several file formats, according to a traditional thesaurus model, into the SKOS representation format. Figure 3 shows the mapping established by the tool between a classical thesaurus model and the SKOS representation model. The Unified Modelling Language (UML) notation is used for the representation of the two models. As it can be observed from the figure, the following transformations are applied:

- A concept scheme is created to represent the source thesaurus.
- Each thesaurus preferred term generates a new concept in the SKOS representation (except if it is not used for classification).
- Each translation derives a new preferred label in the language of the label.
- Each term related by a UF/USE relation (synonymy relation) is converted into an alternative label of the related concept.
- The RT relations between terms are converted to *skos:related* relations between the corresponding concepts. The same happens with the BT and NT relations that are converted to *skos:broader* and *skos:narrower*.
- The description of a term is converted into the definition of its associated concept.
- The concepts whose associated term is marked as TT are included in the concept scheme as top terms.
- Another important item is the URI that must be created for each SKOS concept. It has to be generated by using the information provided by the source format. In the example, the term value used as the preferred label of the concept can be converted into an URI by adding an *http://* prefix.

It is worth noting that this tool can be customized to different source formats (text files, relational databases). It provides a common infrastructure for the parsing of source formats and the writing of SKOS output files (in RDF format). Therefore, the support of a new source format is reduced to the development of a new plug-in component. These plug-ins are simple Java classes which conform to a specific interface and extend the functionality of an abstract class to deal with the specific aspects of each new format.

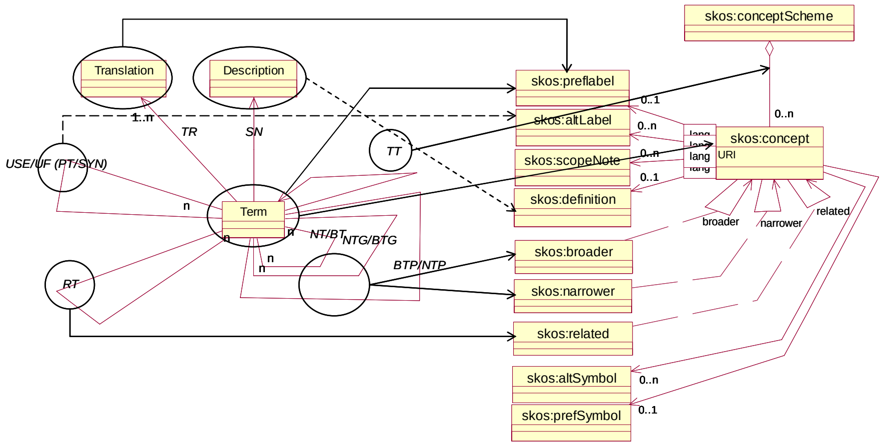


Figure 3: Mapping between a traditional thesaurus model and the SKOS-Core model

Once the thesaurus has been transformed, different statistics about the number of concepts and relations of the source and destination format are computed to verify if the obtained SKOS is correct. SKOS presents many restrictions that have to be validated to assure that the generated file is correct. Among them, the following can be highlighted:

- At least one *ConceptScheme* must exist, and all the contained concepts must refer to a *ConceptScheme*.
- Every concept must have one broader concept except when it is a top term. Then, it must be referenced as *top concept* in the *ConceptScheme* structure and not have a broader term.
- Each concept must have one and only one preferred label for each available language that is unique along the thesaurus.
- All the relations between concepts must reference existent concepts (orphan relations are not allowed).
- The structure of *broader*/*narrower* relationships must not contain cycles.
- The *related* relation is symmetric, so if “A” is *related* with “B”, “B” must be *related* with “A”.
- The *broader* relation is the inverse of the *narrower* one, so if “A” is the *broader* of “B”, “B” must be *narrower* of “A”.

3.2. Extraction of urban clusters

The objective in this step is to extract the concepts related to urbanism from the analyzed thesauri. In order to search urban concepts in the cross-domain thesauri, a set of terms of a thesaurus specialized in urbanism is used as seed for this search. In addition, the relations between the concepts present in the urban thesaurus are used in the next step as a base for the construction of urban domain ontology.

As commented previously, from the available mapping techniques, only the analysis of the values of the properties of the instances is useful. Here, the linguistic similarity between the preferred and alternative labels has been considered for the mapping. The analysis of other properties as definition, scope notes and relations is left for future work.

In the mapping process, every concept of every thesaurus (including the urban one) is compared with every concept of the other thesauri to find equivalences. Two concepts are considered equivalent when at least one of the labels of a concept (preferred and alternatives) is equal to a label in the other concept. Here, the use of multilingual thesauri has the advantage of having labels in different languages to compare (the labels used to describe two concepts may differ in a language but be equivalents in other one). In order to improve the results, plurals, accents and capital letters have been removed. This approach could be enriched with misspellings detection, stemming and word order analysis among others, but given the strict rules used to define the labels used in a thesaurus no much improvement would be expected.

Equation 1 measures the relevance of the mapping obtained between two concepts. The higher the number of labels two concepts share from the total they have, the higher is their equivalence. This equation can be applied to each obtained mapping, but it is used here to analyze the quality of the thesaurus mapping with respect to the urban one showing the relevance that each different knowledge area gives to urbanism (see the experiments in section 4).

$$probabilityOfEquivalence = \frac{2*numberOfMatchedLabels}{totalNumberOfLabelsInTheTwoConcepts} \quad (1)$$

Each set of mapped concepts is grouped into a cluster (group of equivalent concepts), which is identified with one of the URIs of the original concepts. Figure 4 shows a simplified example of the cluster generated for the “Zonas Urbanas” concept (“Urban areas”) mapping different thesauri and considering only the Spanish labels. In the example, it can be seen that the “Area Urbana” concept of GEMET is included in the cluster thanks to the presence of this label in the concepts of EUROVOC and AGROVOC. In addition, the relevance of the mappings is included to show that some of them are stronger than others.

Not all the clusters obtained in the mapping process are useful; many contain concepts not related to urban terminology. Therefore, only the clusters that contain a concept from the urban thesaurus and those with at least a concept directly related (*broader*, *narrower* and *related* relations) to another one in a cluster of the first case are stored. The rest are considered as not relevant to urbanism and they are pruned from the system. To maintain the consistency, the relations of the remaining concepts with the deleted ones are also eliminated.

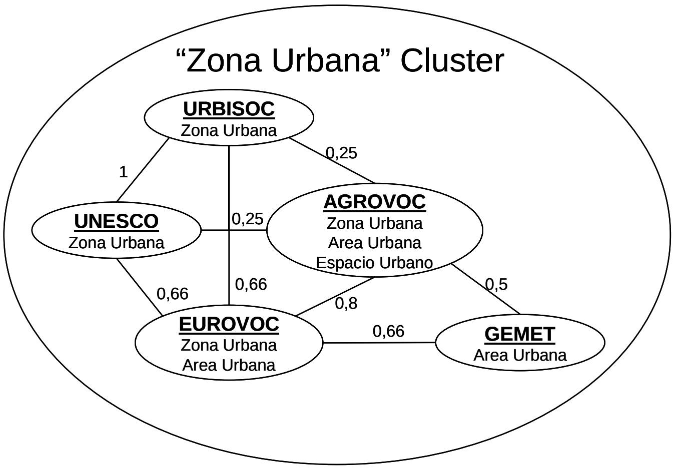


Figure 4: Example of a cluster derived from the mappings

The reason to include clusters that do not directly contain urban terminology is to provide an extension of urban concepts from the point of view of other areas of knowledge.

3.3. Generation of the urban domain ontology

The clusters generated in the previous step describe the urban terminology used in different knowledge areas, but not how this knowledge is inter-related. This subsection describes the process used to relate these clusters generating a network of urban concepts that can be seen as an urban domain ontology.

The relations of the concepts contained in each cluster are used as a basis for the generation of the relations between clusters. Here, in addition to the basic relations (*broader*, *narrower* and *related*), *sibling* relations (*narrower* of its *broader*) have been also considered. Other relations as the *grandparent* (*broader* of its *broader*) or *grandchildren* (*narrower* of their *narrower*) could be considered as well, but they are less relevant and not taken into account in this paper.

Each relation of a concept in a cluster with other concept from another cluster is marked as a relation between the clusters with the type of relation between the concepts. When two or more concepts of one cluster are related to two or more concepts of another cluster, the relation between these two clusters is marked with the different types of the original relations between the concepts, along with the number of occurrences of each different relation type. The marks indicate us the relevance of the relations between the clusters. That is to say, the more concepts inside the two clusters are related, the more relevant is the relation between these clusters. The result of this inter-cluster relation process is a network of interconnected urban clusters.

Given that all the concepts from the urban thesaurus have been included, the obtained network is based on the structure of relations of this thesaurus. But, in addition to this, it includes the relations and relevance derived from the merging of the source thesauri. See figure 5 in the experiment section as an example of the obtained network.

In many situations, it is not interesting to have a network with all the existent relations but only the most important ones. Therefore, a process to prune the less relevant relations has been created. This process receives as input the complete network of concepts and a weight threshold to determine if a relation is maintained. All the relations with a weight below the threshold are pruned. After the pruning, all the clusters that do not have at least one relation with another one are also eliminated.

3.4. Serialization and visualization of the urban domain ontology

For the serialization of the generated structure, we have proposed the use of the Web Ontology Language (OWL) (Bechhofer et al., 2004) and XTM format (Pepper and Moore, 2001). On the one hand, OWL is a widely accepted language for the definition of formal ontologies based on RDF. On the other hand, XTM is a format for the exchange of topic maps with an emphasis on the find-ability of information. We have selected XTM because of its advantages for the visualization and navigation through the generated network of concepts. It can be easily visualized by a wide range of tools compliant with this format. For instance, we have selected the TMNAV tool created in the TM4J project (TM4J, 2001) but other tools could also have been used.

4. Testing the method in the urban domain

The process described previously has been used to generate a network of urban concepts using GEMET, AGROVOC, EUROVOC and UNESCO as thematic thesauri. These thesauri provide a shared conceptualization in the areas of economics, politics, culture and environment: EUROVOC is a multilingual thesaurus covering the fields in which the European Communities are active (it provides a means of indexing the documents in the documentation systems of the European institutions and of their users); GEMET is a thesaurus for the classification of environmental resources developed by the European Environment Agency and the European Topic Centre on Catalogue of Data Sources; AGROVOC is a specialized thesaurus for the classification of geographic information resources (with special focus on agriculture resources), which has been created by the Food and Agriculture Organization of the United Nations; and UNESCO is a general purpose thesaurus for use in the indexing and retrieval of information in the UNESCO Integrated Documentation Network. The different origins and objectives of these thesauri provide different views of the urban terminology they contain.

From the available thesauri about urbanism, URBISOC (Alvaro-Bermejo, 1988) was selected as a basis for the filtering of urban terminology. URBISOC has been developed by the Spanish National Research Council to facilitate classification at bibliographic databases specialized in scientific and technical journals on Geography, Town Planning, Urbanism and Architecture. This thesaurus contains around 3,600 different concepts labelled in Spanish.

These five thesauri have been published in completely different representation formats. UNESCO and AGROVOC are stored in a database format, but each one with a different database management system and a different table structure; EUROVOC is provided in an XML based format with a specific structure of files and XML tags; GEMET can be obtained in an SKOS based format but with some extensions that make it incompatible with SKOS schema; and URBISOC is provided in the web directly in HTML format. Therefore, the transformation to SKOS has been needed for all of them.

As commented in the process description, the mapping system works better when different labels in different languages are available for each concept. GEMET, EUROVOC, AGROVOC and UNESCO are in Spanish, English, and French between others. However, URBISOC is only provided in Spanish. This produces a limitation in the possibilities of filtering urban concepts (only Spanish labels can be compared). Therefore, before using it in the mapping system, the Spanish labels have been translated into English and French using a multilingual dictionary.

Although the use of a dictionary to translate the labels introduces errors caused by synonymy problems it is believed that the specific character of the urban thesaurus limits the problems of polysemy. The use of a multilingual dictionary has only been needed by the lack of a multilingual thesaurus suitable for filtering urban terminology. If a suitable thesaurus were used, the translation step could be skipped.

Analyzing the clusters generated by the mapping system we have found that some clusters with URBISOC concepts did not have any specific connection with urbanism. It was caused by the inclusion in URBISOC of very general terminology not specifically related to urbanism. To solve this problem, we decided to get a subset of URBISOC without general concepts by selecting as a start point the “Urban planning” concept (very frequent in urbanism) and adding recursively all the concepts related by *narrower* and *related* relations. The concepts of URBISOC not found through these relations were discarded. This process reduced the size of URBISOC from 3,609 to 3,091 concepts, eliminating most of the general terminology.

Name	Concepts	Mapped Concepts	Mapped Percentage	Liability
GEMET	5244	834	15.9%	0.611
EUROVOC	6649	935	14.06%	0.394
AGROVOC	16896	772	4.57%	0.454
UNESCO	4424	802	18.13%	0.578

Table 1: Relevance of urbanism in cross-domain thesauri

Using this reduced version of URBISOC, the mapping process produced the results shown in table 1. This table shows the number of concepts of each thesaurus, the number of concepts mapped to URBISOC, the percentage of each thesaurus that has been mapped, and the average relevance of the mappings according to equation 1 (section 3.2). It can be seen that around 1 out of 3 urban concepts (3,091) have been mapped to each thesaurus. The percentage of urban terminology in each thesaurus may not seem very high (18% as maximum), but having into account that they are thesauri specialized in other areas of knowledge, these percentages are quite relevant.

Figure 5 displays a screenshot of the TMNAV tool visualizing part of the generated network. It shows the related concepts around the concept “urban population”. The network contains all the relations that have been found, but as commented in the description of the process and shown later, the less relevant ones could have been deleted to obtain a smaller structure.

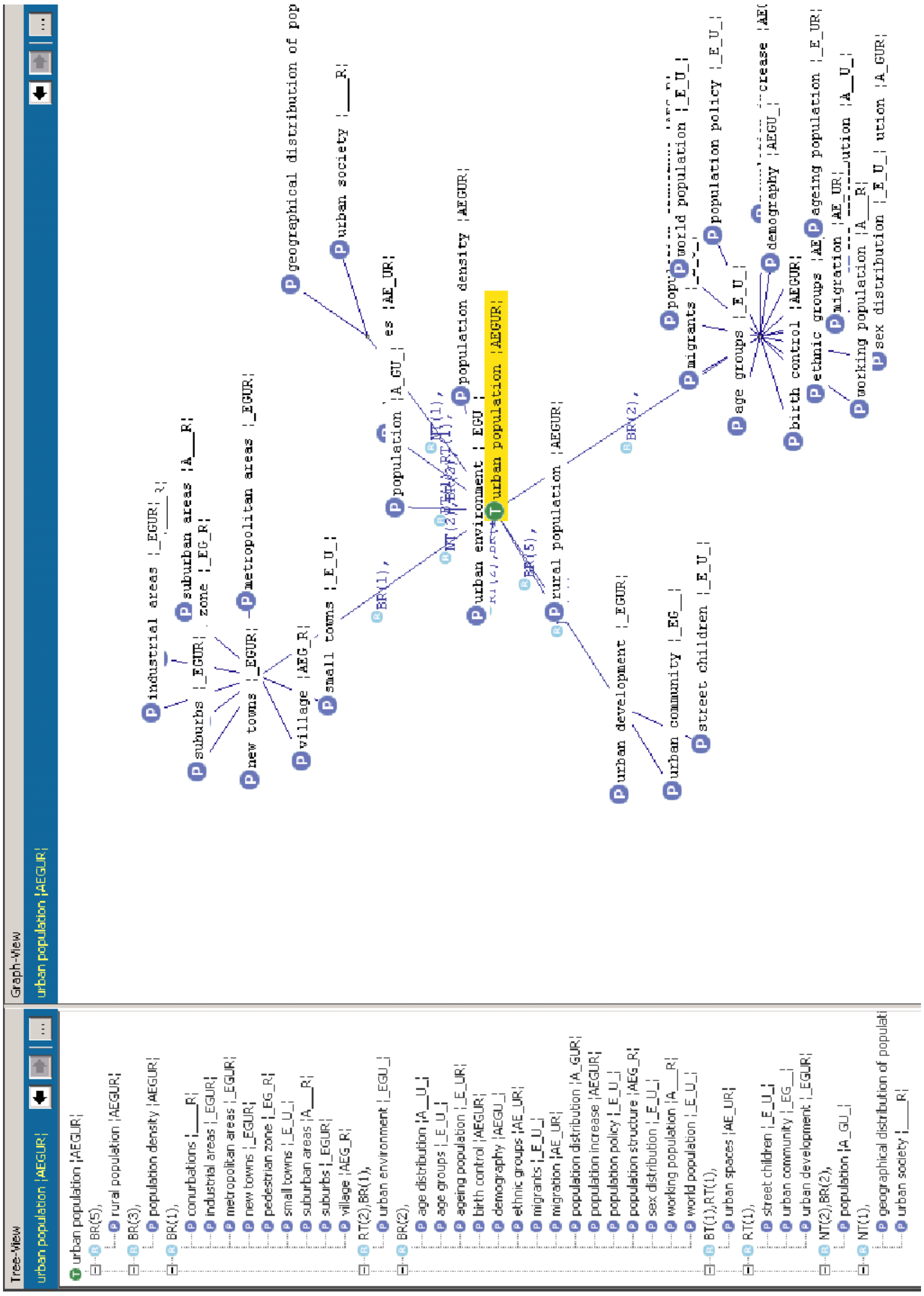


Figure 5: Visualization of a part of the generated urban domain ontology

In the figure, each cluster includes the English label of one of the concepts contained inside (the selection order for the label is UNESCO, GEMET, EUROVOC, AGROVOC and URBISOC), and the initials of the thesauri where the concepts contained in the cluster are from (i.e. AEGUR means that the concepts exist in AGROVOC, EUROVOC, GEMET, UNESCO and URBISOC; _E__R means that the concept does exist only in EUROVOC and URBISOC). The relations are labelled indicating the types and number of the relations found. Here, BT indicates *broader* relation, NT is *narrower*, RT means *related* and BR is used for *siblings*. The total number of relations could also be added but it has been omitted for the sake of clarity. In the example, the "Urban Population" cluster contains concepts of all the thesauri and maintains 2 relations with the "Urban spaces" concept, one of type *related* and another one of type *sibling* (weight 2). This relation is less relevant than the one connecting to the "Rural population" concept where 5 *sibling* relations have been found (weight 5).

Minimum Weight	Clusters	Cluster Size	Relations	Cluster Relations
1	6200	2.28	113888	18.36
2	4341	2.66	42823	9.86
3	2878	3.03	17570	6.10
4	2086	3.24	11333	5.43
5	1353	3.54	5622	4.15

Table 2: Size of the network of urban concepts

The output network contains 6,200 concepts with 5,622 relations of weight 5 or greater, 5,711 of weight 4, 6,237 of weight 3, 25,253 of weight 2 and 71,065 of weight 1. Table 2 shows a summary of the size of the generated network when the less relevant relations are pruned. Each row shows the size of the network that includes all the relations of at least "Minimum Weight" weight. Each row shows the number of clusters, the average size of each cluster, the total number of relations between clusters, and the average number of relations of each cluster.

5. Conclusions

This paper has shown a process to generate a network of concepts related to urbanism. The objective is to use this network as a basis for a first draft of an urban domain ontology. The process uses as input a set of multilingual thesauri from different knowledge areas (e.g., GEMET, AGROVOC, UNESCO and EUROVOC) and a thesaurus specialized in urbanism (URBISOC) to be able to select the urban terminology present in the other thesauri.

The main steps of the generation process are the harmonization of the input formats, the mapping between the concepts to generate clusters of equivalent concepts using linguistic similarity measures, and the establishment of relations between the clusters on the basis of the original relations between the concepts contained in different clusters. Finally, in order to facilitate the visualization and reusability of the generated output, it is transformed into XTM and OWL formats.

In the experiment, we found that URBISOC, the specific thesaurus for the filtering of concepts related to urbanism, contains very generic concepts in the top part of the broader-narrower hierarchy. Therefore, these general terms had to be removed before using it as a filtering mechanism.

As regards the mappings established between the source thesauri to obtain the clusters of concepts, we could observe that urban terminology is a relevant part of the analyzed thesauri (up to 18% in UNESCO). Future work will improve the used mapping techniques by having into account the structure provided by the relations between the concepts.

In addition, we have shown, through the experiment, how to reduce the size of the generated network of concepts by pruning the less relevant relations. This pruning is able to reduce the size of the network from 6,200 concepts to only the 1,353 more related. A future improvement as concerns the relations between clusters is to take into consideration the *grandparent* and *grandchildren* relations between thesaurus concepts. The objective is increasing the relevance values of some of the existent relations. For example, two concepts in a thesaurus can be directly related through a narrower relation, however, in other one they may be related through an intermediate concept.

The urban domain ontology obtained as a result of the method proposed has several advantages in comparison with the thesauri used as source and the thesaurus used for filtering in the following areas:

Consensus and focus: The concepts of the resulting network have been selected by consensus thanks to the mappings among the different sources, removing those concepts that are neither common nor focused on urbanism.

Relations: With respect to the relation structure, the total number of available relations is bigger than the existent ones in each of the original sources. Besides each relation has a weight that indicates its relevance. As future work, the semantics of these relations should be enriched. The information provided by definitions, examples, and naming patterns in the properties of the original concepts should help to refine the current relations (e.g., broader relations could be refined as *part of*, *instance of* or *generalization* relations).

Multilingual support: Thanks to the combination of different sources of knowledge with multilingual support, the output network is enriched with alternative terminology in different languages.

Formalism: Since the output network has been generated using a formal language such as OWL, we have increased its usability, facilitating the work with reasoning engines.

Finally, it must be noted that, apart from serving as a first draft for an urban domain ontology, the generated network of urban concepts can be directly applied in information retrieval systems for resource classification, thematic indexing or query expansion.

Acknowledgements

This work has been partially supported by the University of Zaragoza through the project UZ 2006-TEC-04.

References

Alvaro-Bermejo, C., 1988. Elaboración del Tesauro de Urbanismo URBISOC. Una Cooperación Multilateral. In: Encuentro Hispano-Luso de Información Científica y Técnica. II. Salamanca.

Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D. L., Patel-Schneider, P. F., Stein, L. A., February 2004. OWL Web Ontology Language Reference. W3C, W3C Recommendation 10 February 2004. http://www.w3.org/TR/2004/REC-owl-ref-20040210/

BS, 1985. Guide to establishment and development of multilingual thesauri. BS 6723, British Standards Institute (BSI).

BS, 1987. Guide to establishment and development of monolingual thesauri. BS 5723, British Standards Institute (BSI).

Compatangelo, E., Meisel, H., 2002. Intelligent support to knowledge sharing through the articulation of class schemas. In: Proceedings of the 6th International Conference on Knowledge-Based Intelligent Information & Engineering Systems. Crema, Itally.

Doan, A., Madhavan, J., Domingos, P., Halevy, A., 2002. Learning to map between ontologies on the semantic web. In: The Eleventh International WWW Conference. Hawaii, US.

EEA, 2004. GEneral Multilingual Environmental Thesaurus (GEMET). Version 2004. European Environmental Agency (EEA). http://www.eionet.eu.int/GEMET

EUPO, 2005. European Vocabulary (EUROVOC). European Union Publication Office. http://europa.eu/eurovoc/

Fernández-Breis, J. T., Martínez-Béjar, R., 2002. A cooperative framework for integrating ontologies. Int. J. Hum.-Comput. Stud 56 (6), 665–720.

Gruber, T., 1993. A translation approach to portable ontology specifications. ACM Knowledge Acquisition, Special issue: Current issues in knowledge modeling 5, Issue 2 (KSL 92-71), 199–220.

ISO, 1985. Guidelines for the establishment and development of multilingual thesauri. ISO 5964, International Organization for Standardization (ISO).

ISO, 1986. Guidelines for the establishment and development of monolingual thesauri. ISO 2788, International Organization for Standardization (ISO).

Kalfoglou, Y., Schorlemmer, M., 2002. Information flow based ontology mapping. In: 1st International Conference on Ontologies, Databases and Application of Semantics (ODBASE'02). Irvine, CA, USA.

Kalfoglou, Y., Schorlemmer, M., 2003. Ontology mapping: The state of the art. The Knowledge Engineering Review 18 (1), 1–31.

Lauser, B., Sini, M., Salokhe, G., Keizer, J., Katz, S., 2006. Agrovoc Web Services: Improved, real-time access to an agricultural thesaurus. Quarterly Bulletin of the International Association of Agricultural Information Specialists (IAALD) 1019-9926 (2), 79–81.

Miles, A., Matthews, B., Wilson, M., September 2005. SKOS Core: Simple Knowledge organization for the WEB. In: Proceedings of the International Conference on Dublin Core and Metadata Applications. Madrid, Spain, pp. 5–13.

Miller, G. A., 1990. WordNet: An on-line lexical database. Int. J. Lexicography 3, 235–312.

Noy, N. F., Musen, M. A., 1999. Smart: Automated support for ontology merging and alignment. In: Twelth Workshop on Knowledge Acquisition, Modeling, and Management. Banff, Canada.

Pepper, S., Moore, G. (Eds.), 2001. XML Topic Maps (XTM) 1.0. TopicMaps.org. http://www.topicmaps.org/xtm/1.0/

Stumme, G., Maedche, A., September 2001. Ontology merging for federated ontologies on the semantic web. In: In Proceedings of the International Workshop for Foundations of Models for Information Integration (FMII-2001). Viterbo, Italy.

TM4J, 2001. Homepage of the TM4J proyect. http://tm4j.org/

UNESCO, 1995. UNESCO Thesaurus: A Structured List of Descriptors for Indexing and Retrieving Literature in the Fields of Education, Science, Social and Human Science, Culture, Communication and Information. UNESCO Publishing, Paris. http://www.ulcc.ac.uk/unesco/

W3C, April 2005. Representing classes as property values on the semantic web. http://www.w3.org/TR/swbp-classes-as-values/

An Ontology-Based Intelligent Information System for Urbanism and Civil Engineering Data

Stefan Trausan-Matu[1,2] and Anca Neacsu[1]

[1] "Politehnica" University of Bucharest,
Department of Computer Science and Engineering,
Splaiul Independetei nr. 313, Bucharest, Romania

[2] Research Institute for Artificial Intelligence of the Romanian Academy
Calea 13 Septembrie nr.13, Bucharest, Romania

Corresponding author: Stefan Trausan-Matu,
email: trausan@cs.pub.ro
fax: +40.318.153290
phone: +40.724.985518

Abstract. The paper presents a prototype of an intelligent information system for urban and civil engineering data centered on an ontology. The system will provide intelligent personalized access to concepts from the ontology and to a collection of relevant associated documents indexed according to the ontology's concepts. The declarative knowledge of the ontology and an associated set of production rules may be used for automatic inferences that will enable reasoning for getting intelligent answers to users' queries, under an expert system dialog. The paper uses examples from a first version of the ontology for urban and civil engineering concepts that is in the center of the system. The ontology is developed following an integrated cognitive and socio-cultural approach. It contains a taxonomy of objects structured according to Engestrom's Theory of Activity and Sowa's ontology. This structured development facilitates further knowledge acquisition.

Keywords: information systems, ontology, Semantic Web, expert systems, production rules, Jess

1 Introduction

The number of documents containing data about urbanism and civil engineering is becoming higher every day. Moreover, some of them are changing as new recommendations, regulations, and laws appear or substitute old ones. The majority of this data is now available on the web and they may be accessed both by general search engines like Google (http://www.google.com) or by specific search tools available on particular web sites. However, a major problem is that all these search engines are keyword-based and they are not able to make inferences and to cope with relationships among documents.

A solution to the above problems is to develop an intelligent information system, based on the knowledge of the urban development and civil engineering domains. The knowledge in this base should be both declarative and procedural. Declarative knowledge refers to what is known to exist in the domain, what are the concepts and how they are related. This kind of knowledge is usually constructed around a taxonomy and it may be viewed as an ontology of the domain. In addition to domain knowledge, general knowledge may be also used, e.g. as a top level ontology, containing, for example general concepts describing human activities and regulations.

Procedural knowledge contains rules, recipes about how to use knowledge. For example, a general rule could say that, if you want to do an activity related to an object in a community, you should see what laws or regulations you should respect.

A collection of relevant associated documents (e.g. texts of law, official regulations, etc.) indexed according to the ontology's concepts is also integrated with the ontology.

. The paper continues with a section describing some theoretical concepts about ontologies and production rules. The following section introduces some original ideas about intelligent information systems. Section 4 presents an intelligent information system for urbanism and civil engineering. The paper ends with some conclusions and further development ideas.

2 Ontologies and rules

An ontology is, in the context of intelligent, knowledge-based systems, a declarative knowledge base containing the concepts and the relations that exist in a given domain, it is "a *specification of a conceptualization*. That is, an ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents. This definition is consistent with the usage of ontology as set-of-concept-definitions, but more general. " (Gruber). The name is obviously inspired from philosophy, where it means a "branch of metaphysics concerned specifically with what (kinds of) things there are" (www.shef.ac.uk/~phil/other/philterms.html).

From a knowledge representation perspective, ontologies are semantic networks that state what kinds of concepts exist and what abstraction-particularization relations hold among them. If a concept is a particularization of another concept, it has all the features of the more abstract concept and, maybe, some particular ones, For example, in figure 1 (the Protégé environment - http://protege.stanford.edu - was used for the development of the ontology), the fact that the concept "BridgesAndElevatedHighways" has "Crossovers", "FootBridge", "MobileBridge", "Overpasses", "RailBridge" and "RoadBridge", implicitly enumerates the only possible cases. Moreover, all these inherit properties (e.g. regulations) that belong to "BridgesAndElevatedHighways" or its ancestors.

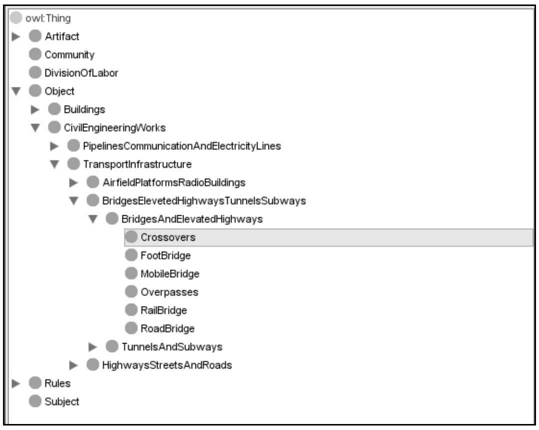


Figure 1 A fragment of the urban development and civil engineering ontology

Ontologies simplify computing in a similar way with Object-Oriented Programming (whose idea has common ancestors with ontologies). For example, an ontology may be seen as a library of concepts that may be used for many applications. Another important resemblance is encapsulation and centralization, which simplify changes: When something changes, it is enough to make a modification in a single place and all the descendants will inherit the new version.

However, ontologies do not cover all kinds of knowledge representation. In addition to declarative knowledge representation, there is a need also for procedural knowledge, saying what to do in a given context. Such type of knowledge may be represented by production rules, which are pairs *condition – action*: IF *condition* holds, THEN PERFORM *action*. Conditions usually contain patterns and variables that may be linked to facts. A production rule system has a conflict resolution strategy that selects the rule that will be applied from the rules that may be applied.

3 Intelligent information systems

An information system stores data, with the goal of providing information starting from that data, when needed. Therefore, we can identify three main features that characterize an information system: the way data is stored, the way new data is included, and the way information is provided as results to queries.

3.1 Storage alternatives

Many information systems are developed around a database. This case has the advantage of efficient implementations but, because data has a very precise and inflexible structure, any change in the conceptual schema and any intelligent dialogue are not possible.

In the context of the huge number of documents available now on the web, it is natural to consider the case of information systems using text as content instead databases. However, even if search engines like Google are able to provide a lot of useful information, a problem is that many times, such engines provide too much information, too much documents as a result to a query. This problem is due to the fact that documents are unstructured and that natural language processing is not able to handle ambiguity and context issues (Winograd and Flores, 1986).

One of the ways to handle the problems of information retrieval from texts on the web is to add explicit knowledge descriptions (metadata) of the content of documents and to integrate this metadata in conceptual frameworks (ontologies) for each domain. This approach is supported by the Semantic Web perspective, which aims at facilitating programs to process texts on the web (Berners-Lee et al., 2001). A first step towards the Semantic Web is to annotate texts with XML (http://www.w3.org/XML/). Using RDF (Resource Description Framework - http://www.w3.org/RDF/) to state facts about web resources is the second step. The next step is the development of ontologies using OWL (Ontology Web Language - http://www.w3.org/2004/OWL/).

3.2 Inclusion of new data

In databases, new data may be included only as new instances, which strictly follow the fixed conceptual model. The case of texts is totally different, practically there are no restrictions. For example, the addition a new document on the web is not restricted by any conceptual model.

Ontology based storage, following the Semantic Web ideas, allows the addition of any document, but requires its metadata annotation. However, the conceptual model is not fixed, it can be changed, by modifying the ontology.

3.3 Intelligent query processing

In order to provide the needed data to various types of users and to different kinds of questions, an intelligent information system should provide several ways of interaction. For example, in the context of the web, a natural way is to browse pages containing useful information. In addition to classical browsing of a fixed structure of web pages, user's profile may be considered for generating a personalized structure of web pages, starting from the domain's ontology (Trausan-Matu et al., 2002). The ideal information system should be, however, able to enter into a dialog with the user, using his/her own language.

Any information act is, in fact, dialogistic. Moreover, as Bakhtin emphasized, any text is a dialog (Bakhtin, 1986). Even if you write something and you put something on the web, this is a potential dialog with the readers of the text.

Different ways of querying in information systems are, in our opinion, different ways of entering in dialog:

a. database query
b. hypertext browsing

c. keyword-based search engine
d. intelligent search engine
e. expert system dialog
f. controlled natural language
g. question answering
h. natural language dialog.

From the above list, only natural language dialog and question answering are, at least for the moment, not satisfactory implemented. All the other ways of information querying are, more or less, possible to implement.

4 An intelligent information system for urbanism and civil engineering

4.1 A socio-cultural ontology for urbanism and civil engineering

The Activity Theory of Yrjö Engeström (1987), emphasizes categories (subjects, objects, and communities) and mediators (general artifacts, social rules and division of labor), see also figure 2. This theory provides a theoretical framework that has been used for developing an ontology for urban development (Trausan-Matu, 2007) that has as basic concepts the components of the above mentioned two group of entities.

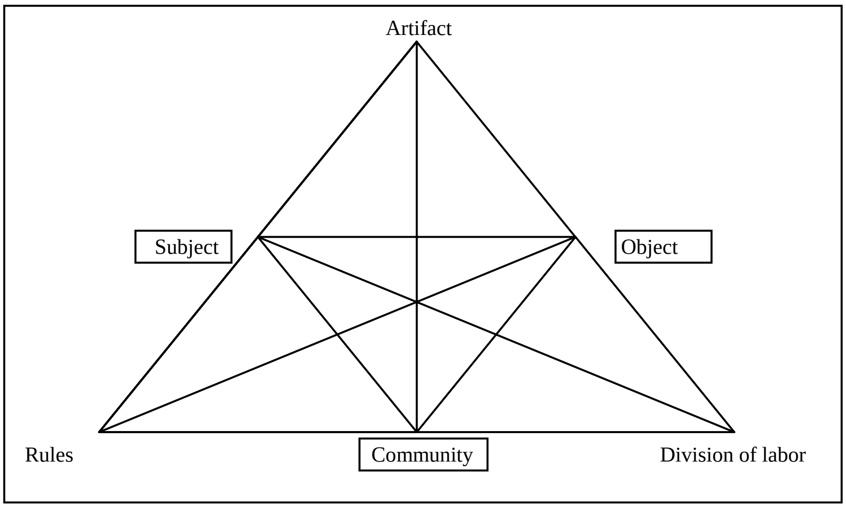


Figure 2 The main concepts of the Activity Theory of Yrjö Engeström

Each of these six entities are a basic concept (or "class") in the socio-cultural ontology (see figure 1). These concepts may have attributes, sub-concepts (that may be also sub-concepts of several other concepts, i.e. multiple inheritance of properties is allowed), and relations with other concepts (see figure 3).

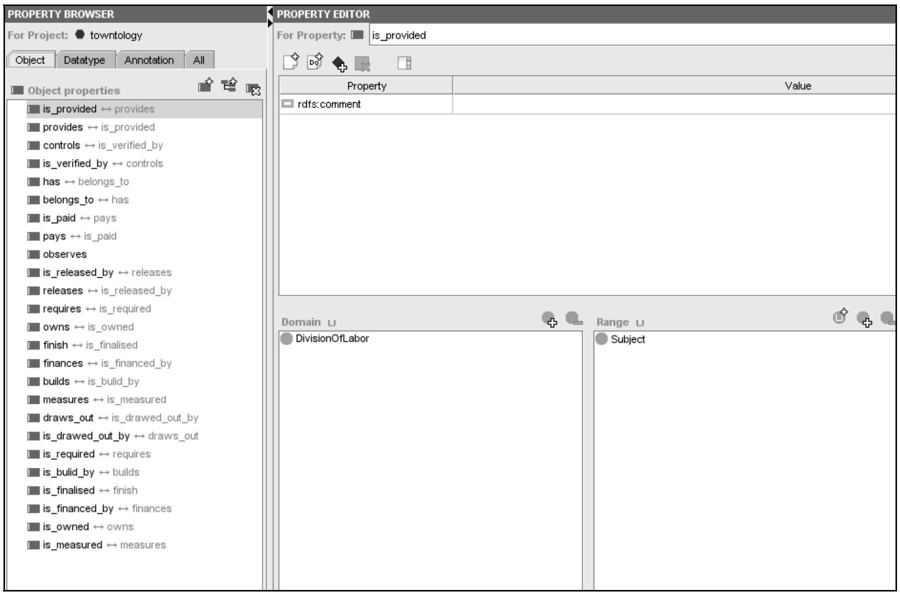


Figure 3 The relations of the urban development and civil engineering ontology

In addition to generic concepts, the ontology contains also individuals (instances). For example, the “Subject” class has 12 instances (see figure 4). One of these, the “LocalAuthority” instance has several relations (“provides”, “releases”, “controls”, etc.) with other individuals.

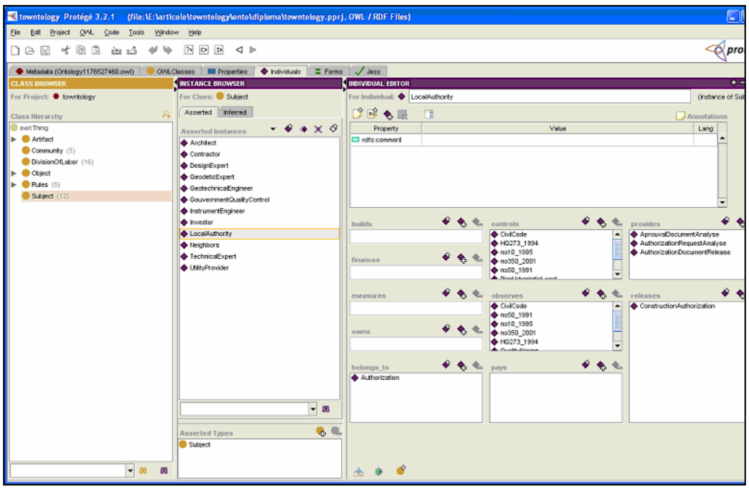


Figure 4 The “LocalAuthority” individual

4.2 Intelligent querying

In the first version of the intelligent information system was implemented an expert system that enters into a dialog with an user, for providing information about topics related to getting urbanism authorizations for new buildings. An expert system is a knowledge-based program in which there is a clear distinction between the knowledge (for example, OWL classes and rules) and the inference engine. The Jess production rule system (http://www.jessrules.com/jess) was used. A program in Jess is a collection of rules that can be matched to the existing data in the working memory. Each rule has a first, matching part, and a second, action one, which modifies the working memory or prints something. A rule may have variables that are linked to values in the working memory using pattern matching. For example, a rule that prints the information that local authorities may provide is below exemplified. In this rule, the variables $?p, $?r, and $?c are matched to all the available data, in the working memory, regarding what the local authority provides, releases and controls.

```
(defrule local_authority
  (declare (salience 1))
  (print go_to_local_authority)
  ?f <- (object (is-a Subject)
                (:NAME "LocalAuthority")
                (provides $?p)(releases $?r)(controls $?c))
  (not (answer ?))
  =>
  (printout t (slot-get ?f :NAME) " provides: " crlf)
  (foreach ?x $?p (printout t "      - "(instance-name ?x)  crlf))
  (printout t "    releases: "  crlf)
  (foreach ?x $?r (printout t "      - "(instance-name ?x) crlf))
  (printout t "    in accordance with: "  crlf)
  (foreach ?x $?c (printout t "      - "(instance-name ?x) crlf))
```

In figure 5 is illustrated a simple dialog that, among others, presents what the "LocalAuthority" can provide, release and control. An important observation is that the data is obtained from the ontology and it may be different if the ontology changes.

```
Do you want to learn what documents you need when building a house? (yes or no) yes
Have you reveive the authorization? (yes or no) no
Do you own the property? (yes or no) yes
Do you have the budget? (yes or no) yes
Did you asked for Urban Certificate? (yes or no) no
Solution: Go to the local authority
Do you want to know more on the subject? (yes/no)
yes
LocalAuthority provides:
    - AprouvalDocumentAnalyse
    - AuthorizationRequestAnalyse
    - AuthorizationDocumentRelease
  releases:
    - ConstructionAuthorization
  in accordance with:
    - no10_1995 (see the web-site: ~#en http://www.pptt.ro/content/legea_nr__101995_privind_calitatea_in_constructii_lege_nr_587_din_29_octombrie_2002.htm )
    - no350_2001 (see the web-site: ~#en http://www.cdep.ro/pls/legis/legis_pck.htp_act?ida=30200 )
    - no50_1991 (see the web-site: ~#en http://www.cdep.ro/pls/legis/legis_pck.htp_act?ida=1322 )
Try again? (yes or no) yes
Do you want to learn what documents you need when building a house? (yes or no) yes
Have you reveive the authorization? (yes or no) yes
Have you paid the taxes? (yes or no) no
Solution: Pay the taxes
Do you want to know more on the subject? (yes/no)
yes
The Investor has to pay
    - AprovalDocumentReleaseTaxe
    - ConstructionTaxe
    - AuthorizationTaxe
    - DesignTaxe
Try again? (yes or no) yes
```

Figure 5 A dialog in the expert system session

4 Conclusions and further developments

Production rules may be used for developing expert systems that can implement intelligent dialogues that provide information contained in ontologies. One important advantage of such intelligent information systems is their flexibility: new rules may be added and the ontology may change.

The system is now under testing and extension, by adding new rules and ontology components. In the next versions of the system, new facilities that use the ontology will be added. For example, the intelligent search and controlled natural language will from another intelligent information system (Trausan-Matu et al., 2006) will be integrated. The ontology will also be used for document search that could be included in the collection (Trausan-Matu et al., 2006).

However, one of the main problems in ontology development is that many times, experts do not agree about a given taxonomy. Moreover, taxonomies may change according to different perspectives of the same person. A solution to these problems is to consider folksonomies instead ontologies.

References

1. Almeida, A., Roque, L., Simpler, Better, Faster, Cheaper, Contextual: requirements analysis for a methodological approach to Interaction Systems development, http://csrc.lse.ac.uk/asp/aspecis/20000003.pdf, retrieved on 23 May 2006.
2. Bakhtin, M.M., *Speech Genres and Other Late Essays*, University of Texas Press, 1986
3. Berners-Lee, T., Hendler, J., and Lassila, O., The Semantic Web, *Scientific American,* May 2001.
4. Engeström, Y., *Learning by Expanding: An Activity theoretical approach to developmental research*, Orienta-Konsultit Oy, Helsinki, 1987.
5. Gruber, T.R., http://www-ksl.stanford.edu/kst/what-is-an-ontology.html, retrieved on 9 June 2007
6. Trausan-Matu, S., Maraschi, D., Cerri, S., Ontology-Centered Personalized Presentation of Knowledge Extracted From the Web, *Lecture Notes in Computer Science* 2363, ISSN 0302 9743, Springer, 2002, pp 259-269
7. Trausan-Matu, S., Bantea, R., Posea, V., Petrescu, D., Gartner, A., Flexible Querying of an Intelligent Information System for EU Joint Project Proposals in a Specific Topic, in *Lecture Notes in Artificial Intelligence* 4203, pp. 290-295, ISSN 0302 9743, Springer, 2006
8. Trausan-Matu, S., A socio-cultural ontology for urban development, in Teller, J.; Lee, J.; Roussey, C. (Eds.), *Ontologies for Urban Development, Studies in Computational Intelligence*, Springer, Vol. 61, pp. 121-130, 2007
9. Winograd, T., Flores, F., *Understanding Computers and Cognition,* Norwood, N.J.: Ablex, 1986.

Semantic search engine for geographic data

Claudia Pegoraro[1], Mauro Velluto[2]

[1] CSI Piemonte, Sistemi Territoriali e Ambientali,
Corso Unione Sovietica 216, 10134 Torino, ITALY
tel. +39 011.316.9093 - Fax +39 011.316.8977 e-mail: claudia.pegoraro@csi.it

[2] CSI Piemonte, Sistemi Territoriali e Ambientali,
corso Tazzoli 215/12B, 10137 Torino, ITALY
tel. +39 011.432.6721 – Fax +39 011.740001 e-mail: mauro.velluto@csi.it

Abstract
A 'useful' knowledge of a data set, referred to a generic domain, allows finding, organizing and integrating data without considering the specific context that justified their acquisition and making them available for transversal applications. The paper refers to the great amount of geographic data managed by CSI Piemonte and it describes the conceptual model proposed to improve their usefulness. It also examines the reasons that lead to the choice of grounding the model on a foundational ontology, providing both the description of the geometrical-geographical feature of territorial elements and their role in specific disciplines. Moreover it represents the base for developing a Semantic Engine for geographical data (some operational examples are schematically described).

Keywords: Geographic Data, Foundational Ontology, Semantic, Search Engine, Civil Protection, Transports.

1. Introduction

A 'useful' knowledge of a reality fraction should allow to gather, organize and integrate data through its content without taking in account a specific discipline or the context that brought to the acquisition of the data.
CSI Piemonte (Consortium for Information System), is a consortium, gathering local governments, that supplies information services to its members. A major duty is the production and the management of a great amount of geographical data that pick up a territorial knowledge which is much more detailed than effectively usable. The wish to reuse information turns out to be quite difficult due to many factors: non homogeneous technical features, data format and architecture, different production aims, last update and quality level; moreover data archives were often designed for specific applications.
To make more usable the territorial information and make it accessible by all potential interested applications the definition and the building of a rigorous and validated

ontology is due, the core of which is a computational ontology that describes the geometric primitives for GIS objects. Of course we guess it would be possible to extend incrementally this ontology and reuse the same primitives in different territorial disciplines.
The title of the article aims at declaring the final pragmatic objective that requires such a large and complex preamble, consisting in the preliminary work on defining our robust ontology, as much as possible validated and universal. Despite this the final objective is also a starting point, in the sense that it would be unsuccessful to build such a transversal and multi thematic ontology without considering the ultimate operative purpose and the specific requirements to be addressed. This comes out also because we would like firstly to detect the real benefits such a transition from our traditional conceptual modelling techniques (ER, UML) to the ontological approach will be able to conceive. [1]
Paragraph 2 will describe some aspects of our knowledge domain and the work approach to geographical data regarding the Piedmont territory. We will pinpoint our aims in developing a semantic search engine and the first step to achieve, that is working at the conceptual level. Paragraph 3 will introduce the fundamental mosaic tiles at the conceptual level, that will be further discussed in paragraph 4, taking in account both real input and theoretical framework. Paragraph 5 will provide the content for the four foundational elements as decisional core for the work, since all the subsequent results will be conditioned by these modelling assumptions. The succeeding sections will provide an application example and then the final considerations and the new aims desirable at the present work stage.

2. Working within the GIS data domain

Nowadays SITAD (Sistema Informativo Territoriale ed Ambientale Diffuso) enables users to move within the great amount of geographic information from the Piedmont Region; this is a web GIS catalogue that allows data and metadata retrieve through categories or lexical queries, mostly getting a panoramic view of the available information through metadata categories exploration or viewing directly specific data layers.
Let’s take as an example the regional road network managed within SITAD by CSI: we are able to find the related information, summarized in metadata forms accomplishing the ISO standard 19115 and including Title, Description, Issue, Keywords, Date (creation, update), Responsible, Origin-description, Distribution, Coordinate Reference System and so on. Moreover let’s suppose we have two users, **A** and **B**:

- **A** enters SITAD with a certain need on the road network; the time required to recover the related information will depend on the number of information he knows before starting the search (e.g. part of the title, issue, keywords). In this case **A** knows very well which data he needs → he looks for and finds the data.
- **B** finds out the existence of the road network data exploring the SITAD catalogue and realizes he might use it . In this case **B** enters SITAD → he finds out by chance interesting data → he realizes he could use the data for his purpose.

Now let’s try to change our perspective and consider a **C** user looking for particular data with a complex need: to produce risk indexes for Civil Protection purposes. It is arguable he would ask: “Which data, considering all available information, may be useful to calculate risk indexes?”. If SITAD was able to understand **C**’s question and had some information on the issue “Civil Protection” he would suggest a well known calculation formula, saying that the Risk (**R**) is the product of the Probability that a certain event occurs (**P**), of the Susceptibility (**S**), that is the predisposition of a certain territory to suffer a damage, and of the suffered Damage itself (**D**):

$$\mathbf{R = P * S * D}$$

Almost all the data may potentially have influence on **R** and among all available data SITAD would certainly extract the “road network”, specifying that roads may be, in consideration of the particular territorial context:

- a target element, if located in the impact area of a disaster: this makes the **D** factor growing;
- a resource, if out of the impact area (e.g. useful tool for transport supplies in a town striken by a landslide): this makes the **S** factor decreasing.

Common sense shows SITAD isn’t able to understand the question at the moment, but what does this exactly mean? **C** user and SITAD shall simply to speak the same language in order to understand each other: this means that the two ones should be essentially able to share **lexicon** and **semantics**. On the technical side it is fundamental to design a multi-disciplinary conceptual model, formalized and codified in order to share information: once this is done **C** will be able to ask complex questions getting back meaningful answers. We are conscious to need an integrated conceptual model within the overall set of our territorial data, to formalise the meaning both if we consider the general cartographic aspects to be represented in a digital map, but also their roles within every specific non geographical discipline. [3] [4] [5] [6] [8]

Getting back and focusing on the road network example, a robust conceptual model, formalized through an ontology, should be able to define:

A. The geometric meaning of a road element, represented by a line or a surface. This data will be related with other cartographic objects through semantic relationships and also through spatial and topological relationships.

B. The meaning of a road element according to one or more thematic disciplines and to one or more validated lexical sources. As far as Civil Protection is concerned, a road element can represent a risk source, the target of a calamitous event or a resource, but if we consider the same road element from the side of another discipline, e.g. territorial or urban planning, it may assume different roles.

3. Designing the ontology

We shall consider four basic constituents to cover the conceptual modelling requirements and describe the knowledge domain in a efficient way: [9]

- **lexicon**: the set of terms covering the overall domain made of several thematic areas;
- **semantics**: the meaning (or meanings) given to each term depending on different sources and different thematic contexts; lexicon and semantics are representing the content reference.

- **ontology schemas** or **ontological maps**: the ontology itself represents the formalized content for the conceptual model and it maps the objects and the set of hierarchical and logical relationships describing the connections among them.
- **computational ontologies**: to reach a reliable ontology the whole conceptual architecture has to be based on primitive definition mapping both the abstract lexical meanings and the geometry primitives in the upper ontology level.

So our purpose is the building of a Geographic Information ontology: the modelling rules will be set apart of every specific thematic context and furthermore will be inclined to avoid the building of a new, original set of computational ontologies; on the contrary they will adopt a validated one, widely acknowledged by the scientific community. This is also going to save the principle that operative ontological models may exist in many applicative situations and be interoperable if an upper level universal ontology is supporting them.

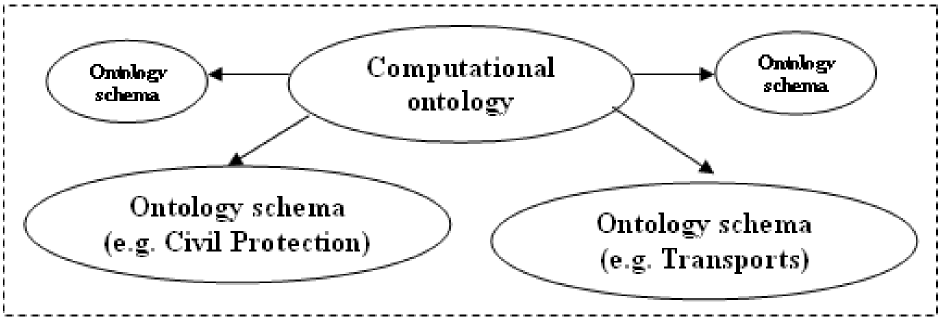


Fig.1 Geographic Information Computational Ontology connections

4. Building the ontology

We tried to dedicate a great care choosing references, at least recognized at the national level, for each one of the four fundamental elements of our ontology:

- For the **lexicon** and the **semantics** of the overall and generic terminology the Treccani Italian Language Dictionary has been chosen; it has also been used to check and validate input coming from the GI specifications. National laws and regulations were put beside the Dictionary to improve detail level definition (e.g. for the Transportation field firstly the Italian Transportation Code and many other technical and scientific texts).
- For the **ontology schemas** the GI specifications from INTESA GIS [13] [14], born for supporting the Italian Spatial Data Infrastructure, were considered. And the focus was initially on the Transportation layer. Attention was paid also to European projects involving the Transportation topic (see EUROROADS and the guidelines of the INSPIRE European Directive [7] [11] [12]), to enlarge the validation of the modelling activity through model comparison. As far as Civil Protection is concerned, we have also considered an ontological schema prepared by the Joint Research Centre in Ispra, commissioned by Piedmont Region as part of the activities to elaborate the Regional

Program for Risk Forecast and Prevention: this schema gives a structure and an organization of the territorial data organic and functional to Civil Protection activities.

- For the **computational ontologies**, the ISO international standards (ISO TC 211 series on Geomatics and other series dedicated to thematic issues) were elected to define the primitives for the upper level ontology: this is consistent with the INTESA GIS data specifications and its adoption of the GeoUML primitive organization, derived from the ISO one.

4.1 Facilities and Computational Ontology structure: OWL, Protégé, GeoUML

OWL-Full is the ontology language we adopted and we used within Protégé, one of the most popular free ontology tools. Automated reasoning on OWL-Full ontologies will later be provided by an inference engine: Jena and Racer are the ones we started to reasoning on.

By the computational ontology side GeoUML is an UML based language, since it includes all the constructs of UML for class diagram specification and the OCL language ("Object Constraint Language") for integrity constraints specification; but, most of all, it is based on the ISO TC 211 standards, since it reuses and reorganized in a simpler way the primitives for geometry representation proposed by ISO 19107 document ("Spatial Schema"), moreover it follows the General Feature Model approach and the rules for UML schemas from ISO 19109 document ("Rules for application schema"). [2]

5. An example from the experimentation phase

Let us try to examine an ontology schema extract, including the upper semantics for a generic Geometric Object (GM_Object in Figure 2) derived from the GeoUML structure and the specific objects kept from the INTESA GIS specifications related to the "road network". The thematic content would be usable in our intention by two thematic disciplines: Transports and Civil Protection, but this approach does not exclude the future enlargement to any kind of objects and its usability by any discipline.

Since a new ontology has always to deal with real data, in our case we shall check one last condition: the mapping of the INTESA GIS objects (presented in the ontological map) with SITAD data and metadata actually available in our regional catalogue.

5.1 Ontology schemas

In figure 2 we can see a taxonomy branch where some of the classes belonging to our domain are included. The schema allows to identify a concrete data (the road network, "Rete_Stradale_Liv1" in the picture) navigating through a model that, starting from the most general and abstract class ("Object"), bounds classes that are on a lower hierarchic level through specialization relationships formalized through the GeoUML approach. For example the class Geometric Point "GM_Point" inherits characteristics and semantics from the class Geometric Primitive "GM_Primitive" that, on its side, is classified as Geometric Object "GM_Object". The model goes from general to particular until it reaches the element "road network".

In figure 3 we can see one of the UML schemas from INTESA GIS containing logical relationships among classes from the Transportation domain [14]; this Transportation schema (with other ones pertaining the “road network” class) is going to be translated in the ontology formalization during our experimentation activity, to join the upper classification ontology (the taxonomy in figure 2) with the thematic lower context.

In figure 4 we can see another of our mosaic tiles, the ontological schema on datasets for the organic management of territorial activities on Civil Protection. Also this schema shall be integrated in the common ontology formalization to be joined with the upper classification ontology and to capture the profile for the “road network” class by side of Civil Protection theme.

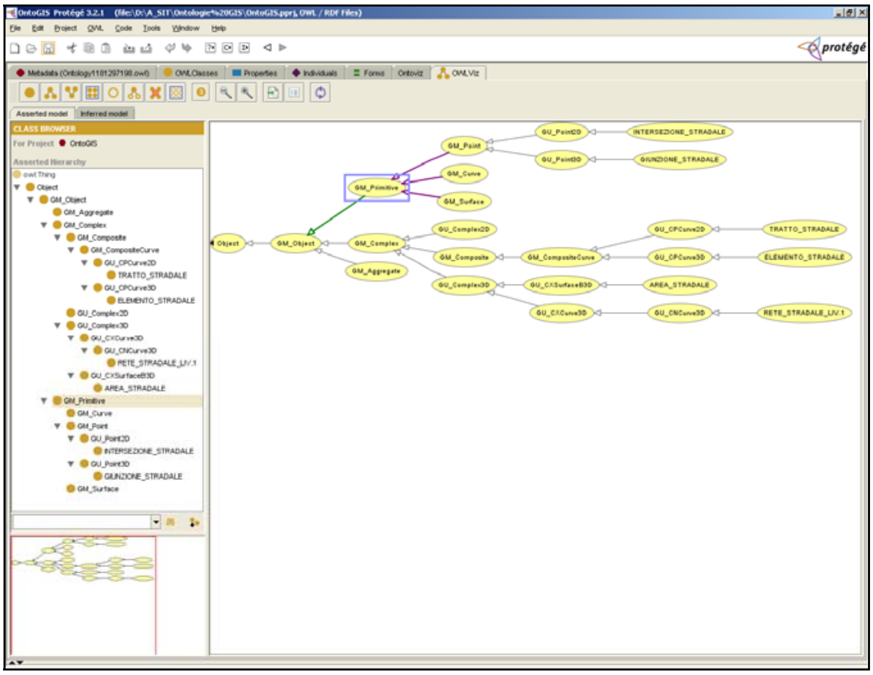


Fig.2 Partial view of the upper ontology schema (screenshot from Protégé)

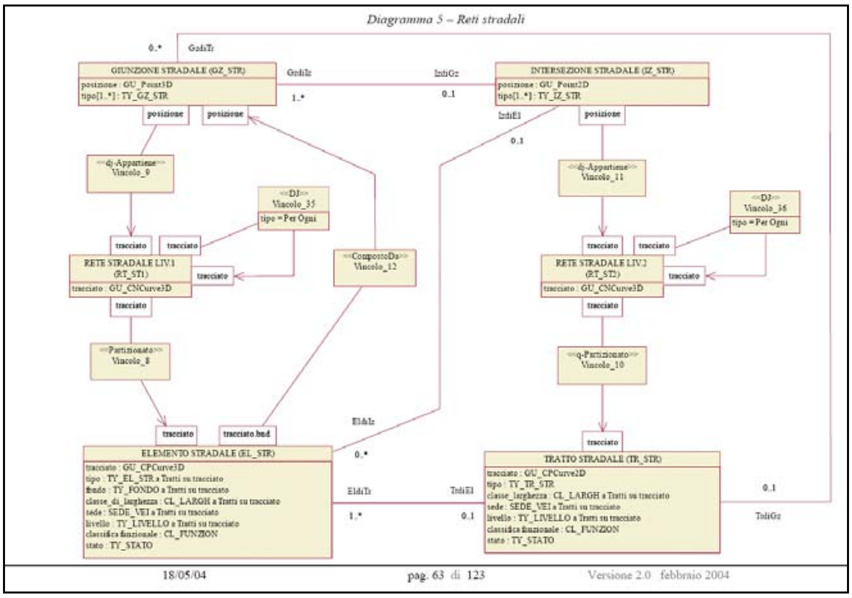


Fig.3 UML schema from the Transportation domain (INTESA GIS specifications)

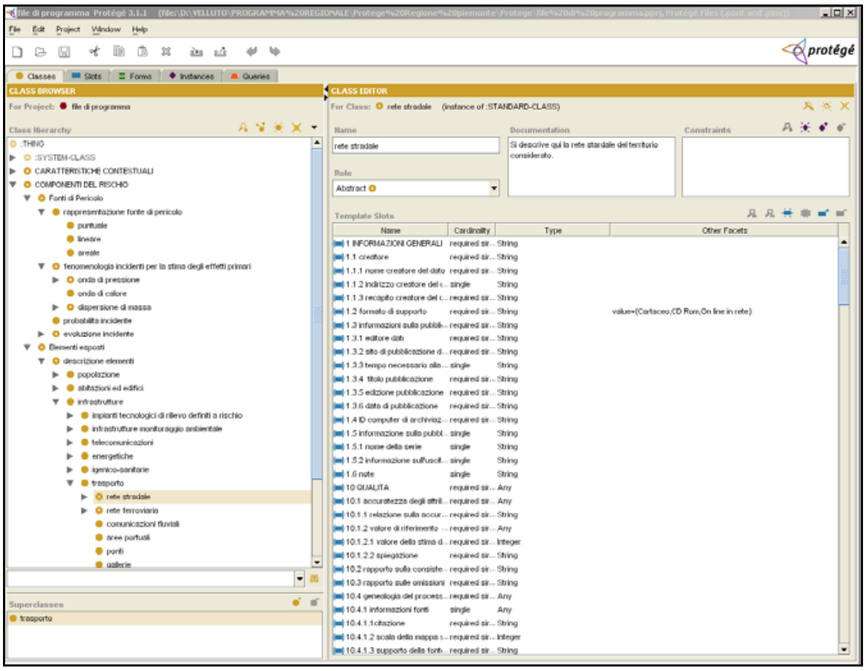


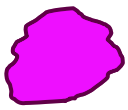
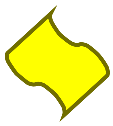
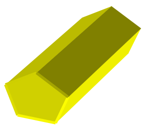
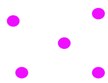
Fig.4 Class hierarchy from the JRC ontology on Civil Protection (screenshot from Protégé)

5.2 Semantic approach for GIS primitives

In our conceptual model upper level classes dedicated to geometry (the most general ones, in the GI computational ontology) were characterized by a precise semantics, related to their geometric nature, derived from the ISO 19107 “Spatial Schema”, that fixed them in a unambiguous way. Afterwards GeoUML took those definitions and built on them the simplified hierarchy we can see in figure 2. The semantic interpretation for basic geometric primitives from these sources was given in Chart 1 and 2.

5.3 A real application case

Let’s consider once again our foundational ontology and let’s focus on the class “road network” in order to make clear the chance to integrate and reuse information through the three ontological schemas presented in paragraph 5.1.

			Graphic example	Name	Description (SEMANTICS)
GM_Objects	GM_PRIMITIVE	2D Euclidean space		**Point**	Basic geographic object made of a single isolated point
				**Curve**	Basic geographic object made of a single line or curve tract
				**Polygon**	Basic geographic object made of a single free shape polygon
		3D Euclidean space		**Surface**	Basic geographic object made of a single free shape surface; the surface may be a plane surface or not, positioned in any way in 3D space
				**Solid**	Basic geographic object made of a single free shape solid, positioned in any way in 3D space
	GM_AGGREGATE	2D Euclidean space		**MultiPoint**	Geographic object made of many (more than one) isolated points. The individual components are allowed to be elementary objects of type Point (note this is not mandatory: only if necessary!)

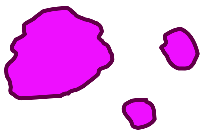
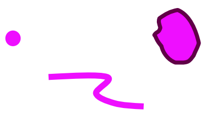
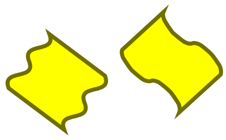
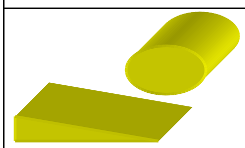
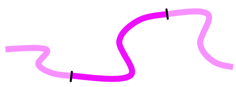
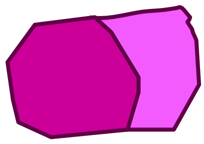
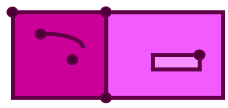
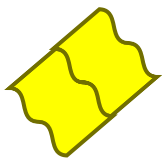
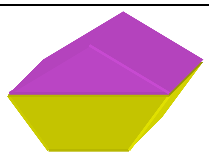
		**MultiCurve**	Geographic object made of many (more than one) separated lines or curves: no contiguity allowed. The individual components are allowed to be elementary objects of type Curve (note this is not mandatory: only if necessary!)
		**MultiPolygon**	Geographic object made of many (more than one) separated polygons: no contiguity allowed. The individual components are allowed to be elementary objects of type Polygon (note this is not mandatory: only if necessary!)
		**Planar Mixed Aggregation**	Geographic object made of many (more than one) separated 2D basic objects: Points, Curves, Polygons with no contiguity allowed. The individual components are allowed to be elementary objects of type Point or Curve or Polygon (note this is not mandatory: only if necessary!)
3D Euclidean space		**MultiSurface**	Geographic object made of many (more than one) separated surfaces: no contiguity allowed. The individual components are allowed to be elementary objects of type Surface (note this is not mandatory: only if necessary!)
		**MultiSolid**	Geographic object made of many (more than one) separated solid objects: no contiguity allowed. The individual components are allowed to be elementary objects of type Solid (note this is not mandatory: only if necessary!)
		Free Mixed Aggregation	Geographic object made of many (more than one) separated 2D and 3D basic objects: Points, Curves, Polygons, Surfaces, Solids with no contiguity allowed. The individual components are allowed to be elementary objects of type Point or Curve or Polygon or Surface or Solid (note this is not mandatory: only if necessary!)

Chart 1 Summary on geometric primitives in the foundational ontology

			Graphic example	Name	Description (SEMANTICS)
GEOMETRY & TOPOLOGY aspects	GM_COMPOSITE	2D Euclidean space		**Composite Curve**	Geographic object made of many (more than one) individual lines or curves which contiguity is requested. The individual components are allowed to be elementary objects of type Curve (note this is not mandatory: only if necessary!)
				**Composite Polygon**	Geographic object made of many (more than one) individual polygons which contiguity is requested. The individual components are allowed to be elementary objects of type Polygon (note this is not mandatory: only if necessary!)
				**Planar Mixed Composition**	Geographic object made of many (more than one) individual 2D basic objects: Points, Curves, Polygons wich contiguity is requested. The individual components are allowed to be elementary objects of type Point or Curve or Polygon (note this is not mandatory: only if necessary!)
		3D Euclidean space		**Composite Surface**	Geographic object made of many (more than one) individual surfaces which contiguity is requested. The individual components are allowed to be elementary objects of type Surface (note this is not mandatory: only if necessary!)
				**Composite Solid**	Geographic object made of many (more than one) individual solid objects which contiguity is requested. The individual components are allowed to be elementary objects of type Solid (note this is not mandatory: only if necessary!)
				Free Mixed Composition	Geographic object made of many (more than one) individual 2D and 3D basic objects: Points, Curves, Polygons, Surfaces, Solids which contiguity is requested. The individual components are allowed to be elementary objects of type Point or Curve or Polygon or Surface or Solid (note this is not mandatory: only if necessary!)

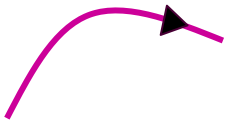
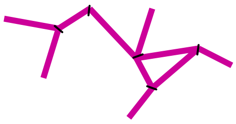
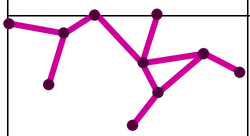
				**Orientable Curve**	Geographic object made of a single line or curve tract with an orientation verse declared on it
				**Graph**	Geographic object made of many (more than one) individual Curves or Composite Curves which contiguity is requested and multiple connection is allowed. The individual components are requested to be elementary objects of type Curve or Composite Curve
				**Edge-Node Graph**	Geographic object made of many (more than one) individual Curves or Composite Curves which contiguity is requested, multiple connection is allowed and points function as nodes. The individual components are requested to be elementary objects of type Points, Curve or Composite Curve

Chart 2 Summary on geometric primitives in the foundational ontology

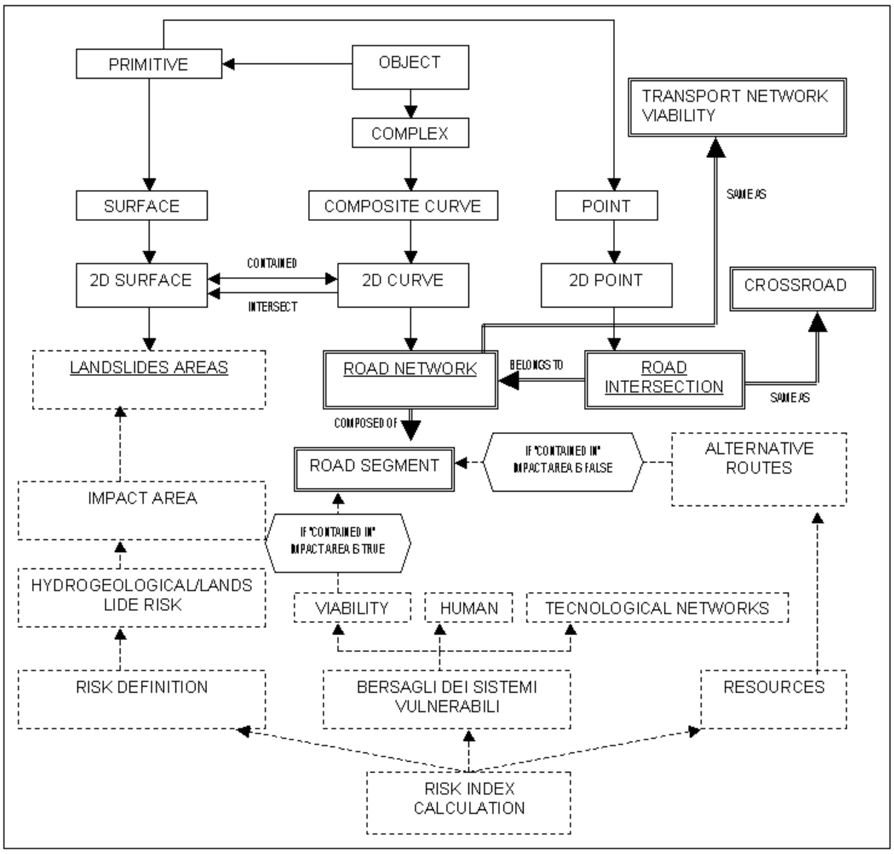


Fig.5 Ontology schema

Figure 5 reports the full line bordered elements referred to the computational ontology, the double lines referred to the Transportation domain and dashed lines to the Civil Protection domain. Some relationships among classes have been expressed and represented as directional arrows (the ones that bind the classes “curves 2d” and “surface2d” belong to the set of topological relationships defined within GeoUML). Two ways arrows represent relationships that allow symmetrical binds (e.g. content in/external), while dotted arrows define synonymous relationships. Hexagonal boxes point out conditional relationships that, in the given case, are solved as ‘true’ or ‘false’ at the upper level of the ontology. Finally underlined names indicate object classes grouping data directly managed by CSI Piemonte thanks to the SITAD catalogue.
Let’s get back to the example in paragraph 2, C user with his own goals that needs to integrate information coming from some specific territorial domains. Moving through the proposed ontology, C user will be able to gather two kinds of simple information:

- Explicit: looking for data about “crossroads” he will be able to find the right location for information in the catalogue even if the original name is “road intersections”.

- Implicit: he can find out which data, among the existent ones over the Piedmont, may be useful to calculate risk indexes for Civil Protection. As an example, “road network” might be useful to calculate the hydrogeological/landslides risk index; then depending on the topological relationship that bind a single element of the road network to the “impact area” it can be considered as a target or as a resource.

6. Conclusion and perspectives

As our main intent was to build a rigorous conceptual model, we have later justified the choice of reference procedures and standards for this task and our examples have shown how such a model can be strategic in order to build transversal ontology applications based on a semantic search engine. In order to assure the implementation firstly we decided to design the model using an UML tool and then import it in a ontology tool (Protégé) to get a schema expressed in OWL-FULL. Then such a schema had to be refined through the definition of all semantics and inference rules. Finally an inference engine (such as Jena) has been allowing us to test the overall model value before moving on to the interface development for data search.
Therefore the ontological formalization as the one described above will enable us to achieve three important objectives and useful deliverables, characterized by complexity increase:

- A **‘reasoning’ vocabulary** that, exploring the model, provides information about GIS and specific thematic disciplines characterized by different complexity and level of detail. It could be useful both as a support to experienced users and as basic didactical tool.
- A **Semantic search engine for geographic data** that, exploring the conceptual model, would enable a generic user to retrieve information in a shorter time **(more rational search)**; to let the user discover new useful data and logical relationships, initially unknown **(increasing data use)**; to extend the range of possible questions, even ‘complex’ ones, that users can ask getting significant results back **(increasing data usefulness)**.
- One or more **GIS applications** that will no longer query a Geographic database through static and predetermined queries, but through dynamic and “deductive” ones, interacting with the conceptual model through the **semantic engine** devices. In this case the answer to the complex questions of **C** user would be accomplished by a pleasant map together with data and information lists.

References

1. Batini C., Ceri S., Navathe S., Conceptual Database Design: an Entity Relationship Approach, Redwood City, The Benjamin/Cummings Publishings Company, 1992

2. Belussi A., Negri M., Pelagatti G., GeoUML: an ISO TC 211 compatible data model for the conceptual design of geographical databases - Internal Report n.2004.21 Dipartimento di Elettronica e Informazione, Politecnico di Milano, 2004

3. Caglioni M., Rabino G. A., Theoretical approach to urban ontology: a contribution from urban system analysis, Workshop of COST Action C21 Towntology, 2006

4. Gruber T. R., A translation approach to portable ontologies. Knowledge Acquisition, 1993

5. Gruber T. R., Toward principles for the design of ontologies used for knowledge sharing - Padua workshop on Formal Ontology, March 1993, later published in International Journal of Human-Computer Studies, Vol. 43, Issues 4-5, November 1995

6. Gruber T. R., What is an Ontology, http://www-ksl.stanford.edu/kst/what-is-an-ontology.html

7. INSPIRE, INfrastructure for SPatial InfoRmation in Europe - http://inspire.jrc.it/home.html

8. Keita A., Laurini R., Roussey C., Zimmerman M., Towards an Ontology for Urban Planning: The Towntology Project – 24th UDMS Symposium, Chioggia, October 2004

9. Masolo C., Borgo S., Gangemi A., Guarino N., Oltramari A., WonderWeb Deliverable D18, Ontology Library (final) - from IST Project 2001-33052 WonderWeb: Ontology Infrastructure for the Semantic Web

10. V.V.A.A., Deliverable D6.3 - Final specification of Road Network Information Model – from EUROROADS Specifications, 2006

11. V.V.A.A., Geographic information — Standards, specifications, technical reports and guidelines, required to implement Spatial Data Infrastructure – Joint Research Centre, Institute for Environment and Sustainability , 2005

12. V.V.A.A., Requirements for the Definition of the INSPIRE Implementing Rules for Spatial Data Specifications and Harmonisation – Joint Research Centre, Institute for Environment and Sustainability , 2005

13. V.V.A.A., Specifiche per la realizzazione dei Data Base Topografici di interesse generale: Il Catalogo degli Oggetti – from INTESA GIS Specifications, 2006

14. V.V.A.A., Specifiche per la realizzazione dei Data Base Topografici di interesse generale: Lo schema in GeoUML delle Specifiche di Contenuto – from INTESA GIS Specifications, 2006

Geo Semantic Web Communities for Rational Use of Landscape Resources

Ernesto Marcheggiani*, Michele Nucci, Giovanni Tummarello, Christian Morbidoni,

Dipartimento di Elettronica, Intelligenza Artificiale e Telecomunicazioni
* Dipartimento di Scienze Applicate ai Sistemi Complessi - Agrur
Technical University of Marche
Via Brecce Bianche – 60131 Ancona (ITALY)
e.marcheggiani@univpm.it, mik.nucci@gmail.com,
{g.tummarello,c.morbidoni}@deit.univpm.it.

Abstract. A methodology based on Semantic Web Initiative standards and technologies in order to improve rational use of urban, rural and semi-natural global cultural landscape resources in periurban fringe expansion areas and in riconnection and ecological rebalancing green areas, is proposed. This paper aims at showing how the need of advanced cooperative annotations and information exchange can be addressed using a paradigm called "Interconnected Geo Semantic Web Communities". Furthermore we aim at designing the bases to implement a Decision Support System to rationally manage resources which makes use of both geographical information advanced technologies such as remote sensing and Geographic Information Systems technique and semantic rules acting on ontologies.
Based on the DBin platform a specific environment has been built to manage interoperable data arising from dedicated specific communities annotations and survey campaigns. The system features "collaborative annotations" and "intelligent" data structures which can be easily personalized easily.

Keywords: Ontologies; Semantic Web; GIS; DBin, Sustainable tourism; Landscape resources

1.0. Introduction

In order to adhere to the European Charter for the Sustainable Tourism, different stakeholders, belonging to thirteen municipalities of the "*Comunità Montana delle Alpi Lepontine*" (Cmal) in Northern Italy, cooperated using an advanced tool for Semantic Web communities, provided with geographical plug-ins which enable a real-time interaction with systems as Google Earth and Google Maps supported by specific ontologies to edit, visualize and share information.

The use cases and associated needs have been highlighted and after that the base tool for this work, DBin (G.Tummarello et al, 2006), a Semantic Web rich client application, has been focused.

DBin enables users to create and experience the Semantic Web by exchanging information structured according to Resource Description Framework (RDF) (RDF Semantics, 2004), standardized by the World Wide Web Consortium (W3C), in peer-to-peer (P2P) "topic" channels. Once sufficient information have been collected locally, rich and fast browsing of semantically structured knowledge becomes possible even offline without generating external traffic or computational load.

DBin has a number of modules to support cooperative tagging and annotations of geographical objects. Different communities of users, e.g. concerned with different kind of geographical objects, can exploit DBin to cooperate in an enriched geo-semantic space. Advanced users, e.g. cultural heritage agencies, can join multiple groups at the same time and use the collective cross domain knowledge which will be relevant not only for planning but also to allow the implementation of Decision Support Systems (DSS) for the rational use of different landscape resources.

2.0. European Charter for Sustainable Tourism in the *Comunità Montana delle Alpi Lepontine* (Cmal) landscape: organization, interests and needs

The case study area belongs to that large geographical region which can be identified with the northern part of Italy, considered as a possible *Mediterranean Megalopolis*. This territory, stretching between the Canton Ticino (Switzerland) and the Italian province of Como, presents many peculiar features. In the so-called “three lakes region” - Lugano, Piano and Como Lakes - a unique urbanistic model has been developed, since the second half of the 19th Century. It deals with that “historical villas landscape”, that occupies fringes degrading from the alpine area toward the lakes coastline. These large territorial fringes can be well divided into parallel parts with the mountainous and hilly ones characterized by rural ecosystems - here the 'alpeggio' (seasonal pasture) practice is still a productive activity that contributes to the persistence of the agro pastoral landscape. Persistence of agriculture in those areas preserves great naturalistic areas ecological functionality, acting as biodiversity reserves and thus subjected to different protection levels.

In the last decades the area has been particularly affected by severe pressure due to both a generalized uncontrolled urbanistic expansion and a relevant touristic flow coming particularly from Northern Europe. It is therefore necessary to cope with these landscape different aspects - scientific, historical, artistic and economical – from a multidisciplinary and holistic point of view which enables a complex scientific image in order to suggest integrated and effective planning actions for a rational use of natural and cultural landscape resources.

According to the Como Tourist office, the area attracts approximately 3 million persons per year, half of which from outside Italy. The area is, in general terms, welcoming and well organized according to traditional standards, but the importance to provide tourists with rich and personalized offers is quickly increasing. This requires integration and sharing of information between a number of local actors using state of the art information technologies. It will be shown how this can be addressed with the paradigm of Interconnected Geo Semantic Web Communities.

In a wider perspective Cmal area is only a part of a broader net of natural and historical elements that constitute the whole local cultural heritage, surrounded by a complex landscape where each stakeholder performs a number of specific tasks, most of which would benefit from cooperatively exchanging geo-spatial semantic annotations.

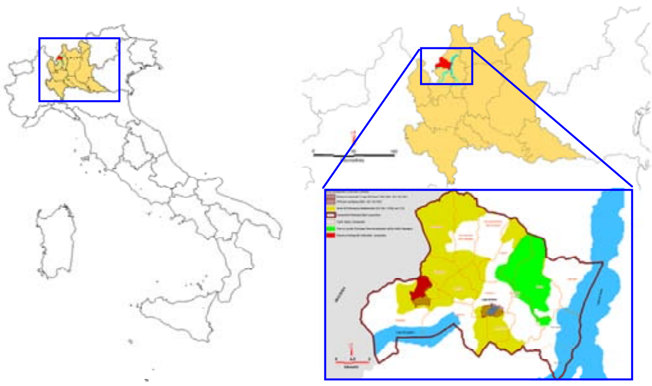


Figure 1. Location of the study area in Northern Italy.

3.0. Geo Semantic Web Communities

The Geo Semantic Web community tool, proposed by the authors, is a specific application of DBin Semantic Web Platform. DBin enables user communities to annotate concepts of common interest and browse information structured by RDF according to ontologies for specific domains of interest.

For this study, DBin has been provided with both geo-enabled visualization and editing modules, as well as with processing capabilities to handle data structured according to geographical and geometric ontologies such as the W3C Basic Geo Ontology (W3C Semantic Web Interest Group, 2003) and RDFGeom Ontology (Chris Goad, 2004).

In this work the term "ontology" refers to a set of classes and related properties, relations and constraints defined using RDF Schema Language (W3C Consortium, 2004a) and/or Ontology Web Language (W3C Consortium, 2004b) for modelling a specific domain.

Geo Semantic Web communities are characterized by typology of knowledge that is exchanged among them. This means that by connecting to a specific community, users will only share and exchange those information which are relevant to the community itself. In DBin this is achieved by means of topic specific P2P groups based on the RDFGrowth algorithm discussed in section 4.0.

In the proposed case study, cultural heritage agencies are interested in annotating the state of cultural patrimony. At the same time parks and woods maintenance are performed by Forest Guards through specific actions, such as dangerous trees pruning or cutting, service buildings and routes maintenance, etc.

The proposed Rich Semantic Web client solution offers to each group a specific environment where such annotations can be edited, browsed and queried to assist decisions and management. Collaborative metadata annotations created within each community, can also point at rich media, posted on the web (e.g. pictures, documents, long texts, etc.). Users who receive annotations could then reply or further annotate it, locally, for personal use or back into public knowledge.

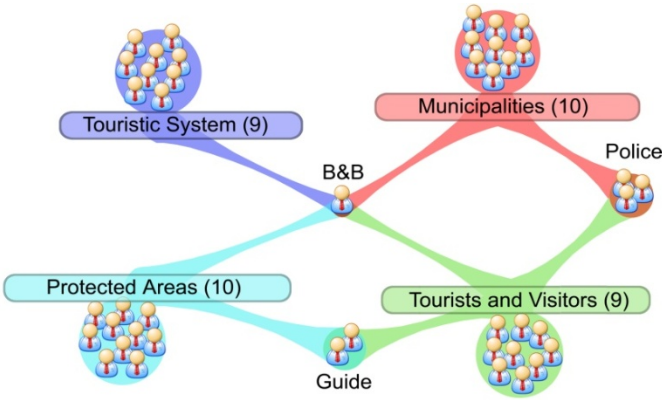


Figure 2. Our reference use case: groups of expert annotate and use their own domain annotations. Experts, by joining different communities can benefits of joint cross-domain knowledge.

Due to the large amount of visitors per year, CMAL area managers - e.g. advisors and majors - have broad interests which go beyond single users community. Taking care of urban area and mobility infrastructure maintenance, dwellers and visitors security, natural and cultural heritage resources conservation, etc., they are interested in performing cross-domain queries, such as unsafe touristic paths or historical building identification. Meanwhile the Local Tourist System (STL) managers are interested in attracting the largest number of visitors per year without caring about the impact of that fluxes on the local natural and cultural resources. STL aims, for examples, at informing visitors about temporary inaccessible naturalistic paths or about B&Bs lying close to naturalistic areas. Tourists are just interested in having the possibility to freely choose accommodation - hotels, B&Bs, country houses, etc. - or places in which it is possible to taste local food, experiencing a safe and positive journey.

It is interesting to underline how the use of web resources in planning holidays in an emergent and increasing market, especially among northern European cultures. Tourism is cross-sectorial, involving a wide range of issues: trade and investment policy, employment and enterprise, public-private partnerships, community and urban planning - land use planning, transportation, etc. - conservation of cultural heritage, protected areas and biodiversity, management of natural resources (water, energy, waste), safety and security, education and workforce development. There are a lot of stakeholders, with their different and sometimes opposing interests and agendas.

In this context, Semantic Web Communities can be seen from a new point of view. If two communities share resources identified by Uniform Resource Identifiers (URIs), object annotations, originally posted in a specific community, are automatically cross posted to other communities, as soon as the annotated objects is of interest to both communities.

Summarizing, actors join either one or more P2P groups to learn and exchange information about topics of interest, thus collecting an heterogeneous amount of structured knowledge expressed in RDF.

In this context the issue becomes how to allow users to interact with RDF data in a natural way. This is a strongly domain- dependent task, as each community deals with different kinds of information and has different needs in terms of browsing and editing capabilities.

The proposed solution is to provide a way for a "group leader" to define and make available to users, a set of domain-specific interaction environments called Brainlets, described in section 5.0.

New Geo Semantic Web Communities can be started at will, with a relative little technological effort, by defining P2P groups for specific topics and “publishing” related Brainlets.

4.0. The RDFGrowth P2P engine: high level overview

RDFGrowth algorithm is used by DBin clients to collect and distribute RDF data within groups of interest. The algorithm is presented and discussed in G.Tummarello et al, 2004; a high-level overview is given here.

Previous P2P Semantic Web applications, such as Wolfgang Nejdl, Boris Wolf 2002, Paul Alexandru Chirita 2004, Min Cai, Martin Frank 2004 and Wolfgang Nejdl 2003, have explored interactions among groups of trusted and committed peers. In such systems, peers rely on each other to forward query requests, collect and return results. On the contrary, the real world scenario of peers, where cooperation is relatively frail, is here considered.

In RDFGrowth peers are certainly expected to provide some external service, but commitment is minimal and in a “best effort” fashion: no commitment in terms of complex or time-consuming operations, such as query routing, collecting and merging, is required from peers.

4.1. Overview and main features

RDFGrowth algorithm enables the creation of P2P groups around topics of interest. When an user joins a specific P2P group, a software agent, based on RDFGrowth algorithm, begins to collect and share only information about the group’s topic. Topic definition is given by the Group URI Exposing Definition (GUED). GUED is an operator which, applied to a RDF knowledge base, returns all and only those resources conforming to a set of semantic constraints (e.g. “hotels and restaurants located within a naturalistic area”). Only the resources extracted by a GUED operator will be shared with peers of a same P2P group of interest.

Peers in the same group exchange among themselves pieces of information extracted locally at each peer by a GUED operator, so that, after a transitory period, each of them will exactly have the same knowledge about the group's topic and each of them will have to store locally the knowledge obtained from the P2P communication channel.

The RDFGrowth approach is particularly suitable in every scenario in which users have necessity to work with off-line data and then, after having processed it, they will have to exchange results within a group of interested people. For instance, this approach could be applied to a geo-annotation scenario in which every user (every peer) could work off-line on mobility devices far from a connectivity point and only after having annotated some locations (e.g. "tourist facilities in a naturalistic area") they will exchange the resultant information among them.

P2P groups and related GUED operators can be defined by editing an XML configuration file. In particular, a GUED operator can be implemented as a set of Semantic Web queries based on SeRQL query language (Jeen Broekstra, Arjohn Kampman, 2004) which selects and extracts information of interest from an RDF knowledge base. Example 1 shows a possible GUED configuration for a group interested on tourist facilities within naturalistic areas in a specific municipality.

```
<group name='TuristFacilities'>
  <gued query='SELECT X FROM {X} rdf:type
        {<http://dbin.org/buildings#touristFacilities>};
         location:locatedIn {Y} rdf:type
        {<http://dbin.org/area#naturalisticArea>}
         WHERE Y = "<http://dbin.org/municipality#13047" ' />

  <default_brainlet name="TouristFacilities"
               uri=http://dbin/brailets/touristFacilities />
</group>
```

Example 1. The Tourist Facilities group configuration.

4.2. Identities and authorship of information

In a potentially large and unregulated P2P community it is important to have information about "who said what", in particular which user is the author of a particular annotation received from the network. To enable this, in DBin a methodology based on the Minimun Self Contained Graph (MSG) theory defined in [13] is used. Such methodology allows each piece of information inserted by users, to be digitally signed at a fine granularity level and in an efficient way. It also assures that authorship information will remain within metadata when they will be exchanged by P2P communication channel. Details about digital signature process are outside the scope of this discussion and can be found in G.Tummarello et al, 2005; only a high-level overview of the basic procedure is given here.

When started up for the first time, DBin clients require the users to provide a valid URI or a nickname (used to automatically generate a valid URI to indentify) which will act as an identifier for the user itself. Then a public and a private keys are generated; the private key is stored locally, while the public key is uploaded to a public server. Every time a user will insert an annotation into the system, the user identifier as well as the URL of the public key will be added to the annotation itself which will be also signed using his/her private key. In this way, after having received a piece of information from a P2P group, DBin clients are able to retrieve the associated public key to identify the author of the annotation, without caring about the provenance of the information itself.

Once the authorship of each annotation can be derived, a variety of local filtering rules can be applied at will. For example, users can build a local policy in order to hide annotations from untrusted authors.

5.0. Brainlets for Semantic Web Communities

Defining Semantic Web communities and working with data coming from specific domains of interest requires dedicated software environments and tools such as data visualization systems, specific data editors, etc. Brainlets can be thought as "configuration packages" setting DBin to operate on a specific domain, providing also needed tools. Brainlets can be viewed by users as full "domain applications running inside DBin". By Brainlets it is possible to define DBin settings for:

- general UI layout, defining also involved UI components and interaction among them;
- ontologies to be use for RDF data and for annotations in a specific domain;
- template for specific domain "annotations";
- templates for readily available "precooked" domain queries;
- templates for wizards to insert new specific domain resources;
- suggested trust models and information filtering rules (e.g. identities of "founding members" or authorities).

Since DBin is base on the Eclipse Rich Client Platform technology, every Brainlet, which technically is an RCP plug-in, can be installed into DBin in a simple way.

New Brainlets can be created by editing an XML configuration file: no programming skills are required. Other than a basic knowledge of the XML, every task involved in the process of creating a new Brainlet are related to "knowledge engineering" aspects such as: selecting appropriate ontologies to describe a specific domain of interest or making queries to select a set of interesting data.

This section describes the main elements involved in creating Brainlets related to a landscape resources planning scenario like that focused on this case study.

5.1. Ontologies and Basic RDF Knowledge Base

In creating a new Brainlet an important step is the choice of the ontologies which formally describe the concepts related to the domain of interest. Once ontologies have been selected, related RDFS/OWL files can be included into the Brainlets itself although they could be placed on the Web.

In the case study, existing ontologies have been used to foster information reuse and interoperability. In particular, the W3C Basic Geo Ontology has been used to represent basic geographic concepts such as points with latitude, longitude and altitude from the WGS84 reference datum specification. In addition to this ontology, RDFGeom ontology has been used to represent more advanced geometric concepts such as: areas, transformations, curves, segments, polylines, boxes, circles, etc. Example 2 shows the use of W3C Basic Geo Ontology to represent the location of a hotel and the example 3 shows the use of RDFGeom Ontology to represent a simple polyline.

```
<buildings:hotel rdf:resources=”http://dbin.org/hotel#1”>
   <geo:lat>55.701</geo:lat>
   <geo:long>12.552</geo:long>
</buildings:hotel>
```

Example 2. representation of the location of a hotel by RDF based on the W3C Basic Geo ontology.

```
<geom2d:Polyline rdf:resource="http://dbin.org/pline#1">
 <geom2d:points>
  <rdf:Seq>
   <rdf:li>
    <geom2d:Point>
     <geom:x>-123.835414</geom:x>
     <geom:y>46.191617</geom:y>
     </geom2d:Point>
    </rdf:li>
    <rdf:li>
     <geom2d:Point>
      <geom:x>-123.821983</geom:x>
      <geom:y>46.189905</geom:y>
      </geom2d:Point>
    </rdf:li>
     … … …

  </rdf:Seq>
 </geom2d:points>
</geom2d:Polyline>
```

Example 3. representation of a simple polyline by RDF based on the RDFGeom and RDFGeom2D ontologies.

5.2. Navigation of RDF Resources

The way resources are presented and browsed is crucial to interface usability and effectiveness in finding relevant information. RDF data structure representation is based on a graph model. Today many graph-base visualizers exist but the problems related to the usability and usefulness of this approach are well-known: large graphs with a lot of nodes are difficult to render and navigate. Normal users are familiar with a folder-like structure so, in DBin, it has been decided to implement a Resources Navigator based on a flexible and dynamic tree structure. This approach fits better with respect to the number of resources. Resource Navigator has a completely configurable semantic multiple branches folder tree structure, each one is able to contain specific resources with respect to criteria specified by the Brainlet author.

In this case study, different Navigators have been defined. The Tourist Facilities Navigator, shown in figure 3 has been configured to have more branches, each one can organize different kind of tourist facilities; Municipalities Navigator is shown in figure 4.

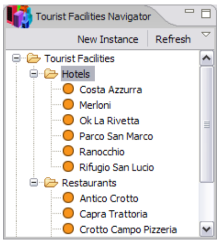


Figure 3. Tourist Facilities Navigator.

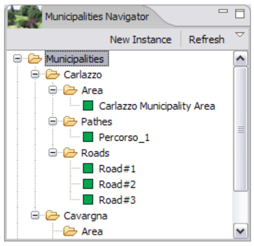


Figure 4. Municipalities Navigator.

5.3. "Precooked queries"

While creating a Brainlet the author can identify for the domain of interest some queries which can be frequently used to fulfil relevant use cases. The author of the Brainlet can define these queries, called in DBin "pre-cooked queries", within Brainlet XML configurations file.

The pre-cooked queries will be ready-to-use by the end-users of Brainlets, who will only have to "fill in the blank boxes" on the dedicate window (figure 5).

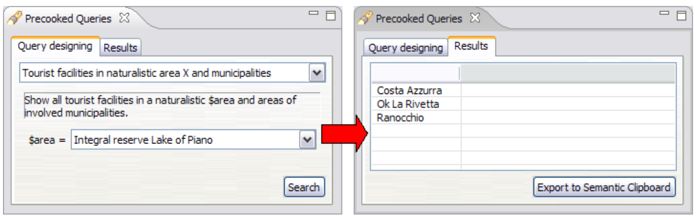


Figure 5. An example of a precooked query.

For the case study focused here, such a query could be: "find all touristic facilities in the naturalistic area "X" and areas of municipalities involved". The results of precooked queries can be viewed in a tabular way or can be sent to a dedicated visualizer (e.g. visualization plug-ins based on Google Map/Google Earth) by the "semantic clipboard" which is a DBin clipboard for heterogeneous data semantically structured.

5.4. URI Wizards

In the Semantic Web, each resource is identified by a URI. To foster knowledge interoperability within communities, a methodology is needed to avoid different users to choose different URIs to identify the same concept. This can be achieved by defining procedures to assist users in assigning identifiers to new instances. These procedures can be defined by the author of a Brainlet and can be encoded into the so called

"URI wizards". By XML configuration it is possible to define customized URI Wizards to assign meaningful URIs to each specific class of resources.

In the case of geographic objects, a simple "URI wizards" can be configured to obtain URIs from their natural aspects. For example an identifier for a touristic facility can be derived from its address as well as an identifier for a naturalistic area can be obtained from the coordinates (latitude and longitude) of its centroid.

Solutions for minting identifiers on Semantic Web are still in their infancy. The URI wizard approach offers flexibility to accommodate future methods as they will be made available and/or reach popularity.

5.5. Custom domain dependent annotation templates

In DBin, end-users can add new properties to a resource of interest. Brainlets use ontologies to assist users in this operation, suggesting which properties can be associated with a resource. However for some specific domains of interest, the author of a Brainlet could have the necessity to define annotations with a complex structure. This type of annotation, which in DBin is called "complex annotations", can be defined using an ad-hoc ontology.

When a user selects a resource to add it a new complex annotation, DBin is able to determine which type of annotations can be applied to the specified resource and provides a dynamic wizard.

For example, in the Brainlet developed for the use-case considered here some, geographic complexes annotations have been defined. Figure 6 shows one of the defined geographic annotations and related wizard which permits to users to fill geo-spatial information such as latitude, longitude and altitude of a resource (e.g. a touristic facilities).

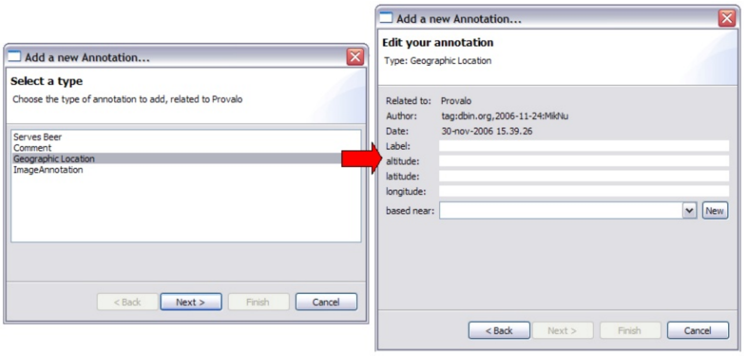


Figure 6. An example of a geographic complex annotation and related wizard used to insert geo-spatial information.

5.6. Brainlets and Geographics Plug-ins

Brainlets can also use additional plug-ins to address domain specific needs. For the considered case study specific geographic plug-ins, which permit a real-time interaction with Google Maps an Google Earth, have been used. The plug-in based on Google Maps has been used to show to end-users geographic information like the location of a building as well as it has also been used as “annotation tool” by which an end-user can add new geo-spatial annotations only selecting a resource and pointing in its real location on the map. The plug-in which enables real-time interaction with Google Earth has been used to show complex geographic information (such as regions, paths, etc.) or, for example, to show the results of a complex pre-cooked semantic query.

In figure 7 the Google Maps-based module is showing the geo-location of a selected resource (the location of a tourist facility) while in figure 8 it is possible to see Google Earth showing the result of a cross-domain pre-cooked query (in this specific case: “show all tourist facilities within a naturalistic area and the related involved municipality).

Figure 7. Google Maps-based module showing the location of a resource basing on geo-spatial information structured by RDF.

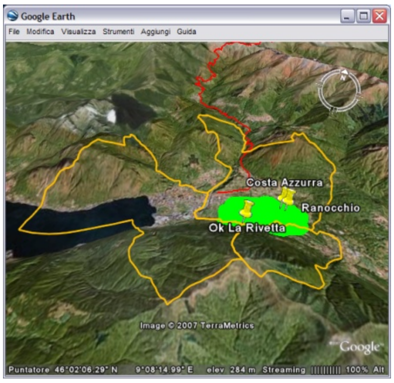


Figure 8. Real-time interaction with Google Earth which shows the results of a cross domain pre-cooked query.

5.7. Social Model

Brainlets, by providing an aggregation medium for ontologies, users and data representation structures, are therefore good catalyst of overall semantic interoperability process. As users gather around popular Brainlets for their topic of choice, the respective suggested ontologies and data representation practice will form an increasingly important reality. If someone decided to create a new Brainlet or Semantic Web application in general which could target the same user group as the said popular Brainlet, there would be an evident incentive in using compatible data structures and ontologies.

6.0. Related Works and Conclusions

Tourism is one of the world's largest economic sectors. Sustainable tourism development can be a powerful tool for economic growth, natural and cultural resources conservation. While tourism represents an important development opportunity for many countries and communities, it can also have very negative impacts, such as disrupting and harming the socio-cultural authenticity of local communities, and threatening natural and cultural heritage.

Communication has a huge role in supporting sustainable tourism development. Communication can create and facilitate a system that allows stakeholders to exchange opinions and arrive in a rational planning at consensual solutions. Effective use of communication tools can also link products to markets, and can contribute to visitors' safe and positive experiences (Communication and Sustainable tourism, 2006).

In general, the advantages that ontologies provide for geographic information processing include the enhancement of communication, systems engineering, and interoperability (F. Reitsma, K. Hiaramatsu, 2006). Many approaches to the visualization of semantically annotated data have been proposed in literature. RDFGravity exploits graphs as a visualization tool of RDF triples (see Sunil Goyal, Rupert Westenthaler). In E.Pietriga 2002, an environment for both RDF browsing and authoring is discussed. Welkin (SIMILE a) is an RDF visualizer based on elliptical zooming of the connections among resource, while Longwell (SIMILE b), implements the idea of faceted browsing of RDF data. MIT Haystack (Quan, Dennis and Karger, 2004) is a tool for development of Semantic Web applications, also focuses on interface organization (layout and functionalities).

While pro and cons can be argued for each specific approach, it is clear that user interface issues are complex with no clear single solution. DBin responds to this by the Brainlet UI interface, which enables a topic specific "mash-up" of different visualization paradigms.

DBin stands out as an end user application which provides Semantic Web capabilities and a plug-in based open source, rich client approach.

The combination of these two, effectively enables efficient extension and integration into domain specific tools and ontology based browsing, querying and searching information. The addition of proper geographical annotations handling components enables "Geo Semantic Web Communities", which can be quickly started by domain experts.

Especially when compared to full featured Web GIS (as the one in Mahesh Rao et al, 2006) or specific software, the support for geo-spatial use cases is today still limited to relatively simple objects, in particular to those that have a well specified location, that is, those that can be approximated with a point.

The purpose of this work, however, it is not that of matching the state of the art specific tools. The potential of Semantic Web based annotation communities seems unprecedented in terms of flexibility and for the ease of integration of information across communities. This is what the project seeks to explore and turns to be particular important when information is useful to support decision in whole landscape resources participatory planning (Marcheggiani et al. 2007).

It is clear how the full potential of Geo Semantic Web Communities and of the paradigm in general is to be explored and validated. This is however the case for the whole Semantic Web initiative and with respect to this, the contribution of the DBin platform and the applications that have been shown here is to represent, arguably, the most tangible instantiation of such technologies available for actual use today.

DBin is distributed under the Gnu Public License and the default distributions includes all the geographic visualization and editing plug-ins that have been used for the screen-shots. Further documentation and compiled executables can be downloaded at http://dbin.org, where it is also possible to find a few minutes screen demo which shows the whole DBin paradigm and platform.

References

Chris Goad, 2004. RDFGeom, http://www.mapbureau.com/rdfgeom1.1/index.html, September.

Communication and Sustainable tourism, 2006. Proceedings of the global e-conference and summer speaker series on the role of development communication in sustainable tourism, May 29 - June 9, 2006, on www.devcomm-co.

E.Pietriga, 2002. Isaviz: a visual environment for browsing and authoring RDF models, Proceedings of the 11th International World Wide Web Conference, Honolulu, Hawaii, 2002.

G.Tummarello, C.Morbidoni, J.Petersson, P.Puliti, F. Piazza, 2004. RDFGrowth, a P2P annotation exchange algorithm for scalable Semantic Web applications, P2PKM Workshop, Boston, 2004.

G.Tummarello, C.Morbidoni, M.Nucci, 2006. Enabling Semantic Web communities with DBin: an overview, Proceedings of the International Semantic Web Conference, Athens GA, USA, 2006.

G.Tummarello, C.Morbidoni, P.Puliti, F.Piazza, 2005. Signing individual fragments of an RDF graph, World Wide Web Conference, Chiba, Japan, 2005.

F. Reitsma, K. Hiaramatsu, 2006. Exploring GeoMarkup on Sematic Web", 9th AGILE Conference on Geographic Information Science, Visegrad, Hungary, 2006.

Jeen Broekstra, Arjohn Kampman, 2004. SeRQL: An RDF Query and Transformation Language, 2004 ISWC 2004, Hiroshima, Japan.

Mahesh Rao, Guoliang Fan, Johnson Thomas, Ginto Cherian, Varun Chudiwale, Muheeb Awawdeh , 2006. A web-based GIS Decision Support System for managing and planning USDA's Conservation Reserve Program (CRP), Environmental Modeling & Software, 2006.

Marcheggiani E., Bocci, M., Colantonio, R. Galli, A., 2007. Syntethic Indicators for an Integrated Landscape Planning Model, in Methods and Instruments for the Scientifically Sound Use of Landscape and Socio-Ecological Indicators in Planning, Landscape Online Ed. (landscapeonline.de), in press.

Min Cai, Martin Frank , 2004. RDFPeers: A Scalable Distributed RDF Repository based on A Structured Peer-to-Peer Network, Proceedings of the World Wide Web Conference, New York, USA, 2004.

Paul Alexandru Chirita, Stratos Idreos, Manolis Koubarakis, Wolfgang Nejdl, 2004. Publish/Subscribe for RDF-based P2P Networks, Proceedings of The International Semantic Web Conference, Heraklian, Greece, 2004.

Quan, Dennis and Karger, David R , 2004. How to Make a Semantic Web Browser, Proceedings International WWW Conference, New York, USA, 2004.

RDF Semantics, 2004. W3C Recommendation, http://www.w3.org/TR/rdf-mt/.

SIMILE a. Welkin, http://simile.mit.edu/welkin/.

SIMILE b. Longwell, http://simile.mit.edu/longwell/.

Sunil Goyal, Rupert Westenthaler. RDF Gravity - RDF Graph Visualization Tool, http://semweb.salzburgresearch.at/apps/rdf-gravity/index.html.

W3C Consortium, 2004a. RDF Vocabolary Description Language: RDF Schema, W3C Recommendation, 10 Febrary 2004.

W3C Consortium, 2004b. OWL Web Ontology Language Overview", W3C Recommendation, 2004, http://www.w3.org/TR/owl-features/.

W3C Semantic Web Interest Group, 2003. Basic Geo (WGS84 lat/long) Vocabulary, http://www.w3.org/2003/01/geo/.

Wolfgang Nejdl, Boris Wolf, 2002. EDUTELLA: A P2P Networking Infrastructure Based on RDF, Proceedings of the World Wide Web Conference, Honolulu, Hawaii, 2002.

Wolfgang Nejdl, Wolf Siberski, Martin Wolpers, Alexander Löser, Ingo Bruckhorst, 2003. SuperPeer Based Routing and Clustering Strategies for RDF Based Peer-To-Peer Networks, Proceedings of the Twelfth World Wide Web Conference, Budapest, Hungary, 2003.

Exploring Ontologies of Historic Landscape Characterisation: Towards an approach for recognising the impact of incremental change to historic legibility in urban areas

Stephen Dobson
Department of Landscape, University of Sheffield,
Arts Tower, Western Bank, Sheffield, S10 2TN, arq06sd@shef.ac.uk
Tel: 07876 247235
Fax: +44 (0)114 275 4176

Abstract. Historic Landscape Characterisation is a GIS-based mapping approach for documenting the character, time-depth and degree of historic visibility present in the modern landscape. It is an approach that was pioneered by English Heritage in the mid-1990s and is currently over halfway through completing full coverage of the English landscape. Philosophically its aims are in accordance with those expressed by the European Landscape Convention (the Florence Convention 2000) as the programme of characterisation recognises that all landscape has history and the fabric of the past is interwoven with every aspect of the present. Obviously the degree to which this is visible or 'legible' varies greatly from place to place and the depth of time that is represented may be relatively short, but the subtle influence of underlying spatial patterning and land-use still can have a significant influence on the character of the environment that we live in today.

If we are to consider that all landscape has historic character and that the legibility of the past is an important resource, to be respected throughout the process of modern change and new development, then we need to ensure that planners, developers and architects have the necessary tools to treat historic legibility with sensitivity. Historic Landscape Characterisation is a mapping method that has been influenced by other characterisation practice common in Europe, such as Landscape Character Assessment and is designed to be complimentary. Information regarding the presence and visibility of historic character in all parts of the landscape provides an important tool and, whilst this has been developed over many years for rural areas, its application in urban environments is relatively new. There is also a relative lack of critical research regarding its use in anticipatory planning and the design process and very little technical research into its full potential as a modelling tool in GIS.

This paper aims to theoretically explore an ontological framework for the attribute of legibility in HLC and proposes possibilities for creating a 'legibility matrix' across urban areas for use in planning and development. The traditional 'buffer zone' approach for protecting the overall context surrounding 'special' historic remains is less applicable when considering the general concept of legibility since this is intrinsically part of all modern surroundings. By exploring the relationships between historic features and their legibility, it may be possible to create a GIS-based tool to illustrate impact on the matrix as a whole if one or two of its parts were to be removed through development.

Emerging archaeological ontologies

The field of archaeology has regularly employed various taxonomies to conceptualise the relationships of artefacts, contexts, features, phases, and sites within the broader landscape. Such definitions are used both in the recording process and in post-excavation interpretation and are inevitably defined by a variety of attributes that are established to best meet analytical requirements. For those archaeologists working with data from complex urban sites or when there is an increased need for data interoperability, formal ontologies are emerging as a means to tackle these multifaceted models and interrelationships. With regard to field archaeology, the University of York has initiated a project entitled, 'New approaches to post-excavation on complex sites' (NAPEX) which was defined to address the challenges of urban deposit modelling and to build a knowledge-base from a variety of specialists. This project aimed to; "encourage approaches in which the various classes of archaeological data – stratigraphic, architectural, artefactual, ecofactual, sedimentological etc – can be better integrated." (Perring, Roskams and Vince 2001). Chunxia Zhang et al (2002) explore the structure of an archaeological domain-specific ontology more fully and present an important model for the sharing of knowledge through these means. Here, the conceptualisation of the world is based on the need to understand and reconstruct past events through their material remains and subsequent deposition. Categories are related in this ontology through the relationships of 'is-spec' and 'part-of'. These define whether a category is specified by another and/or which category it is part of. The 'part-of' relationship class is outlined as an example in figure 1.

Figure 1: 'Part-of' Relationship Classes (Chunxia Zhang et al 2002, 215)

1. Integral-object/Componant. For example, "The wall is part if a carcass".
2. Area/Place. For example, "A tomb is part of a graveyard".
3. Activity/Feature. For example, "The design is part of a piece if pottery".
4. Collection/Member. For example, "An artifact is part of a complete set of artefacts".
5. Mass/Portion. "A block of accumulation is part of an accumulation layer"

As defined by the three interpretive axioms used in this project, the domain-specific ontology presented here aims to provide a "formal conceptual specification and taxonomic structure of archaeological objects" (Chunxia Zhang et al 2002, 215). Whilst an archaeological ontology is particularly important for furthering our understanding of the past in complex urban landscapes, this paper aims to examine which categories and relationships may enable other domains to interrogate the data to aid decision-making regarding the future. 'Future-oriented' (Rippon 2004) stakeholders (planners, architects, designers) have the responsibility not only to plan and design landscapes of the future but to do so sympathetically and with consideration to our inherited historic landscape. Historic Landscape Characterisation is explored here as a possible means of providing additional categories and relationships so as to support axioms of impact assessment and resource management.

The problem with creating tools that help define capacity for change is that value judgements will inevitably be required, with regard to the historic environment, in order to form the basis for decision-making. This need for a system of ranking will often favour the oldest, rarest, most important and 'the best' which obviously results in specific sites, monuments or remains being favoured over others. This process is at risk of devaluing the wider historic landscape and subsequently provides one of the main drivers for change present in the landscape philosophy of the European Landscape Convention and in the programme of Historic Landscape Characterisation. This leads to the questions of; 'how might the holistic treatment of historic environments work in practical terms?' and 'can a domain-specific ontology for archaeology be linked to and therefore used by future-oriented stakeholders assessing the contemporary past?'

The contemporary past

Acts of recognition and understanding create two parallel contexts for historic remains. The first is relict and regards its original place in time and its meaning to the society that created it. The second is as an appropriated, contemporary existence that now belongs to modern society and thus has new and alternative (or additional) meaning.

> "Material culture of past societies which, by virtue of being appropriated by archaeologists in the act of discovery, becomes their material culture too" (Edgeworth 2003, 108)

In this case, whilst landscape features will have their various origins in the past, they are now appropriated by us as modern, contemporary features that may take on new meaning through their existence in a modern setting. Architectural features, spatial alignments of streets and boundaries will all point to the land use and architectural decisions of the past and it is the interplay between these and more recent architectural and spatial forms that can define the historic character of a city.

Historic Landscape Characterisation

Historic Landscape Characterisation (HLC) is an approach that aims to record the past in its contemporary setting and to document historic landscape features as parts of the modern environment. Characterisation as a general tool for managing change in the landscape was originally derived from Conservation Area Legislation (Clark, Darlington and Fairclough 2004) which came into being from the Civic Amenities Act of 1967. This was originally introduced as listed building legislation was deemed inadequate for the protection of character in areas of urban development. A conservation area in this respect was defined as; "an area of special architectural or historic interest, the character or appearance of which it is desirable to preserve or enhance." (*Town and Country Planning (Listed Buildings and Conservation Areas) Act 1990)* 'Character' is referred to here as distinct from 'appearance' and was therefore not considered to be equivalent. Character, as defined in this act, is a concept or a perception that contains more or alternative meanings than 'appearance'.

> "The special character of these areas does not come from the quality of their buildings alone. The historic layout of roads, paths and boundaries; characteristic building and paving materials; a particular 'mix' of building uses; public and private spaces, such as gardens, parks and greens; and trees and street furniture, which contribute to particular views - all these and more make up the familiar local scene."
> (EnglishHeritage:http://www.english-heritage.org.uk)

The assessment of character has therefore been central to the management of change in the urban and rural landscapes from the introduction of this Act until present and has applied a well developed set of tools – the most significant being Impact Assessment, Landscape Character Assessment and, more recently, Historic Landscape Characterisation.

HLC as a practical application was initiated in Cornwall in the mid-1990s and since then there have been around 30 projects dedicated to its original aims and carried-out by a variety of county-based teams within a flexible remit for implementation. Initially, one of the main aims was the achievement of an integrated approach to conservation and environmental planning but, as HLC becomes adopted by an increasing variety of stakeholders, perhaps these aims are still being formed and also in-formed both by current practice and past experience.

In the early 1990's the British Government White Paper "Common Inheritance" suggested that best practice might include the identification of 'special' areas for inclusion on a national selective register however, emerging perspectives on sustainable development lead English Heritage to adopt an emphasis on the 'character' of landscapes rather than merely highlighting its 'best bits'. The key premise for the development of HLC is in the changing nature of our environment and in particular the extent of this change that is visible in the landscape. Earlier systems that were primarily based on designation and development control failed to consider the whole environment equally in favour of a protectionist philosophy highlighting the importance of restricting change in 'special areas'. Since the whole environment can contain time-depth legibility (that is our ability to 'read' the past in our contemporary landscape), we are immersed in an environment that contains important information regarding the various stages that make up our palimpsest landscape. In terms of historic character, sustainable growth and development can only be achieved if these considerations are formally embraced in our methods and practice. In this case, the adoption of a broader definition of 'historic landscapes' leads us to require a more holistic set of tools and perspectives that can enable us to appreciate not only the importance of all elements equally, but also the means of their creation – the diversity of human agency. When we experience the remains of the past we are primarily looking at the products of change and so it is important that we continue to facilitate change but in a sustainable manner that neither fossilises context-less 'special' places nor erases heritage all together simply because it fails to match a particular set of criteria based upon levels of national importance.

Legibility then can provide a temporal anchor and sense of transition without which we would effectively create a continual present that is forward looking but has no link to the past. This is not to value nostalgia but is a much more functional perspective that recognises the need for temporal orientation and respect for the historic environment as a resource. In this sense, the presence of historic legibility is a characteristic in the landscape that can be valued above its absence and so provides us with the logical basis for assessing capacity for change. The preservation of legibility will require further categories and relationships to be added to the archaeological ontology presented by Chunxia Zhang et al and will focus on describing how time-depth is *perceived* and so what are the characteristics of spatial forms that enable the urban landscape to be 'read' as to its varied historic origins. By mapping such relationships ontologically it is feasible that the removal of one relationship will alter the state of others and so highlight where landscape change may affect our continued ability read the past.

A conceptual influence on how this may be achieved is the early, experimental use of Time-Depth Matrices which were explored in the pilot HLC project conducted in Cornwall (Herring 1999). Matrices were used to here to illustrate the primary form of evidence that may be expected when regarding various archaeological periods within the landscape. The attribute of legibility in this case is supported by the expected form that it would take, such as; extant landscape features, subsurface archaeology, circumstantial evidence, documentary evidence or palaeo-environmental (figure 2). Matrices were not openly adopted by all subsequent HLC projects but stand as a generic and potentially useful tool for the non-specialist stakeholder when assessing the nature of legibility within particular historic character types. By exploring the concept of time-depth matrices further, but increasing their scope to map the way in which legibility can be determined for a specific (rather than generic) historic feature then we potentially create a tool for urban designers to assess the historic impact of their actions on the 'readability' of the past in the present.

Returning to the work by Chunxia Zhang et al, an additional 'Legibility' category could be related to 'Archaeological-Culture' through the 'part-of' relationship class of: 'Component' (figure 3). 'Collection/Member' ('part-of') related categories for 'Legibility' then define how legibility is constructed. For example:

1. *Spatial* – Legibility of the past through form rather than fabric (ie modern road alignments)
2. *Material* – Legibility through recognition and interpretation of fabric
3. *Contiguity* – Legibility based on the close spatial proximity of evidence

4. *Fragmented* – Legibility based on spatially dispersed 'pockets' of evidence
5. *Plan* – Legibility detected predominantly through plan or aerial views
6. *Perspective* – Legibility detected predominantly at 'eye-level' from within the streetscape

The categories above aim to demonstrate how archaeological and historic remains have modern perceptual characteristics as well as those defined by their usage, production purpose and past meaning.

Case Study: The 'Crofts', Sheffield (UK)

The 'Crofts' is an area close to the heart of the city centre of Sheffield. As the industrialisation of Sheffield increased over the last 200 years, the 'Crofts', as for many other areas of the city, became home to worker's housing, steelworks, furnaces and various cutlery and tool manufacturing works (figure 4). In 2004, Archaeological Research and Consultancy at the University of Sheffield (ARCUS), conducted an archaeological desk-based assessment prior to development (ARCUS 2004). The report observed that historically, the area was part of the 18th century expansion of the city and now consists mainly of a few 20th century works buildings but gained its name from the medieval field enclosures predating its more recent development. Whilst its industrial past is of obvious importance in furthering our understanding of the city's evolution, the full significance of the area perhaps lays in the fact that earlier phases of history have become fossilised in the street layout of the Crofts and thus presents us with a valuable historic perspective and depth of time that is not immediately obvious from archaeological material alone.

In the medieval period, this part of the city was the site of a large open field called the 'Town Field' which had been enclosed by the mid 17th century. The small enclosures or 'Crofts' are displayed in maps from both Harrison (1637) and Gosling (1736) and include names such as White Croft, Pea Croft, and Holis Croft. Whilst no archaeological material evidence remains of the medieval period due to subsequent industrialisation (ARCUS 2007), the shape and form of enclosures are still present in the modern street layout with individual streets still bearing the names of the original crofts they outline. Figure 5 shows the 'Crofts' in 1945 and whilst the landscape is one of commercial and small-scale industry, the alignment of buildings and streets around Hollis Croft, White Croft, Bakers Lane and Solly Street (originally Pea Croft) all point to a medieval land use that is not widely appreciated but is arguably an extremely valuable asset that could help frame future regeneration strategies.

> "Most people in Sheffield today are largely unaware of the 'Crofts' as an area with distinct identity. It is rarely a destination in itself, and it is bypassed by the arteries of the modern city. Instead, the place is defined by an older identity as a slum – an identity which has been used to characterize the 'Crofts' since the 19th century." Belford (2004, 166)

Whilst the archaeology of this area is predominantly 18th to 20th century, historic legibility includes a much earlier agricultural phase. Therefore, if we are to consider the ‘readability’ of the past as an important resource we must protect it through additional means to the more traditional and ‘fabric-based’ evaluation that is often applied to archaeological and architectural features. In the example of the ‘Crofts’, preservation of historic legibility was achieved, albeit somewhat serendipitously, through development that followed a pre-existing grain or footprint of the past. By employing the suggested ‘part-of’ categories presented in this paper with a more traditional domain-specific ontology, we might begin to model the process of interpreting the past in the present. For example, the legibility of these medieval croft enclosures could be described as being reliant on a combination of ‘contiguous’, ‘spatial’ forms that are experienced in ‘plan’. The ‘perspective’ category is rejected since modern buildings hinder our reading of time-depth at the street-level as is ‘material’ since the alignment of boundaries are retained only by modern fabric rather than medieval. In this simple example, building height and style might be assumed to have less impact on the preservation of medieval legibility of the ‘Crofts’ than changes to street layout or the renaming of streets.

Summary

To the urban designer, the knowledge that an ancient field boundary, for example, is visible through a combination of streets and walls is vitally important to the assessment of capacity for change. The boundary may not physically remain, but its *form* may still be present in the alignment and configuration of modern fabric. In this case it may be perfectly acceptable to change the fabric of the streets and walled boundaries, but arguably, it would be beneficial to the historic character of the area to preserve the legibility of the past by retaining their spatial alignment. Building heights, for example, may also be influenced by the extent to which historic legibility relies on perspective or ‘street-level’ perception rather then aerial or plan views. A formal ontology is proposed here as a method for structuring and managing the inherent complexity associated with the various means by which we perceive the past in our contemporary surroundings. This is presented as a tool to aid the preservation of our very means of reading the past in a living, urban landscape. By linking legibility categories to an existing HLC GIS, the proposed ontology would theoretically support an ability to search for areas where the modes of perceived historic legibility might be compatible with the type of planned change. By examining the ‘Crofts’ we may appreciate that change brought about by industrial expansion was ultimately responsible for preserving the footprint of a medieval past and therefore helps retain a region’s ‘lines of life’ (Cullen 2006, 111). It is hoped that this paper provides some indicators as to how the process of archaeological and historic interpretation may be linked to existing characterisation map resources for assessing not only the impact of change but also to help influence its design.

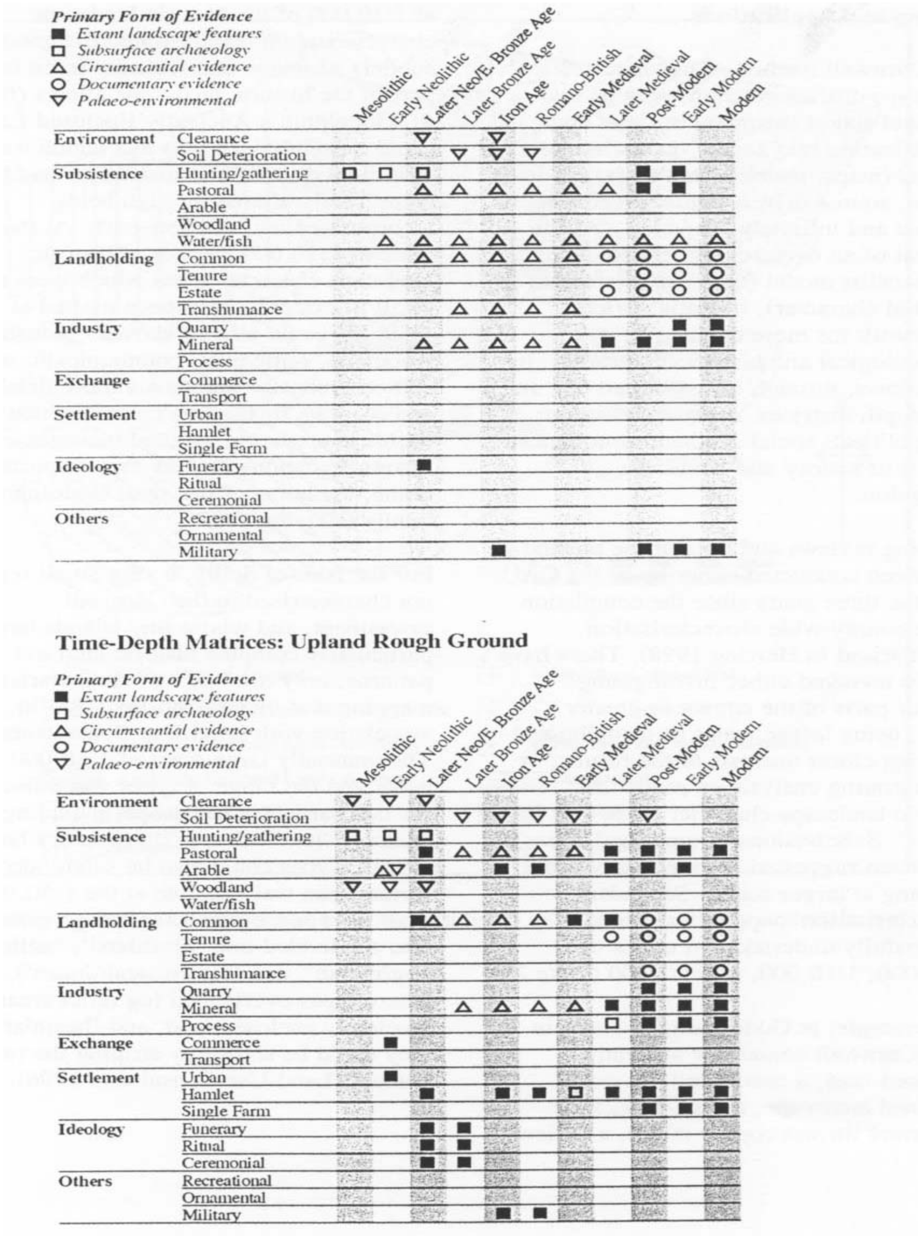


Figure 2: Cornwall Historic Landscape Characterisation Time-Depth Matrices (Herring 1999)

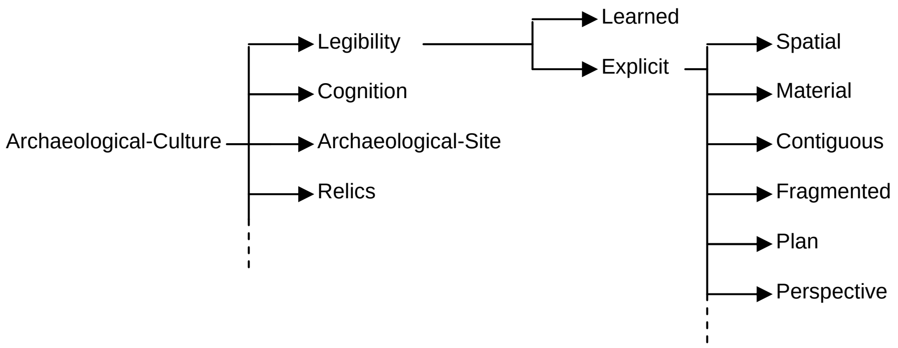


Figure 3: 'Legibility' in the Backbone Structure of Archaeological Categories. Based on: Chunxia Zhang et al (2002, 219)

Figure 4: A typical street-scene of The 'Crofts' in the 19th century
(Source: Sheffield Local Studies Library)

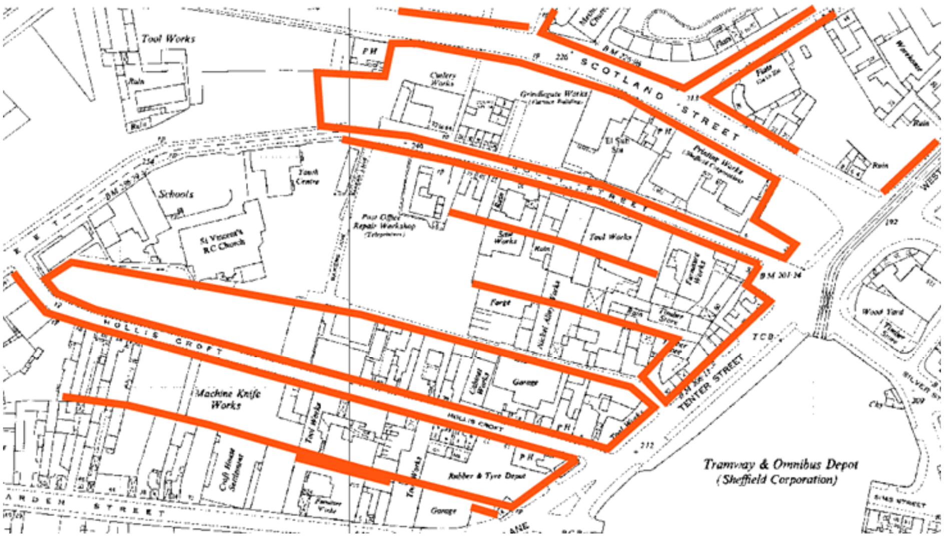


Figure 5: The 'Crofts' in 1945 with 1736 croft enclosures overlay (Source: Ordnance Survey County Map Series 1:2500 (1945) Landmark Historic Map Data, University of Edinburgh)

References

ARCUS, 2007 *Archaeological Evaluation and Mitigation of Land Off Tenter Street, Sheffield Projects 814e, f & g* ARCUS Assessment Report 814g.2 August 2007

ARCUS, 2004 *Archaeological Desk-Based Assessment of Land at Tenter Street, Sheffield, South Yorkshire* ARCUS Project Report 814.1 May 2004

Belford, P. 2004 'Urban Industrial Landscapes: Problems of Perception and Protection' in (eds) D. Barker and D. Cranstone *The Archaeology of Industrialization: Papers given at the Archaeology of Industrialization Conference, October 1999* Maney 2004

Chunxia Zhang, Cungen Cao, Fang Gu, Jinxin Si 2002 *A Domain-Specific Formal Ontology for Archaeological Knowledge Sharing and Reusing* in 'Practical Aspects of Knowledge Management: 4th International Conference, PAKM 2002 Vienna, Austria, December 2-3, 2002 Proceedings' Lecture Notes in Computer Science Vol 2569/2002, Springer, Berlin pp213-225

Clark, J., Darlington, J., & Fairclough, G. 2004, *Using Historic Landscape Characterisation*, English Heritage and Lancashire County Council, Preston

Cullen, G 2006 *The Concise Townscape* London: Architectural Press

Edgeworth, M 2003 *Acts of Discovery: An Ethnography of Archaeological Practice* BAR International Series 1131

Herring, P. 1999 *Cornwall: How the Historic Landscape Characterisation methodology was developed* in (ed) G. Fairclough 'Historic Landscape Characterisation "The State of the Art"' English Heritage, London

Perring D, Roskams S & Vince A 2001 *New approaches to post-excavation on complex sites: the Project Design* (http://www.york.ac.uk/depts/arch/napex/napex_intro.htm)

Rippon, S. 2004 *Historic Landscape Analysis: Deciphering the countryside* Council for British Archaeology

Using an Ontology-based Model for Knowledge Representation in Rural Landscape *

Antonia Cataldo[1] and Antonio M. Rinaldi[2]

[1] Dipartimento di Pianificazione e Scienza del Territorio - Università di Napoli Federico II, 80125 Piazzale Tecchio, 80, Napoli, Italy
[2] Dipartimento di Informatica e Sistemistica - Università di Napoli Federico II 80125 Via Claudio, 21, Napoli, Italy
{cataldo,amrinald}@unina.it

Abstract. In the field of the territorial planning the problems related to the management of information increase everyday. This process needs of more and more effective and efficient Knowledge Management techniques. Such techniques lead to a suitable organization (and therefore to a more detailed analysis) of the phenomena, which in turn suggests how to choose the actions to perform. At the same time, the interoperability between heterogeneous subjects at different levels of the planning process and belonging to multiple disciplinary fields becomes an unavoidable element. Information represents a strategic resource for the subjects in charge of territory planning: it is in fact raw material, job tool and final product. In a complex research scenario such as the territorial science, some fundamental elements, useful to analyze and resolve several problems, can be represented using a framework based on an ontological approach. Amongst the different territorial planning fields, we chose as case study the issue of landscaping, in particular the one related to rural landscapes. This choice is motivated by the cultural and scientific innovations, which led in recent years to a completely new interpretation of landscape, as a good to be preserved (being it an expression of the ecological and social transformations on the territory resulting from the community activity). The paper aim is to resolve conceptual misunderstanding and semantic ambiguity and, on the other hand, we need a precise and accurate description of our knowledge, i.e. of the terminologies representing our concepts, which influence goals and entail actions to be implemented. Due to those considerations we propose a suitable ontology-based model implemented in a system to manage ontologies.

Keywords: Rural Landscape, Knowledge Representation Formalisms and Methods, Ontologies, Semantic Networks, OWL, WordNet

1 Introduction

The problem of defining new methodologies for Knowledge Representation (KR) has a great influence on information technology and, from a general point of view,

* Introduction and conclusions are by both authors; sections 2 and 5 are by Antonia Cataldo; sections 3 and 4 are by Antonio M. Rinaldi.

on cognitive sciences. In particular, the wide spreading of communication technology forces the modern information systems to manage large amount of data. New techniques have been developed to solve those problems. Some of them are based on ontologies to reduce conceptual or terminological mess and to have a common view of the same information. The ontological aspects of information are intrinsically independent from information representation, so the information itself may be isolated, recovered, organized and integrated with respect to its content. A formal definition of ontology is proposed in [1] according to which *"an ontology is a formal and explicit specification of a shared conceptualization"*; *conceptualization* refers to an abstract model of a specific reality in which the component concepts are identified; *explicit* means that the type of the used concepts and the constraints on them are well defined; *formal* refers to the ontology propriety of being "machine-readable"; *shared* refers to the fact that an ontology captures the consensual knowledge, accepted by a group of persons. We also consider other definitions of ontology: in [2] *"an ontology defines the basic terms and relations comprising the vocabulary of a topic area, as well as the rules for combining terms and relations to define extensions to the vocabulary"*. This definition indicates the way to proceed in order to construct an ontology: i) identification of the basic terms and their relations; ii) agreeing on the rules to arrange them; iii) definition of terms and relations between concepts. From this perspective, an ontology includes not only the terms that are explicitly defined in it, but also those one that can be derived using defined rules and properties. Thus an ontology can be seen as a set of "terms" and "relations" among them, denoting the concepts that are used in a specific domain. In this paper we investigate on the use of ontology in the territorial planning fields using a suitable ontology-based model implemented in a system to manage ontologies. The landscape is a central aspect in our approach for the interesting discussions about this theme in the research community and for its intrinsic complexity in the territorial planning field. For these reasons it is an ideal candidate for a formalization using ontologies. Amongst the different in this field of interest we chose the one related to rural landscapes. The cultural and scientific innovations led in recent years to a completely new interpretation of landscape, as a good to be preserved (being it an expression of the ecological and social transformations on the territory resulting from the community activity). From this point of view, the landscape has the distinguishing peculiarity of not being characterized by just one category of elements (physical, natural, historical,...), rather representing the totality of them. Amongst the wide-ranging Italian landscape conformations, the rural one not only represents the most important component in terms of surface, but also a system of great importance both from an environmental and from a cultural and architectonic point of view. It indeed represents a basic joint element between the human activity and the environmental system, where human abilities and skills represent the continuous research of a global eco-systemic balance. Once understood that the process of understanding the landscape is priority and propaedeutic for the definition of actions for its preservation, management, and planning, we need to individuate those homogeneous territorial contexts, which contain highly related and characterizing factors. To this aim we must try to resolve conceptual misunderstanding and semantic ambiguity and, on the other hand, we need a precise and accurate description of our knowledge, i.e. of the terminologies representing our concepts, which influence goals and entail actions to be implemented. A global and shared planning, by means of complementary factors and synergies (i.e. shared knowledge), allows the territorial entity to increase its own auto-organizing skill, in order to evolve.

The paper is organized as follows: in section 2 we describe our choice about the concept of landscape; in section 3 we define and describe our model using OWL [3] a standard language to describe ontologies; in section 4 is described the system architecture and the approach used to extract information from a general knowledge base to arrange ontologies; a use case of the proposed model and strategy is in section 5; conclusion and future works are in section 6.

2 The concept of landscape as sorroundings

The recent cultural innovation about the landscape topic shows that the process of landscape identification must be supported with the central and active role of asset (landscape itself) which must be defined and protected regarding to its general context. After the European Landscape Convention [4], undersigned after a long and hard discussion between state members, the landscape planning action based on passive preservation has come to an end. Pursuing a suitable landscape planning action, the public authorities should define a set of *"general principles, strategies and guidelines that permit the taking of specific measures aimed at the protection, management and planning of landscape"*; from this assumption the landscape management is the basic element *"to ensure the regular upkeep of a landscape, so as to guide and harmonize changes which are brought about by social, economic and environmental processes"* [4]. Since the *"landscape means an area, as perceived by people, whose character is the result of the action and interaction of natural an/or human factors"* [4], we argue that:

- it is very important to understand how every inhabitant perceives the landscape;
- the landscape key features come from natural and/or cultural factors, because the landscape evolves during the time from natural forces and/or from human actions;
- the landscape is a whole of natural and cultural elements, which must be simultaneously considered and completely related.

Once observed that the landscape is a key element both for the individual and social well-being and, for this reason, aiming at quality goals implies a detailed definition of the characteristics that local people expects for its life environment, once it has been identified and described in its general characters (aesthetical, perceptive and eco-systemic ones). A better landscape quality corresponds to a more effective social organization, assures the individual and collective wellbeing, increases the places' capacity of attracting investments by developing their territorial competitiveness. Hence managing the landscape means considering the landscape as an entity related both to the economical and social needs of the population and to natural processes. The landscape becomes the output of an ongoing process in which many entities are involved. The approach to landscape planning is renewed by the characterization and the definition of *landscape values*. These considerations are taken into account in several national and international laws and directives defining various perspectives to the analysis of the landscape. For example in Italy, the "Accordo tra il Ministero per i Beni e le Attivit Culturali e le Regioni e le Province autonome di Trento e Bolzano sull'esercizio dei poteri in materia di paesaggio" (published on the Italian Official Gazette of 19.04.2001) invites the Regions to start management policies that take into consideration, apart from landscape recognized as owning exceptional values, also the entire territory. This directive permits the application, for each part of the territory and in according to the landscape's features and values, of suitable preservation tools and innovation forms for requalification which can create a diffused landscape quality [5]. To this aim, the planning tools must be provided with a "cognitive, prescriptive and pro-active content", and the procedures of preservation, enhancement and requalification must be set on the basis of the "level of integrity and significance of landscape values". Therefore we consider fundamental and relevant the phases of landscape learning, assessment of its conservation state and definition of the quality of its values [6]. As required by the cited Agreement, the following "Codice dei beni culturali e del paesaggio" completes a regulatory landscape coding which, at least in its principle enunciations, seems one of the most innovative in Europe. This new Code introduces in Italy an innovative definition of the landscape concept compared with the previous laws; for the first time, it highlights the centrality of the anthropized, built and managed landscape. These considerations clearly show the extreme complexity of the concept of landscape, which is to be seen as the whole of several aspects (natural, anthropo-cultural and perceptive). The landscape recognizing is obtained not only by means of elements identification (climatic, environmental, physical, historical), but also identifying the spatial and temporal relations between them. *"The landscape recognition is obtained by the interaction between physical-biologic elements and human activities, considered as part of the environmental historical definition process and it can be defined as the complex combination of objects and phenomena related between them by mutual functional relations, as well as locations, in order to be a whole"* [7]. The ability of clearly recognizing the constituent features of a landscape represents one of the most important factors from the observers. A landscape poor of "recognizable" elements or characterized by a messy overlapping of several components (natural and/or anthropological) or from an excessive fragmentation of the same ones, could "disturb" an immediate perception; it is usually considered a landscape poor "of values". On the other hand, a landscape in which the complexity is clearly structured and noted comes positively appreciated. The value of a landscape essentially is in the possibility that its structure and identity are clearly read.

2.1 Methodologies for landscape reading and rendering

With the increasing attention about the landscape, the questions related to evaluation and modalities of reading the specificities of territorial patrimony are assuming more and more importance, because it is not simple to assign precise rules to outline the landscape. The knowledge discovery is the first, basic and necessary phase for every preservation policy, innovation and requalification of the landscape; only through a suitable cognition of landscape values it is possible to manage themselves changes. The different ways of reading the landscape have always been subordinate, till now, to the planning requirements; they emphasized of just an aspect of the landscape respect to another one. The Italian experiences in the '70 and the '80 are a meaningful example: when, as a result of the institution of parks and natural reserves, for pursuing purpose of protection and government of nature, environmental were more preserved than the cultural and historical ones. n an analogous way in several European Countries the formalization and specification of landscape components have different roots and characteristics: for example in north Europe, the attention is put on natural and ecological problems; in others countries other aspects, as the aesthetical and architectonical ones, are meaningfully considered; in other ones economic, productive, and recreational elements are taken more in account.

In [8] the author argue that from the international experiences of landscape reading we can define six fundamental typologies:

- *Reading through atlases and lists:* it is used in the large scale in order to characterize and circumscribe different landscape units; in this way the territory is subdivided in homogeneous and adjacent areas. This experience is used in several European Countries at a national level (Slovenia, Norway, France and England) or for regional initiative or local administrations (as the preliminary or complementary regional studies for Regional Territorial Plans in Italy) or, finally, from research agencies (Italy, France).
- *Reading through natural features and environment:* this typology is related to the disciplines of earth and biological sciences. These, considering abiotic and biotic factors, examine the natural and environmental aspects of life places, paying particular attention to the systemic relations between living organisms and environment. The landscape ecology scope is to know natural assets features relate to anthropological dynamics, in order to have an ecosystem balance. Therefore the landscape is the result of an organic and complex ecosystem, comprehensive of people and their activities. We can found this approach in Germany in '70 years, but it has been diffused in other European Countries (i.e. Austria). The environmental approaches also consider how the user perception regards to nature: the botanical interest in some species of trees or cultivations for their high value of biodiversity, it is often related to an historical and cultural interest hidden by the species themselves. It is the case of terracing, historic orchards or vineyards and old tracks. From this point of view we argue that the quality of places is not only tied to the quality of the environment, but also to its historical and cultural features (in Germany, in '80 years, it is the Kulturalandschaft).

– *Reading through visual perception:* it is used to comprehend the morphologic characteristics and design of considered places and, on the other hand, to know the physical persistence and the historical and cultural meanings. It is a methodology used in many European Countries (Great Britain, France, Holland, Spain, Poland), it sources from the American experience, which considers the landscape as a discipline.
– *Reading through historical and cultural persistences:* its aim is to give a punctual knowledge of the historical patrimony (it is a methodology diffused at an international level during '60 as result of a series of Charters and Dispositions); it has found its formalization with the developing of cultural assets lists and recently with the analysis of historical systems. This last methodology (mainly used in Great Britain) is used to find relations (physical, functional, visual) between system parts.
– *Reading through symbolic meanings:* it is used to describe places that are valuable for communities (places of memory) even if they are lacking of specific goods. They are places described in pictorial or literary works, places where particular historical events took place, or places where important people in political and cultural history stopped.
– *Reading through aesthetic and material characters:* they not only define the object in itself (design), but also the composing materials, and the used techniques. It is a reading typology used in Italy (especially in the provincial plans level) Great Britain, and in the alpine area (Germany and Austria), for the description of rural assets.

Every reading methodology is built on specific cultural tendencies and it is subordinated to the specificities of the different territorial realities. The various contributions about the main reading methodologies and landscape assessments used in the European Countries underline the remarkable methodological elaboration and testing worked out during these last years, but also (as already for the definition of the term landscape) the urgent need of a terminology and basic criteria for information exchanging, allow comparing the different experiences, thus yielding to mutual enrichment, crucial for a general growth on these topics, both from a cultural and an operational point of view [5]. Today, even if codified methods for studying and describing the landscapes are increasing, the need for a new methodology for reading the landscape, which take in account the conceptual evolution of it, the multi-disciplinarity of this matter, and the interoperability among the different points of view of researchers. For those reasons the use of conceptual and formal models, as the ontology-based ones, is a useful tool for knowledge managing and sharing in a complex scenario as the landscape planning.

2.2 Rural landscape reading

Rural landscapes are the expression of the correlation between human activities and environment, where the human facility is showed by a continuous search of balance. They express the functional evolution, occurred in the course of time, linked to the working technique, dwelling ways, natural dynamics and social conditions. Those are landscapes produced by a long adaptation until reaching ecological essentiality and stability given by man-nature compromise. The Italian rural landscapes are characterized by several architectural productions being different in forms and typologies, but linked by a common author: the farmer-architect. By using the stones taken from the fields or the being more easily found and economic material, he unconsciously designed a landscape with such a high historical value that it has become the symbol of the local culture. The functionalism leading the different architectural typologies doesn't stress only the formal component, but points out the strict connection between the function linked to rural activities and saving exigencies and rational management, reproducing a model that can be defined *unconsciously sustainable* [9]. Knowing and understanding the landscape matrices is the first step toward their appraisal and conservation. The forms of rural landscape issue not only from the territorial physical structure, but also from the rules entailing the use of social power to transform the territorial structures following each other in the course of time. Apart from the productive activities, we should consider various elements of a man-made landscape (historical centres, architectural models, building structures). Historical built heritage should be investigated starting from its identification on a territorial scale, since it cannot be divided from its context landscape scheme. It is unconceivable to investigate settlements and buildings as close systems to be preserved and/or safeguarded without considering their ambit condition. Indeed, geo-morphological aspects, social and cultural dynamics, economic and political situations, criticalities and vulnerabilities are all part of such analysis. Also physical and anthropological factors contribute to the definition of landscape, or at least to the perception of it. Therefore, the reading of rural landscape concerns not only the visual experience or the morphological, geographical and naturalistic aspects, but gives back also historical and cultural meanings, attributed by communities which live in and use those places. After all, the rural landscape reading is mainly referred to three closely connected aspects:

- the historical point of view and the landscape development during the time;
- the landscape shape;
- the functional-anthropological aspect (use of the ground, takeovers, infrastructures,).

From a general point of view, the methodologies used to read the characteristics of a place start from a low detailed scale level. However in this way we can find only the presence and the role of the elements in the general places morphology; on the other hand, other elements (i.e. terracing, hedges, tracks) structure the landscape not only with their design (formal character) but also with their materials, or particular constructive techniques and traditional cultivations. A field represents with its shape and color the function of "field"; but with its function it is also to be specified the cultivation method and every useful cognitive aspect, and the way it is related and inserted in the context [10, 11]. The landscape shapes give back the objects story. Every object, once recognized, assumes the role of sign and it must be interpreted. The signs are the representation of the reality (signifier) - Eco [12] asserted that *"the sign stays for something else which is its object"* - and their interpretation gives back the signified (the sense and the value) attributed to the object-element. The reading of the traditional landscape signs must not have the scope of a reconstruction (restoration) of a given landscape, but of understanding the meaning of the signs themselves to insert them in the planning actions like live objects fitted in the context and in the currently needs. In a nutshell, for representation and interpretation of landscape elements we argue that two different dimensions should be integrated: an objective and a subjective one, giving to the perception (consequently to the subjectivity) a crucial relevance and, in particular, a great influence to its characteristic and clear relationships with the environment and the territory. To this aim, we need of a flexible framework used on several territorial realities, therefore a method for integrating all the undertones in the complexes and peculiar territorial systems.

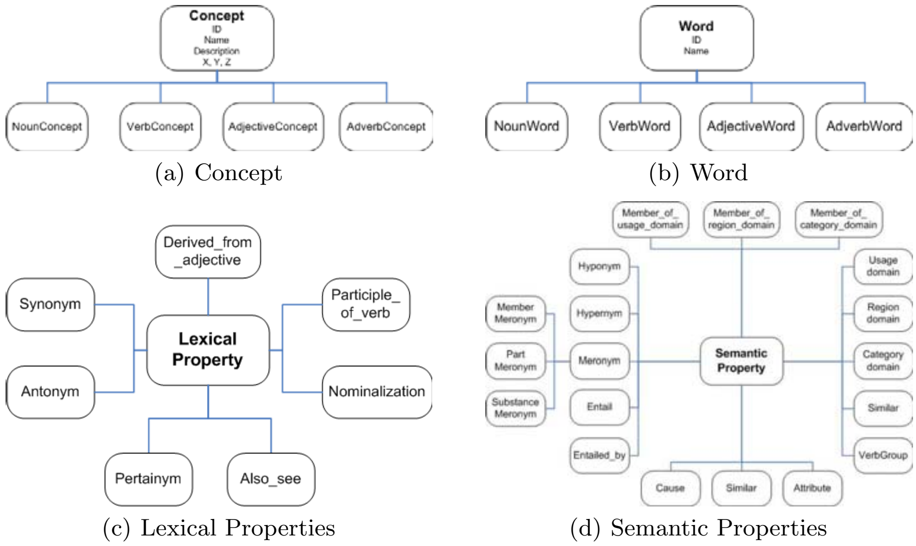
3 Ontology Representation

The aim of our paper is to define and implement a model for knowledge representation using a conceptualization as much as possible close to the way in which the concepts are organized and expressed in human language and use it on a specific conceptual domain as the rural landscape. We use WordNet [13] as general linguistic knowledge base. All information in WordNet is organized using linguistic properties. The basic unit in WordNet is the synset, a logic set of words related by the synonymy property. Each synset is a concept in WordNet. All the synsets are related to the others by pointers that represent linguistic properties. Two kinds of relations are represented by pointers: lexical and semantic. Lexical relations hold between word forms; semantic relations hold between word meanings. Examples of those relations are hypernymy/hyponymy, antinomy, entailment, and meronymy/holonymy.

Now we are in the position of defining our model; it is composed by a triple $< S, P, C >$ where:

S is a set of objects;
P is the set of properties used to link the objects in S;
C is a set of constraints on P.

In this context we consider *words* as objects; the properties are *linguistic relations* and the constraints are *validity roles* applied on linguistic properties with respect to the considered term category. In our approach the knowledge is represented by an ontology implemented w.r.t. a semantic network. A semantic network can be seen as a graph where the nodes are concepts and the arcs are relations among concepts. A concept is a set of words which represent an abstract idea. One of the most important progress in the KR applications derives from proposing [14], studying [15] and developing [16] languages for knowledge representation. Even if those languages have several differences they share some



(a) Concept (b) Word

(c) Lexical Properties (d) Semantic Properties

Fig. 1. Model components

common aspects based on the specification of objects (concepts) and the relationships among them.

In the last years several languages have been proposed to represent ontologies. It is the authors' opinion that OWL is the best language to reach our purpose in terms of expressive power. Therefore we describe the semantic network implementing the ontology in OWL using the defined model. In particular we use the DL version of OWL because it has enough effectiveness to describe the ontology. The DL version allows the declaration of disjoint classes which are used, for example, to assert that a word belong to a syntactic category. Moreover it allows the declaration of union classes used to specify domains and property ranges to relate concepts and words belonging to different lexical categories.

3.1 The proposed model

We formally describe the ontology schema and the corresponding semantic network representation using OWL. Every node, both concept and word, is an OWL individual. The connecting edges in the semantic network are represented as *ObjectProperties*. This properties have some constraints that depend on the syntactic category or on the kind of property (semantic or lexical). For example the hyponymy property can relate only nouns to nouns or verbs to verbs; on the other hand a semantic property links concepts to concepts and a syntactic one relates word forms to word forms. Concept and word attributes are considered with *DatatypeProperties*, which relate individuals with a pre-defined data type. Each word is related to the represented concept by the ObjectProperty *hasConcept* while a concept is related to words that represent it using the ObjectProperty *hasWord*. These are the only properties able to relate words with concepts and vice versa; all the other properties relate words to words and concepts to concepts. Concepts, words and properties are arranged in a class hierarchy, resulting from the syntactic category for concepts and words and from the semantic or lexical type for the properties.

Concepts and Words In figure 1 the hierarchies used to represent the objects of interest in our model are shown. Figures 1(a) and 1(b) show that the two main classes are Concept, in which all the objects have defined as individuals and Word which represent all the terms in the semantic network. These classes are not supposed to have common elements therefore we have defined them as disjoint. The class Word define the logical model of the word forms used to express a concept. On the other hand, the class Concept represents the word meaning related to a word form. We can see that the subclasses have been derived from the related categories. There are some union classes useful to define properties domain and codomain. We define some attributes for Concept and Word respectively. In particular Concept has: *Name* that represents the concept name; *Description* that gives a short description of concept; *X, Y, Z* that localize a concept in a 3D space. On the other hand Word has *Name* as attribute that is the word name. Moreover for all elements we define an *ID* within the WordNet offset number or a user defined ID.

Table 1. Model features

(a) Properties

Property	**Domain**	**Range**
hasWord	Concept	Word
hasConcept	Word	Concept
hypernym	NounsAnd VerbsConcept	NounsAnd VerbsConcept
holonym	NounConcept	NounConcept
entailment	VerbWord	VerbWord

(b) Property features

Property	**Features**
hasWord	*inverse* of hasConcept
hasConcept	*inverse* of hasWord
hyponym	*inverse* of hypernym; *transitivity*
hypernym	*inverse* of hyponym; *transitivity*
cause	*transitivity*
verbGroup	*symmetry* and *transitivity*

(c) Model constraints

Costraint	**Class**	**Property**	**Constraint range**
AllValuesFrom	NounConcept	hyponym	NounConcept
AllValuesFrom	VerbConcept	hyponym	VerbConcept
AllValuesFrom	NounConcept	attribute	AdjectiveConcept
AllValuesFrom	AdjectiveConcept	attribute	NounConcept
AllValuesFrom	NounWord	synonym	NounWord
AllValuesFrom	VerbWord	synonym	VerbWord
AllValuesFrom	AdjectiveWord	synonym	AdjectiveWord
AllValuesFrom	AdverbWord	synonym	AdverbWord
AllValuesFrom	VerbWord	also_see	VerbWord
AllValuesFrom	AdjectiveWord	also_see	AdjectiveWord

Properties The semantic and lexical properties are arranged in a hierarchy (see figure 1(c) and 1(d)). In table 1(a) some of the considered properties and their domain and range of definition are shown.

Constraints The use of domain and codomain reduces the property range application; however the model so far described does not have a perfect behavior in some cases. For example the model does not know that even if the hyponymy property is defined on the sets of nouns and verbs, if it is applied on the set of nouns it has as range the set of nouns, otherwise if it is applied to the set of verbs it has as range the set of verbs. In table 1(c) there are some of defined constraints and we specify on which classes they have been applied with respect to the considered properties; the table shows the matching range too. Sometimes the existence of a property between two or more individuals entails the existence of other properties. For example, being the concept dog a hyponym of animal, we can assert that animal is a hypernymy of dog. We represent in OWL this characteristics by means of property features. The table 1(b) shows several of those properties and their features.

4 System architecture

The proposed model has been implemented in a software system able to create, manage and manipulate ontologies. This software is linked with WordNet which, using the previously described model, is completely mapped into OWL. Many information systems use a knowledge base to represent data in order to satisfy information requests. In our approach we use an appropriate algorithm to extract from WordNet a domain semantic network; this net provides a general representation of our domain of interest. The system has several modules which implement its basic functionalities. The *system interface* shows to the user the ontology catalog stored in the ontology repository (i.e. a relational DB) by means of an appropriate software module called *OntoSearcher*: OntoSearcher performs a syntactic search or a browsing in a directory structure arranged by arguments to the aim of finding an ontology relevant to the user interest. When OntoSearcher finds a suitable ontology, the *OntoViewer* builds a graph (a semantic network) to represent the ontology. In the following we describe the algorithm used to build dynamically the semantic network. If the user domain of interest is not in the Ontology Repository, she can build an ontological domain using WordNet. This step is performed by the *OntoExtractor*; it enables some functionalities to built an ontology represented by a semantic network. We propose a dynamic construction of the semantic network using an ad hoc algorithm which take into account the WordNet structure. The network is built starting from a domain keyword that represents the context of interest for the user. Moreover, every domain keyword may have various meanings (senses) due to polysemy propriety, so a user can choose its proper sense of interest using the tool interface. We then consider all the component synsets and construct a hierarchy, only based on the hyponymy property; the last level of our hierarchy corresponds to the last level of WordNet one. After this first step we enrich our hierarchy considering all the other kinds of relationships in WordNet. Based on these relations we can add other terms in the hierarchy obtaining an highly connected semantic network. Clearly, even if a knowledge base could be large and detailed, it will never give us a high level of specialization for every existing knowledge domains. Our approach tries to give a solution to this problem. In fact a user can interact with our system in order to create a first ontological knowledge representation or create a new one using the *OntoEditor* module. The OntoEditor functionalities allow a user to modify the ontology structure as a whole adding new terms and concepts in the network, linking terms and concepts using arrows (lexical and semantic properties), deleting nodes and arcs. We used Java technologies to implement the system and in particular the interaction with the semantic network is obtained by means of Java 3D libraries.

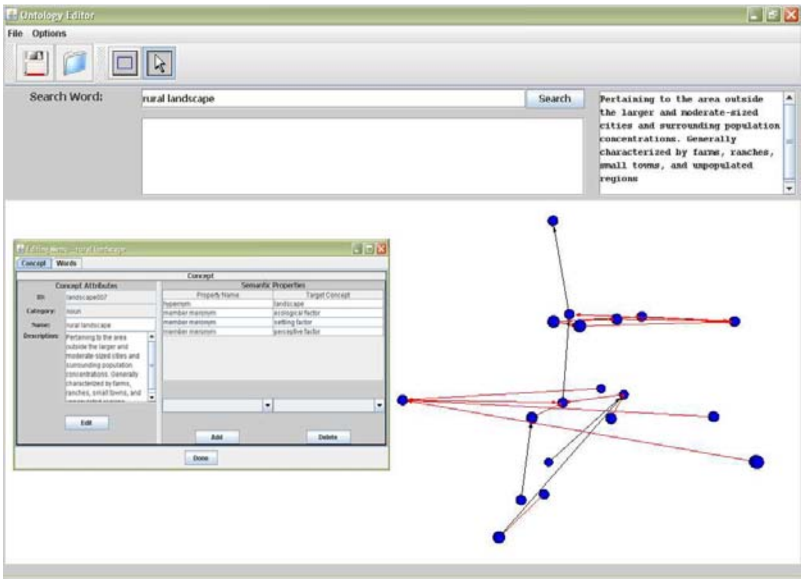
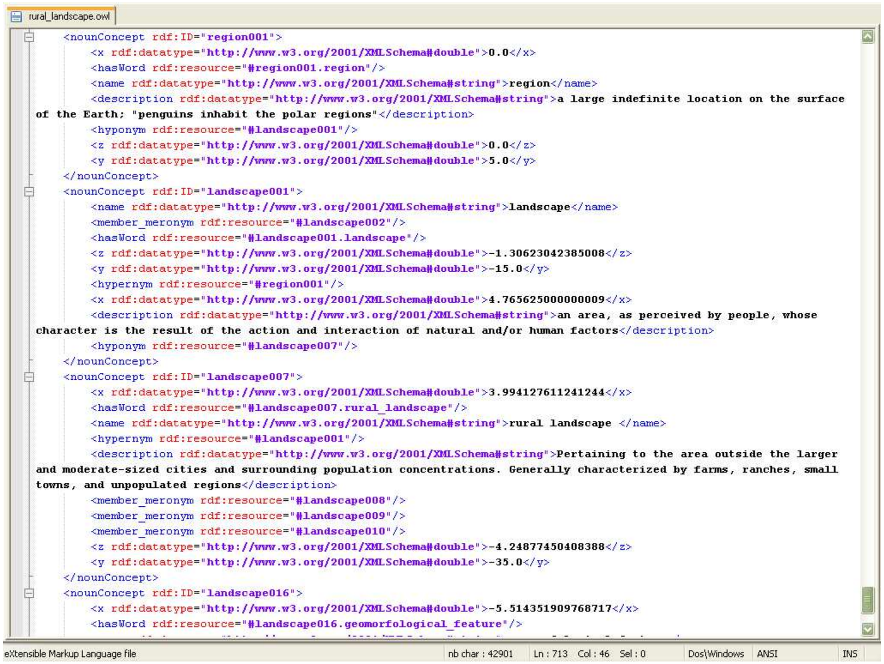


Fig. 2. Tool interface

5 Use case example

In this section we show a use case of our tool. The recent national and international legislative, regulatory and implementing initiatives described in the previous sections put in evidence the need for defining a detailed methodology for representing knowledge about landscape. The knowledge definition tasks start from several elements often poorly related. This factor increases the system complexity introducing an high heterogeneity. In the ontology *pre-consensus* step we define our glossary using well-know knowledge sources as [17, 4, 18, 19], providing a more detailed description of the variables, since it was affected by several particular situations about our field of interest. A specific ontology about *rural landscape* has been created ex novo using this glossary. The related elements and phenomena have been individuated by applying disaggregating and re-aggregating processes to the rural landscape components. The rural landscape image has been described by means of its natural (ecological) factors, built-up (settling) factors and visual and perceptive factors. All these macro-categories have been arranged in classes, which in turn have been divided in features and variables. Successively a domain ontology about *landscape* is extracted from WordNet following the steps described at the end of the section 4 starting from the keyword *region(sense3): a large indefinite location on the surface of the Earth*. The process to extract an ontology from WordNet starts from an interaction with the user inserting a specific term (i.e. region) by means of the user interface and choosing the proper sense reading the description of the related concepts. The system gets the right sense and build the ontology following the steps described in the previous section. We choose to link the new proposed concept of landscape (*an area, as perceived by people, whose character is the result of the action and interaction of natural and/or human factors* [4]) directly to the region synset, using it as bridge for ontology merging. Therefore we expand this ontology giving more information about the rural landscape (see figure 2). Using our tool we have the OWL representation shown in figure 3. We notice that we give a knowledge expansion of WordNet knowledge base inserting new synsets about the territorial planning research field.




```xml
<nounConcept rdf:ID="region001">
    <x rdf:datatype="http://www.w3.org/2001/XMLSchema#double">0.0</x>
    <hasWord rdf:resource="#region001.region"/>
    <name rdf:datatype="http://www.w3.org/2001/XMLSchema#string">region</name>
    <description rdf:datatype="http://www.w3.org/2001/XMLSchema#string">a large indefinite location on the surface
of the Earth; "penguins inhabit the polar regions"</description>
    <hyponym rdf:resource="#landscape001"/>
    <z rdf:datatype="http://www.w3.org/2001/XMLSchema#double">0.0</z>
    <y rdf:datatype="http://www.w3.org/2001/XMLSchema#double">5.0</y>
</nounConcept>
<nounConcept rdf:ID="landscape001">
    <name rdf:datatype="http://www.w3.org/2001/XMLSchema#string">landscape</name>
    <member_meronym rdf:resource="#landscape002"/>
    <hasWord rdf:resource="#landscape001.landscape"/>
    <z rdf:datatype="http://www.w3.org/2001/XMLSchema#double">-1.30623042385008</z>
    <y rdf:datatype="http://www.w3.org/2001/XMLSchema#double">-15.0</y>
    <hypernym rdf:resource="#region001"/>
    <x rdf:datatype="http://www.w3.org/2001/XMLSchema#double">4.765625000000009</x>
    <description rdf:datatype="http://www.w3.org/2001/XMLSchema#string">an area, as perceived by people, whose
character is the result of the action and interaction of natural and/or human factors</description>
    <hyponym rdf:resource="#landscape007"/>
</nounConcept>
<nounConcept rdf:ID="landscape007">
    <x rdf:datatype="http://www.w3.org/2001/XMLSchema#double">3.994127611241244</x>
    <hasWord rdf:resource="#landscape007.rural_landscape"/>
    <name rdf:datatype="http://www.w3.org/2001/XMLSchema#string">rural landscape </name>
    <hypernym rdf:resource="#landscape001"/>
    <description rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Pertaining to the area outside the larger
and moderate-sized cities and surrounding population concentrations. Generally characterized by farms, ranches, small
towns, and unpopulated regions</description>
    <member_meronym rdf:resource="#landscape008"/>
    <member_meronym rdf:resource="#landscape009"/>
    <member_meronym rdf:resource="#landscape010"/>
    <z rdf:datatype="http://www.w3.org/2001/XMLSchema#double">-4.24877450408388</z>
    <y rdf:datatype="http://www.w3.org/2001/XMLSchema#double">-35.0</y>
</nounConcept>
<nounConcept rdf:ID="landscape016">
    <x rdf:datatype="http://www.w3.org/2001/XMLSchema#double">-5.514351909768717</x>
    <hasWord rdf:resource="#landscape016.geomorfological_feature"/>
```


Fig. 3. Owl representation of rural landscape ontology

6 Conclusion and future works

The design, implementation and reuse of existing ontologies is a non trivial task. When we want to use different ontologies, they must be combined in same way; this task is performed either with an ontology integration or leaving them separated. In both cases the ontologies must be aligned i.e. they must be in a condition of mutual compatibility. In this paper we have proposed an approach to solve this problem; we define a simple and general model, taking into account a linguistic approach considered as the natural communication way between human agents. On the other hand the use of formal models for knowledge representation could represent a necessary starting point in several research fields and in particular in the territorial planning one. From a general point of view we have an evolution during time of the concepts; this is a cause of *knowledge obsolescence*, so there is the need for a continuous updating. We note this problem in WordNet, in fact in this knowledge base the concept of *landscape* is related only to a *visual appearance* dimension. Moreover in all fields of knowledge the research innovation allows the definition of new concepts. For example the concept of *rural landscape* does not exist in WordNet; this lack needs of a knowledge expansion. This is a preliminary study, other issues are to be investigated. We are implementing a further specialization of our ontology tacking more and more into account the perceptive aspects of local communities. On the other hand we are thinking about the definition of an algorithm for automatic ontology integration; a solution to the ontology mismatch problem and the design and implementation of a distributed system for information sharing based on our knowledge representation model.

References

1. Gruber, T.R.: A translation approach to portable ontology specifications. Knowl. Acquis. **5**(2) (1993) 199–220
2. Neches, R., Fikes, R., Finin, T., Gruber, T., Patil, R., Senator, T., Swartout, W.R.: Enabling technology for knowledge sharing. AI Mag. **12**(3) (1991) 36–56
3. Dean, M., Schreiber, G.: OWL Web Ontology Language Reference. Technical Report http://www.w3.org/TR/2004/REC-owl-ref-20040210/, W3C (2004)
4. European Community: European landscape convention (2000)
5. Cecchi, R.: Reading the landscape. In Scazzosi, L., ed.: Reading the Landscape. Gangemi, Roma, Italy (2002) 7–8
6. Guido, M.R.: A contribution to the knowledge of landscape identification and assessment. In Scazzosi, L., ed.: Reading the Landscape. Gangemi, Roma, Italy (2002) 13–16
7. Calzolari, V.: Storia e natura come sistema. Argos, Milano, Italy (1999)
8. Scazzosi, L.: Reading the Landscape. Gangemi, Roma, Italy (2005)
9. Callioni, P.M.: L'uomo e il territorio nel paesaggio rurale italiano. Technical report, Roma, Italy (2005)
10. Turri, E.: Semiologia del paesaggio italiano. Longanesi, Milano, Italy (1990)
11. Turri, E.: La lettura del paesaggio. In Zerbi, M.C., ed.: Il paesaggio tra ricerca e progetto. Giappichelli, Torino, Italy (1994)

12. Eco, U.: Segno. Isedi, Milano, Italy (1974)
13. Miller, G.A.: Wordnet: a lexical database for english. Commun. ACM **38**(11) (1995) 39–41
14. Minsky, M.: A framework for representing knowledge. Technical report, Cambridge, MA, USA (1974)
15. Brachman, R.J.: On the epistemological status of semantic networks. In Findler, N.V., ed.: Associative Networks: Representation and Use of Knowledge by Computers. Academic Press, Orlando (1979) 3–50
16. Brachman, R.J., Schmolze, J.: An overview of the Kl-ONE knowledge representation system. Cognitive Science **9**(2) (1985) 171–216
17. Colombo, A.G., Malcevschi, S.: Handbook for environmental impact assessment indicators. landscape indicators. Volume 5. Centro VIA Italia, AAA and Fast, Milano, Italy (1999)
18. European Community: CORINE Land Cover Project (2000)
19. OECD-Organization for economic co-operation and development: Macrothesaurus. www.oecd.org (2007)

Ontologies for urban regeneration opportunities and weaknesses for their development within city cohesion policies

Francesco Rotondo
Department of Architecture and Town Planning, University of Bari Polytechnic – via Orabona 4, 70126, Bari, Italy
Corresponding author. Tel.: +39 347 5957508; fax: +39 80 5963348.
E-mail address: f.rotondo@poliba.it (F. Rotondo).

Abstract. This paper outlines the construction of an urban regeneration ontology (OUR). An initial discussion on the aspirations of urban regeneration and of the theoretical background and possible definitions of an ontology is followed by a description of an experience in which a prototype of an ontology for urban regeneration is developed. Using this prototype as a starting point, the experience has in turn led to the clarification of particular user needs and requirements. The prototype was developed by "Towntology software", supported by COST Action C21 funding, by the LIRIS Laboratory at the University of Lyon. It has been tested in Italy in the Apulia Region office for Spatial planning and Environmental Management, as described in the paragraph "OUR (in) practice". Initial tests have shown interesting results and offered suggestions of possible user needs and requirements that are discussed in the paragraph "User needs and requirements for ontology specification". The reported experience is not yet over, hence not all of what may be commented upon has been written or observed. Nevertheless these initial results can be considered useful in tracing possible future directions for research.

Keywords: Urban Regeneration; Ontology; User needs

1. Introduction

The European Union (EU) Community Strategic Guidelines 2007-2013 place particular emphasis on the specific needs of certain zones, such as urban and rural areas. The guidelines encourage an "integrated approach" towards cohesion policy, not only stimulating growth and creating jobs, but also pursuing certain social and environmental objectives.

Furthermore, the European Parliament, in its report on the urban dimension within the context of enlargement[*], welcomed the incorporation of sustainable urban development in cohesion policy.

---

[*] Report on the urban dimension in the context of enlargement, rapporteur: Jean Marie Beaupuy, A6(2005) 0272 on 21.9.2005.

The success of the URBAN Community Initiative[†] is in no small measure due to such an integrated approach. URBAN has targeted social and economic cohesion in parallel, thus removing barriers to employability and investment whilst at the same time focusing on social and environmental goals. The mobilisation of a broad range of partners with different skills has underpinned this approach. Forthcoming EU urban regeneration policies attempt to consolidate these successes with new policy initiatives directed towards the regeneration of deprived urban areas, with the JESSICA[‡] policy a case in point. In this context, ontologies could play a significant role in developing and managing these new policies thereby strengthening integration, sharing ideas and increasing knowledge of problems specific to urban regeneration. Furthermore they could, in the context of the European Union, present a multilingual tool capable of demonstrating concepts, shared definitions and the relationships between them.

At present, institutions dedicated to the management of regeneration policies at all levels, whether EU, regional or municipal, often demonstrate real difficulty in terms of interpreting the language used by an architect, a planner, an ecologist or an economist due to discipline-specific terminology. Urban regeneration may therefore mean different things in different disciplines. Ontologies could be a useful tool in ordering, integrating and making transparent a range of possible meanings associated with a policy. It would appear to be particularly useful within the European context, where the coherence of different actions, even within the same field, is often difficult to establish whether at a European or municipal level.

2. Significant elements of urban regeneration and its multiple dimension

Urban regeneration is an integrated urban policy approach mixing multiple dimensions: economic, social, cultural, spatial and environmental. New urban planning and design methods replaced rationalist architectural codes and conventions by locating some key points which, when seen alongside the *Leipzig Charter on sustainable European cities*[§], echo the aspirations of urban regeneration policies and strategies. This may be expressed, for example, in the an upgrading of the physical environment and encouraging sustainable urban transport, the strengthening of the local economy and labour markets, or in the promotion of proactive education and training policies for children and young people in deprived urban areas.

[†] The Community Initiative URBAN II (2000-2006): Communication from the Commission to the Member States of 28 April 2000 laying down guidelines for a Community initiative concerning economic and social regeneration of cities and of neighbourhoods in crisis in order to promote sustainable urban development (URBAN II), C(2000) 1100 of 28.04.2000.

[‡] Joint European Support for Sustainable Investment in City Areas

[§] the *Leipzig Charter* is a document by the Ministers responsible for urban development policy of the EU member states, edited in its final draft version on 2 May 2007, available on line at:
http://europa.eu.int/comm/regional_policy/index_en.htm.

The recommendations of the Charter summarize strategies put forward by the New Urbanism movement (Dutton, 2000) or those already declared in the *New Charter of Athens* (2003) by the European Council of Town Planners (2003) and embrace:
The creation of high quality public spaces and their reinforcement within city structure;
The improvement of energy efficiency in buildings and the modernization of infrastructure networks favouring a compact city form;
The use of greenery as a bio-infrastructure to enhance the sustainability of cities; The encouragement of mixed-use neighbourhoods, buildings and blocks (shops, offices, apartments, and homes on the same site), with a wide diversity in demographic make-up in terms of age, income level, culture, and race; The promotion of well-conceived social housing policies, with suitable and affordable housing; The participation in urban policies which lead to a better level of education and training contributing to achieve their ambitions and to ensure equal opportunities on a long-term basis.
All of these elements are of crucial importance to deprived urban neighbourhoods not only in reducing inequalities but also in preventing social exclusion and improving the physical environment. Indeed, new EU initiatives, JASPERS[**] and JESSICA, and several particular measures of the European Regional Development Funds (ERDF) will support, from 2007 to 2013, exactly such policies within urban regeneration.
The multi-dimensional nature of urban regeneration processes encouraged by the European Union is therefore evident when seen in the context of the concrete objectives of urban regeneration itself and the support of specific European policies which target those objectives. To a region implicated in such European policy making it is therefore of primary importance to manage the multi-dimensional nature of the problem, by drawing upon different skills and competencies and sharing the same words and objects whether physical, economic or social (EC, 2006). Ontologies could be a potential way of organizing this complex and multifaceted task, as we attempt to outline in the following paragraphs.

3. A possible Ontology for Urban Regeneration (OUR)

According to Gruber (1993), an ontology is an explicit, formal and shared conceptualization of a particular domain. The conceptualization process represents the attribution of unambiguous meanings to terms defining knowledge in that precise domain (domain ontology). Guarino (1998) defines an ontology as a set of logical axioms designed to account for the intended meaning of a vocabulary.
A domain ontology for urban regeneration is therefore expected to express the viewpoints and satisfy the informational needs of multiple stakeholders and interest groups, including, yet by no means limited to, town planners, environmental agencies, municipalities, police departments, owners and sellers of real estate, third sector associations. These actors use different jargons and pursue different, occasionally conflicting tasks, even if they manage similar or related domains. Reports of the

---

[**] Joint European Resources for Micro to Medium Enterprises

experience of ontology development in many fields of application[††], underline that different jargons and informational needs are hard to accommodate in a consensual ontology. In the case of OUR, it is not strictly necessary to arrive at a unique definition of a term, if, as may be the case, the obtaining of a common definition proves impossible. Rather, it is sufficient that all agents involved in the same or similar activities are at least informed and have the possibility of knowing others definitions. For example, in an "Objective 1" region such as Apulia in Italy, it is necessary that the Environmental Management Department, the Regional Planning Office, the Transport Bureau and the Economic Planning Bureau are aware of other definitions thus avoiding a scenario in which each of them promotes different and possibly incoherent funding policies on the same urban regeneration objective. Such a situation is not simply theoretical. Consider, for example, the different funding for the construction of cycle paths within the previous phase of cohesion policies (2000-2006), which may be considered one of the objectives of urban regeneration. Promoted by the Transport Bureau, this particular objective may take on a more functional image, useful in increasing the possibilities for movement in an urban context. This clearly does not correspond with an altogether different definition of a cycle path associated with leisure and nature, which could be built with natural materials such as compact sand and bordered by green hedges, perhaps intended as tourist routes through the countryside. Indeed, such a vision including all of the above elements was conceived during the same policy phase by the Environmental Management Office. In the same period the Regional Planning Office promoted urban regeneration processes in which it funded the same objective, in this case encouraging an alternative method for commuters to reach the work place, schools etc. With this in mind, an ontology for urban regeneration related to regional geographical information systems has the potential to begin forming a common platform for spatial policies. It would also perhaps avoid a repeat of such ineffective communication and a lack of knowledge, especially within an organization of any considerable size (Apulia region boasts more than 3000 employees). Indeed, this has already been highlighted by various experiments carried out by the Italian Ministry of Innovation through its centre for Informatics in Public Administration (CNIPA in the Italian acronym; www.cnipa.it), in a project entitled "ARIANNA", available at: http://arianna.diviana.net/. According to Barone and Di Pietro (2005), ontologies could introduce administrative simplification and efficiency, better control of public expenditure and the simplification of policy management. Moreover, a domain ontology could demonstrate all of the different definitions of a given policy and its corresponding objective, in a semantic integration of a certain number of databases, and with the software Towntology, produced by the LIRIS, Institut National des Sciences Appliquées de Lyon. Such methods also offer the possibility of using images to represent concepts, thereby creating a visual dimension which can potentially offer a different pathway towards the same objective.

[††] Examples of ontologies are available at
http://protege.stanford.edu/download/ontologies.html

4. OUR (in) practice

With reference to the case of the Apulia region, as well as representing a large number of situations especially within the “Objective 1” regions of the EU, in the case of OUR, we have started to create the ontology from the point of view of a town planner. The 110 terms identified in describing the domain were then submitted to other agents, identifying alternative definitions of concepts and related objects, illustrating them with images and showing their relationships in a dynamic chart which changed its representation according to the interests of the agent managing the ontology.

The chosen agents are all interested in developing urban regeneration policies at a regional level and they are directly involved in the elaboration of the specific regional measures to apply the European programmes such as the next JESSICA, or what is referred to as the Operative Program in the ERDF.

A working group of five professionals was established including a civil engineer working within the field of public service utilities, a biologist specialised in ecology, an architect specialised in urban planning, an agronomist and an economist within the field of structural fund management. They were guided through the process by the author and one of his students.

Fig. 1. The meeting of the working group involved in study in the Apulia Region.

The survey was conducted using the well known SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis method, in an attempt to establish whether the ontology could be a useful tool in their public administration offices, whether the prototype used basic words and definitions and whether there was any conflict or disagreement regarding such definitions and relationships.

Following an outline of the meeting the participants were engaged in the analysis of the ontology.
The “Editor” page within the software provides the space in which definitions and relations between terms may be entered, as shown in the figure 2.

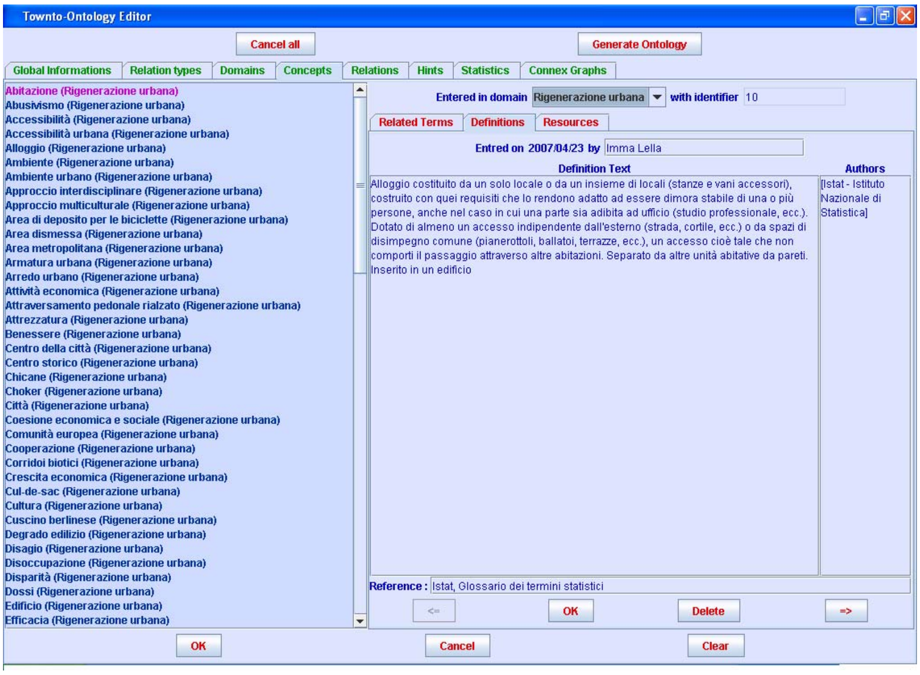


Fig. 2. The Towntology editor offers the possibility of adding as many definitions as necessary simply using the arrow keys. In this example, the definitions of terms, for example, “abitazione” (house), are in Italian.

A total of 110 definitions within the urban regeneration domain have so far been entered into the OUR, ranging from somewhat abstract or complex terms such as “urban decline” or “social and economic cohesion” to a definition of concrete objects such as “cycle path”, “chicane” or “green corridor”. For each definition there is the possibility of indicating a reference and a URL with a link to a corresponding image (as shown in figure 3), therefore attempting to pinpoint the disciplinary knowledge at the base of the term, meaning that the knowledge base may be enlarged by users from different fields.

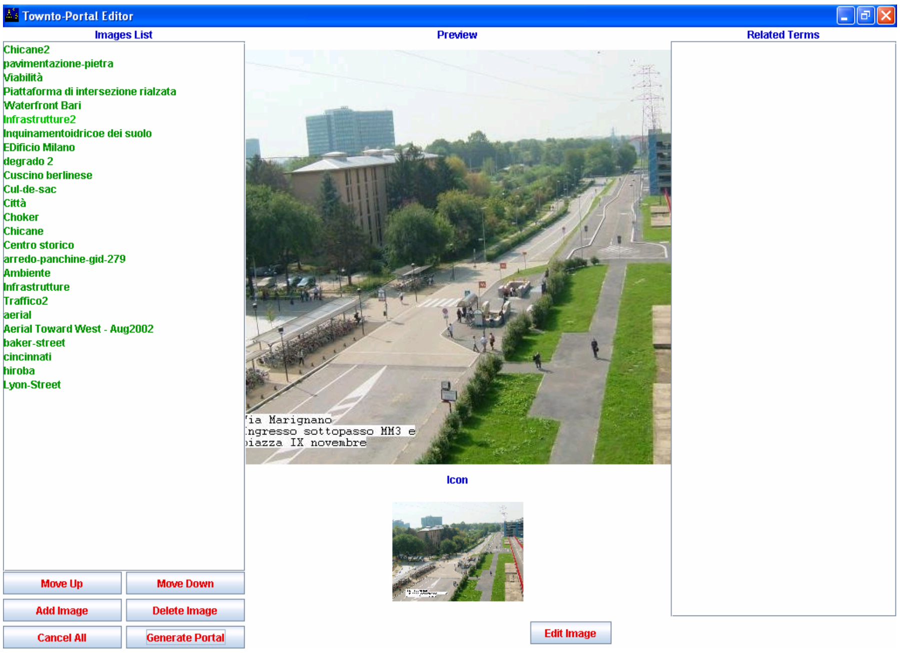


Fig. 3. A typical scene from the Portal Editor showing an image of a complex space with a pedestrian precinct, a pedestrian crossing, a bus stop and a green hedge.

Having edited the ontology, the Townto-Browser offers the possibility of surfing the ontology to reveal relationships, the level of integration of particular terms and their general value, as is shown in figure 3.

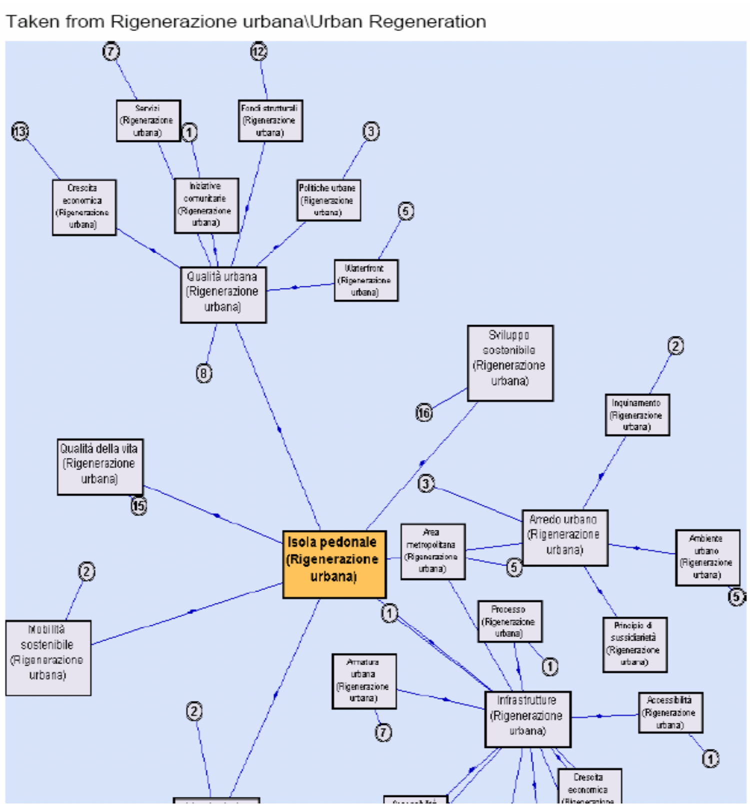


Fig. 3. A typical display of the townto-browser showing the term "isola pedonale" (pedestrian precinct) and other related terms.

Since its conception and creation, OUR has been tested with the groups described above who work within different departments of the Apulia Region. This has resulted in 25 definitions of the original 110 being modified and the addition of a further 9 definitions of existing terms, demonstrating no agreement on those terms. The most striking difference observed has been between the definitions inserted by the Transport Department and those made by the Environmental Management Department.

Leading on from this, the next step could be the integration of the ontology into a Geographic Information System (GIS) for the Apulia Region (currently under construction) linking terms to actual examples from around the region focusing on each city where regional departments are involved in planning urban regeneration policy.

Having a unique domain ontology could reduce the possibility of overlapping decision making and could offer the possibility of integrated policy making by sharing the same knowledge base. Indeed, with regards to the ontology running on a GIS, it may be possible to highlight locations where urban regeneration policies either have been, or are in the process of being applied, with all the resulting terms and relationships.

5. User needs and requirements for ontology specification

After this first, yet significant, experience in collecting impressions from participants it is possible to outline some user requirements:

- OUR could perhaps be of most use if used as an integrated tool within usual policies and policy making, rather than as an exceptional instrument;
- In order for OUR to be effective it has to be available on the web or at least on the intranet of the public office or institution involved;
- A unique multitask interface could be developed within the "Towntology" software with which the user would simultaneously be able to locate the list of terms (possibly with a multilingual description), their definitions, their relationships and any associated imagery where applicable.

The availability of multiple on-line ontologies takes on a particular importance, especially when considering EU policies given that:

a) Public organizations are predominantly divided into a range of departments with a high level of specialization yet a low level of integration. If ontologies were to become an integrated tool which could be applied to even standard policy, or better still if applied on a GIS, public organizations could potentially arrive at a greater integration of policy content.

b) Ideally, OUR would be available on-line or at least on the intranet of the organization in question, as its value is determined by the possibility of being utilized by anyone involved in urban regeneration regardless of their physical working location**.** In this way the glossary will grow and every definition could be discussed and eventually shared in a unified way.

c) EU cohesion policies are frequently multifaceted and complex, often with various possible implementation choices, deriving in part from the particular characteristics of the nation in which it is applied.

In the case therefore of single large-scale organizations, ontologies could lead towards a better cohesion in the way that different member states may apply the same EU policy.

6. Conclusions

Although yet to be completed it may be, considering the initial results of the experience, possible to assume that OUR is potentially a tool which could foster improved communication between stakeholders. Possible future directions for research in the field of ontologies for urban regeneration with reference to EU policies could be a compared evaluation between ontologies as seen within different languages and cultures as, for example, with a French urban renewal ontology, as has been developed by the EDU Laboratory in Lyon (Berdier, Roussey, 2007), alongside another in English thereby making an ontology available in the official language of EU. A step beyond this would perhaps be the conception of a more extensive experiment involving EU offices, in which regeneration policies are developed and managed. As the user requirements highlighted by the case of the Apulia Region demonstrate, the possibility of using OUR in practice is strictly related to the wider diffusion of ontologies within public administration routine. From the first definition by Gruber (1993) of an ontology in the sphere of Artificial Intelligence, only within the last few years have we seen some experiences. The greater the increase in the availability of data sets, the more an ontology lends itself to being a useful instrument in providing clear definitions and corresponding relationships within a specific domain.

Acknowledgements

I gratefully acknowledge the cooperation of officials from departments of the Apulia Region Departments for having participated in the OUR experiment and for having offered their invaluable suggestions, especially with regards to user needs. It should be added that the views expressed in this paper do not necessarily reflect those of the scientific group working with the EU COST Action C21 which sustained the project.

References

1. Barone, A., & Di Pietro, P. (2005). "Ontologia. Elemento Chiave per la Costruzione della Società della Conoscenza". *IterLegis 5-6*, 9-15.
2. Berdier, C., Roussey, C. (2007). "Urban Ontologies: the Towntology Prototype towards Case Studies". In J. Teller, J. Lee & C. Roussey (Eds.), *Ontologies for Urban Development* (pp. 143-155). Berlin: Springer.
3. European Council of Town Planners. (2003). *The New Charter of Athens 2003.* Firenze: Alinea.
4. Dutton, J. A. (2000). *New American Urbanism: Re-forming the Suburban Metropolis*. Milano: Skira.
5. European Commision, Commission Staff Working Document (2006). *State Aid Control and Regeneration of Deprived Urban Areas. Vademecum*. Document available on-line at: http://europa.eu.int/comm/regional_policy/index_en.htm.
6. Gruber, T. R. (1993). "A Translation Approach to Portable Ontologies". In *Knowledge Acquisition*, 5(2), 199-220.
7. Guarino, N. (1998). "Formal Ontology in Information Systems". In N. Guarino (Ed.), *Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998* (pp. 3-15). Amsterdam: IOS Press.

Colour plan for urban design

Mauro Ceconello, Mario Bisson, Cristina Boeri, Giorgio Vignati

Politecnico di Milano, Dipartimento INDACO, Via Durando 38/A, 20158 Milano, Italy
mauro.ceconello@polimi.it, mario.bisson@polimi.it, cristina.boeri@polimi.it,
giorgio.vignati@polimi.it

Abstract. The aim of this paper is to describe a research project concerning the colour planning in urban environment. The first part draws the research held in cooperation with the urban furniture departement of Milan City Council; the main goal was to collect the chromatic data of urban furniture inside a sample area in the historical centre of Milan and to use virtual reality to evaluate the proposed actions. There are some important issues that deal with colour in urban design: from the practitioners point of view it refers to how to choose a colour and why in submitted projects, and from city managers side how to evaluate different choices.
Some issues concern how to choose colour organization, how to describe the shared information and knowledge and how to valuate the consistence of colour projects.
Then we have investigated the potential of ontology representation in our application domain. We describe our experiences and considerations using this information technology tool as more expressive than the classic relational database; the goal is to be able to extract from the ontology representation knowledge guidelines and methodologies in order to verify coherence of the projects concerning either urban furniture or public space design.

Keywords: colour, 3D visualization, urban planning,

1. Introduction

The research focus is an ontology-based approach for colour planning in urban design. Starting from a project carried out for Milan City Council we had identified three different steps. The first related to the survey of the chromatic component in a central area of the city to provide a tool for the use of colour in urban planning. The second action, in relation with the former, aims to test the effectiveness of virtual reality VR techniques to manage urban visual and descriptive data with an effective and user-friendly approach. At last we had considered ontology based technologies for knowledge description, identifying operating tools and a layout hypothesis. Last but not least we propose an integration, still to be investigated, between ontology and tridimensional project representation capable of displaying and exploring data in a VR environment.

2. Managing the chromatic component of the urban furniture in Milan

The importance that the chromatic component covers in the architectural and environmental design revealed herself, also in Italy, in a strong attention towards all urban planning and design actions through the colour plan tool.

The Colour Piano of Torino on the late ‘70s, represents the first attempt in Italy to give a rational answer to the façade restoration problem on urban scale on the basis of an objective historical documentation.

The colour plans drawn up later in other towns were set up, till today, as a useful town planning tools to recover colour tones, original materials and manufacturing techniques of historical buildings, to ensure the maintenance and the conservation in time, contributing to define a coherent image of the town.

In this context the Urban Furniture department of Milan city Council entrusted a study to the Colour Lab of the Indaco Department of the Polytechnic of Milan focused on the management and the planning of the urban furniture chromatic component. Those elements and handmade objects, with their own characteristic, connote and characterize the public space and represent the "urban detail”.

The first part of the research deeply analyses the chromatic component in all furniture elements present in the sample area. The analysis and the mapping of the existing chromatic component was synthesized in a database aimed both at highlighting its qualities and as organization tool of the urban furniture characteristics. Our database, drawn in Fig. 1, is concerned about different kinds of information: visual features (picture and drawing), historical, cultural and technical data (materials, productive technology, maintenance cycle) and the coding of the perceived colour needed to manage the chromatic component.

In detail, the database manages the technical aspects related to the chromatic component using:

- NCS chromatic notation for all colours present in the object;
- other notations such as RGB and CMYK obtained by conversion from NCS;
- CIE and RAL codes, when provided by companies.

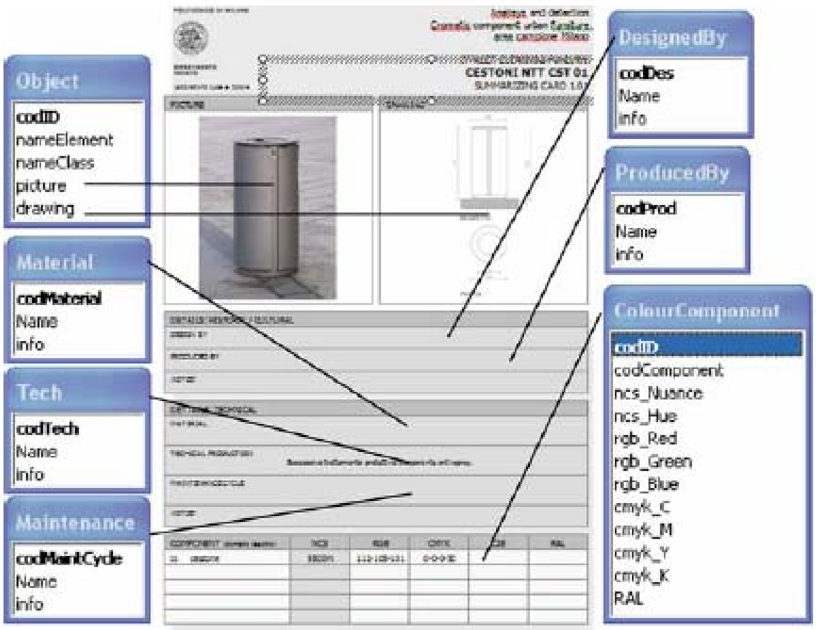


Figure 1. Report example and source data tables.

The decision to use the chromatic system NCS - Natural System Colour - in the subsequent planning phase is due to its representation as a common and shared language and as a support to visual data, through a real reference colour sample. The NCS is internationally recognized as well as the most diffuse colour standard in Europe. The information on RGB and CMYK is supplied with the purpose to offer a useful reference for colour representation in displaying and printing.

A plan which regulates the use of colour in the urban furniture elements, must also consider the relations inside a chromatic map. Therefore we arranged some synthetic reports to point out the perceptive chromatic dominant characterizing the furniture elements and the chromatic relationships inside elements belonging to the same functional group.

At last we believe extremely useful to create a visual representation of the whole range of the detected colours during the research in relation with the 1950 colours available in the NCS chromatic system: with this colour table it is possible to read how mostly used colours in the urban furniture could define some homogenous areas, both concerning hue and chiaro-scuro, compared with others barely used.

The study highlighted some data that form a necessary tool to perform a cognitive investigation; a preliminary action to define a colour plan to control the use of colour in the elements of urban furniture.

It's opportune to remark how a chromatic plan project concerning urban furniture components had to be considered as a detailed colour plan into a wider chromatic planning of the city.

3. Virtual urban design: a new way to explore the city

Once that the colour analysis had been carried out, a further step was to use VR applications to evaluate the correct use of chromatic variations through the design process.

The city possesses numerous pieces of information that are either descriptive or visual. In the first case, the widespread use of territorial information systems - GIS has made the management of the data related to the various features that make up the area of land considerably easier. In the second case, the use of digital models is a powerful instrument of representation and a valid form of help that enables people to see and appreciate both the current state as well as the hypothetical modifications even before the latter come into reality. Furthermore, it allows for the correct evaluation of outcomes and effects before a real intervention effectively occurs. This is even clearer in the case of territorial representation for which the level of detail (which is the fundamental guarantee of the 3D GIS simulations) adds an irreplaceable merit to design projects (Brail & Klosterman, 2001). Indeed, it is an instrument which is of the utmost efficiency since, a part from representing a particular environment; it enables us to freely navigate as well as to observe features of interest from a variety of viewpoints. Therefore, its indispensability can be best understood when a high level of detail and a total level of adherence to reality are mandatory. Moreover, this hyper-realistic model allows us to reach further and get to the very construction of virtual environments (Raper et al., 1993; Kraak et al., 1999).

The use of 3D models has been for years the object of investigation in the field of territorial representation as well as the inter-connected field of event simulation. With the introduction of programmes for automatic design, the usage of numerical models has attempted to replace their physical and studio homologues (Batty et. al., 1998). Consequently, the use of simulators that are more or less automatic has tried to create the right conditions to generate urban environments. Nevertheless, these programmes, based upon pre-defined libraries, find it hard to adapt to the morphological features of most Italian and European cities. Many Italian cities like Florence, Milan or Rome call for detailed and well-articulated models with particular characteristics that cannot easily be standardized.

Figure 2. Visualisation of Piazza Cordusio in Milano; model with textures rendered with the virtual theatre.

That means that volume extrusion models must be enriched with details and particular features which make them more similar to their real counterpart. This operation has been made possible by the creation of maps and facades of buildings (as well as photographs) which have been chosen for such experimental purpose. The photographs are used to create the necessary texture to apply to the models, a part from being particularly helpful in the definition of size and construction details. The last outcome in the modelling phase – the so-called "detail" phase – is, therefore, a digital object which reproduces its own corresponding figure in reality both in form as well as in appearance. Although the various steps have been subjected to a methodological approach, the time involved in each one of them should be taken into account as well. Indeed, in order for the research to be fully carried out, it is absolutely essential to evaluate the feasibility of the process as a whole as well as to revise the process step by step in order for the later to be as rigorous and scientific as possible.

A greater level of definition can be achieved by using 3D laser scanning systems which enable the acquisition of whole portions of buildings as clouds of points which can be later modelled into forms. Technological processes like the aforementioned, in addition to digital photogrammetry systems have been extensively used at the Virtual Prototyping and Reverse Modelling Laboratory at the Polytechnic of Milan and they have enabled us to perfect the definition of digital models, as required. However, at this point, we have to underline the importance of exactly defining the scope and the level of detail which are deemed necessary for our modelling. Indeed, since the models are simplifications and schematizations of reality, there is always a gap between reality and the level of detail in the actual model. If the level of detail is not sufficient, we run the risk of losing the main interactions and so the model will turn out to be incomprehensible and useless. On the contrary, if the level of detail is excessive, the model will become much too complicated and will end up being just as indecipherable as in the previous case. Therefore, the definition of the level of detail is one of the most important planning steps that have to be undertaken when these models are utilised.

The fundamental characteristic of digital models, apart from the photo-realistic simulation of reality, is that they are able to act as true virtual prototypes endowed with behavioural and performance similarities. They make us observe, simulate and analyse the design project (as well as its behaviour) in a much better way than offered by similar technologies, both from visualization and database point of view.

The further shift from the hyper-realistic model to the systems of virtual reality enables the planner to validate the planning decisions in relation to urban re-organization (including new forms of intervention in the pre-existing urban landscape) by using sophisticated equipments and the aid of stereoscopic visualisation (Ceconello, 2003). The models created in the previous phase are imported in suitable software that has been studied for the digital visualization of re-constructed environments on a large screen in a virtual theatre (Fig.2). At this stage, an interactive real-time navigation through the virtual re-construction clearly indicates all the potential problems that a planning operation may bring about in any particular urban area while offering the best type of solution in the most suitable context. It is easy to visualize the different proposals in their natural sequence and to single out the most appropriate. A further merit stems from the fact that results and planning activities are now provided for the benefit of a non-specialized public. In addition, because the outcome can be export in many formats, a variety of packages and applications can be utilised and link; for instance with GIS databases.

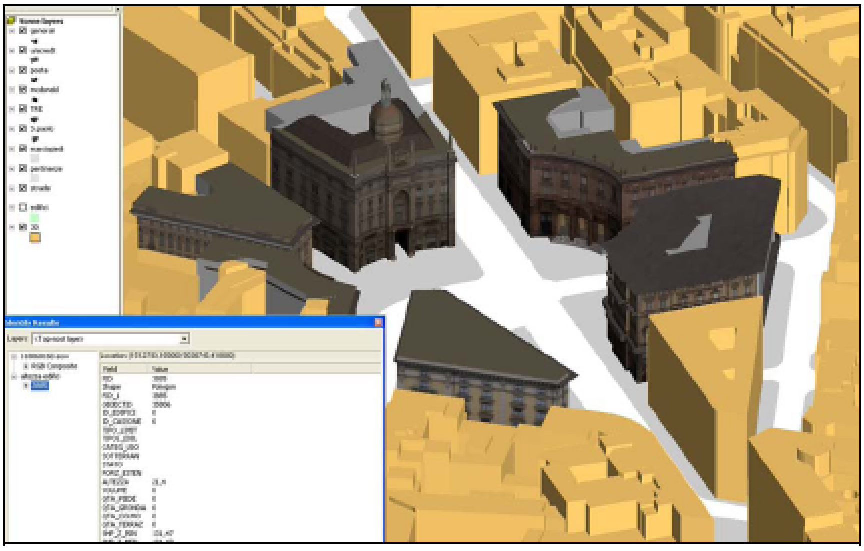


Figure 3. Query of a 3D model with textures in a GIS.

3.1 Experimental Results

In order to verify the aforementioned process, we have decided to create a model of Piazza Cordusio which is situated in the historical centre of Milan. The basic idea was to use the virtual reconstruction in order to evaluate the re-design of the site without doing drastic intervention by trying to understand if colour could help to improve the appearance and the perception of the site. Elements such as traffic lights, benches, lampposts, just to mention a few, can generate visual noise. For this reason, a purposive sampling of all the objects and a study of the colour map of the area was carried out.

Starting from digital cartography, the model was created while adding more accurate information from drawings and plans of the buildings. In order to add details on the façades, a lot of pictures were acquired with a digital camera in order to help for instance, to better define balconies, ornaments, and complex shapes. While creating the model, it was decided not to use laser scanners and photogrammetric techniques because, at that stage of the project, the focus was to validate the whole process from the urban planner point of view, without excluding the use of more accurate and sophisticated methods once its applicability has been verified and demonstrated. Once the urban scene has been loaded into the virtual environment, the user is able to navigate into the scene while visualizing the three-dimensional elements in stereo. By simply pressing a key, various scenarios can be studied such that the best proposal can be identified. Since the models are in the VRML format, they can be easily exported into other applications for complementary results. Finally, the textured models were integrated in the municipality GIS. In the present case, the GIS integration was completed with the ArcGis 9.1 environment (Fig.3). By following this approach, it is possible, in the long run, to create a digital model of the entire city.

4. Some issues, remarks, reflections

Throughout the research performed for the municipality of Milan some problems and remarks emerged and make us dealing with new research issues. Particularly for what it concerns the analysis and the survey of the chromatic component, all the opportunities allowed by digital technologies were carried out both for the support of data management to develop a knowledge base about colour regulation for public space and as a designer tool for surveys and directions supporting the project.

In fact, there are some important issues that deal with colour in urban design: from the practitioners point of view it refers to how and why choosing a colour in submitted projects, and from city managers side how to evaluate different choices. In other cases it concerns how to describe the shared information and knowledge and how to valuate the consistence of colour projects. Our activity had been carried out in a pragmatic way, referring to our training and know how.

We had investigated municipality purposes together with the survey process and user requirements, and summarized with some keywords all the remarks and main critical points that information systems could, on our opinion, fulfil:

- knowledge structure must be shared;
- choices, rules, and constraints must be explicit;
- rules consistency and coherence must be verified;
- all requirements and needs must be modifiable;
- tools have to be easy to use and interactive.

According to these requirements and keywords we investigated on-line scientific literature (Pohl, 2000) and exploited two relevant and attractive research areas: the former relates to techniques and tools for knowledge modelling and processing, the other deals with visualization both for design scenarios and as an interface for knowledge base. A brief survey highlighted the increasing use of buzzword like *semantic* and *ontology* with different meanings for knowledge representation (Gruber, 1995), and *virtual urban design* to describe and evaluate visual aspects. As we faced a variety of subjects, we had bounded our study analyzing two topics with a bottom-up approach: to build ontology for colour application, and to experiment a 3D virtual environment to interact with city scenarios.

Our main goals are:

- to define information system architecture according to the last research directions;
- to test ontology and visual representation using only open source software and tools;
- to verify effectiveness both from the city manager and practitioners point of view.

Now we describe the building of the ontology and we present some remarks that came out during our work. We used Protégé as one of the widespread ontology editor software and largely known in this research area. This developing environment is supported by a dynamic an authoritative scientific community and by high level tutorials and free documentation.

4.1 Building the ontology

As first attempt we imported an entity-relation schema from an existing database built as schetch up in Fig.1. The translation from database schema to ontology is fast, but the result is less expressive than ontology representation. The classes (objects conceptualization) include also the bridge tables that have no meaning in real world but only in relational schema. Although this approach could be useful if already exist a database developed and filled in, we have chosen to give up this activity and decided to build ontology according to guidelines references (tutorials). We used an OWL plug-in available in Protégé, as this is the W3C standard shared among the community with guarantee of maintenance and evolution. Furthermore this environment can be linked to applications of automatic reasoning, necessary to work with the ontology to provide logical inferences. In the ontology literature we have found the ontology definition often schematized as taxonomy + (relations, rules, constraints) = ontology; then we started to define our taxonomy. The first step was to identify the classes (abstractions about the world) and the relationships among them to build the class hierarchy (these are known as IS-A relationships).

From the study concerning urban furniture components, the classes arise in natural way.

With a visual editor, as Protégé and its wizards, the ontology building activity is user friendly also for newcomers. Other classes arise from other data objects concerning the urban components (producedBy, DesignBy, Material, MaintenanceCycle…). Figure 4 illustrates a fragment of ontology and the graphical visualization obtained with Jambalaya plug-in; the visualization provides an alternative representation that let the user understand the concepts in a more intuitive way.

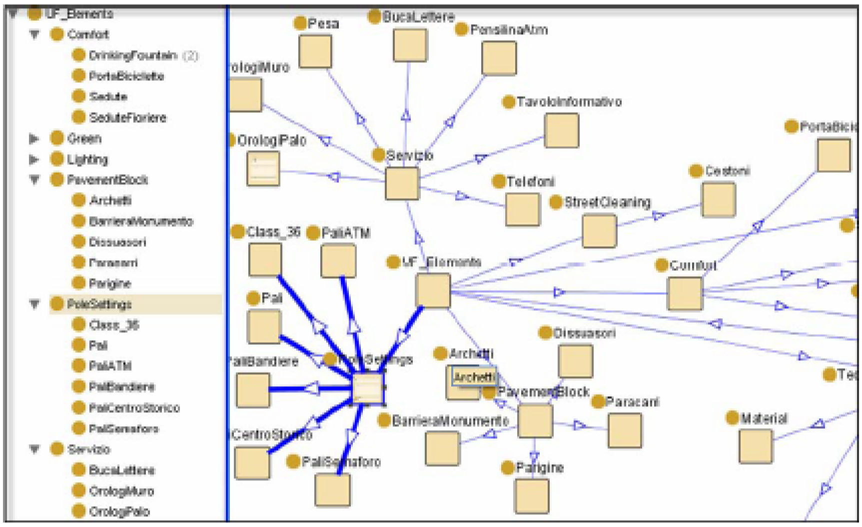


Figure 4 Example of visualization using Jambalaya plug-in.

The ColourDomain class represents the information about the component colour. The NCS class store the information about the standard NCS. The main problem is how to organise a standard. We think that it is necessary to build an independent domain ontology of colour, that should deal only with the description of colour systems (standard as NCS, Munsell etc.), its linguistic description and all specific topics and issues around colour. However our goal was to regulate the colour in public spaces, then we assumed a naïve definition of a ColourDomain class. We expanded only the NCS colour system class describing (at present) 1950 colours. We can define a property hasNuance and hasHue with string values.
Another class is TownConcepts that conceptualize the town objects and their relationships as the main goal of the Towntology Project (Keita, et al, 2004); in our case we have supposed to use an existing ontology developed and managed, because in the real application different ontologies from different implementation will be link together. We have identified a possible solution starting from the available tools and the OWL features implementation:

- Create a forbidden colours list. It defines a class that allows to constraint hasColour property of a class. So if we want to state: "in the city all street objects must not use the colours colour1, colour2…" we can express this statement defining a class with a list of individuals (forbidden colours) and bind the hasColour property of street class.
- Define a list of relationships rules as "the colour of information pole must be different from the colour of traffic light pole". We must define a class "colour of traffic light pole" and put a constraint to hasColour property of traffic light pole.

Then we have to analyze a generic statement corresponding to a rule on colour objects, building a set of rules in DL (description logic) for describing *primitives* or *define* classes, as a set of necessary and/or sufficient conditions statement, after defining and use *property restriction* to relate the colour property to these classes. These basic statements can be made more complex using all set operators available in OWL-DL language. This formalisation allows to validate the consistence of ontology, using a reasoner as Racer, Pellet, Kaon2, and to exploit new knowledge later (by inferences), implicit but hidden in ontology representation.
From the city manager point of view these formalized rules allow to reach a non contradictory plan. On the other way, using a similar tool, the practitioner can query the system (maybe in a semantic way in the future) to obtain a support for a project according to the colour plan.
At present we are filling in the ontology with objects, properties, individuals and setting simple test rules to verify our approach.

5. Future work and conclusions. Visual Urban Design and Ontology

Our exploration in the technologies previously described, let us carry out some remarks on future research. We think of ontology as a tool for practitioners to draw up a non contradictory colour plan and evaluate projects, but for this purpose we must be

able to define a simple but realistic colour plan as a set of statement describing rules and constraints; we argue that probably we will not be able to describe in an analytical way every single part of the plan either for a limit of the available tools or to the impossibility to organize the plan into several detailed actions.
Once that all surveys and analysis had been done, and the colour plan had been implemented, it is our opinion that virtual prototypes are a good tool to evaluate the most relevant scenarios in real-time on a cognitive basis, but to be more effective and useful they have to be supported by a knowledge-base representation. As a matter of fact, the planning action needs a support either in the analysis or in the presentation of the results. In the normal course of activities, the use of GIS supports decision making is extremely useful for managing planning tasks.

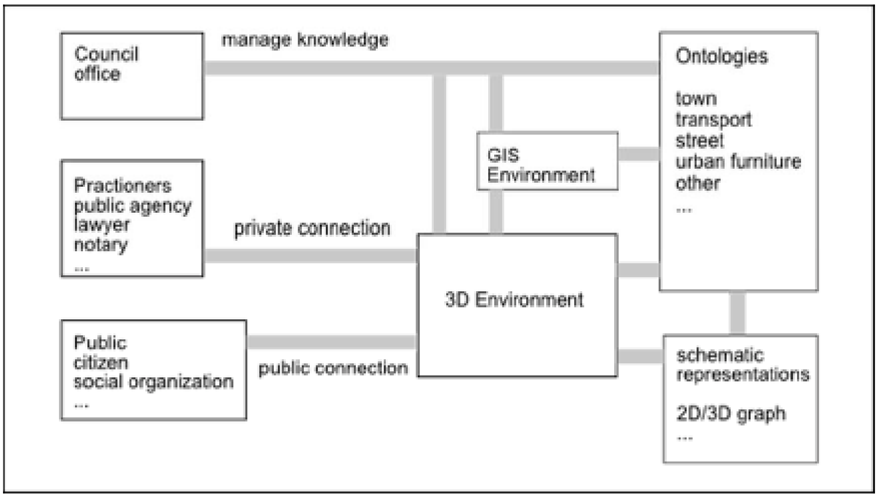


Figure 5. Scheme describing an integration of activities through 3D environment.

Virtual reconstruction and presentation of possible intervention on the fabric of a city, gives the city manager the chance to visualize its impact. Therefore our intention is to suggest an integration of these representations, ontology and visualization techniques, in the normal practice of a planning office in a City Council. (Fig.5) The project can be considered a test bed to verify a process in which these items could be put together in order to create a "best practice" in urban planning.
The research is currently in an ongoing phase and, as we speak, all the preliminary investigations supporting the colour plan and the visualization techniques had been applied to the historical centre of Milan, and evaluated positively both by technicians and city managers employed in urban planning. At present, we are evaluating the ontology application, but we need more data available to do a meaningful test. However from this case study, there is an opportunity to initiate more complex activities in order to evaluate various strategies and go in depth either in the basic topics or in the application tools.

References

1. Batty, M. (2007), Model cities, *CASA Working Paper 113*. London: Centre for Advanced Spatial Analysis, University College London.
2. Batty, M., Dodge, M., et. al. (1998), Modelling Virtual Urban Environments, CASA Working Paper 1. London: Centre for Advanced Spatial Analysis, University College London.
3. Brail, R., Klosterman, R. (2001), Planning support systems: integrating geographic systems, models, and visualization tools. Redlands, Esri Press.
4. Ceconello, M., (2003), Strumenti e tecniche di visualizzazione. In M.Gaiani (Ed.), Metodi di prototipazione digitale e visualizzazione per il disegno industriale, l'architettura degli interni e i beni culturali. Milano, Polidesign.
5. Ceconello, M. Paquet, E. (2007), Virtual urban design, in Vric Laval Virtual 2007, Proceedings 9th Virtual reality International Conference, 18-20 April, Laval, France, ISBN 2-9515730-6-5, pp. 177-183.
6. Gruber T. R., (1995),. Towards principles for the design of ontologies used for knowledge sharing. International Journal of Human-Computer Studies, 43:907–928.
7. Keita, A.K., Laurini, R., , Roussey, C., Zimmerman, M. (2004), Towards an Ontology for Urban Planning: The Towntology Project. In Proceedings of the 24th UDMS symposium, Chioggia, October 27-29.
8. Kraak, M.J., Smets, G., Sidjanin, P. (1999), Virtual reality, the new 3D interface for geographical information systems. In: Camara, A., J. Raper Spatial Multimedia and Virtual Reality. London, Taylor & Francis, pp. 130-136.
9. Pohl, J., (2000), Transition from Data to Information, CADRC Technical Paper, Nov.27.
10. Raper, J., Kelk, B. (1991), Three-dimensional GIS. In: Maguire, D.J., Goodchild, M., Rhind, D.W. Geographical information systems: principles and applications. Essex, Longman scientific and technical, pp. 299-317.
11. Raper, J., McCarthy, F., Livingstone, D. (1993), Interfacing GIS with Virtual Reality technology, Proceedings Association for Geographic Information conference, Birmingham, 3, 25, pp. 1-4.

Web references

Protégé.stanford.edu

Generating Urban Morphologies from Ontologies

Luca Caneparo[*], Mattia Collo[*], Davide Di Giannantonio[**], Vincenzo Lombardo[**], Alfonso Montuori[*], Stefano Pensa[*]
[*] High Quality Laboratory - Territorial Integrated Project, Politecnico di Torino
[**] Department of Computer Science, Università di Torino

Abstract. The paper presents the ongoing research work on a software system for supporting the exploration of the numerous and often interrelated factors that can affect the urban design.
The present implementation of the system supports the simulation of different urban scenarios in relation to the uniqueness and constraints peculiar to a design and a site.
The paper considers our ongoing research work to formally represent the implicit and explicit knowledge used by means of ontologies. The ontology semantic system administrates a set of rules and relations among urban entities. To this aim, we are dealing with different issues: all the factors involved in the urban design process cross various domain knowledge, from different competencies and sources; the knowledge is both semantic and procedural.

Keywords: urban ontology, generative systems, multi-agent model.

1. Introduction to urban design

Numerous and often interrelated factors potentially concur to urban design (Alexander, 1996). The number of these factors is recognized large, furthermore they are often ill-defined, and can assume different meanings relating to the urban and cultural context. They often play in interrelated ways: "depending on how it is acted upon by other factors and how it reacts to them" (Jacobs, 1961). Jane Jacobs also stated: «Cities happen to be problems in organized complexity, like the life sciences. They present "situations in which a half-dozen or even several dozen quantities are all varying simultaneously and in subtly interconnected ways."»

To model and represent these factors and their interconnections we are implementing an ontology semantic. We are not committed to model each and every factor, but to support the designer in defining her/his significant factors and their interconnections, then freely exploring the dimensions of the space of the design solutions generated by the system.

The dimensions of the solution space increase with exponential law, thus the system is implementing:

- Ontology-based to orient the system knowledge towards a domain interpreted through rules of inference and logic approach;
- Generative system for exploring the very large number of design solutions (Garey and Johnson 1979).

This approach substantially differs from an optimization process, where the space of the solution is explored by a minimization algorithm in order to recognize the solution that best fits the given criteria. With the proposed methodology, the designer can explore the high-dimensional space of the solutions generated by the system according to the given constraints and criteria, in order to find the solution/s that best fit the implemented criteria but also further ones that the system is not considering, while the designer is. For instance the designer can be considering aesthetical criteria or ones relating to the site or the milieu etc. These additional criteria possibly are not implemented in the system explicitly.

2. Characters and complexity of urban design

We explore the urban design by process, considering it an ill-posed domain, bounded on one side by the requirements and on the other the strategies used by the designer, in order to define a frame for the complex and interconnected factors ranging from, for instance, principles, visual representations, numerical data, mediations, analysis, as well as quantitative requirements, urban regulations and standards.

We propose an ontology semantic system approach to bridge the representation of the factors and the designers (the end-users of the system).

The generative system supports the designers in simulating different urban alternative scenarios in relation to the uniqueness and constraints peculiar to a project or a site.

These factors produce different configurations relied with typological and measurable inputs, and generate a large number of solutions.

For exploring the high-dimensional space of design solutions the system offers both tools for searching it and for structuring it.

The system implements strategies and technologies for searching the high dimensional space of design solutions and for defining and presenting structures in the space of solutions. We do expect structures in the space of the solutions to be a promising methodology for more effectively supporting the designer in recognizing the "neighbourhood", where s/he expects to be the design solutions, then to explore this neighbourhood to evaluate the small variations in the constraints and criteria that defines them. To make the system more flexible we improve the semantic of relations with the ontology language. Thus the final aim of the generative system is not defining one, the best fitting solution, but to drive the designer towards one or more "neighbourhoods" in the space of the design solutions, where s/he can gather suggestions or directions in designs that otherwise s/he will possibly not have been considering.

3-D scenarios allow urban designers to explore the space of design solutions defined, for instance by typological, quantitative and performance factors. The simulation makes it possible also to structure the units in relation to different typologies (e.g. type and quantity). Thus we can explore the multiple simulations of scenarios to understand what happens to the system when we change the quantity of a specific typology for a site. Obviously this change affects the interrelated types.

We can verify the relations between the building placement and the ambient factors evaluating the daylight comfort for each building and unit in relation with its neighbourhood (Cheng et al. 2006). Our generative system is hybrid: a software seamlessly integrating the generator of a large number of design solutions and the browser for searching and structuring the high-dimensional space of the design solutions, according to variable and customizable factors defined by the designer. We constructed a system of rules and relations using ontologies. This approach allowed a flexible customization of the system to various requirements and situations. Indeed we could build a set of rules changeable to different requirements. Using the ontology logic and semantic the complex a set of rules has been intelligible by digital systems.

To this aim, we are dealing with different issues: all the factors considered cross various domain knowledge, from different competencies and sources; the knowledge is both semantic and procedural.
We are experimenting different strategies for the formalization of the ontology:

- top-down, starting from the general notions (meta-design in Bazzanella 1971) to the scale of the building, discredited in *voxels*;
- bottom-up, starting from the analysis of cases, defining several layers of ontologies (after Bradshaw 1992), each referring to a specific case, and then representing the overlapping information without having to build a general, wide domain ontology;
- a middle approach is possible (Uschold 1996), but we found less references to present design practices.

3. Related works

We are recognizing two principal directions of research in ontologies dealing with our approach to urban design.

The first explores ontologies as tools to ensure the cooperation between several urban actors and interoperability between different urban data bases. A system based on ontology is promising, because it allows dealing with the semantic dimensions of the urban design. An ontology can be defined as a semantic graph regrouping concepts (Keita, Laurini et al. 2004).

This research developed a prototype system -*Towntology*- able to navigate the ontology displaying its graphical form and query it. Several approaches were used for the construction of the ontology. They decided to work first on street planning and on mobility. Already more than 800 concepts (French language) are introduced in their ontology, with 21 types of relations between concepts. This study describes the difficulties of creating this ontology, and the way (tools) to solve the problem. The solution is to lay out specific ontologies for each actor (with each field), and a global ontology. Thus any request and data exchange must be done via these ontologies. The research works within the framework of the Towntology project aims at writing urban ontologies in a visual form. This visualization is a way to be explored towards a co-operative system. From the descriptive point of view of ontologies, an extension of XML is designed and used.

In this direction Metral's (2006) study aims to contribute to a better communication between the various actors involved in an urban planning process due to the complexity of the urban planning process. The research pones three goals:

Semantic Integration of the objects (data and documents) related to an urban planning project to create an urban knowledge base. Indeed urban data and documents (spatial data from GIS, 3D city models, texts, maps, plans, pictures) are represented in their system by an ontology with concepts and instances related to the data and documents so as to present an integrated view of the project.

The second purpose is to develop an interface to visualize the content of the knowledge base. They identify visual objects to define from the objects of the knowledge base. At the same time they describe objects of various kinds: documents, 3D objects but also abstract concepts to delineate representations for non geometric entities.

Third aim of this investigation is the adaptation of the interface to user profiles and centres of interest in order to access to the urban knowledge according to different viewpoints of various actors (urban planners and designers, politicians, citizens, communities, etc.) with differences in terms of knowledge background and vocabulary. They set an adapted interface, a system of visualization of what is useful or relevant to the user exploratory interface to understand an urban planning project in a personalized way.

The second approach defines a use of ontologies to exchange information between computers and to make the systems flexible for different domains of use.

Decision Support Systems (DSSs) (Schevers, 2006) can help clients, by presenting design solutions using virtual reality, and by offering relevant feedback such as costs, energy usage, distances and density. Using these systems, clients and stakeholders can run ‘what-if’ scenarios on different design solutions and see the consequences of their changes.

To make sure that a DSS is suitable, it needs to cope with the project situation at hand. This research investigates DSSs that can be built 'on-site' and that are flexible enough to model project-specific issues. A custom-built DSS can be tailored for each project. Therefore this study explores the use of ontologies. The use of an ontology enables further reuse of software components and software models. With this approach, several parts of software components can be reused and put together to quickly develop a customised DSS.

4. Ontologies for urban design

4.1 Representing knowledge

The design process needs an adequate communication to relate with the stakeholders, sharing ideas with the team, until the choice of the way to present the project. The knowledge representation of the system improves the research and definition of a language capable of expressing the specific knowledge and information used during the design process. From this perspective we can consider a model -system of representation- as a language capable to share and represent information. A model is not a ‘picture’ of the problem, but rather a device for the attainment or formulation of knowledge about it (Kaplan, 1963). Indeed, sometimes the most important outcome of the modelling process may not be the model itself, but rather the insight we gain as we struggle to articulate, structure, critically evaluate, and agree to it (Moore & Agogino, 1987). Therefore the purpose of the process is not only a proper representation of a system but rather how this process can help us to better understand the domain of knowledge represented in it.

“Conceptual modelling is the activity of formally describing some aspects of the physical and social world around us for purposes of understanding and communication. Such descriptions, often referred to as conceptual schemata, require the adoption of a formal notation, a conceptual model in our terminology. Conceptual schemata can also be used to communicate that common view to newcomers, through a variety of graphic and linguistic interfaces.” (Mylopoulos,1991)

To support the digital tools for urban design being flexible to different types of designs, they must rely on a highly customizable and versatile language. This versatility can be pursued by a system grounding its components and attributes using an ontological model. The ontology represents a way to describe a standardized system of information and relations expressed in unambiguous vocabulary. The quest for large-scale knowledge sharing provides a strong motivation for common ontologies. Careful design of the ontology and the careful definition of the terminology also benefit developers and users of the system (Bradshaw et al., 1992).

Ontology, in the philosophical meaning, is a systematic account of existence as perceived by humans. “The more traditional term is Aristotle's word category, which he used for classifying anything that can be said or predicated about anything." (Sowa, 2000) “. One of the main roles of ontologies is to disambiguate, providing a semantic ground, a consensual conceptual vocabulary, on which one can build descriptions and communication acts” (Gandon, 2002).

The flexibility of ontology language structured on a consensual conceptual vocabulary makes it an ideal tool to share and operate with knowledge in A.I. domain. Klein (2002) defines an ontology as conceptualization of things in a computer interpretable format. By relating machine-interpretable information to each other, inference support can help consistency and can support further interoperability (Gruber, 1995).

4.2 Ontologies as urban systems

We agree with Coppola (2005) in transferring the design activity to the domain of A.I. is appropriate to clarify the design process. The process can be declarative or procedural. Declarative knowledge operates to verify operations, it reaches a more efficiency in terms of technological examination to allow a description of the computational process. On the other hand procedural knowledge works on the syntactic level, it improves more quality intended as propriety and suitability verification in relation with the domain to allow a description of design process.

We are dealing with two major aspects:

- transcription of a formal ontology, which represent urban design process;
- definition of a generative design methodology based on A.I.

To delineate the formal ontology we have to clarify the procedural paradigm of the system. We should define the process which describes the design (not its morphology). It is central for this stage comprehend the value of inference determined by the single choices along the design process. Therefore it is important perceive how each alternative can affect the further ones.

The exploration of tools for urban design process reckons the urban forms as organisation of spaces ordered by a system of relations and hierarchies making in their rapport a sort of language. A semantic with rules and structures.

The study has been divided in three principal steps:

- analysis of real cases to infer examples for defining the typologies;
- identification of some simple typologies which refer for all considerations about urban units;
- analysis of real existing buildings constituted of a mix of different typologies, according to these examples we define the relationships among typologies as parts of complex systems.

The semantic model includes specific information to arrange different units. For each typology are defined dimensional parameters as width, depth and rise. These values define dimensional ranges, that is categories within the dimensions varying between a maximum and a minimum, directing towards best fitting values. This way the designer can easily decide, further to her/his requirements, the optimal parameters that the system must pursue. This approach assures high flexibility and adaptability to different situations, whether external inputs defined by clients or economical constraints. Thus the designers can simulate and exploring a wide range of different solutions in terms of typology and dimensions.

Next pass has been defining the spatial relations which identify a building system. In this perspective the research has been oriented in analysing *simple* typologies infer from *simple* buildings. So a series of relations has been determinate for describing less complex systems which can be straightforward to be defined and represented. A number of typologies have been described from the semantic and conceptual point of view. Hence it is a set of relations which describe how each typology works outlining different groups of activity holding in diverse typologies (Coppola, 2006).

The research starts from the study of urban milieus, to highlight the organised and interconnected among space and built, parts of the city and functions. We thoroughly consider the structure of buildings, as a formalised system of relationships among instances. The analysis of real cases aims to explore a sort of system of rules which defines the base for generating a variety of buildings, differing among them, but pertaining to a common instance or class.

Therefore the relations between parts of the buildings have been analysed with the definition of an ontology model. With this kind of approach, it becomes possible describing connections among *simple* functions, i.e. unitary one. We define a set of conditions specified by the role of each part, in relation with the others and the whole building. This allows the ontological system to classify and organise parts and relations. This approach shows how each single part / function of a building must follow different rules according with its properties and instances.

The ontology has a high level of adaptability for describing multiple configurations and conditions. Thus it can express behaviours which are part of similar solutions but which produce different formal results. Through the ontology language we can manage quantitative value, which describe the typology dimensions according to other typologies. It is possible to explore relations occurring among typologies depending on quantities and dispositions. For instances with a fixed quantity of residential surface required, quantities and relations between elements present in the system will change according to the residential typology chosen.

As we consider the urban system the field of application of our research, this system can suggest us rules for development indeed as rules for alteration. The system can be highly customizable, to define real and specific cases. On the other hand it can demonstrate flexible in the research of solutions, to allow the exploration of creative solutions, far from ordinary results, suggested by the practice. Of course it permits to control all the input and verify possible irrelevant condition in imputing or organising data. Also is possible to rethink all the rules of growth which generated insignificant circumstances in way to revise the strategies of design during the design process.

5. System implementation

5.1 Ontologies representation

Now that the purpose and the scope of our ontologies have been stated, we need to explicitly represent in some formal language the conceptualisation captured. Typically, this will involve (Uschold, Gruninger 1996):

- committing to the basic terms that will be used to specify the ontologies (*e.g.* class, instance): they will form the meta-ontology of representational terms used to express the main ontologies;

- choosing a representation language, capable of supporting the meta-ontology;
- specifying the ontologies using the chosen formal language.

Following these guidelines we discovered that, at least for the ontology about the building typologies, a frame-based knowledge model is the best way to formalize the domain of interest.

We chose the Protégé software [Protégé], an open source knowledge-representation system, because it supports constructing and storing frame-based domain ontologies and also because it is widely known as a flexible, well-supported and robust development environment. It has been used in the past years in several successful projects and has also been used to implement a transcription of the Suggested Upper Merged Ontology proposed by IEEE (SUMO).

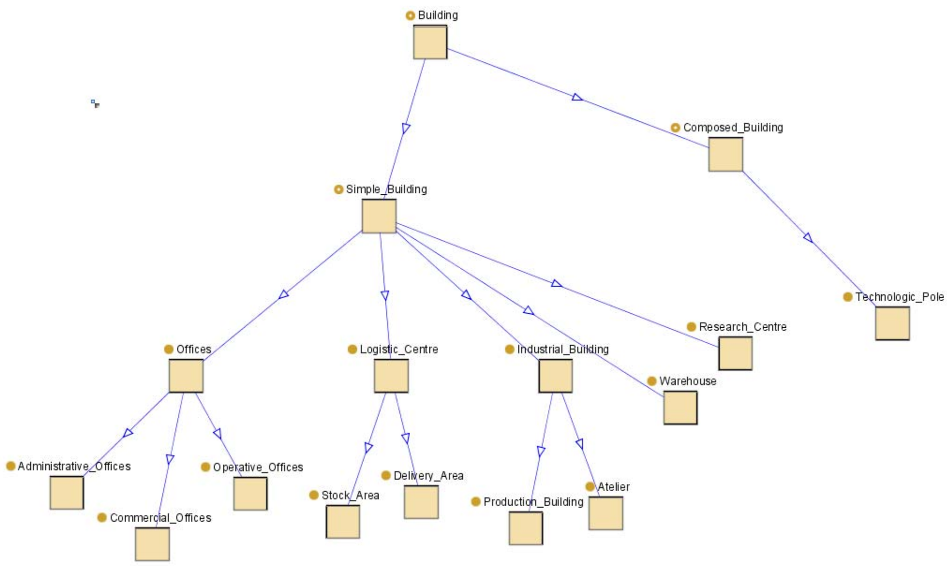


Fig.1 The building typologies ontology coded in Protégé.

The Protégé meta-ontology defines classes, slots and facets. Classes are concepts in the domain of the discourse and constitute a taxonomic hierarchy: the concept of *inheritance* let us to state that if a class A is subclass of a class B, then every instance of A is also an instance of B. Slots describe properties of classes and instances and are defined independently of any class; when they are attached to a frame, like a class, they can have a value. Facets are the way to specify constraints on allowed slot values; they include cardinality constraints, restrictions on the value type (*e.g.* integer, string, instance of a class), minimum and maximum value for a numeric slot, and so on. The restrictions defined in the facet apply to a single attachment of a slot to a class frame.

Using the Protégé meta-ontology as knowledge model for our building typology ontology, the hierarchy of the classes represents the ISA (subsumption, specialization) relations between the concepts of the domain. Other types of relations like PART-OF (aggregation), RELATED-TO (general association), and so on will be coded with slots; a substantial benefit in defining a formalization of spatial relations comes from a recently presented taxonomy of topologic, directional and proximity relations that we plan to introduce in our ontology (Mele et al 2007).

5.2.1 Interpreting ontologies

Elementary our generative system consists of the representation of the plots, rules and typologies that describe the methodology for aggregating the units in buildings/s, and evaluation parameters. The complete set of rules was defined by using the ontology model.

Protégé provides an interface which other applications can use to access knowledge bases. These applications need not use or display any of the Protégé user interface.

The programmatic interface to Protégé projects and knowledge bases is through the edu.stanford.smi.protege.model.Project class in the protege.jar file. This class has a getKnowledgeBase() method that provides access to the contents of the knowledge base. We enforced a closer mapping between our generative system and the Protégé model by importing in our inferential engine the original class structure of the ontology model. Protégé provides several mechanisms that generate such classes (Knublauch,2003). For example, you can directly generate Java classes for it with the BeanGenerator. We exported our model to UML and then generated Java classes with the Poseidon for UML tool. Poseidon product uses Velocity as a template tool in order to generate Java code from UML diagrams that you design using their tool.

5.2.2 Multi-agent model to develop morphologies from ontologies

We define rules as mandatory and non-violable set of constraints, which describe the urban regulations and the geometrical features of the plots. Besides, we consider on the same rank the area covered by the building and the layout of the site fitting the requirements of the design. Combining these parameters we were able to address the development and the distribution of the units into building volume/s on the site. The space of the design solutions generated is directly related to the mandatory rules, e.g. from the urban regulations as maximum height, distance from other buildings, etc.

Despite of a number of rules, the space of the design solutions is still high-dimensional. The space of the design solutions is generated by a multi-agent model. Agent-based modelling offers a flexible way to study the behaviour of mathematical models. In agent-based models the time evolution of all the system emerges from the level of the individual agent’s action and behaviour.

Agent-based modelling has demonstrated effective in economic, social and natural sciences, and design (Gero and Fujii 1999), especially when it is not possible to define an analytic solution to the problem.

We designed and implemented the multi-agent model in Java Swarm environment (Minar et al. 1996). The input to the multi-agent model is generated by the relationships between “local” actors (e.g. the owners of the land, dwellings etc.), “global” actors (investors, public decision-makers etc.) and their interactions (market, social, etc.). We can simulate different urban dynamics at the building scale considering as input the following factors: typology, plot edge, total building volume, front and depth dimensions, floor number and height, number of buildings, minimal distance between buildings, minimal distance from the edge

The multi-agent model generates and evaluates the suitable solutions to the input problem which satisfy the hard constraints imposed by the input factors, considering as evaluation parameters the following: solar availability on building façade (irradiance performance), ratio between floor gross surface and open surface, ratio between floor gross surface and façade surface, ratio between floor gross surface and roofing surface, construction cost, land value.

After Hillier and Hanson (1984), we map the plots on a three-dimensional grid. For simplicity, we use a cubic grid (for instance, 1 by 1 by 1 m). On the other hand this simplification has demonstrated flexible enough to represent concave plots (e.g. in historical cities) or buildings with non Euclidean envelope.

We define a building agent for each building. Each building agent starts from a building with minimum, suitable layout and tries to optimize the layout adding new units-cells to the building. The starting seed location of building/s is random. We decided to use an exhaustive research strategy to exhaust the space of the design solutions, reducing pruning rules to a minimum. In this way we were able to generate a large number of initial, random seed solutions to explore unexpected solutions and relations between the units.

Since each building-agent can create a proper schedule, it is important that all of these activities are coordinated in a logical way. A top level agent schedules the action of each building-agent. The top level agent also manages the user interface, scheduling the updating of the graphical display (see figure 2).

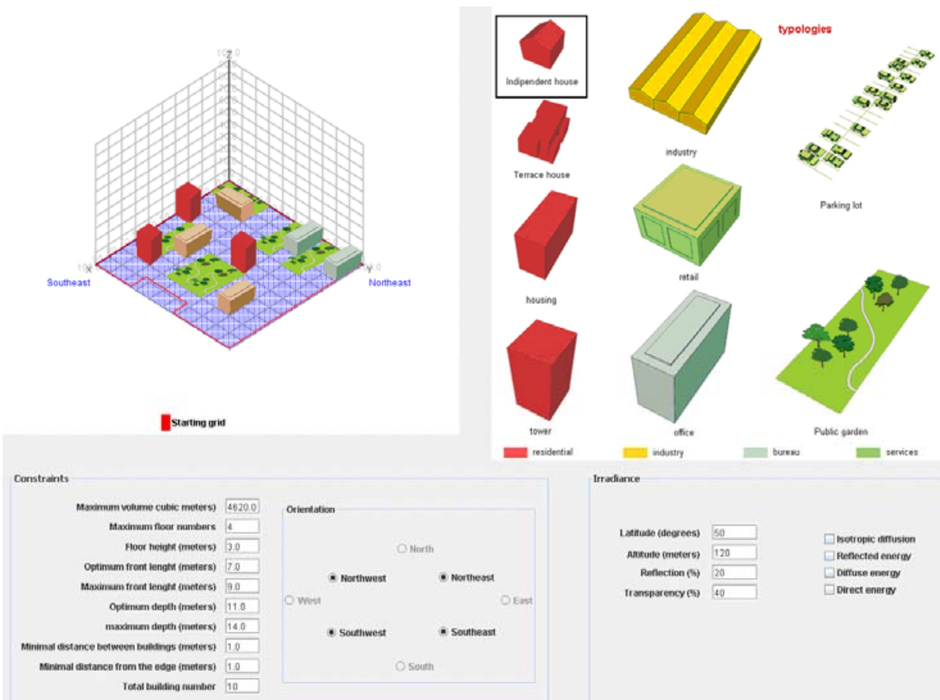


Fig.2 The graphical user interface of the generative system

This multi-level integration of schedules is typical of Swarm based models, in which the simulation can indeed be thought of as a nested hierarchy of models, in which the schedules of each agent are merged into the schedule of next higher level (Johnson 1999). The simulation proceeds in discrete time steps. As the simulation proceeds, the building agents update their state and report their state to the observer top-level agent. Auxiliary agents that facilitate the work of the building agents, can be instantiated if more than one typology is present.

The building agent uses a cellular automata to add cells to the building and communicates with the other agents in order to respect the hard constraints on distances between buildings. A cellular automaton is a collection of cells on a grid of specified shape whose state evolves through a number of discrete time steps according to a set of rules based on the states of neighbouring cells (Weisstein 1999).

Two of the most fundamental properties of a cellular automaton are the type of grid on which it is computed and the number of distinct states (usually represented with different colours) a cell may assume. We defined the cellular automata used by the building agents on Cartesian grids in 3 dimensions with binary cells representing the occupied/not occupied state. The user can decide the dimensions of each cell (for example front by depth by height = 1 m by 1 m by 3 m).

In addition to the grid on which a cellular automaton lives and the states its cells may assume, the neighbourhood over which cells affect one another must also be specified. In this case the most natural choice could seem the "nearest neighbours," in which only cells directly adjacent to a given cell may be affected at each time step. Instead, we decided to use the 3-d generalization of the von Neumann neighbourhood (a diamond-shaped neighbourhood), because in this way the automata update algorithm was less computationally expensive than the nearest neighbours one and able to accomplish the requested task.

The automata add a new cell in the position which maximizes the score. When the first floor of the building is completed the automata proceeds to complete the remaining floors. Each time a cell is added the building agent communicates to the top-level agent the new state of the automata. This process continues for each automaton until the total volume requested for the site is reached, as checked by the top level agent.

6. Case Project: Mirafiori Area in Turin

We have experimented the methodology and system in various case projects. Here we present a project at the urban scale: the possible evolution scenarios for the Fiat Mirafiori factory in Turin, Italy. The case project outlooks the conversion of a portion of the factory, since the Fiat Company has dismissed a slice of about 310,000 m2, of the total, huge area: about 3,000,000 m2. In the historical evolution of the site, the considered industrial area has become an integral part of the urban milieu of Turin.

The Mirafiori project may become an opportunity for the city to explore a number of possible scenarios both on the expectations on the production models and on the perspectives for the metropolitan area inserted in a larger network of cities. Thus we consider the case project as the act of exploring advanced indications, representing them beforehand by next-decades scenarios. The aim of these scenarios is not foreseeing future urban morphologies and policies, but to expose factors that can contribute to the shaping of a large area: making these factors visible, so that the various actors, involved or with responsibilities in the process, can recognize and weight them.

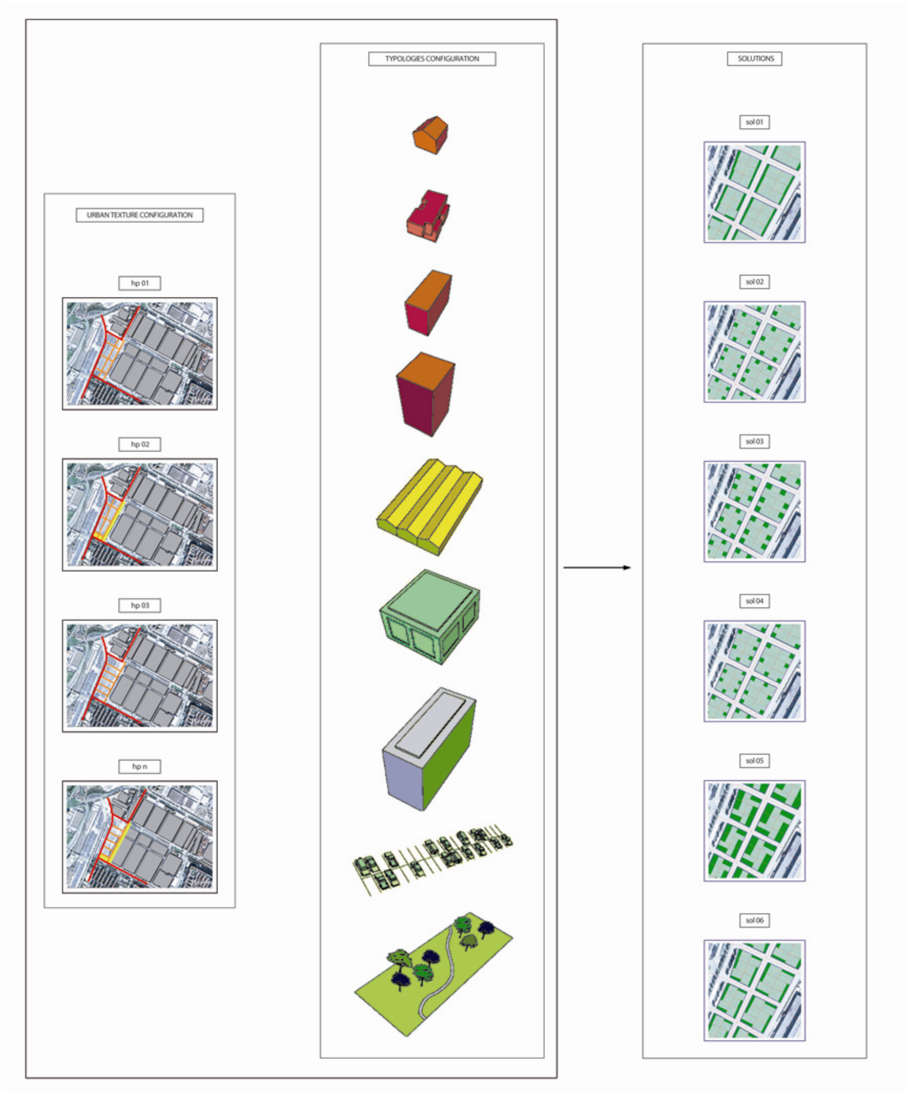


Fig.3 Matrix representation of a set of future scenarios.

The scenario making for this project has been a way of understanding the dynamics of the area and city, consequently trying to identifying the leading factors that can drive the dynamics.

The simulation of these scenarios with our generative system aims exploring:

a) the hypotheses regarding the future use of the area, as the result of an open decision-making process, aimed to define a strategic plan for renewing the site (Spaziante 2006);

b) the urban morphology as the subject of the scenario generation and evaluation, to support the designers', planners', decision-makers' and citizens' structuring and describing of the relations between the factors that shape the morphology at the urban scale.

The simulations of the four scenarios with our generative system allow us making visible the relationships among alternative destinations, typologies, and variable volume, height, distance between buildings, plot edges etc. Using ontologies we built the set of relations between parts of the system. The ontologies have ensured the easy experimentation of different sets of alternatives. Different scenarios were implemented and simulated to explore the respective space of the design solutions.

7. Conclusions

The system aims supporting the early evaluation of alternative solutions from the initial phases of the urban design process. Especially the capability to rapidly generate alternatives is conceived to support design and planning practices, particularly during decision-making, because it can quickly represent alternatives and their interrelations (e.g. scenarios) into morphological, three-dimensional outcomes. The capability to explore urban design in 3D models allows the designers and decision-makers to verify the process at the full extent: it demonstrates a powerful tool for widening the discussion and participation among designers, planners, decision-makers and citizens.

According to our experience, often this meeting sticks the discussion into opposite positions that, in principle, can hardly find a balance. Instead the high-dimensional spaces of the design solutions generated by the system, starting from the ontological description of the some factors, can be used to systematically explore alternative and creative options.

Generating the large space of alternatives and visualizing them turns out really useful to outline both the interrelation between the factors and the borders between the possible and the unsuitable.

Ontologies can provide a specification of domain information by representing concepts and relations, this specification helps in building semantic information systems to support activities. Indeed we do expect ontologies to straight the interpretation and writing of the factors by designers, planners and decision-makers, allowing them to quickly and interactively change the concepts and their relations, for generating 3-D urban scenarios and exploring the space of design solutions.

References

1. Alexander, C. (1996) A City is Not a Tree, in Le Gates, R & Stout, F. (Eds) The City Reader, London: Routledge.
2. Bazzanella, L. (1971) *Strutturazione programmata di spazi elementari*. Torino: Istituto di elementi costruttivi.
3. Bradshaw, J. M. Boose, J.H. Shema, D.B. Skuce, D. and Lethbridge, T.C. “Steps Toward Sharable Ontologies for Design Rationale”. Proceddings of the AAAI-92 Design Rationale Capture and Use Workshop, AAAI-92, San Jose, CA, July 1992.
4. Cheng, V. Steemers, K. Montavon, M. & Compagnon, R. (2006). *Urban Form, Density and Solar Potential.* (6-8 September 2006) PLEA2006 The 23rd Conference on Passive and Low Energy Architecture, Geneva, Switzerland.
5. Coppola, C. (2005). *Computer e creatività per l’architettura, intelligenza artificiale e sistemi formali,* Firenze: Alinea editrice
6. Coppola, C. (2006). *Dalla Progettazione Assistita All’ontologia Dell’architettura.*

Unpublished paper
7. Gandon F. "*Distributed Artificial Intelligence and Knowledge Management: Ontologies and Multi-Agent Systems for a Corporate Semantic Web*", PhD in Computing, INRIA and University of Nice - Sophia Antipolis, November 7, 2002
8. Garey, M. & Johnson, D. (1979). *Computers and Intractability*. San Francisco: W.H.Freeman.
9. Gero, JS and Fujii, H, (1999), A computational framework for concept formation for a situated design agent, *in* K. Hori and L. Candy (eds), *Proceedings Second International Workshop on Strategic Knowledge and Concept Formation.*., Iwate Prefectural University, Iwate, pp. 59-71.
10. Hillier, B and Hanson, J. (1984) *The social logic of space*, Cambridge University Press, New York.
11. Jacobs, J. (1961). The *Death and Life of Great American Cities*. New York.: Random House.
12. Johnson, P. (1999) *Swarm User Guide*, University of Kansas.
13. Kaplan, A. (1963). *The Conduct of Inquiry*. New York: Harper and Row.
14. Keita A., Laurini R., Roussey C., Zimmerman M. (2004)"Towards an Ontology for Urban Planning: The Towntology Project". In CD-ROM Proceedings of the 24th UDMS Symposium, Chioggia, October 27-29, 2004, pp 12.I.1
15. Klein, M. (2002). Combining and Relating Ontologies: An Analysis of Problem and Solutions, Vrije Universiteit, Amsterdam, Netherlands
16. Knublauch, H. (2003) An AI Tool for the Real World: Knowledge Modeling with Protégé, JavaWorld June 20, 2003, http://www.javaworld.com/javaworld/jw-06-2003/jw-0620-protege.html?
17. Mele, F. Calabrese,A. Licenziato,L. Minei,G. Sorgente, A. Talamo, O. “Spatial reasoning for 3D visualization and reconstruction of architectural artefacts” AIIA WORKSHOP FOR CULTURAL HERITAGE, A Satellite Workshop of the 9th Congress of the Italian Association for Artificial Intelligence University of Milan Bicocca 20 september 2005
18. Métral C. (2006) An Ontology-based System for Urban Planning Communication, COST C21 Towntology Workshop Geneva - November 6-7, 2006
19. Minar, N, Burkhart, R, Langton, C and Askenazi, M (1996) *The Swarm Simulation System: A Toolkit for Building Multi-agent Simulations*, Web paper: http://www.santefe.edu/projects/swarm/, Santa Fe Institute.
20. Moore, E.A. & Agogino, A.M. (1987). INFORM: An architecture for expert-directed knowledge acquisition. *International Journal of Man-Machine Studies*, 26, 213-230.
21. Mylopoulos, J. (1991). Conceptual modeling and Telos. Technical Report DKBS-TR-91-3. Department of Computer Science, University of Toronto, Toronto, Ontario, Canada, November 1991.
22. Protégé (2007).The Protégé Ontology Editor and Knowledge Acquisition System. 7June. 2007 <http://protege.stanford.edu/>
23. Schevers H., Veldman D., Boulaire F.,_Droaemuller R., Ontology-Based Demand Support Systems For Urban Development, Clients Driving Innovation: Moving Ideas into Practice (12-14 March 2006) Cooperative Research Centre (CRC) far Construction Innovation
24. Spaziante, A (2006), *Ricerca sul sito di Mirafiori, fase 1 – Linea C*, Technical Report LAQ-TIP, Politecnico di Torino.
25. Sowa, J.F. (1991). Toward the expressive power of natural language. In J. Sowa (Ed.), *Principles of Semantic Networks.* San Mateo, CA: Morgan Kaufmann.
26. SUMO (2004). Suggested Upper Merged Ontology. 7June. 2007 <http://www.ontologyportal.org/>
27. Uschold, M. Gruninger, “Ontologies: Principles, Methods and Applications”. In Knowledge Engineering Review, June 1996, Vol. 11, N° 2. p 93-155
28. Weisstein, EW (1999) *CRC Concise Encyclopedia of Mathematics*, CRC Press, Boca Raton, FL.

Elaboration and application of an ontology in a process of architectural project

Emmanuelle Pellegrino Jeanneret
Centre de Recherche en Architecture et Architecturologie (CRAAL)
Chemin de Rovéréaz 11
1012 LAUSANNE
empellegri@bluewin.ch

Abstract[1]**:** It is a question of making a demonstration of the potential use of an ontology of architecture during the development of the various phases of a process of project and showing thus how an ontology can be concretely operational in this field of application. First, the relevance of such an ontology is examined. Second, some of the stages of the process of project where a consultation of the ontology allows a designer to progress are examined. These stages concern as well the classification of the elements of the project as their composition. By opening as much as by closing possibilities, the ontology can frame the process and help gradually the designer to find solutions. An ontology that will not only enables a designer to operate conceptual choices but also to carry them out.

Key words: Architecture ; ontology ; classification ; composition ; process ; language ; model ; reference ; inference ; rule.

1. Introduction

It is a question of making a demonstration of the potential use of an ontology in the development of the various phases of a process of architectural project. We aim to build an ontology at the interface between processes of classification and processes of composition, which, articulated, are producers of meaning.

At this interface, the ontology is conceived as a piece of a computational system of project that will assume all the drawing functions: descriptive, prescriptive and speculative [02]. The computational systems able to help the designer in his reasoning are those which integrate algorithms [03].

Endowed with inference rules, the computer possesses a certain independency in solving problem of conception. The ontology is conceived at the same time like a referential structure, in what it is informative and descriptive, and like an inferential structure in what it affords to produce inferences between concepts modelizing

[1] This article is registered following a first article published in 2006 [01]. We will not return on the questions which have been already treated there.

reasoning of authors. Offering a range of choices, it is thus a structure where preferences will be brought out[2].

In architecture, there are many kinds of buildings and many styles, attached to periods or authors. During his career an architect deals with clients who have very varied requirements. However, he does not have a sufficiently large memory to contain all the data useful for the exercise of his profession. He often needs to consult reference books to answer the various requests.

The ontology can become an essential tool. A tool of wide-range, in what it will classify a vast field of concepts covering decades of development of architectural and urban thought. A tool going down until a level of detail in what it will come down from general concepts at large architectural scales (for example a type of structure, "with column"), to concepts at small scales (the "capital" of the "column").

Architecture is anchored in times, styles, manners. Constructive techniques are not the same today than yesterday or than tomorrow. Architecture reflects the new methods of production, the discoveries in the technologies applied to materials, and also the lifestyles in constant evolution. But these processes are interactive. New architectural concepts have allowed and will allow the exploitation of techniques as much as they will create their development. New ways of life will upset the plans of architecture as much as new plans will upset the practices in the habitat. Our ontology aims to unroll the diversity of the concepts which have marked out the history of the city and of the building; it is an Encyclopaedia which will give an account of our knowledge.

Urban form is a fabric made up of various structures which have been gradually added the ones to the others all along the centuries. According to different dynamic, city is prone to a constant evolution and becomes of this fact unceasingly more complex. The concepts implemented in the form of the buildings as in the form of the networks contribute to explain the urban form. The increasing complexity of the urban form moreover returns into an ontology. In order to understand a given period, it is necessary to know the concepts in usage in this period. And the concepts do not evolve linearly; some ones appear, disappear or sleep and then reappear [04].

Next chapter gives some elements to understand the application of an ontology of architecture such as we have conceived it.

2. From classification to composition

In order to understand the role which can hold an ontology in the process of project, we will expose in detail the fundamental mechanisms of the architectural design in the passage from classification to composition. Classification treats of the grammatical questions of a project. Since Ch. Alexander, many researchers have used graphs to decompose the process of project [05]. Some software, such as space allocation systems start from graphs to build a geometric organization of the plan [06]. The relations networks which link the rooms together are completely different according

---

[2] Our ontology is built in Protégé, in OWL Language. Based on the principles of the predicative formal logic, Protégé allows the description of concepts and relations semantically rich. Protégé is developed at the University of Stanford and Manchester.

to architectures and contribute to a large extent to define their style. In general the designer works from a diagram of basic relations. In the passage to a complex form, this basic diagram can be developed in thousand manners.

It is a question of solving step by step questions of connexity (course), adjacency, and order (succession) between rooms. Horizontal circulations, i.e. various paths to go from a room to another one are defined. By an immediate crossing, when two rooms are at the same time related and contiguous, or through one or several rooms which separate them. It is also a question of determining, if necessary, vertical circulations. The access to another level can be more or less direct, according to the instrument (helical staircase or ramp), or its position, more or less central/peripheral in relation to the served place. Gradually an order emerges, which is defined from the relative position of the rooms; this position is then ordered. The first figure shows an order of succession of the rooms around the stay different in the top and in the bottom diagram, while the relations of connexity are unchanged.

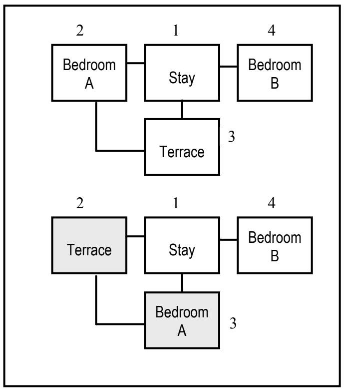


Fig. 1 Relations of connexity and order

This grammar is made up by successive approximations; during the process of project, many phases of going and return between geometrical forms, their measurements and the form of the relations are performed. The modification of the form or of the dimensions of a room can involve in return a modification of the relations, of the adjacencies and of the transition from a room to another one. The reverse is also true. While the designer refines the form of the relations, the form of the building emerges gradually with geometry and relevant measures, consistent with the choices he has set. At the beginning the range of the choices are thus largely open; there is infinity of possible solutions.

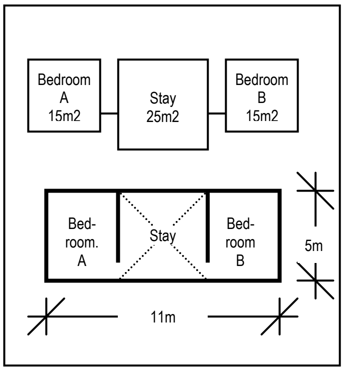


Fig. 2 From diagram to composition

Constraints reduce the field of the choices. Constraints of site (view, orientation, topography), regulation constraints (right of way, right of view, admissible volumes), constraints related to rules of style. Let us take a very simple example. A diagram of three rooms must be integrated into a geometry (Fig. 2).

Condition n°1: the global form is a rectangle or a square.

Condition n°2: 55m2 maximum available surface (regulation constraint).

Condition n°3: stay is a square (style constraint).

Condition n°4: rooms are rectangles with at least 2 ml on the small side, or squares (functional constraint).

Imagine now that the designer is dissatisfied about the size of the stay and that he wishes to increase its surface. Capturing 36m2 for the stay, he modifies his diagram. In this new geometrical configuration (Fig. 3), the fourth condition is not met. Rooms do not have, at least, a width of 2ml on their small side. In the passage from classification to geometry, a deformation is thus possible up to a certain point; beyond, it is necessary to reconsider the classification to find a solution. For example, a relation of direct connexity can be added between the rooms (Fig. 4). Or the designer can give up the third condition, and draw a rectangular stay (Fig. 5), or modify again the surface of the stay, bringing it back to 30m2.

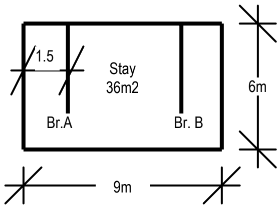


Fig. 3 Snd. geometrical configuration

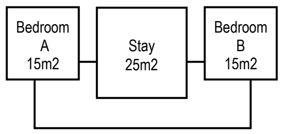


Fig. 4 Adjonction of a direct connexity relation

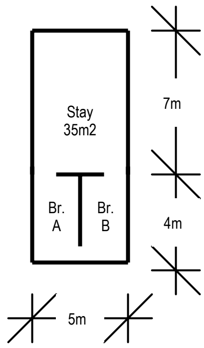


Fig. 5 Variant with a rectangular stay

When he is conceiving, the architect must thus find a compromise between various constraints, which are not always compatible. As instrument which compiles knowledge, ontology can therefore help him to find solutions. The choices of the designer will be conscious, and not intuitive. Not only because he will be able to question the ontology, but because in fact, still much more, the ontology will question him. The ontology is thus an intelligent tool and of "interactive" use. It will lead the designer to take various paths; the classes of objects represent crossroads.

The system of project in which the ontology functions questions the designer and proposes him choices, examples and paths (Fig. 6). The problematic of paths can find a technical solution in the field of Artificial Intelligence [07].

In Protégé, the architect will also frame the development of his project by questioning the ontology through the function "Queries". For example, if he wants to know the list of the modern houses having more than ten rooms, or an interior patio.

Today ontologies can therefore partially describe topological relations. For example, MADS [08], developed by geographers, modelizes this kind of relations. MADS proposes a conceptual model able to modelize spatial data between physical objects[3].

[3] MADS, like most of ontologies, mainly aims to provide a better communication between different actors intervening in the urban form; but there is no projective ambition.

3. Form of the classification in OWL

Protégé enables us to describe classes of objects. The classes are described according to a taxonomic logic founded on relations of inclusion and exclusion (Fig. 7). But an ontology which would only offer a hierarchical classification would not have any interest. If architecture uses many classifications, it is to integrate them in sentences. In the spoken language, the act of speaking consists in forming sentences to communicate something. It is necessary to understand the act of projecting as the equivalent of a speech.

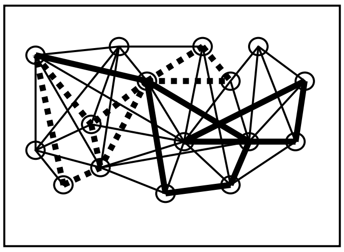


Fig. 6 Paths in a network

Like any language, architectural language is founded on an economy [09]. The economy is based on the principle of substitution. In the language, paradigmatic elements can take place in a syntagm in a given place and enter the composition of infinity of syntagms.

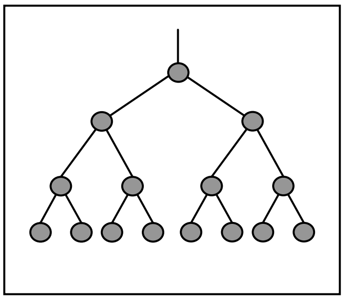


Fig. 7 Classification of concepts according to a hierarchy

In the syntagm "the wall is made of wood", the paradigmatic element "made of wood" indicating a characteristic of the wall (an extension of the class of comprehension "wall"), in fact its substance, can be replaced by another element taken in the same paradigm (Fig. 8). But the same paradigmatic element can also enter the syntagm "the table is made of wood", i.e. to be a characteristic of another class. The principle of the language is that with few elements one can say infinity of different things. Such combinatory allows substances, qualities and quantities to be declined.

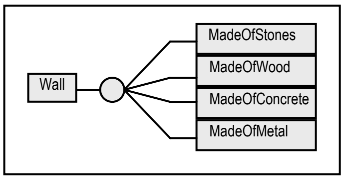


Fig. 8 Order of substitutions

In OWL, classes can be set in relation through properties. In the process of project, the description of the concepts organized in a tree structure is doubled by a structure of concepts in network, according to an operational logic (Fig. 9).

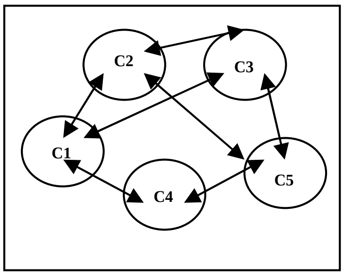


Fig. 9 Classification of concepts according to an operational logic

Protégé allows thus to articulate a finished number of paradigmatic elements in a non finished number of syntagms. In architecture, this operational logic is necessary at the scale of the components of the building.

The project of architecture articulates in a complexity dimensions that treat of materials, colours, contents, forms; all these dimensions are in relation to the position of the components in a set.

Operational logic allows to connect rooms with furniture, elements of architecture with materials, materials with dimensions, forms with contents or to indicate the rules to carry out a concept step by step.

Our ontology is divided in three general branches [01] (Fig. 10). The first describes instances of houses. The second describes the meta-concepts (of a meta-language) with the help of which it is possible to speak about another language; in fact about the geometrical language of the plan. The third describes meta-concepts connoted in languages of authors. Some concepts, such as "plan libre", enter the composition of the language of different authors. They do not have the same meaning. This branch connects meta-concepts with authors. Properties linking classes to other classes allow the various levels of the classification to be connected. An instance of house is described through its rooms or through its characteristic architectural elements. Moreover, according to authors concepts do not always have the same meaning; the third level indicates thus to which author a concept is linked, and to which other concepts this author connects it in his language (Fig. 11). These links are in fact inferences. Inferences describe the reasoning carried out by an author, to define the elements of his language.

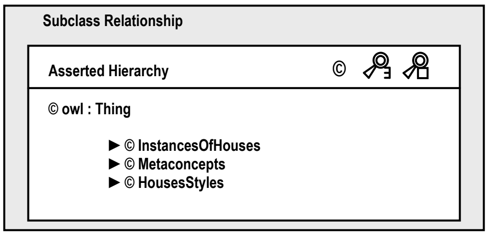


Fig. 10 General form of the classification in OWL

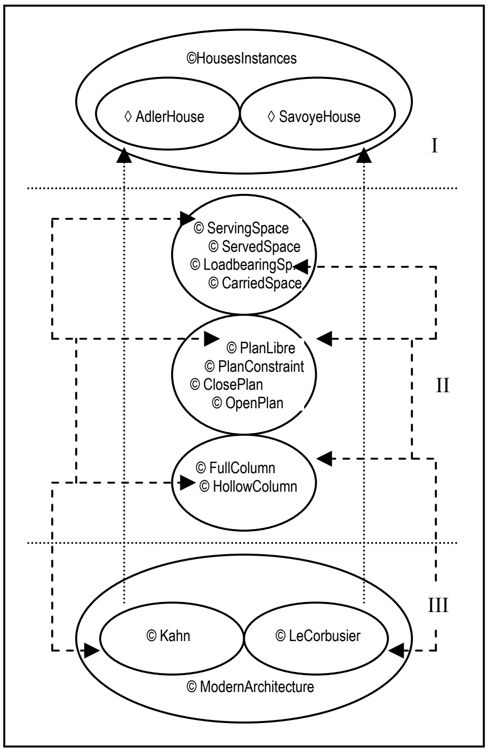
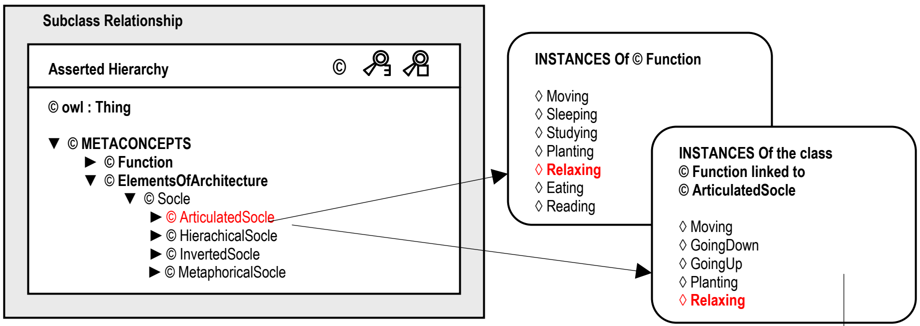


Fig. 11 General functioning

By appointing superclasses, Protégé opens thus the possibility for classes to belong at the same time to different superclasses.



Instances can be linked via properties (Fig. 12). The designer is informed that the socle is dislocated in various spaces. It can be used notably to move on horizontal level, to change level, to shelter, to enter, to plant. The ontology thus returns the form of a socle to contents (galleries, staircases, terraces, porches, and so on); therefore, since it indicates how to produce meaning, the ontology is actively involved in the design process.

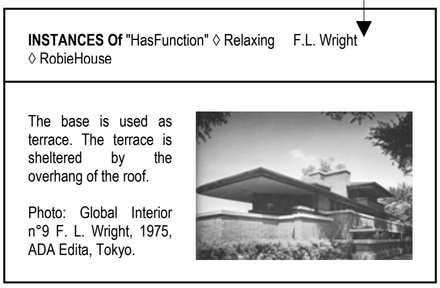


Fig. 12 Relations between instances

4. Inferences and inference engine

In Protégé, properties function like constraints. While existential conditions *describe* what enters the composition of a class (concepts and/or characteristics), universal conditions *define* what enters there. In the project of architecture, the range of the choices is gradually tightened by constraints. Making it possible to lay constraints, our ontology will be adequate to assist a process of project. Universal conditions enable us to give a definition of the elements constitutive of a house of a Master, to enumerate the concepts-keys of a given architectural language.

The closure axiom allows the "necessary and sufficient" components entering the composition of a house to be defined. It implicitly opens the possibility of creating variants of a model. The axiom describes a structure. This structure admits variations, but within a delimited framework. A variant of the villa Savoye with a different order of superposition of its composition principles would not be anymore a variant of the villa Savoye (Fig.13). It would be something else.

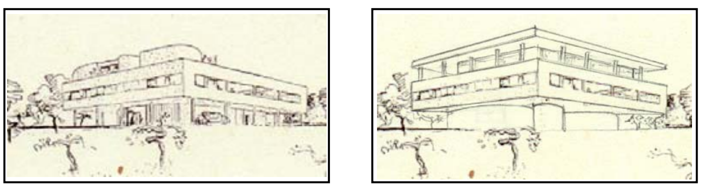


Fig. 13 On the left : Villa Savoye, original version [10] - On the right : modified version

But the interest of an ontology is also the opening towards new compositions. And transgression is the basic mechanism of invention. Very often Masters explicitly invent concepts against an established order. "OpenPlan" is a transgression of the classic plan interpreted as "ClosePlan", and "DecomposedPlan" is a transgression of traditional and modern plans interpreted as being "Composed". Any invention is anchored in a norm; when the object of the transgression enters the norm, it is in its turn transgressed.

From classes defined by a closure axiom, Racer performs automatically classification. A new object can be classified like subclass of one or several classes. Racer infers thus automatically hierarchies between concepts. One can also visualize in a tree form the hierarchy of concepts drawn by the operational logic, on levels of depth defined by the user himself.

Actually Racer carries out deductive inferences[4]. If the class of the "PuristHouses" is defined at least by windows "EnBandeau", and the class of the "BrutalistHouses" at least by "BétonBrutDeDécoffrage", a house which would have at the same time both characteristics would be reclassified like subclass of the class of object of "PuristHouses" and that of "BrutalistHouses". However, that leads to a non-sense. Purism and brutalism are opposite trends. To avoid such a reclassification, classes can be disjoined in Protégé. Or it could be a question of opening a new class, by means of the "and-and" concept of Venturi[5] [11].

One part of the conditions is inherited. For example, any house is defined by rooms (at least one) and by elements of architecture (walls, columns…). If one declines a subclass of the class of the houses, these conditions are inherited.

With the closure axiom, we can define Wright's Willets house like an object articulating a finished number of architectural elements (Fig. 14). Each one of these conditions is thus necessary to project in the "Prairie" style.

Each room consists of architectural elements, openings (doors, perhaps windows), flagstones, and so on.

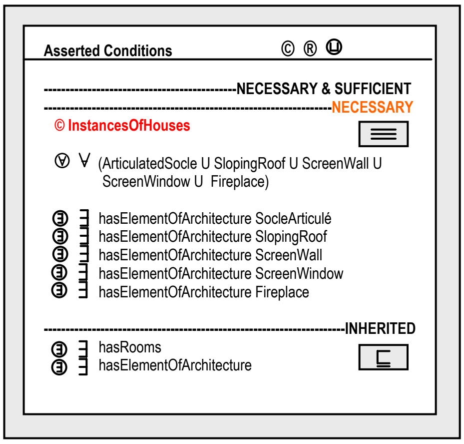


Fig. 14 Conditions

Therefore, if a "Bedroom" of a "PrairieHouse" has walls and windows, it inherits the elements of the class of objects "Prairie", i.e. "ScreenWall" or "ScreenWindow".

[4] In our research, we wish to carry out abductive inferences. Actually the researchers of Stanford are developing the SWRL (Semantic Web Rule Langage). Different built-in apply this language in Protégé; with these built-in we search to carry out abductive inferences from classes of objects.

[5] R. Venturi opposes the "and-and" complexity to the "or-or" simplicity of functionalist architecture.

In order that such heritages can be creators of invention, we seek to establish abductive inferences, which are part of the mechanism of innovation and invention.

5. Descriptive logic and geometry

In a plan of architecture, furniture qualifies the room by assigning it a function. Small icons indicate the place of the furniture in the plan. Informing about the destination of the room, making it explicit, furniture has an indicative function.

Geometry of spatial composition and logic of descriptive classification do not function in the same way [12]. In descriptive logic, the tree draws inclusive relations; B and C are included in A if B and C are A (Fig. 15). In spatial geometry, furniture is included in rooms (Fig. 16). But it does not share common characteristics within the meaning of descriptive logic. And, according to architectures, inclusions are variable. In Wright, for example, one can change nothing; rooms are conceived with the furniture. Wright draws everything, carpet, tables, chairs [13]. In Mies Van der Rohe, there is a functional block in the centre of the room, with kitchen, bathroom and WC. The rest of the space remains undifferentiated, undivided; each one can lay out the table, the bed, the armchair according to his desire.

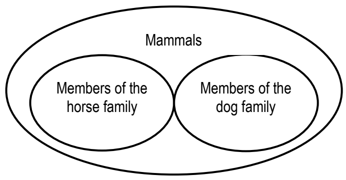


Fig. 15 Descriptive logic

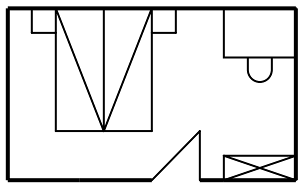


Fig. 16 Geometry

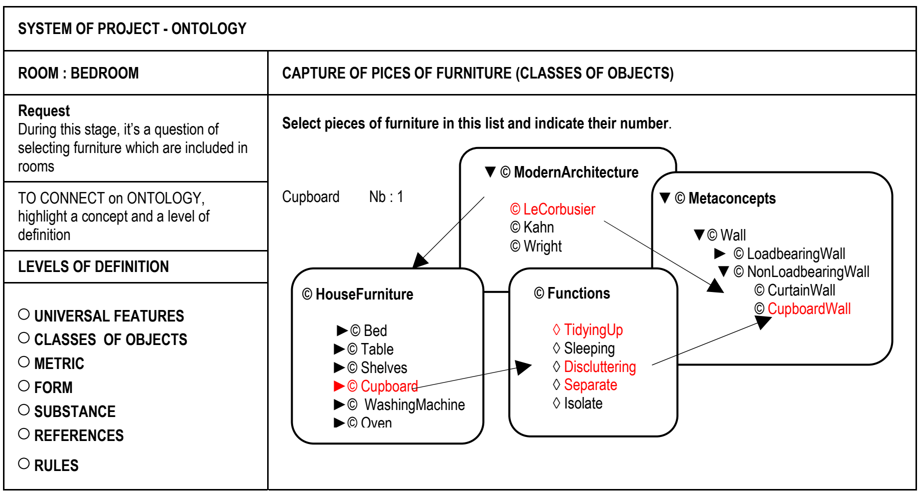


Fig. 17 Capture of furniture

In Le Corbusier, cupboards are arranged in walls. According to his reasoning, the disposition of the furniture in the centre of the room represents the norm, disposition that clutters it up. In order to avoid clutter, Le Corbusier proposes to transgress this norm. He invents the concept of "WallCupboard". In Protégé, a class of objects "HouseFurniture" will be linked to a class of objects "HouseRooms" via a property "isIncludedIn" (Fig. 17).

The choice of a piece of furniture can involve a return on the metric or on the form of the building in project.

6. Assistance for data capture

Our graph editor (GED)[6] copes with the articulation of the classification and of the composition of the elements of the project. GED gives an account of the stages of classification in form of graphs and diagrams [14]. Graphs and diagrams contain vertices, modelizing the various rooms, and arcs modelizing the spatial relations between vertices. Diagram is a familiar work tool for architects of today. This kind of representation allows the problems to be decomposed, by layers, in sub-problems.

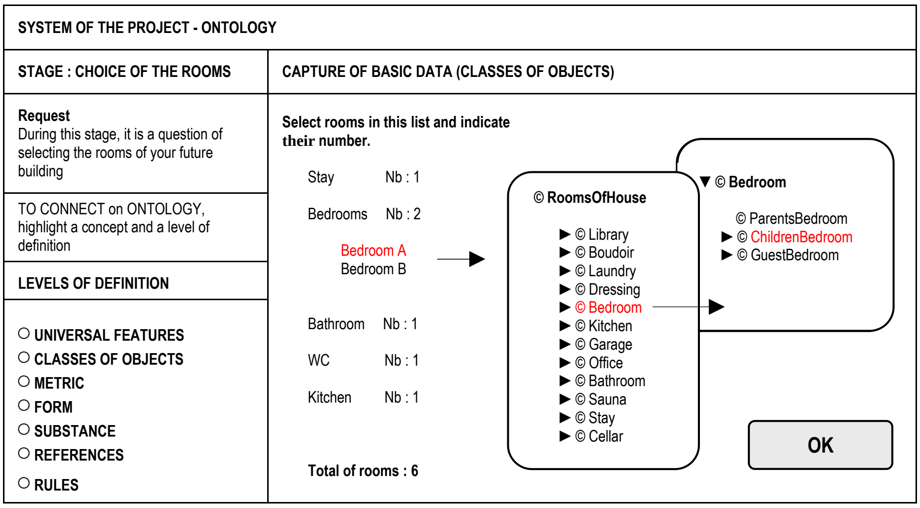


Fig. 18 Capture of basic data

Ontology can guide the designer in the classification stages. The first stage of a project consists in establishing and capturing a program (Fig. 18). It is a question of selecting the principal rooms of the future building. The ontology can frame the work of the architect through a list of rooms, general or which describes rooms attached to a particular style. The ontology can also inform about the contents of a specific building, like a museum or a hospital. Or the designer can directly capture a diagram

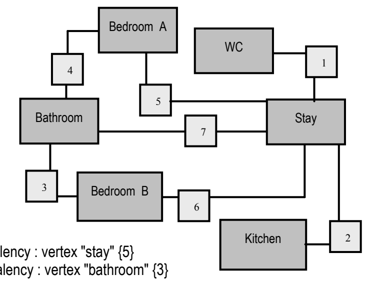
[6] Software made by the CRAAL, within the direction of D. Coray and P. Pellegrino. Later, the ontology could be connected with GED; choices made in the ontology would be deferred in form of vertices and relations in diagrams and graphs.

of a house. He will gradually modify the form of the relations to adapt it to a new context or will seek to integrate a form of grammar in figures of composition of another style.

7. Assistance for capture relations

A series of rules will frame the path of the designer when capturing relations, in particular through the graph editor. GED allows properties of a graph to be calculated. As each architectural language has got its own grammar, the graphs of a house of Mario Botta are different from those of Marcel Breuer, and do not have the same properties. These properties can be measured by rates. For example, a symmetrical graph will have a very high rate group of automorphism, since one can invert a great number of arcs without modifying the graph. Thus a graph in a classical architecture, which is by definition symmetrical, must have a very high rate of automorphism. Calculations carried out with GED, during the process of project, will be able to inform the designer about the degree of resemblance of his project with a style of reference. Values of reference will be inscribed into ontology.

Figures 19 and 20 present another example: the calculation of the valency. The valency counts the number of relations linking each room with all others rooms. It thus informs about the degree of privacy of the rooms in the house. In the ontology, the valency will be gathered together with the following rule: "*for any room with a valency degree equal or higher than 3, it is advised to introduce a room with distributive characteristic*".

<table>
<tr><td colspan="2">SYSTEM OF THE PROJECT - GED (GRAPH EDITOR) – CALCULATION OF THE PROPERTIES OF A GRAPH</td></tr>
<tr><td>CALCULATION OF THE VALENCY</td><td>CAPTURE OF THE RELATIONS OF CONNEXITY
Inspired from: P. Pellegrino, D. Coray, E. P. Jeanneret, S. Cirilli, E. Nanchen, 1995, La conceptualisation de l'espace en architecture, CRAAL-FNSRS, Genève-Berne.</td></tr>
<tr><td>In this page you can proceed to the calculation of the valency starting from your diagram, with GED.</td><td rowspan="3">Result of the analysis of the properties of the diagram (graph), its valency
Bedroom A {1} : 2
Bedroom B {2} : 2
Stay {3} : 5
Kitchen {4} : 1
Bathroom {5} : 3
WC {6} : 1


1st vertex with the highest degree of valency : vertex "stay" {5}
2nd vertex with the highest degree of valency : vertex "bathroom" {3}</td></tr>
<tr><td>Adjonction of rooms with distributive characteristic
For rooms with a high valency degree, we strongly advise to introduce a room with distributive characteristic</td></tr>
<tr><td>RULE
For each room with a valency degree of value "3" or higher, a vertex with distributive characteristic will be add.</td></tr>
</table>

Fig. 19 Capture of relations of connexity and calculation of the valency

The application of this rule adds two more rooms, which are distributive spaces, and two new relations. All the relations of the diagram are modified, in terms of degree of connexity and order. This rule has thus contributed to complex the graph. And the ontology can also help to qualify these spaces.

Seeking to be informed about different ways for gathering rooms together, the designer will be able to find, under the concept of "Filtering", an extract of the theories of Ch. Alexander. A page will describe the principal manners of assembling a room with another, in more or less long durations and distances.

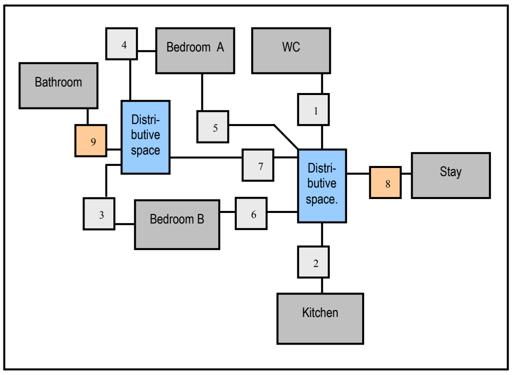


Fig. 20 Transfer of the relations on a diagram

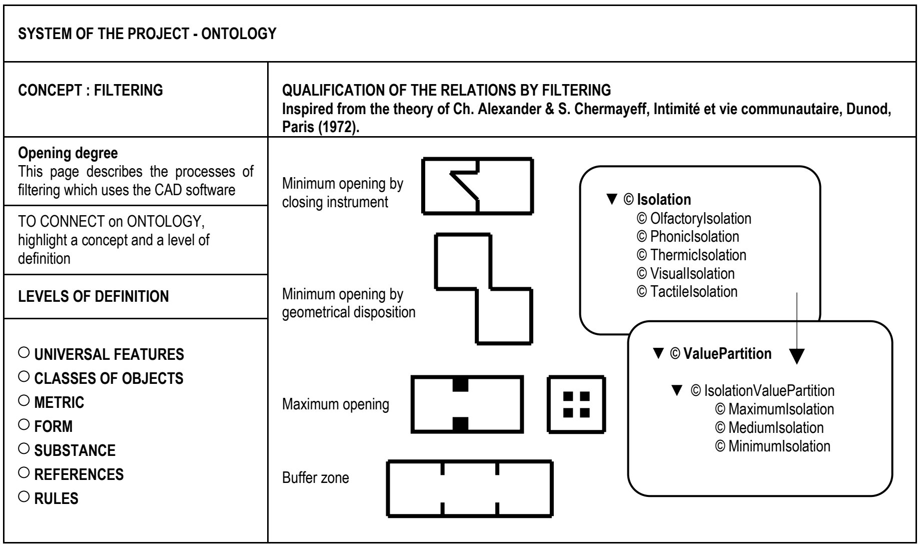


Fig. 21 Qualifications of the relations by filtering

The "filtering" concerns the degree of isolation, of intimacy of a room. The filtering can be measured by values reported to feelings, relative to the human five senses. In term of visual intensity (which returns to more or less transparent materials, letting more or less filter the light or the imprint of bodies and objects, or to the framing of the landscape, which is different through a large window or through a judas hole), of phonic, thermic or olfactory intensities (which return to more or less thick walls or to the use of certain materials more or less insulating), of tactile contact (serving hatch).

Values can be allotted ("Value Partitions") to classes in Protégé. According to the position of a room in a local or a global set, the needs for insulation are variable. If one wishes a maximum degree of insulation for a restroom, if the restroom is contiguous with the stay which is also used as music room, then it is possible to build thick walls and a buffer zone between the rooms. But if the device is too heavy, perhaps it better worth to modify the position of the room in the overall configuration. A reasoning by abduction will bring an adapted solution (Fig. 21).

8. Assistance for composition

A building is not only composed of relations; it is also composed of Euclidean forms, measures and substances. In the articulation between contained form and containing form, the building takes meaning. According to the position or the place taken within a set, forms and dimensions find their meaning. It was seen that the designer proceeds by abduction to carry out these adjustments.

Consequently, one cannot draw without change from the ontology concepts which are strictly a matter of form, because during the process any room, any element and any relation expressed in a diagrammatic form is built in a measured form and "embodied" into a substance. Therefore, form, metric and substance are into narrow correlation (Fig. 22). Many parameters (freely selected and determined at the end of the path) cross thus in the design of each object of the project.

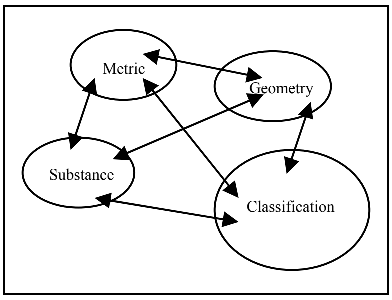


Fig. 22 Form, metric, substance

Let us take a simple wall. Does it delimit an interior space and an exterior space? If it is the case, what is the climatic context? The technology applied to a reinforced concrete wall in Stockholm is not completely the same one than in Algiers, would this be only relative with the quantity and the quality of the insulating material. Or does this wall delimit interior spaces? In this case, which kind of privacy and insulation is researched (see Fig. 15)? Is it a load-bearing wall? What is the weight of the load? What length is the wall? What is the material? Thus, once again, the ontology is a tool which can frame the project. According to situations it declines a range of choices, initially towards a particular class of object (© wall), then towards one object of this class, which, taking into account its situation in an overall configuration, will be able to adapt to the various constraints. Each element of the project must be put in congruence with the others [15].

In the ontology, the questions of form return to a series of concepts; notably to descriptors of the architectural project (which are iterative geometrical entities) and to rhetorical figures anchored in styles. TOP (Taxis Oriented Project)[7] is a software which allows, starting from a stable element, by using geometrical operations, the figures of an architectural composition to be composed.

---

[7] Software made by the CRAAL, within the direction of D. Coray and P. Pellegrino.

Geometrical operations are mediated by rhetorical figures of the natural language applied to architectural language. Descriptors and figures are memorized in the ontology. The ontology will inform the designer about which descriptor to use and which stylistic figures to use for assembling them. The figure 23 [16] presents an example of complexification with a palladian descriptor and operations such as rotation, reflection, intersection, and suppression. So as the choices operated in ontology could be deferred automatically in TOP, we intend to interconnect them.

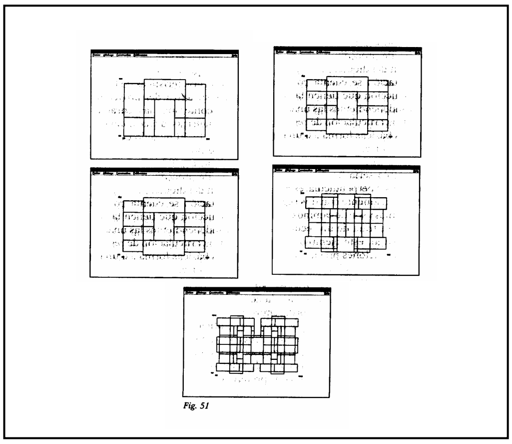


Fig. 23 Complexification of a descriptor with TOP

8.1 Assistance for the definition of a metric

In architecture, metric is a fundamental data of the project[8]. The manner "of giving measures" to rooms takes part to a large extent in defining a style. Metric can be influenced by exogeneous or endogenous factors. In general, indicative measures are given to the architect at the time of the establishment of the program.

The ontology must therefore give an account of rules concerning the metric which return to characteristics such as being "larger or equal", "smaller or equal", "of the same size than", "abnormally large" (metaphor). And existing architectures can be described in m2 of surface. In Protégé, it is possible to deal with metric data; the value "integer" allows a class to be associated with numerical data.

In figure 24 the "Hall" is set in relation with dimensional specifications via a property "HasMetric". The designer learns that in Kahn's architecture, the hall is as large as the stay ("Hall" = "Stay"). When integrating into a Euclidean geometry, the application of this rule of composition will probably involve a reclassification of some vertices of the diagram. The diagram is equipped with indicative surfaces. In order to visualize the effect of the metric on the diagram, surfaces are brought back to their square root.

[8] Actually, TOP treats with geometry. It will be developed so as to be able to treat of measures.

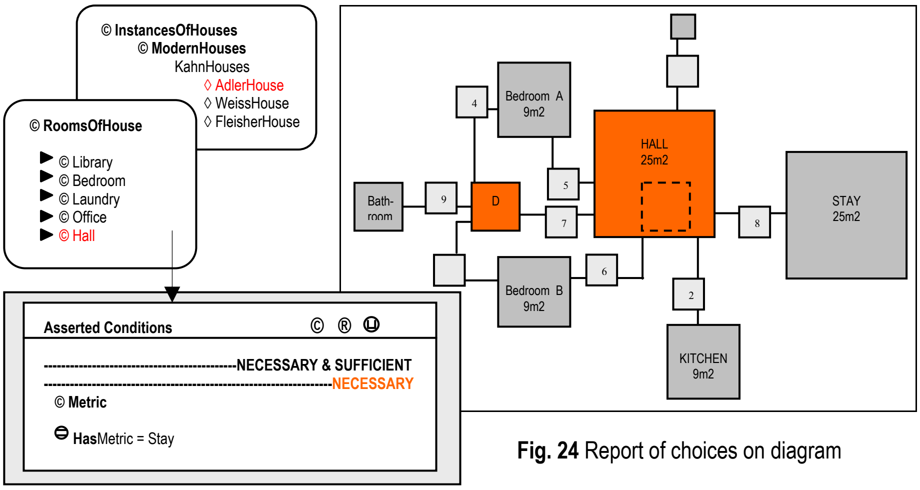


Fig. 24 Report of choices on diagram

8.2. Active assistance towards Euclidean geometry

How is it possible to frame concretely the process of project when setting an Euclidean geometry? In figure 25, the global form of the future building is questioned in its relation to the context. In the ontology, each descriptor returns to a form of context. The ontology describes archetypical manners to consider the relation of the building with its site.

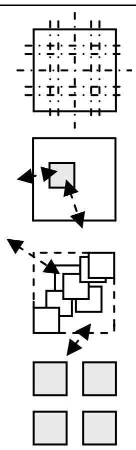
SYSTEM OF THE PROJECT - ONTOLOGY		
CONCEPT : ARCHETYPE	General page : COMPOSITION AND CONTEXT	
This page describes the archetypes which uses The CAD software	 **ARCHETYPE I** Reference form	**PLAN CONSTRAINT** Axes articulate text to context.
TO CONNECT on ONTOLOGY, highlight a concept and a level of definition	**ARCHETYPE II** Non-contextual form	**PLAN LIBRE** RELATION FULL / VOID The play between full and void articulates text to context.
LEVELS OF DEFINITION	**ARCHETYPE III** Ultra-contextual form	**OPEN PLAN** INTERACTIONS INSIDE / OUTSIDE
○ UNIVERSAL FEATURES ○ CLASSES OF OBJECTS ○ METRIC ○ FORM ○ SUBSTANCE ○ REFERENCES ○ RULES	**ARCHETYPE IV** Textual form	**DECOMPOSED PLAN** EXPLOSION OF THE UNITY OF COMPOSITION The context as something else than the site in its presence.

Fig. 25 Forms of the context : archetypes

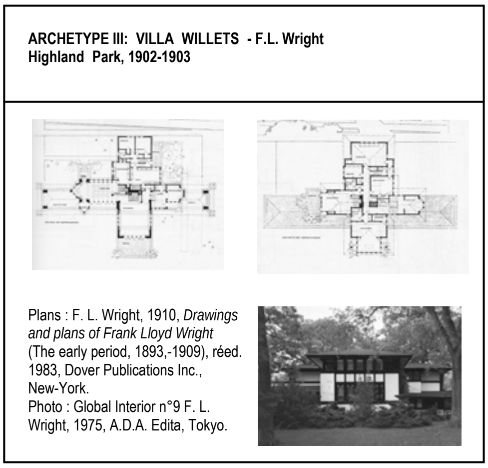


Fig. 26 Visualization of a house instance

These archetypical forms return to instances. The figure 26 describes an instance representing the third archetype. Rules of inference attached to each archetype will make it possible for the designer to engage a process of development towards the form.

Let us imagine that the designer finally chooses the second archetype (Fig. 27). The field to which belongs the "smallest common descriptor" is linked to a classification via a property of content, which returns it to the range of the rooms. If the descriptor is a SCD, then it contains "ColdSpaces". If the descriptor is a BCD, then it contains "HotSpaces"; at this stage of the process, the choices are targeted at the level of depth "P1" of the tree. The designer will choose ahead in the process, in the classes of objects "HotSpaces" and "ColdSpaces", the objects appropriate to his project, at more raised levels of depth, "P2", "P3".

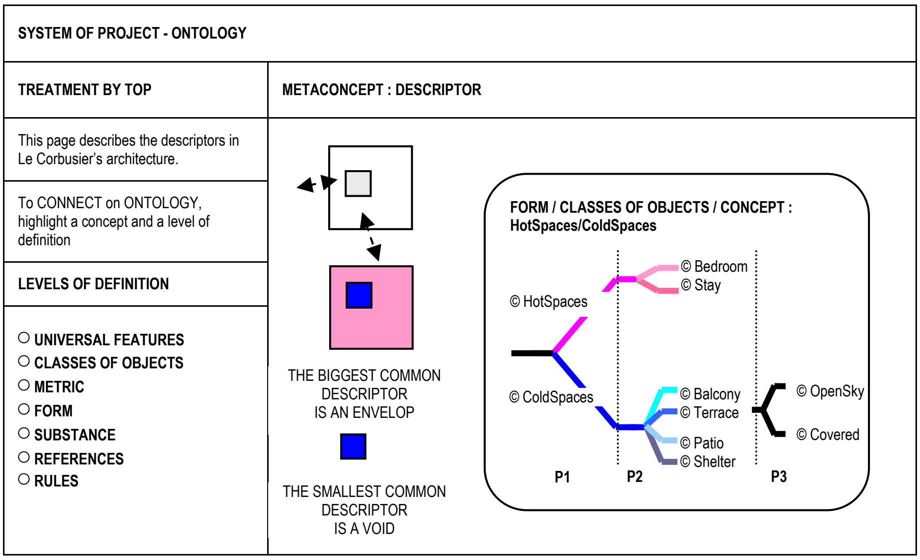


Fig. 27 Descriptor

10. Outlines

These various examples have shown how an ontology can be a useful instrument for who conceives architectures. Protégé offers an adequate structure for the development of the process of project, adequate with the reasoning carried out by the human brain. Basically the act of projecting consists in inventing by setting out constraints, in order to build a single object, whereas at the beginning of the process choices are only limited by a program and a site.

Our ontology is able to inform about all the dimensions of a project, at global and local scales. By its encyclopaedic content, it is an open structure. The ontology gives rules of languages; returning a concept to another, the work of the designer is framed. The designer can enter the problematic of his project anywhere and decide to solve initially functional, dimensional, formal, or expressive dimensions. Moreover the ontology allows the designer to work at the double level of innovation and invention.

This ontology aims to become a high-performance tool of the project.

References

01. Pellegrino Jeanneret E.: Modelization of the conception and conception of the model in architecture, Springer Verlags, Berlin (2006)
02. Estevez, D. : Dessin d'architecture et infographie, CNRS Editions, Paris (2001)
03. Terzidis, K: Algorithmic architecture, Architectural Press (2006)
04. Camacho-Hübner E., F. Golay: Preliminary insights on continuity and evolution of concepts for the development of an urban morphological process ontology, Springer Verlags, Berlin (2006)
05. Alexander Ch.: La synthèse de la forme, fr. transl. 1971, Dunod, Paris (1964)
06. Kalay, Y. E.: Architecture's New Media, The MIT Press, Cambridge, Massachusetts, London (2004)
07. Scappapietra, S., et alii, MADS, modèle conceptuel spatio-temporel, EPFL
08. Russel, S., Norvig, P. : Intelligence artificielle, Pearson Education, Paris (2006)
09. Prieto L. J.: Pertinence et pratiques, Minuit, Paris (1975)
10. Le Corbusier, Les Editions d'architecture, Artémis, Zürich (1973)
11. Venturi R.: De l'ambiguïté en architecture [1976], fr. transl. Dunod, Paris (1966)
12. Pellegrino P. : Le sens de l'espace, Tome IV, Le projet architectural, Anthropos, Paris (2007)
13. Castex J.: Le printemps de la Prairie House, Mardaga, Bruxelles (1985)
14. Pellegrino P., Coray D., Jeanneret Pellegrino E. and alii: La conceptualisation de l'espace en architecture, CRAAL-FNSRS, Genève-Berne (1995)
15. Pellegrino P.: Le sens de l'espace, L'Epoque et le Lieu, Livre I, Anthropos, Paris (2000)
16. Pellegrino P., Coray D., Pellegrino Jeanneret E.: Arquitectura e informática, Gili, Barcelone (1997).

Expressing Urban Development Concerns within a Domain Ontology

N. Athanassopoulou[1]

[1] University of Westminster, Watford Road
Northwick Park Harrow HA1 3TP, London, UK
N.Athanassopoulou1@student.wmin.ac.uk

Abstract. A regional survey, with the continuation of the study of [31] for the much discussed King's Cross area in London, provides first answers to the question 'when', 'why', 'who', 'for whom' and 'how' is space produced and used. It is proposed to model space for land management and planning purposes with a methodology of ontology development that relies on Patrick Geddes *Civics: as an applied sociology,* acknowledging the relationship between the social and the technical. A paradigm follows with references to Athens and three preliminary taxonomies on the urban and rural planning and real property rights that apply in Greece. The approach contrasts with and draws on strengths of currently known information system models for the Cadastre, borrows from experimentally work in semiotics and pays particular attention on the spatial and geographical aspects of the entities under study.

Keywords: Ontology, Urban Planning, Rural Planning, Cadastre, Greece, London, Patrick Geddes.

1 A regional survey at King's Cross in London

It is argued that London has being a planned city but that it was not until 1963 that an attempt was made to match metropolitan plans with a corresponding government structure [15, p.93-94] and that *The London Plan*, finalized in the year 2004, marks an important change of direction since a new growth is heralded as a sign of the capital's vitality as an 'exemplary sustainable city'. In this plan King's Cross is distinguished for having [35, paragraph 5.37]: «the best public transport accessibility in London. This will improve further with the completion of the Channel Tunnel Rail Link (CTRL), Thameslink 2000 and the Cross River Tram. Construction of the CTRL will release 20 hectares of underused land. Its central location and unique public transport accessibility offer particular scope for high-density business development, as well as housing.»

1.1 The geographic survey

The rural character of the area around the station in the 18th century (e.g. Rocque's map) was strengthening by a new canal scheme promoted that was to be named Regent canal. The famous architect John Nash would incorporate it as a decorative feature in his design of the homonymous park with the most important endeavour, amid this complex of actions, being the new Regent street, conceived in 1811, the year in which the future George IV formally assumed power as Prince Regent, to match the grandiose style of Napoleon's Paris with the vast undertaking of connecting the terraces of houses, villas, markets and landscaping of the park to the palace of Carlton House in St Jame's. The report of John Nash was thorough and it was bold. The new street would provide «a boundary and complete separation between the Streets and Squares occupied by the nobility and gentry, and the narrow streets and meaner houses occupied by mechanics and the trading part of the community» [6, p. 64 – 69].

Much later, in 1904, Patric Geddes presents the view that every scientific survey involves a geographic and historic exploration of origins and social types and characteristically differentiates the English and French mentality according to the social character of farmer and hunter (open houses/villages - fortified and big walls for the city) [11, p. 137 – 138]. He insisted that designs that do not take into account the *genius loci* fail, as did John Nash's plan for Regent street [37, p. 114], most probably referring to the act of segregation performed, seen as unnatural for the English tradition.

At the junction of Gray's In Road, Petonville and King's Cross Road, in 1836, was erected the monument which resulted in the name of the locality being changed to King's Cross. It depicted the four patron saints of Britain and at the top was a statue of George IV and according to [16] it provoked such unfavourable comment that the statue was taken down in 1842 and the rest three years later. The serious frictions that were taking place at this period of change are also illustrated with graffiti in the nearby area.

1.2 The historic survey of recent time

The station was built in 1851-2 and when it opened it was the biggest station in England [16]. The registered land use in the years between 1931-1935 (Dudley general survey in [7]) is that of «chief urban area», which means that at the time the greater area of King's Cross was well integrated, yet not necessarily with a loss of its newly acquired social character and role. Indicatively are noted some of the trade unions that were found in the area and may relate to this past: at Euston Road the Unisar, Public Service Union, and at Chalton road the National Union of the Rail, Maritime and Transport workers.

1.3 Citizenship and housing

The 'unsavoury' reputation of King's Cross was established before the construction of the railways and it is in the edge of Somers Town, the old slum, and the so-called opportunity area on the other side, where the British Library is standing today (erected in 1998). For the sociological condition at King's Cross as a place to live, Charles Booth's map of London poverty [5] at the end of 19th century was compared with the contemporary axial map for London [17]. On the west side of the railway lands, the Camley street, is unclassified in Booth's survey, being at the time a coal ward and depot area and on the axial map it appears in light blue colour and with the bounding streets highly unconnected to the surrounding region. The two areas of Barnsbury in the east and Maiden Lane in the upper north, of the railway lands, are also not well integrated in the urban fabric, evoking the question, especially for the first case, whether the character of the public housing residential area that is infamous curiously here withstands time. The Regent's canal acquires with the new redevelopment plan [18] three bridges for improving the connection to the surrounding region and for revitalising its walking channels, though they may prove insufficient at smaller scale.

1.4 The applied sociology of the present

In the *Development Plan for Greater London* of the year 1963 the roads in the area were examined by way of easing pedestrian movement with the relocation of a mixture of uses that allow human interests to interact with the place and not only pass through and by it. Such an interaction was found at Judd Street and its surrounding quarter, while from descriptions in the 'Survey of London' [33] it is noticeable that this land is private. Where the major works for the Human City are going to take place, the land owners are the London and Continental Railways. Close to the railway trucks there is available space and this availability sustained the planning policy impetus for a large-scale development on the strategic position of King's Cross.

1.5 The present and future of King's Cross

In 1988 the London Underground Consortium (LRC) presented a master plan with a large-scale urban park (25 acres) at its heart and two alternative solutions for the lower, triangular area. This plan also proposed a mixed use development and a great glazed vault for the CTRL on the assumption that the CTRL would come into King's Cross station underground. It is probable that the decisions have changed due to the lethal fire at King's Cross underground railway (tube), a year before the proposal's submission. After the terrorist attacks of the year 2005 the reality requires to be studied more thoroughly in respect of further transportation opportunities and alternatives. The canal opens possibilities of movement that can foster the inhabitants' sense of well being, boost their confidence and drag them outside a segregated shell, to the opportunities that the city offers (vocational, recreational, educational, cultural, tourist etc) to citizens of Europe as an answer to a long persisting planning question.

Under this optic, the re-examination of the master plan of 1988, with the park, the pastoral scene dominating King's Cross Central becomes also relevant. Which of the two proposals, or a third one to be examined, better suits the *genious loci* of King's Cross and what is the incipient future that emerges?

Fig. 1. The proposal of 2004 with representational troops like the white swans, a sign of calmness and cleanness amid modernity. The proposal of 1988 foresees as a central function the 25 acre park in the shape of a heart, a sign of pulsing organism that controls the circulation of people and functions.

1.6 Literature of civics and incipient character

Around the hub of a *pluriethinic city* which is forecasted to attract 60 million of passengers every day by the year 2020, comparable to the number of passengers at Heathrow airport (Argent King's Cross, [18]), was and is still added much to the urban form, urban type and urban space's perception via published maps and images, but the essence of the character of the place remains baffling. Is this a lie or a failure to understand it?

A vertical section of London's life inside the houses of King's Cross Central, and one is that from the roof depicted on Mike Leight's movie [23], can enhance the conclusions on the type of social classes present in the area and the greater immigration patterns that do not necessarily relate to the previous imperial character of Great Britain. In Camden Borough almost one third of residents come from black or minority ethnic groups, one in ten of the population is Muslim, and there are significant new and refugee communities, including from Somalia [4, p. 26]. It is interesting to examine the current practice in relation to the relief work of Patrick Geddes in Cyprus, because he had not anticipated that the Armenians refugees might not want to stay there permanently [34, p. 55].

2 Land management in the context of regional surveys

The current document places in the centre of interest the realization of regional surveys, for enabling urban and rural planning and programming. As an initial categorization for the sitting and development of human settlements A. Aravantinos [1, p. 32 – 35] recognizes the following criteria:

(a) The climate-natural environment. The human quest for a natural environment that will not be against the physiology of his organism and which in the same time will satisfy him psychologically is one of the primary reasons that make him chose certain places and repel others.

(b) The material for the coverage of immediate needs. The commerce and the different communication and transfer networks have reduced the significance of this criterion. However the negligence of these basic relations of humans and nature is being paid off from the residence of the city (e.g. cases of embargoes due to wars).

(c) The defence. The strategic, especially the defending, geographical position was in the past an essential criterion for the development of a settlement.

(d) The communication. The need to secure communication between the settlement and other, more distant places drives the installation as closer as possible to a vantage point in relation to the traffic network, terrestrial, maritime, aerial.

(e) The pre-existing human endeavours. To certain extend the parameter of chance affects the location of the first human endeavors (e.g. movement of populations) and the later sitting of settlements.

(f) The distance from other settlements. For cases that do not appear to have any particular physical advantage in relation to adjacent sites the explanation maybe in the structure of the greater area and in particular in the combinatorial position of other small and big settlements of the geographical region in which it belongs.

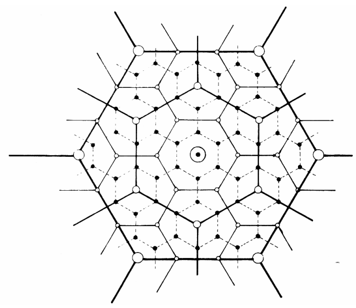


Fig. 2. The framework of hierarchical settlements and their space of influence according to Christaller (1933).

2.1 Areal unit of reference and historical, social, legal framework.

As an areal unit of reference the guidelines of UNECE [8] characteristically see as suitable the land parcel, though this unit excludes the possibility to examine the elements that semantically make up an urban form, from within and not only from without (descriptively), but also globally and back and forth in history (following Henri Levefre's distinction in [21, p. 37]). It is proposed to use a *quasi object* that encompasses the intimate relationship between the *social processes and the spatial form*, central to the method of regional report with the intention to yield an overall perspective of an area's social ecology and very schematically of a relation of the type:

PLACE… WORK… FOLK

That originates from the simple biological formula:

ENVIRONMENT… CONDITIONS… ORGANISM

That had to be applied and defined by the social geographer to become:

REGION… OCCUPATION… FAMILY-type and development

And which with slight variations lead to Le Play's simple phrasing in the beginning [11].

This practice was seen as able to face the threat of planning being concerned exclusively with a simple ordering in the physical environment Hellen Miller [27, p. 174 – 182]. Patrick Abercrombie, who later became responsible for the development of the County of London Plan, was much influenced and his plan (until the making of the third generation of New Towns) is about a cellular city where the basic element of planning is not the vaguely defined urban space, but the green background in which individual communities, each consisting of smaller cells or neighbourhoods appear as islands of urban development. It is also very interesting to notice that according to Geddes what was seen as being lost in the development of London were the villages and hamlets.

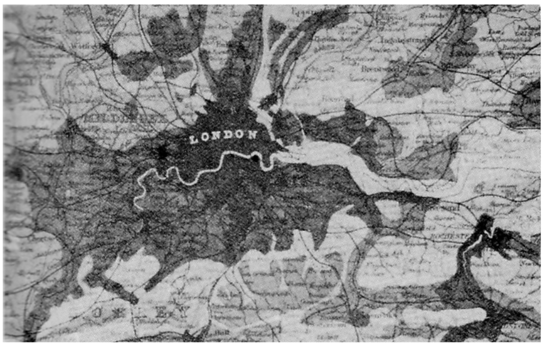


Fig. 3. We are making no plea for over-centralisation; on the contrary we are inclined to think that many ganglia maybe needed to maintain the health of so vast and multi-radiate a body politic… (after [12, p. 28-29]).

Peter Halls links and analyses of the County of London Plan see the: «insistence of Geddesian survey methods to tease out the elusive community structure of London, that metropolis of villages. Then there is a brilliant combination of Perry's neighbourhood-unit principle with Stein's and Wright's hierarchy […] in a solution that imposes order on the world's least orderly great city; but in a way so natural that no one would notice […] The County Plan used the new road system, to create a cellular London: the new order was to be implicitly organic. The plan of 1944 met several changes but as Peter Hall puts it, it somehow survived and London is one of the few places where it is possible to see the Howard-Geddes-Mumford vision of the word made actual. [13, p. 187].

2.2 The CIAM and Team10

The IV international congress of modern architecture (C.I.A.M.) examined the situation of 33 cities from the national teams each of which encompassed 3 plans-maps: (a) one with indications of the areas of housing, working and recreation, (b) one with the traffic network and (c) one with the city and its larger region that represented the four functions together [20, p. 8]. The outcomes of the analyses, as presented in the «Athens Charter» include, among others, the observation (No 11) that what is lost by the city's development are the green spaces, that only constructions with some height could satisfy with success major legal requests and (No 28) the challenge for the authorities to establish a 'regime of land' (No 29).

Geddes has influenced the work of Le Corbusier [20, p. 14], but there exist another decisive difference between the two movements in their pursuit of a therapy to the ills of the city. In the first case the ultimate unit was that of the *cell of neighbourhood*, while in the second the *unit of residence of plausible size* (e.g. Unité d'Habitation in Marseilles, 1947-1952) with a parallel adoption of *a normative model in planning* that is easily recognizable and decomposable (e.g. theory of the 7Vs road system). As George Candillis says [20, p.125 – 132], the recognition of the four major functions in city planning was something outrageously new in the period. In the end of the war those responsible in Europe were not ready to face the implacable problems or redevelopment and in the perplexity they discovered the small book of the «Athens Charter», which became the 'handbook of the good planner'. Millions of towns and villages in France, England, Germany, Italy acquired their regional plan with the schematic and non well understood criteria of it.

In the congress of the Aix-en-Provence, in 1953, was presented an opposition to the conformist policy of the C.I.A.M., primarily from the young architects and the next congress, the 10th, was to be organised by them who later were self-called: Team10 [20, p. 126]. According to Volker M. Welter [38] the concept of a heart of the city was part of the intellectual baggage Team10 was prepared to shed; Peter Smithson's impression on the meeting of the VIII C.IA.M. was about a formalistic implantation of a heart into modern urban context and of a similar formalism growing in the standardized council housing without any regard for the particular conditions. For an alternative the Smithsons turned to the Valley Section, but in a manner and purpose very different from that originally presented by Geddes.

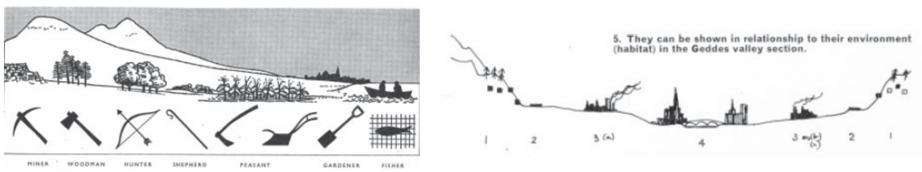


Fig. 4. Geddes' Valley Section and adaptation of different types of housing to the local conditions at the various levels of the valley, by Allison and Peter Smithson (after [38])

2.3 Comments on the development of modern Athens

According to [26, p. 17, ref.: P. Zepos, Ownership per floors, Athens, 1931] the development of the institution of *horizontal ownership* (an exception to the roman *superficies solo cedit* rule) relates to the concentration of population to towns with limited possibilities for expansion (e.g. Genova and Naples in Italy, Cyclades in

Greece), due to narrowness in enclosed with walls or protected areas (e.g. medieval German towns, Grenoble in France, Corfu in Greece). In the capital city the rate of population growth was dramatic. From 453.037 people in 1920, it soared to 801.622 in 1928, in Athens and Piraeus and the development of housing happened in a manner and in a social, economical, historical and political framework that lead to the *super intensive exploitation of the urban land*. [26, p. 16 – 71].

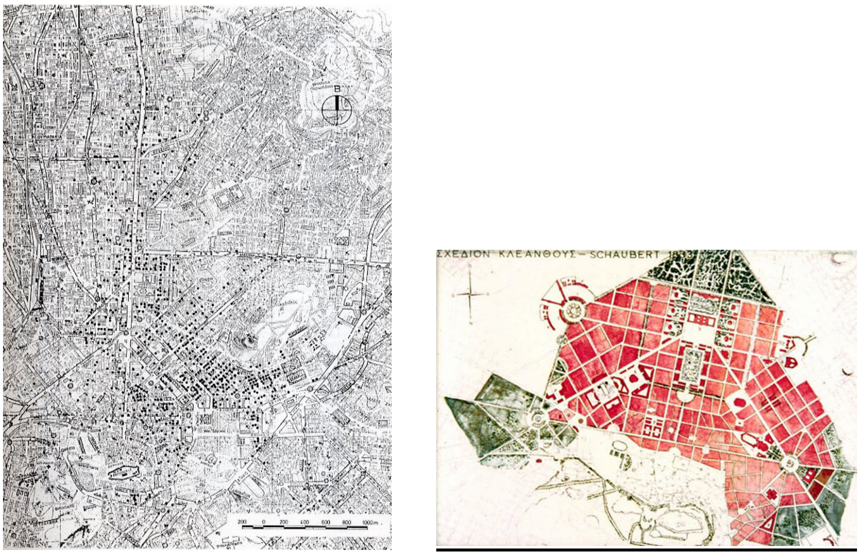


Fig. 5. With dots, is shown the development of multi-storey buildings of Athens between 1919 – 1941 [after 26, p. 115] that follows closely the Kleanthous and Schaubert plan, in 1833 (with the exception of the quarter of Agora, Plaka, Psiri) and expands mainly to the north and west.

Of prime importance is however the damage that results from the use of the legal instrument of usucaption outside the core of the city. As I. S. Makris highlights [25, introduction] it is the massive change of the land use character from that *to be used by displaced persons* (i.e. allotments) to that of *immiscibly residential* that makes today so pressing the need for developing a new agricultural legislature in Greece. The writer examined two recent and typical claims of *ownership* with the legal instrument of *usucaption*, though without access to the complete set of records and the spatial data. The first one is about an inheritance where the argumentation develops around the successive acquisition of the ownership. The compulsory law 431/1968 does not recognize that the beneficiaries *possess* (in the meaning of *usufruct*) the allotment unless the possession is *factual* and thus this land maybe acquired by *usucaption* also from third parties. The second case is about a land use character change. Here the decision calls again for the compulsory law 431/1968 and the argumentation seems to contradict the Greek civil code which with its article 1055 prohibits the acquisition of rural land of the type of allotment via *usucaption* (whether foreseen or unforeseen). The reference to the agricultural code and the compulsory law, which comes from the years of the dictatorship in Greece (1967-1974), enters to an amount of detail that overlooks why the current legal framework for sustainable development treats the urban and rural land as one whole (Law 2508/1997). By assuming that the land parcel was disseminated by the Greek state to be used by refugees, it can further be argued that the intention was not to support a policy (in its civic meaning) for 30 or the most 50 years that does not amount to half the life-cycle of a generation. According to [22, p. 154 and 377] «the Fund for Refugee Assistance was set up hastily to deal with the emergence in 1922-25. Meanwhile, however, the Refugee Settlement Commission (RSC) was established in 1923 by the Greek government in agreement with the League of Nations as an autonomous supra-national body to administer the permanent settlement of urban and rural refugees in residence and productive work. It launched an epic enterprise and was dissolved in 1930, when its funds were spent».

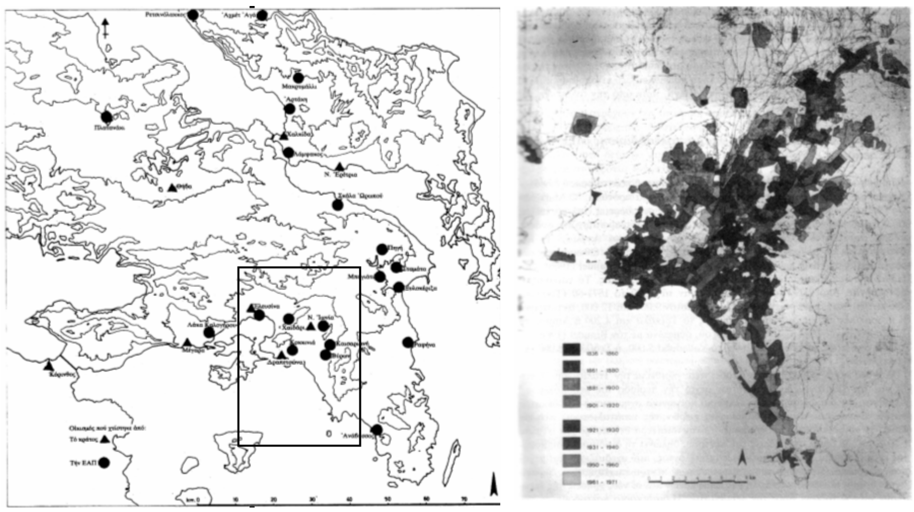


Fig. 6. On the left, the refugees settlements in Attica, 1923-26. Greece (League of Nations, 1926, unique map. After [22, p. 157]. On the right, the expansion of Athens plan 1836-1971 [22, p. 49].

An area in Northern Athens that is becoming highly urbanized was selected for examining the spatial factor. What was found interesting in the allotment plan, which is until date used for eliciting the land parcels' boundaries, are the natural, almost free hand lines, on top of the diagram. According to the official regulations of surveying works of the year 1954, the cadastral details to be included in a (cadastral) diagram are, among others, the resolving lines of cultivations and different types of land and specifically of fields, grasslands/pastures, diverse plantations, heaths and sandy/barren areas, swamps, lakes, bushes, forests and other trees [29, p. 26]. For the forest and mountainous lands the same regulations [29, p. 30] specify that perimetric surveying is usually taking place to end with the boundary of the agricultural area of the village to be used for the animals grazing. The conclusion drawn was that these lines depicted the limits of a forest, or of a land to be used for grazing. A contemporary image of Google Earth verified the first assumption.

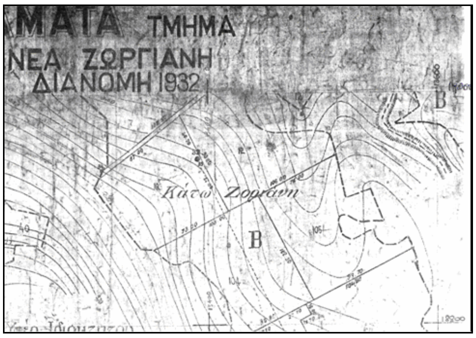


Fig. 7. Extract of allotment plan of the year 1932 in northern Athens with splines around the forest. Original scale 1:2000.

3 Ontology development with presentation of the first taxonomies

Taxonomies allow to hypothesize that though it would be extremely surprising to find today an *agricultural holding*, with its common meaning, untouched from the 19th century at an area like King's Cross, it may not be entirely surprising to locate a type of *allotment*, that seems to have been common also in the British tradition after the war, though not necessarily as an adjunct to apartments in building blocks, as it happed in central Europe [1, p. 189], or as a piece of land disseminated in rural or forest areas, as in Greece in the early 20's, but for the relevant purpose of land to be used for agricultural work to betters one's every day living conditions.

3.1 Real property rights and cadastres

Different law systems analyzed in [9] reflect the status of major legal families and the current study intends to focus more on the impact of Germanic law in civil law systems with roman roots, like the Greek one and on the England and Wales' traditions.

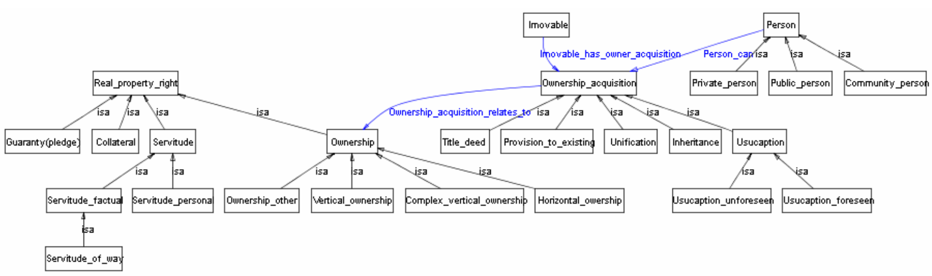


Fig. 8. Extract from real property rights taxonomy with emphasis on ownership, based on interpretation of Greek Civil Law.

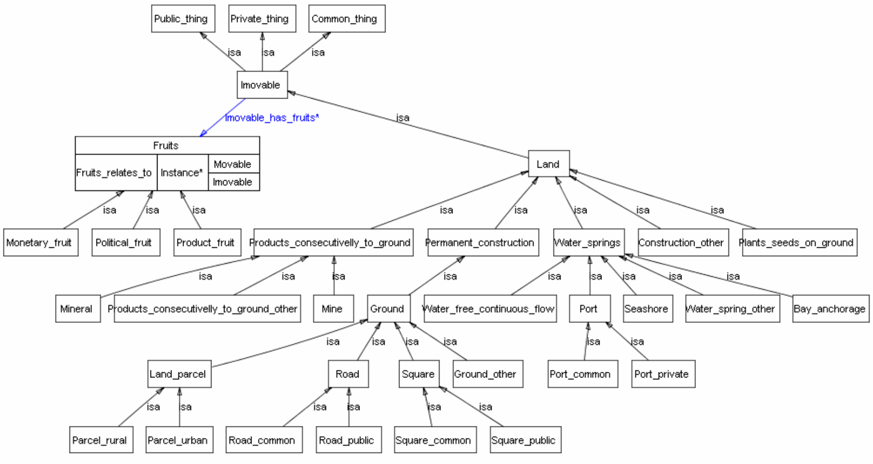


Fig. 9. Extract of property rights taxonomy, based on interpretation of Greek Civil Law.

3.2 Rural Planning

Rural planning as a perspective of land use is primarily meant to provide an understanding framework for the process of producing food, feed, fiber and other goods by the systematic raising of plants and animals. As a perspective of the triad, Work Place and Folk is meant to provide an understanding of the production and change of rural space.

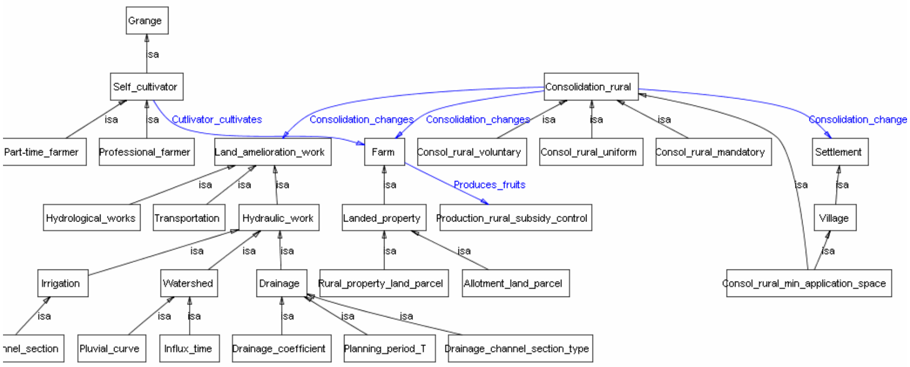


Fig. 10. Extract from the taxonomy for the consolidations, based on interpretation of Greek agricultural code and a typical environmental impact assessment and hydraulic works study.

Fig. 11. Extracts from the taxonomy for the topographic works in rural areas.

3.3 Urban Planning

Different directions in planning exist worldwide that include the application of methods like the: Mixed Scanning, Disjointed Incrementalism, Urban Management, Corporate Planning, Participatory Planning, Advocacy Planning, Negotiation – Mediation and Self help and other [1, p. 103 -109]. The ontology with its code provides a framework of land use planning activities at different community levels of the built and social environment.

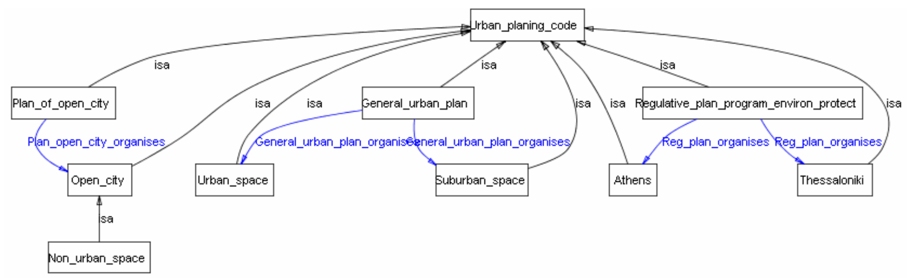


Fig. 12. High level taxonomy based on the Greek law for the sustainable development of the cities and settlements of the country [19].

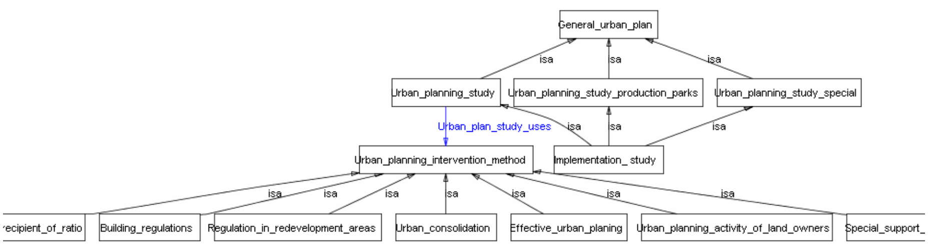


Fig. 13. Focus on the general urban plan [19].

3.4 Related work

The explicit treatment of levels of reality, purpose and perspectives of this reality proved to lead to a system that fails to look holistically on the way settlements are developed, as shown on the following figure.

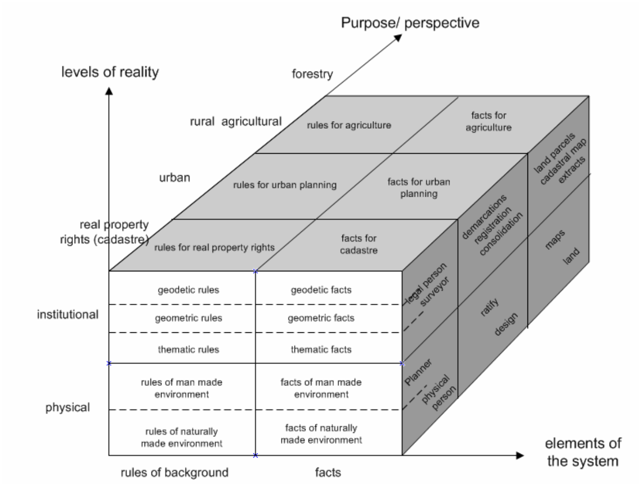


Fig. 14. A model of a Land Information System reality (extension from [10], [3])

3.5 Geometry

The ontology of geometry can encompass the theory of sets [24, [36]. A landed property can thus be represented, according to some scale and theme of interest, as:

<parcel, polygon> or
<parcel, point> or both
<parcel, polygon>, <parcel, point>

Landed property (here the term ‘landed property’ expresses the relationship: Real World – Model) after a consolidation is expected to benefit in the best possible way from the new development (here the term ‘landed property’ expresses the relationship: Model – Computer representation) and questions of proximity are posed. By representing these lands as polygons, or as unions of polygons and the irrigation system as a union of lines, which in turn are produced as a chain of line segments, the process of *interpretation*, *inference* and *integration* that constitutes Y. Newman's [30] *transgradience* can be applied, with the section and symmetrical difference of the two sets being closer to what by intuition is understood as pertinent, for example, to the *spatial relation of touch*.

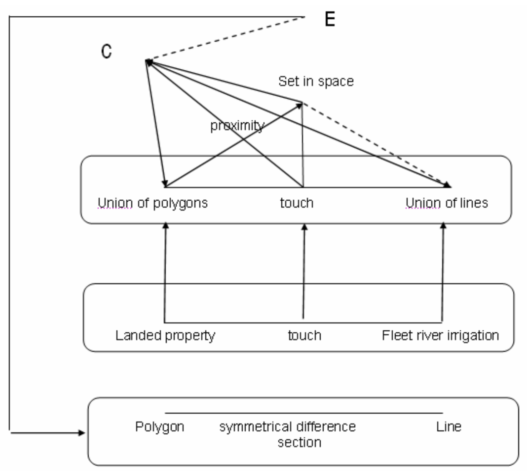


Fig. 15. Higraph formalism for moving to the higher order of conceptualization (adaptation from [30])

3.6 More on representation

M. F. Worboys [39] proposes a notation that distinguishes between the:

- *Identity indiscernibility*: The representation fuses the identities of objects in the target domain (e.g. a collection of distinct trees (citrus tree, olive trees, pines etc) that are merged into a single map object)
- *Typological indiscernibility*: The representation fuses the types of objects in the target domain (e.g. when all these distinct trees are presented as distinct map elements that are assigned the type tree.

3.7 Conflicts and constraints

The integration of the three ontologies of Rural Planning, Urban Planning, Real Property Rights is expected to reveal conflicts, but it should be possible, in most cases, to reach answers from the objectives set by the holistic sustainable development. Consensual type conflicts are expected to lead to constraints as illustrated with the experimental usage of Gottlob Frege's classic notation of *concept writing* [32, p. 10].

All buildings are physically connected to a land-parcel

Paraphrased to:

For every x, if x is a building, it is false that for all y, if y is a land-parcel, x is not physically connected to y

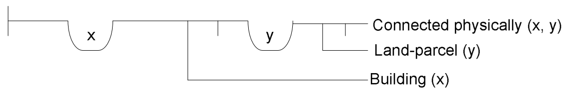


Fig. 16. Constraints with the use of sentential calculus

The instances of houses in the ontology are subject to the defined constraint. What would happen if the function of the 'physical connection' is substituted with that of an abstract one, like the 'ownership'? Or if the term 'building' is substituted with the term 'house'? The relation: 'house' is a 'building', is valid if all real world objects in the system are buildings and a method for treating successfully exceptions at different levels is also needed.

4 Conclusions and future work

The need to examine closer and more thoroughly the real life condition of villages, neighbourhoods, towns, cities, metropolis and their relation to the current and often problematic legislature was presented with a first approach of Athens and London and a first review of the elements that pertain to the three ontologies of urban, rural planning and real property rights. The focus in England and Greece, shows that the examination of different legal systems and planning traditions can help, instead of hinder, the elicitation of formal answers to questions that reveal in land management that are receptive of automation. However, the suitability of Patrick Geddes social ecology and schemata for the purposes of urban and rural planning and its presentations need to be further understood (e.g. their fashion of usage for defining the main subjects of a regional survey of a cultural landscape) with examples from both countries, especially in Greece, as a further means to judge its suitability for utilization in *Geographic Information Systems* that rely on *sentential calculus* for the modeling and production of spatial forms. In this line of work, there obviously exist an interest on the theory of fractals for utilizing the idea of *self similarity in the way variation is repeated at one scale to another* to model spatial patterns that are not 'easily noticed'.

Acknowledgments. The author would like to thank Prof. Ray Harris with whom the communication always brought something exciting, like the closer studying of London and its traditions. Prof. Myra Spiliopoulou and Dr. Epaminondas Kapetanios for the idea to write this article to address the particular community. The former colleagues at Ktimatologio S.A. for the development of so serious problematization on the Greek reality and especially Ms. Amaryllis Deligianni who brought into light the intriguing nature of usucaprion.

References

1. Aravantinos, I. A.: Urban Planning, Issues from Theory and Practice, Symmetria Editions (in Greek) (1990)
2. Barker, F., Jackson, P.: The History of London in Maps, Barrie & Jenkins, London (1990)
3. Bittner, S.: An agent-based model of reality in a cadastre, PhD thesis, Institut for Geoinfornation and Landesvermesung, Wien (2001)
4. Camden's sustainable community strategy 2007-2012, Camden Town Hall, March 2007, London (2007)
5. Charles Booth's survey of London poverty 1898-99. Available at: <http://booth.lse.ac.uk/>
6. Davis, T.: John Nash. The Prince Regent's Architect, Country Life Limited, London (1966)
7. Digimap map collection by EDINA. Edinburgh University Data Library for UK tertiary education and research community. Available at: <http://edina.ac.uk/digimap>
8. Economic Commission of Europe, UNECE: Guidelines on real property units and identifiers, (2004)
9. European University Institute (EUI): Real Property Law and Procedure in the European Union, (2005) Available at: <http://www.iue.it/ >
10. Frank, U. A., Bittner S., von Wolf A.: The Structure of Reality in a Cadastre
11. Geddes, P.: Civics: as applied sociology, Paper read before the Sociological Society at a meeting in the School of Economics and Political Science (University of London) (1904). Available at: <http://www.gutenberg.org/etext/13205>
12. Geddes, P.: Cities in evolution. An introduction to the town planning movement and to the study of civics, Ernest Benn Limited, London (1968) (first published in 1915).
13. Hall, P.: Cities of tomorrow, Backwell Publishing, (2002), (first published in 1988)
14. Hellenic Cadastre legislature and interpretation documents, in Greek. Available at: <http://www.ktimatologio.gr>
15. Hardy, D.: London: The Contradictory Capital, in David L.A. Cordon, ed.: Planning Twentieth Century Capital Cities, Routledge, (2006).
16. Hibbert, C., Weinred, B.: London Encyclopaedia, Macmillan London, (1983)
17. Hillier, B.: The common language of space: a way of looking at the social, economic and environmental functioning of cities on a common basis. Available at: <http://www.ucl.ac.uk/>
18. King's Cross redevelopment plan. Available at: <http://www.kingscrosscentral.com/>
19. Law 2508/1997: Sustainable development of cities and settlements of the country and other arrangements, The statutory office, (1997), in Greek. Available at: <http://www.minenv.gr/1/13/131/13108/g1310827.html>
20. Le Corbusier: The Chart of Athens, Ipsilon Editions, (1987), in Greek.
21. Lefebvre, H.: The production of space, Blackwell (1991)
22. Leontidou, L.: Cities of Silence. Working-class colonization of Athens and Piraeus 1909-1940, Politistiko Idrima ETBA, (2001) (first published in 1989), in Greek.
23. Leight, Mike: High Hopes, Film Four International in Association with British Screen, Portman Productions Ltd (1988)

24. Lelek, A.: Introduction to sets and topology, Trochalia editions (original title: Zbiory, edited in 1964) (1992) (in Greek)
25. Makris, I.: Agricultural Legislature, Sakoulas Ed., (2001), in Greek.
26. Marmaras, M.: The urban multi-storey building of interwar Athens, Politistiko Idrima ETBA, (1991), in Greek.
27. Meller, H.: Patrick Geddes. Social evolutionist and city planner, Routledge, (1990).
28. Ministry of Agriculture, Direction of Land Reforms & Synergism, Division of Consolidations: Analysis and guidelines for the implementation of the new law 674/1977 «For Land Consolidation etc», The statutory office, Athens, (1981), in Greek.
29. Ministry of Agriculture, Direction of Topographic Works: Regulations of topographic works, The statutory office, Athens, (1954), in Greek.
30. Neuman, Y.: A theory of meaning, Information Sciences 176 1435 – 1449 (2006)
31. Ross, R., Mapping changes: Visualizing King's Cross, MSc dissertation in Modernity, Space and Place, Dept. of Geography, University College London (2005)
32. Sowa, J. F.: Knowledge representation. Logical, philosophical and computational foundations, Brooks/Cole (2000)
33. Survey of London. Volume XXIV. King's Cross neighbourhood. The parish of St. Pancras. Part IV, London county council (1952)
34. Susan, P. P.: Faith in History. Armenians rebuilding community, Smithsonian Institution Press, Washington and London (1997)
35. The London Plan, (2004). Available at: < http://www.london.gov.uk/mayor/strategies/sds/index.jsp>
36. Tarski, A.: Introduction to logic and the methodology of deductive sciences, Oxford University Press (1965)
37. Welter, V. M.: Biopolis. Patrick Geddes and the City of Life, MIT Press (2003)
38. Welter, V. M.: Post-war CIAM, TeamX, and the influence of Patrick Geddes, Conference on TeamX between modernity and everyday (2003) Available at: <http://www.team10online.org/>
39. Worboy, M. F.: Communicating geographic information in context, Meeting on Fundamental Questions in GIScience, Manchester (2001)

The main goal of the Action C21, of the European programme for Cooperation in the field of Scientific and Technical Research, is to increase the knowledge and promote the use of ontologies in the domain of Urban Civil Engineering (UCE) projects, in thewith a view of to facilitating the communications between information systems, stakeholders and UCE specialists at a European level.

On 17-18th October 2007, in Turin, Italy, the Action, which is known informally as "Towntology", convened a successful Workshop to address emerging issues in the field.
The emphasis was on the developing a deeper understanding of how ontologies work in practice with a view to informing the development of future ontologies and conceptual tools that will make communication between different urban development disciplines easier.

The theme of the Workshop was set by Working Group 3 of the Action. This Group is interested in the socio-technical issues that emerge during the development and use of ontologies in organisations. It recognises that the introduction (and revision) of ontologies will impact on working practices within organisations. So WG3 seeks to examine the use of ontological frameworks and systems in practice and to find out how actual use differs from designers' intentions and what we can learn from these anomalies to design better ontologies in future.

This volume presents the contributions (accepted after a reviewing process) to that workshop, in order to capture the essence of the meeting.

ISBN 00-0000-000-0

ESCULAPIO