CONCEPTUAL MODELLING OF ENVIRONMENTAL KNOWLEDGE: DYNAMISM AND CONTEXT[1]

PILAR LEÓN ARAÚZ
ARIANNE REIMERINK
Universidad de Granada

RESUMEN

Las bases de datos terminológicas (BDTs) deberían reflejar estructuras conceptuales que guardaran cierto parecido con el modo en el que los conceptos se encuentran representados en la mente humana (Meyer et al. 1992), lo que incluye la representación de la variación contextual. En EcoLexicon, una BDT medioambiental, se han contextualizado tanto las redes conceptuales como la información gráfica de acuerdo a ciertas restricciones basadas en roles y dominios. Los dominios restringen las relaciones activadas por los conceptos en función del comportamiento de sus referentes en el mundo real. Por otro lado, la recontextualización según rol es independiente del dominio y ofrece nuevas redes conceptuales en forma de clases genéricas. Cada red recontextualizada es además enriquecida con imágenes prototípicas de cada dominio y rol.

Palabras clave: contexto, dinamismo, recontextualización, conocimiento medioambiental, BDT, información gráfica

ABSTRACT

Terminological knowledge bases (TKBs) should reflect conceptual structures in a similar way to how concepts relate in the human mind (Meyer et al. 1992), which includes the representation of contextual variation. In EcoLexicon, an environmental TKB, conceptual networks and graphical information have been contextualized according to role-based and domain-based constraints. Domain membership restricts concepts' relational behaviour according to how their referents interact in the real world. Semantic role recontextualization, on the other hand, is domain-independent and provides new conceptual networks in the form of upper-level conceptual *classes. Each recontextualized network is provided with images that are most prototypical.*

Keywords: context, dynamism, recontextualization, environmental knowledge, TKB, graphical information

1. INTRODUCTION

EcoLexicon[2], a multilingual terminological knowledge base (TKB) on the environment, provides an internally coherent information system which aims at covering a wide range of specialized linguistic and conceptual needs. TKBs should reflect conceptual structures in a similar way to how concepts relate in the human mind (Meyer *et al*. 1992), which includes the representation of contextual variation. Context includes external factors (situational and cultural) as well as internal cognitive factors, all of which can influence one another (House 2006: 342). This view goes hand in hand with the perception of language as a kind of action, where the meaning of linguistic forms is understood as a function of their use (Reimerink *et al*. 2010). In the linguistic community, all approaches seem to coincide in defining context as a dynamic construct (Austin 1962; Gadamer 1995; Sperber and Wilson 1986, 1995). However, term bases are often restricted to generic-specific and part-whole relations, whereas conceptual dynamism can only be fully reflected through non-hierarchical ones, such as the notions of movement, action and change, which are directly linked to human experience and perceptually salient conceptual features.

This paper focuses on the representation and recontextualization of conceptual information in dynamic networks and graphical resources in EcoLexicon. Recontextualization is carried out according to contextual domain (Section 2.1) and semantic role (Section 2.2). Conceptual recontextualization affects not only the conceptual networks, but also the complementary information provided, such as images (Section 3).

2. CONCEPTUAL RECONTEXTUALIZATION

In knowledge modelling, concepts are very often classified according to very different dimensions (shape, function, colour, etc.). Multidimensionality (Kageura 1997) is commonly regarded as a way of enriching traditional static representations, enhancing knowledge acquisition through different points of view in the same conceptual network (León Araúz and Faber 2010). As is well-known, the more relations that users are able to activate through a particular concept, the more knowledge they are likely to possess for the domain. In such a wide domain as the environment, multidimensionality increases the number of possible relations activated by specialized concepts, since it is also intimately linked to the semantic roles concepts may play. In a process-oriented domain (Faber *et al*. 2006) the same concept may act as an AGENT or a PATIENT, as an active PROCESS or a RESULT. For example, the concept WATER can be either an AGENT (in the process of EROSION) or a PATIENT (in WATER TREATMENT), which implies that WATER can be related to other concepts through the conceptual relation *causes* as well as *affected_by*. However, the environmental domain has caused a great deal of information overload, which ends up jeopardizing knowledge acquisition.

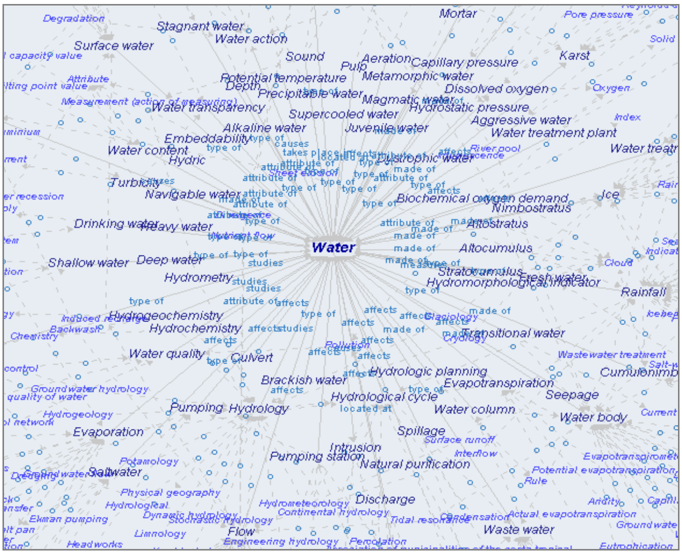

Figure 1. Information overload in the network of WATER

This is especially due to the fact that multiple dimensions are not always compatible but context-dependent. Although concepts are entrenched cognitive routines which are interrelated in various ways facilitating their co-activation, they actually retain enough autonomy that the execution of one does not necessarily entail the activation of all of the rest (Langacker 1987: 162). This is the case of certain concepts such as WATER (Figure 1), which have such a low degree of specificity that they can be involved in a myriad of events. For instance, even though WATER subtypes, such as PRECIPITABLE WATER, DRINKING WATER and NAVIGABLE WATER, all represent the same facet *function*, strictly speaking, they are not coordinate concepts, because they belong to different environmental paradigms that rarely coincide, if ever, in time or space.

Yeh and Barsalou (2006) state that when situations are not ignored, but incorporated into a cognitive task, processing becomes more tractable. In the same way, any specialized domain reflects different situations in which certain conceptual dimensions become more or less salient. As a result, a more believable representational system should account for recontextualization according to the situated nature of concepts.

2.1 Domain-based constraints

The environmental domain has been divided into different contextual domains according to corpus information and expert collaboration: HYDROLOGY, GEOLOGY, METEOROLOGY, BIOLOGY, CHEMISTRY, CONSTRUCTION/ENGINEERING, WATER TREATMENT/SUPPLY, COASTAL PROCESSES and NAVIGATION. Domain membership restricts concepts' relational behaviour according to how their referents interact in the real world. For instance, CONCRETE is linked to WATER through a *made_of* relation. Nevertheless, this proposition is irrelevant if users only want to know how WATER naturally interacts with the landscape or how it is purified of contaminants. Consequently, the proposition CONCRETE *made_of* WATER should only appear if users select the CONSTRUCTION/ENGINEERING context.

Domain-based constraints are neither applied to individual concepts nor to individual relations, since one concept can be activated in different contexts or use the same relations but with different values. Constraints are instead applied to conceptual propositions (León Araúz *et al*. 2009). As a result, when constraints are applied, WATER only shows relevant dimensions for each contextual domain. In Figure 2, WATER is only linked to propositions belonging to the context of ENGINEERING/CONSTRUCTION.

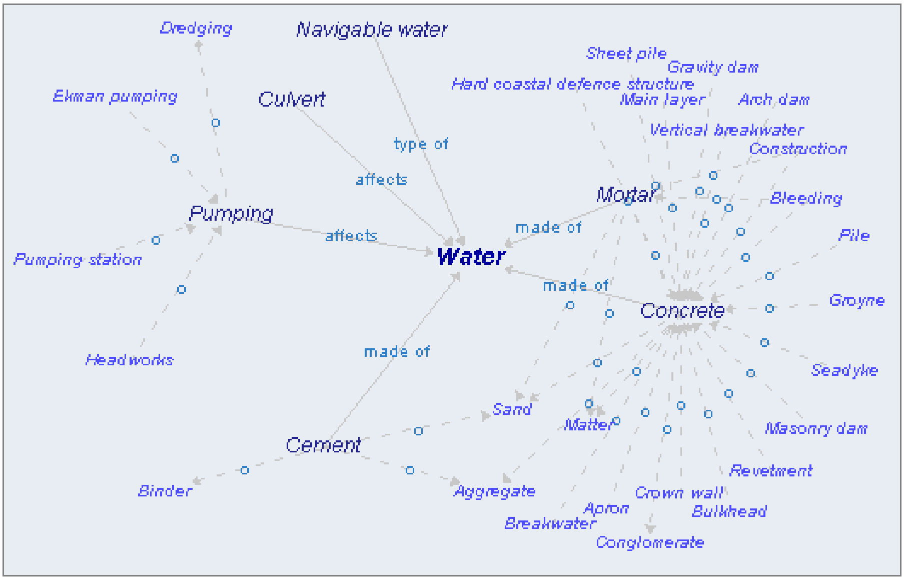

Figure 2. WATER in ENGINEERING/CONSTRUCTION

However, in Figure 3 the GEOLOGY domain shows WATER in a new structure with other concepts and relations. The number of conceptual relations changes from one network to another, as WATER is not equally relevant in all contextual domains. Furthermore, relation types differ too, which also highlights the changing nature of WATER's internal structure in each case. For example, in the ENGINEERING/CONSTRUCTION contextual domain, most relations are *made_of* and *affects*, whereas in the GEOLOGY domain, *causes* and *type_of* stand out. *Affects* is also shared by the GEOLOGY domain, but the arrow direction shows a different perspective, since in geological contexts WATER is a much more active AGENT than in

ENGINEERING/CONSTRUCTION, where the concept is more subject to changes (PATIENT). Finally, WATER is not always related to the same concept types. In ENGINEERING/CONSTRUCTION, WATER is only linked to artificial entities or processes (PUMPING, CONCRETE, CULVERT), while in GEOLOGY it is primarily related to natural ones (EROSION, GROUNDWATER, SEEPAGE).

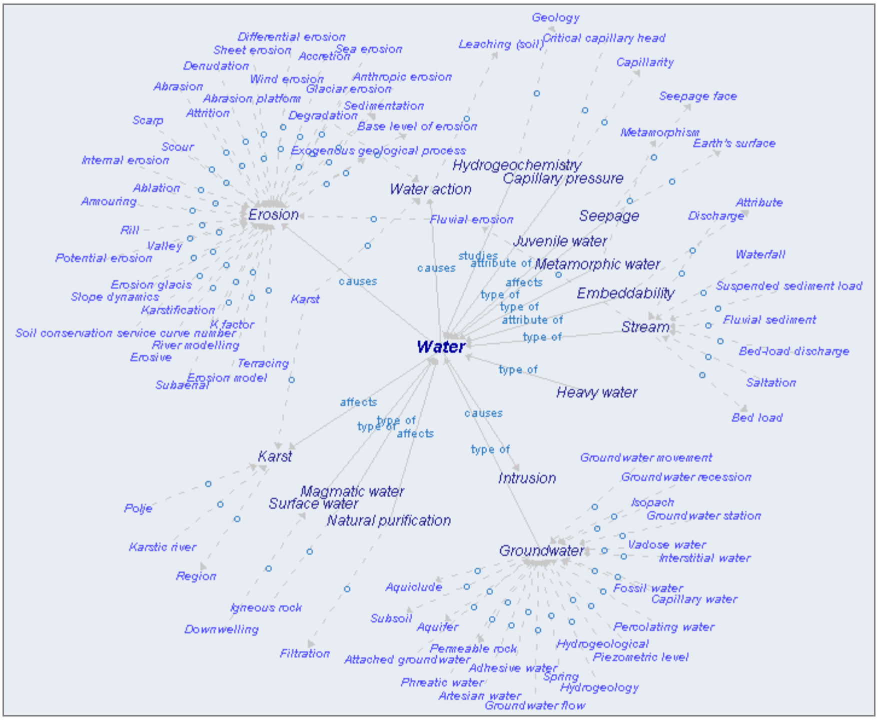

Figure 3. WATER in GEOLOGY

2.2 Role-based constraints

Semantic role recontextualization is domain-independent and offers new networks in the form of upper-level conceptual classes. In this way, users can visualize how concepts like WATER behave either as an AGENT or a PATIENT in all kinds of events. This highlights certain relational constraints associated to the natural aspect of concepts and not to those of their referents. Thus, role-dependent networks will be characterized by a certain type of relations.

Interestingly enough, hierarchical relations are invariable parameters (León Araúz and Faber 2010). Entities may *have parts* or be *part of* other wholes whether they are AGENTS or PATIENTS, but that is not the case for non-hierarchical relations. If an entity behaves like a PATIENT it cannot *affect* anything, as it would then become an AGENT. Prototypically, a PATIENT can only activate its inverse relation, *affected_by*.

Role-based relational constraints are applied to individual concepts according to their own perspective in a given proposition. For example, in WATER CYCLE *affects* WATER, WATER is a PATIENT. This kind of constraints can only be applied to the first hierarchical level, since they are focused on a particular concept and not its whole conceptual proposition. In the next figures, the overloaded network of WATER (Figure 4) is restricted according to the AGENT role (Figure 5).

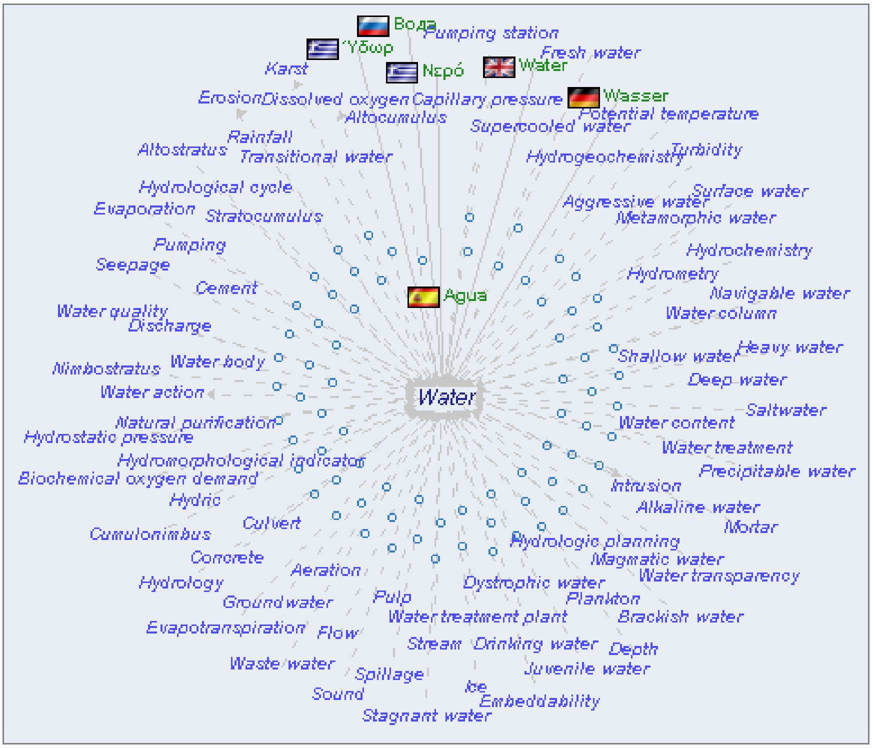

Figure 4. Role-free network of WATER

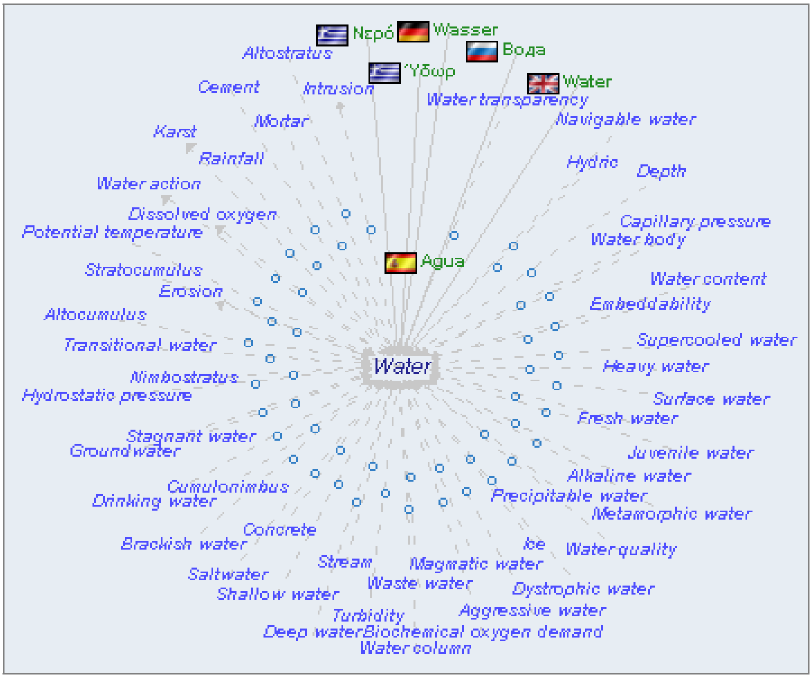

Figure 5. Agent-based network of WATER

Actually, role-based domains by themselves are not sufficient to recontextualize knowledge in a meaningful way. In the role-free network, WATER appears linked to 72 concepts, whereas in the role-based one, WATER is related to 50. Despite the difference, the concept still appears overloaded, especially once the second hierarchical level is displayed. However, contextual domains, although usually dominated by one role, restrict relational power of versatile concepts in a more quantitative way.

2.3 Intersection of role- and domain-based constraints

A new recontextualization can take place with the intersection of role- and domain-based constraints. For example, WATER can be framed as an AGENT (Figure 6) or a PATIENT (Figure 7) or even both (Figure 8) within the HYDROLOGY domain.

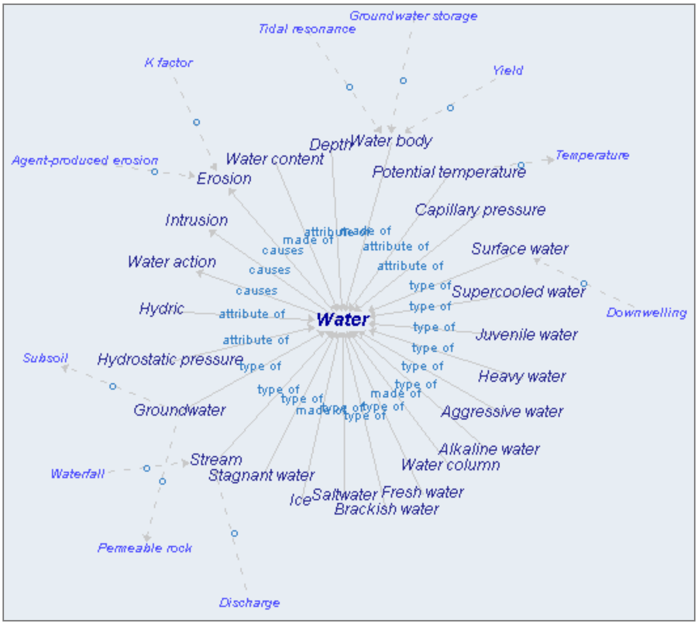

Figure 6. WATER as an AGENT in HYDROLOGY

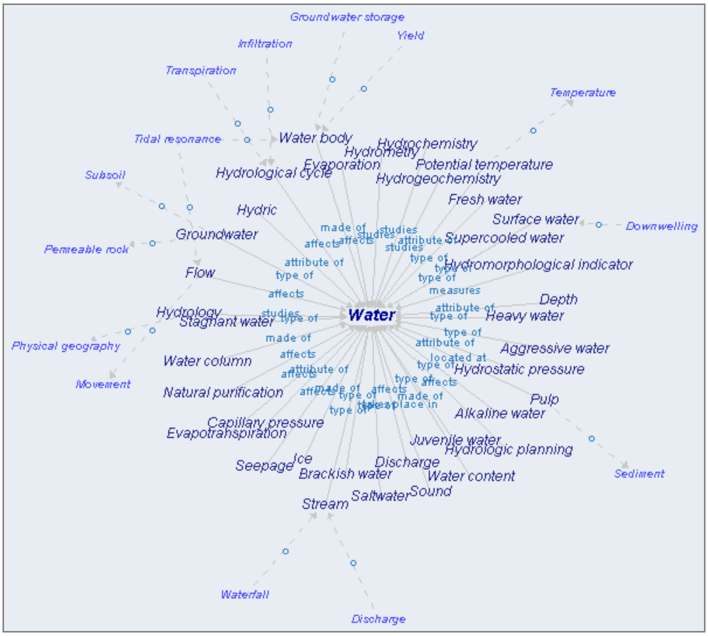

Figure 7. WATER as a PATIENT in HYDROLOGY

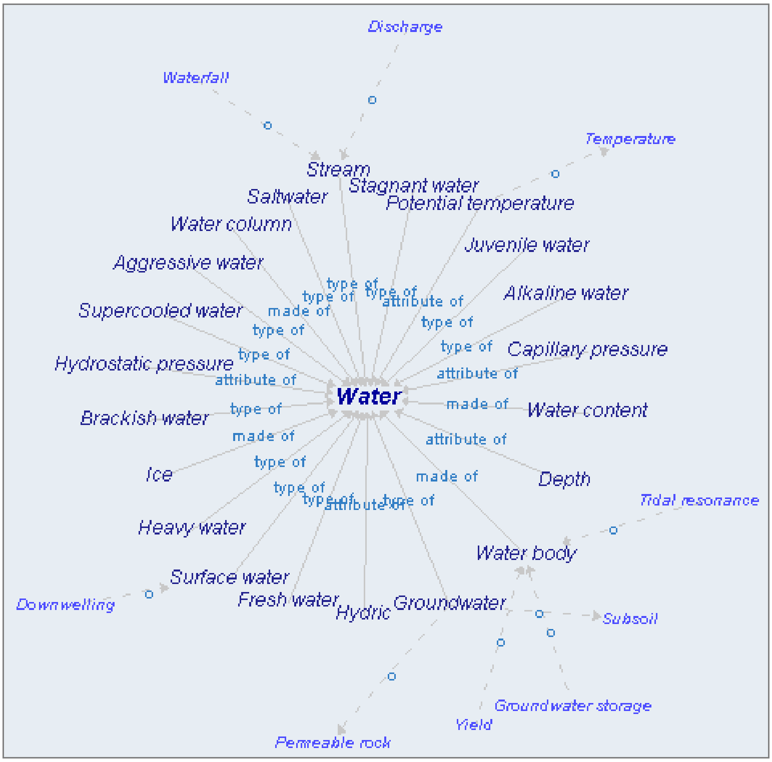

Figure 8. WATER as an AGENT and PATIENT in HYDROLOGY

Now, the first level appears constrained according to different roles in a particular contextual domain, which at the same time applies for the second level. It is worth noting that Figure 8 only shows hierarchical relations (*type_of, attribute_of*, *made_of*), because these are the only ones shared by concepts that can be AGENTS or PATIENTS. In Figure 6, however, the representation adds the relation *causes*, typical of AGENTS, and in Figure 7, it adds propositions where WATER is *affected_by*, *measured*, *studied* or *located_at*.

3. VISUAL RECONTEXTUALIZATION

Conceptual recontextualization affects not only the conceptual networks, but also the complementary information provided, such as images. For each contextual domain a prototypical domain image is selected. The image in Figure 9 represents the HYDROLOGY domain depicting the WATER CYCLE as the process through which WATER continually circulates between the earth and the atmosphere, the prototypical event of the domain. The geographical background of the image shows a high degree of iconicity, which makes real world entities easier to identify. Among others, we can identify the mountains, ocean, and sky with clouds. In order to show where each process of the WATER CYCLE takes place, labels such as ‘Condensation’, ‘Transport’ and ‘Precipitation’ are inserted in the image. The dynamism inherent to this cyclical process is added to the image with arrows, which graphically show the direction of the different processes of WATER EXCHANGE, involving any body of WATER and the AIR (e.g. SNOW, RAIN, and WATER VAPOUR).

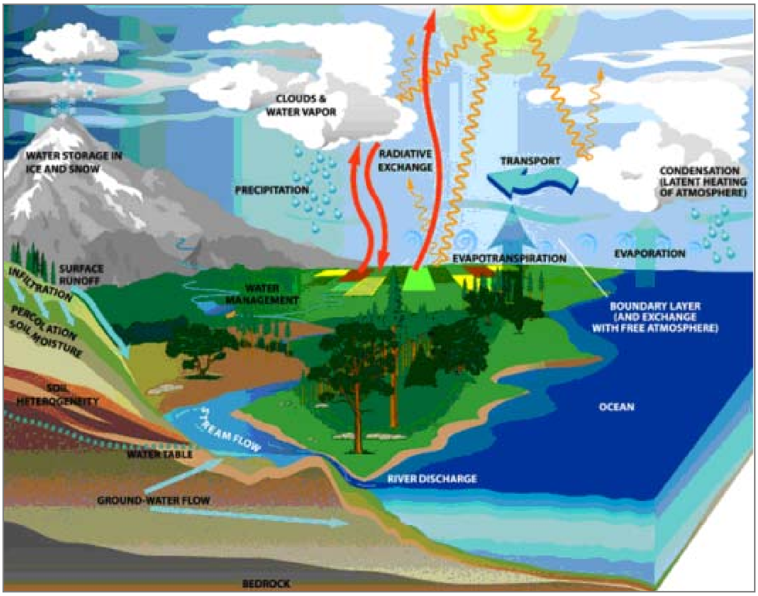

Figure 9. Prototypical image for hydrology

Figure 10 shows the prototypical event image of the WATER TREATMENT domain. It depicts a WATER TREATMENT PLANT and shows how WASTE WATER is cleaned through several processes and turned into a reusable effluent.

The WATER TREATMENT image is less iconic, as many elements of the water treatment machinery are illustrated with more abstract figures, such as the SCREW PUMP (illustrated with the fat black lines in screw form in a rectangular area) and the SCREEN (a grid in a rectangular area) at the top. There is no iconic background in this image, although it is obvious that each part of the process takes place in a different location. In addition, the top-bottom sequential order of this image and the arrows clearly show there is a beginning and an end to the process.

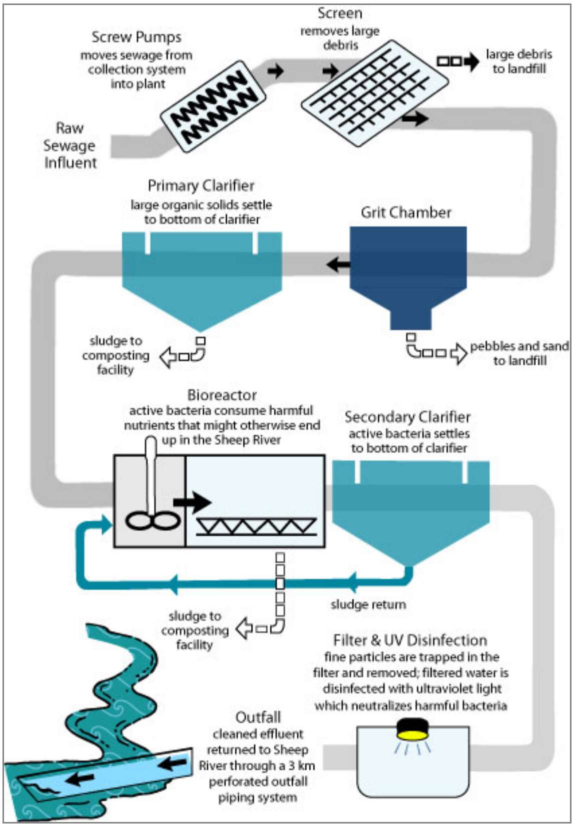

Figure 10. Prototypical image for WATER TREATMENT

The differences may be due to the fact that the WATER TREATMENT process is a human-induced process and therefore has a specific function. This can be seen in the higher degree of abstraction in the image and the increase in textual information. Even though both images depict a cyclic process involving WATER, they clearly frame the concept in two different processes that are incompatible in space and time.

Generally, prototypical domain images are all representations of dynamic events showing movement and change in time. Both HYDROLOGY and WATER TREATMENT prototypical event images describe a complete cyclical process with WATER at its centre and combine iconic and abstract visual clues to describe it as an AGENT or PATIENT respectively. Thus, WATER could be considered the prototypical concept within these two domains. This is why the WATER CYCLE and WATER TREATMENT images also correspond to the image shown in the WATER entry when framed in each domain and role. However, in the GEOLOGY domain, WATER is not as predominant. Therefore, its prototypical image does not only focus on WATER, but also on rocks and magmatic processes. In Figure 11, WATER is partially represented as one of the agents in the cyclical process of geological formations. On the right upper side, WATER is depicted similarly to the image of the WATER CYCLE, but here the focus is on WATER as the agent of WEATHERING, TRANSPORT and DEPOSITION.

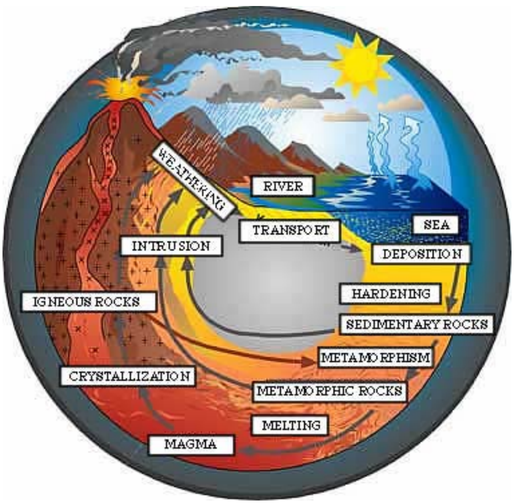

Figure 11. Prototypical image for GEOLOGY

Therefore, when framing WATER within the GEOLOGY domain, another more explicit and concrete image must be shown in order to focus on the active role of the concept in that particular domain (Figure 12).

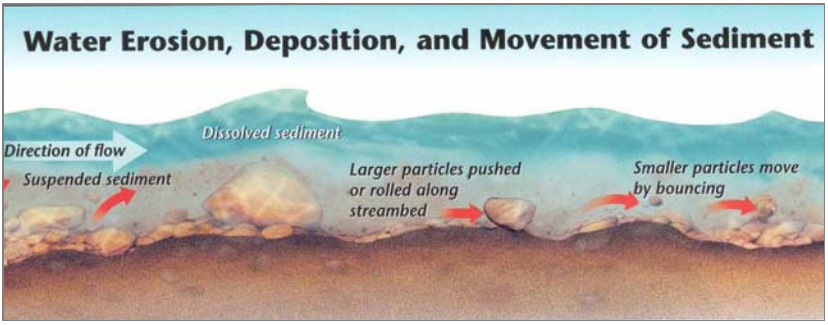

Figure 12. WATER in GEOLOGY

Agents need a procedural image to be represented as such. The image in Figure 12 is characterized by (1) iconicity of the natural elements involved in the process (water, bottom, particles of different sizes); (2) abstract visual clues (arrows with different colours) representing the dynamism of the process based on time, and (3) textual labels and explanations guiding the user through the image.

4. CONCLUSIONS

For process-oriented and multidisciplinary domains such as the environment, conceptual contextualization provides a qualitative criterion for the representation of specialized concepts in line with the workings of the human conceptual system, as well as a quantitative solution to the problem of information overload. On the one hand, conceptual dynamism is the main cause of multidimensionality in this domain; therefore the conceptual contextualization of different entries is performed according to role-based domains and contextual domains. On the other hand, visual contextualization is carried out 1) representing each contextual domain through a prototypical image, depicting the prototypical event of the domain, and 2) representing several images for one concept, based on the relation type of the concept with each of the domains in which it participates.

NOTES

[1] This research has been funded by project FFI2008-06080-C03-01/FILO, from the Spanish Ministry of Science and Innovation.
[2] http://manila.ugr.es/visual

REFERENCES

Austin J. L. 1962. *How to Do Things with Words*. Oxford: Clarendon.

Faber P., Montero Martínez S., Castro Prieto M.R., Senso Ruiz J., Prieto Velasco J.A., León Araúz P., Márquez Linares C., Vega Expósito M. 2006. “Process-oriented terminology management in the domain of Coastal Engineering”, *Terminology*, 12(2): 189-213.

Gadamer Hans G. 1995. *Truth and Method*. New York: Continuum.

House J. 2006. “Text and context in translation”, *Journal of Pragmatics*, 38: 338-358.

Kageura K. 1997. “Multifaceted/Multidimensional concept systems.” In Wright S.E. and Budin G. (eds.). *Handbook of Terminology Management: Basic Aspects of Terminology Management*. Amsterdam/Philadelphia: John Benjamins, 119-32.

Langacker R. W. 1987. *Foundations of Cognitive Grammar: Theoretical Prerequisites*. *Vol 1*. Stanford: Stanford University Press.

León Araúz P., Faber P. 2010. “Natural and contextual constraints for domain-specific relations.” *Proceedings of Semantic Relations. Theory and Applications*. Malta.

León Araúz P., Magaña Redondo P., Faber P. 2009. “Managing inner and outer overinformation in Ecolexicon: an environmental ontology.” *Proceedings of the 8th International Conference on Terminology and Artificial Intelligence*. Toulouse, France.

Meyer I., Bowker L., Eck K. 1992. “COGNITERM: An experiment in building a knowledge-based term bank”, 159-172.

Reimerink A., García de Quesada M., Montero Martínez S. 2010. "Contextual information in terminological bases: a multimodal approach", *Journal of Pragmatics*, 42-7.
Sperber D., Wilson D. 1986. *Relevance: Communication and Cognition*. Cambridge, MA: Mitt Press (2nd edition 1995. Oxford: Blackwell).
Sperber D., Wilson D. 1995. Postface to the second edition of *Relevance: Communication and Cognition*. Oxford: Blackwell.
Yeh W., Barsalou L. W. 2006. "The situated nature of concepts", *American Journal of Psychology*, 119: 349-384.