AN ONTOLOGY-BASED ENRICHED EBOOK FOR TEACHING STATISTICS

Jean-Marc Meunier, Samuel Szoniecky

AN ONTOLOGY-BASED ENRICHED EBOOK FOR TEACHING STATISTICS

Jean-Marc Meunier, Samuel Szoniecky
Laboratoire Paragraphe, Université Paris 8, France
Contact email : jmeunier@univ-paris8.fr

With the development of Internet and the use of digital pedagogy, the production of open educational resources grows very quickly. Sharing these resources has become a major challenge both from an economic point of view, but also for the dissemination and sharing of practices among teachers. The problematics are psychologic (help for learning), ergonomic (usability, acceptability) but also documentary (reuse resources) and editorial (enhancement of digital archives). Based on an ontology developed by University of Oxford and translated in French, we developed a digital archive of open educational resources for teaching statistics. We present and discuss e-book enrichment scenarios in undergraduate statistics courses. We present how this ontology-based approach allows thinking enriched e-books as an adaptive and collaborative learning device.

INTRODUCTION

Teaching statistics requires the coordination of a functioning logic (the theoretical basis of statistics) and a use logic (what test to choose and how to implement it concretely). The first presupposes the acquisition of a very abstract knowledge, the second presupposes the very fine particularization of this knowledge (Richard, 1983). This coordination is based on the structure of knowledge, in particular the organization of concepts and their properties. These properties can be structural or functional (procedural-related). For example, knowing which properties a sample must have in order to be able to apply a parametric test is the first logical aspect. Knowing how to structure a data table according to the software used, or how to adapt the procedure depending on whether you are working on individual data or a distribution, is a matter of the second. This articulation of knowledge and know-how through the structure of knowledge offers a formalization of the notion of competencies (Paquette, 2007) and at the same time offers ways to develop a logical chain of learning. Above all, it makes it possible to design new types of teaching materials that make it possible to break away from strict linearity of reading and thus offer its adaptability to the diversity of learners and situations in which teaching resources are consulted. It is this possibility that we propose to explore in this article.

FORMALIZE THE KNOWLEDGE STRUCTURE TO INSTRUMENTALIZE IT

This knowledge structure can be formalized in a domain ontology. It is the description of an area of knowledge using a set of hierarchical and interrelated concepts. It is based on a class taxonomy and a set of inference rules. Within the context of a project called Ontostats (http://www.ontostats.univ-paris8.fr/), supported by the French Ministry of Higher Education, Research and Innovation, we translated into French the ontology of STATO statistical concepts developed by the University of Oxford (Gonzalez, Rocca-Serra, Burke, & Sansone, 2014).

The interest of such an ontology is to allow the sharing of knowledge, not only about concepts but also about the relationships between them. Ontology is a vocabulary used to label concepts and the relationships between them. The separation of labels and concepts allows for a multilingual description of the domain of knowledge. Relationships are class inclusion relationships, but also properties or attributes that can take multiple values. They allow us to deduce what is usually implicit in the discourse and contribute to the development of decision support systems.
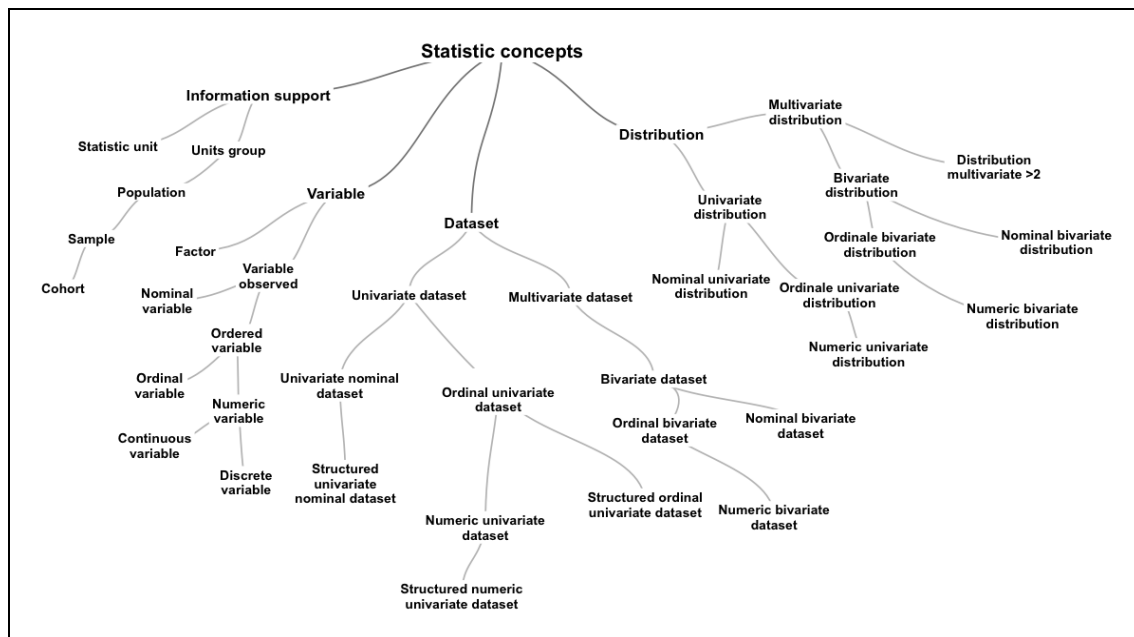
Figure 1 An extract from the ontology of statistical concepts (Meunier, Szoniecky, & Lamolle, 2017)

Ontologies also have many advantages for indexing educational resources compared to using a norm (Bouzeghoub & Elbyed, 2006). However, ontologies are promising especially for building learning aids. They provide a conceptual model for the study of learning disabilities. An ontology can be used to design training paths by adapting the structure of teaching sequences (Guo & Chen, 2007), to design resource recommendation systems based on relationships between concepts (Shen & Shen, 2005), and also when ontology is associated with learning traces analysis, course personalization (Castro & Alonso, 2011) and learner evaluation (Romero, North, Gutiérrez, & Caliusco, 2015).

The STATO ontology has been used to index a set of open educational resources, notably those developed by French thematic digital universities. These resources have been collected in a digital archive using Omeka-S. In addition, the archive offers each identified teacher the opportunity to work on his or her own collection of resources and to propose a personalized editorialization. It is on this last feature that our enriched digital book project is based.

TOWARDS A NEW CONCEPTION OF A DIGITAL TEXTBOOK

The book is a medium through which a teacher can express and structure knowledge at some time in order to enable a learner to acquire it. The appropriation of this knowledge through the reading of a textbook depends, of course, on the objectives of reading, the level of knowledge, the activities motivating reading and the knowledge related to the use of the reading device itself (use of glossary, search engine, navigation function). The teacher, when developing the pedagogical design and editorializing the educational resources, makes choices and places them in an original production that takes into account the context, domain and audience.

Armatte (2010) distinguishes between four approaches to the teaching of statistics: (i) a formal, hypothetical-deductive approach (focused on the teaching of axioms); (ii) an inductive or experimental approach, centred on simulations and data mining; (iii) a problem-oriented approach based on the logic of the social field or the problems of daily life; (iv) the historical approach consisting in presenting to students the consequences of problematization that led to the development of a particular method. We can add a fifth, procedural one, which consists in focusing on the when and how to do, therefore very dependent on the analysis tool, particularly software, used during teaching. The expertise coordinates its five approaches in a more or less important way, but it is rare to have the time, within the same course, to present them all. Depending on the discipline and the role that statistics play in it, a teacher may prefer one to another. In the same way, a student, depending on his or her previous background, may be more or less sensitive to one

or the other. Giving students a bridge between these different points of view is one of the challenges and not the least of them to allow access to these different points of view.

However, the current digital versions of the statistics textbooks because, for the vast majority of them, they are transpositions of the paper version offer very few advantages. For many actors in the field, e-books are nothing more than "the combination of a classical book structure, or rather the familiar concept of books, with characteristics related to the electronic environment" (Landoni, 2003). It does not offer the possibility of varying points of view and linking them to a more general body of knowledge, so that the book is enclosed within itself. To overcome this enclosure, it is necessary to imagine a new form of writing which according to Clément (2000) is characterized by three essential aspects: hypertextuality, its distribution on the network and finally its dynamic and multimedia character.

The first characteristic makes it possible to overcome the linearity of the discourse. If this is a valuable help for beginners, it can be a brake on learning as skills increase. It can even be blocking when the book is used in problem solving or synthesis activities on a given theme (Roy, 2017). With hypertexts, it becomes possible to browse a document in a non-linear way. However, we must be careful. Many studies in cognitive psychology show that a badly controlled hypertextual structure can be a major barrier to learning, but well used, an opportunity (Vörös, Rouet, & Pleh, 2011). We propose to rely on an ontology to allow the transition from a thematic structure, the table of contents, to a structure oriented towards the relationship between concepts and procedures, to see them at different levels while maintaining an overall coherence. Ontology becomes the guide to navigating and working on course content.
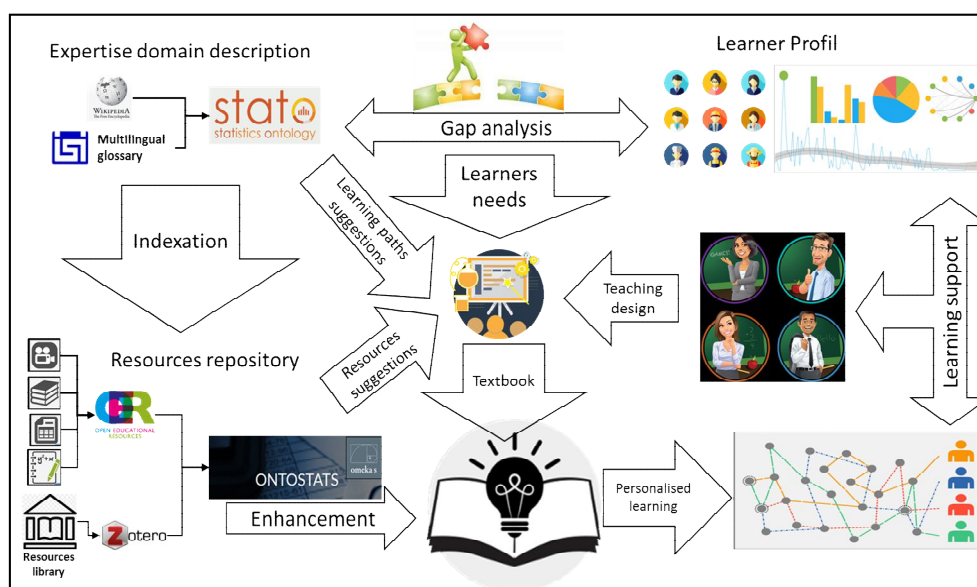

Figure 2 Principles of Enrichment of an e-book from an ontology

The second characteristic is distribution on the network. This is a reality that all teachers now know. A course can no longer be limited to itself. Students as needed complement it or compare it with other resources on the Internet. We may regret it and develop defensive strategies, or take advantage of the opportunity for diversity to enrich our course or offer students a personalized learning experience. This involves rethinking the way a course is written to take full advantage of computers by inserting the text into a larger set of documents that, if correctly indexed, make it possible to associate with a book additional or alternative open source and editorial resources that correspond to the teaching objective. The concept of document discussed here is very general as it concerns texts or multimedia documents, exercises, datasets, notes or synthesis of students' lectures, as well as forum or frequently asked questions or information stored in the ontology itself (definitions, multilingual lexicon and relations with other concepts). From this point of view, the digital manual is not just one of several documents on a learning management system (LMS), it is the LMS itself. Such a conception of the textbook makes it a dynamic object that can be adapted according to the needs of the learners, or even adapted to each

of them to offer them a personalized reading and learning experience. Figure 2 summarizes the relationships between ontology, resource repository and manual development.

CONCLUSION AND FUTURE WORK
The platform will be in production in the autumn of 2018, at the same time as the realization of a first prototype manual. At the same time, we have initiated two studies. The results of the first one on the use of open and editorial educational resources and electronic books should allow us to refine the specifications for such an electronic book. The second one aims to identify the profiles and needs of the learners in order to personalize the learning paths. This includes identifying a minimum number of questions to assess the learner's state in terms of expert knowledge and providing him/her with the most relevant resources and activities to help him/her progress.

REFERENCES
Armatte, M. (2010). Le rôle de l'histoire dans l'enseignement de la statistique. *Statistique et Enseignement*, *1*(2), 23‑47.
Bouzeghoub, A., & Elbyed, A. (2006). Ontology Mapping for Learning Objects Repositories Interoperability. In *Lecture notes in computer science* (p. 794‑797). SpringerSpringer, Berlin, New York.
Castro, F., & Alonso, M. A. (2011). Learning Objects and Ontologies to Perform Educational Data Mining (p. 532‑536). CSREA Press.
Clément, J. (2000). Le e-book est-il le futur du livre ? In *Les savoirs déroutés* (p. 129‑140). Lyon, France: ENSSIB / Association Docforum.
Gonzalez, A., Rocca-Serra, P., Burke, O., & Sansone, S.-A. (2014). Stato - Statistics Ontology. Consulté 13 février 2018, à l'adresse http://stato-ontology.org/
Guo, W.-Y., & Chen, D.-R. (2007). An Ontology Infrastructure for an E-Learning Scenario. *International Journal of Distance Education Technologies*, *5*(1), 70‑78.
Landoni, M. (2003). Electronic Books. In J. Feather & P. Sturges (Éd.), *International Encyclopedia of Information and Library Science* (2nd edition, p. 168‑171). London: Routledge.
Meunier, J.-M., Szoniecky, S., & Lamolle, M. (2017). Apports d'une ontologie à l'apprentissage des statistiques. Présenté à Colloque Francophone International pour l'enseignement des statistiques, Grenoble, France. Consulté à l'adresse http://cfies2017.sfds.asso.fr/
Paquette, G. (2007). An Ontology and a Software Framework for Competency Modeling and Management. *Educational Technology & Society*, *10*(3), 1‑21.
Richard, J. F. (1983). Logique du fonctionnement et logique de l'utilisation. Consulté à l'adresse http://hal.inria.fr/inria-00076356
Romero, L., North, M., Gutiérrez, M., & Caliusco, L. (2015). Pedagogically-Driven Ontology Network for Conceptualizing the e-Learning Assessment Domain. *Journal of Educational Technology & Society*, *18*(4), 312‑330.
Roy, E. (2017). *Statistiques mieux « dévorées » à l'aide d'un ebook ?* (master thesis). Université Paris 8, Saint-Denis, France.
Shen, L.-P., & Shen, R.-M. (2005). Ontology-based learning content recommendation (English). *International journal of continuing engineering education and life-long learning*, *15*(3‑6), 308‑317.
Vörös, Z., Rouet, J.-F., & Pleh, C. (2011). Effect of high-level content organizers on hypertext learning. *Computer Human Behavior*, *27*(5), 2047‑2055.