Big Content in an Educational Engineering Approach (教育学工程方法视角中的大内容制作)

Colpaert, Jozef
Universiteit Antwerpen
(安特卫普大学)
jozef.colpaert@uantwerpen.be

Abstract: Language teachers find themselves under increasing pressure to use technology, and to adapt their pedagogical approach accordingly. The first question they are confronted with is: which technology to use, why, how, when and where, and which justification to provide for this choice. A second, more intricate, question is: which content to use, why, how, when and where, and which justification to provide for this choice. This article provides a report on our research on content for language learning in the last thirty years. After defining the concept of content in a more ontological way, we provide an overview of all data and content types that have become available recently. We present educational engineering as a method for making a justifiable choice, but at the same time we point out a number of issues associated with certain content types. The development of content for interactive language courseware is extremely labor-intensive, and it lacks reusability. We explain our attempts to work out a generic model for content structuring, and justify why Open Data seems to become a promising alternative at this point.

Keywords: Content for language learning, educational engineering, Open Data, Big Data

1. Introduction

When we look at the history of education, we can see that content has not really been an issue for many centuries. The classic example in Europe would be Aristotle, who was a *peripatetic* lecturer. He walked about as he taught under the colonnades of the Lyceum in Athens. Oral transfer of knowledge, ambulatory education, combined with illustrated texts on papyrus or parchment, later complemented by bas relief sculptures, stained glass and paintings have been the traditional educational media for centuries. Blackboards and wall maps were typical nineteenth century educational artifacts. The textbook as we know it dates from the early twentieth century, and the sixties saw the appearance of radio, television, tape recorder and the stencil machine in the classroom.

On other continents like Asia, this evolution has been somewhat different, but since the end of the twentieth century, and especially since the emergence of the Internet, a broad revolution has taken place in education worldwide, transforming learning and teaching fundamentally. These changes entail significant challenges for learners, teachers, content providers (authors), publishers, researchers and policy makers. In this article we want to show why content is currently a challenging issue for language teachers worldwide, and we will explore a strategy based on our own thirty years of experience in the field. This overview entails a certain amount of self-referencing we hope the reader will forgive us.

2. Data, information and content

Language teachers find themselves under increasing pressure to use technology. This pressure is reflected in the pervasive but persuasive terminology they are confronted with: *blended learning*, *digital pedagogy*, *flipped classrooms*, *digital natives*, *21st century skills*, *virtual learning environments* and *serious games*. Each of these terms has been coined some day by a scholar who wanted to give a name to a largely unknown phenomenon, but we all started to use these terms with our own acceptation and purpose in mind. The latest cry is *Big Data*. What does this mean?

Big Data originally had a precise definition. It refers to data sets that are so large or complex that traditional data processing applications become inadequate, and that should be analyzed by powerful computers. The term also refers to the use of predictive analytics or certain other advanced methods to extract value from data in the form of patterns, trends, and associations, regarding human behavior and performance. However, more and more scholars in the field of language learning and teaching started to use the term to indicate the sudden emergence of enormous amounts of content sources which might be useful for language learning and teaching, accompanied by new phenomena such as the Semantic Web, Open Educational Resources (OERs), Massive Online Open Courses (MOOCs), and Learning Analytics.

Let us first come back to our terminology and refine our ontologies. We can define Data as tokens, characters, symbols stored on data carriers, and transmitted as messages that can become Information or Content. Shannon and Weaver (1963) were the first to define Information as a measurable concept, using a mathematical model. The amount of Information in a message is defined as the extent to which the uncertainty (or Entropy) is reduced by the message on the receiver's side. Communication is defined as the exchange of Information. Applied to education in general, and applied to educational technology more specifically, Information in a system reduces uncertainty on the learner's side by providing permanent and just-in-time Information on what to do and how to do it. In the other direction, the system becomes more certain about the learners' needs, achievements, level, attitude etc. by using data retrieved from his/her behavior.

But what is Content exactly in this respect? If we look at it in the context of instructional design and educational technology, specifically Computer Assisted Language Learning (CALL), we can define Content from a pragmatical point of view as Data which can be used in a meaningful way for language learning and teaching, and which can be expected to have an effect on learning. The most traditional content we know are the texts being used in textbooks. Series of characters, symbols, images that we try to remember, transcribe, fill in or translate as tasks or exercise types. This textbook content has become more interactive in the digital age. The appearance of content can adapt itself to different circumstances, the teacher can edit it, and it can return feedback on the learner's actions.

So Data are series of raw tokens which can turn into Information when they reduce uncertainty, or into Content when they contribute to learning. The Medium (Colpaert et al., 2012) is the carrier of the data. Technology can be defined as any routine or tool based on some kind of specific knowledge to perform an operation or treatment on data on a specific medium.

3. The Big Data tsunami

So *Big Data* is the latest blurred ontology. The term is persuasive as it lets us believe that there is a revolution going on in the world of pedagogy, that we should radically change our behavior as (language) teachers, and that all data are possibly useful. In fact, how much of these data can become Information and how much can become Content?

It is true that the range of data sources has expanded considerably in the last decade. Big Data stands for the overwhelming availability of accessible data worldwide, developed for educational purposes or not, and fans out into several relevant phenomena. They all imply different roles for authors, teachers and learners.

We can classify as Information any data stored as Teaching Information, Learner Information and Research Information. *Teaching Information* groups all documents which are supposed to guide and support the teacher, such as educational programmes (e.g. Common European Framework and national standards), curricula descriptors, course objectives and teacher guidelines. *Learner Information* groups all data that contain Information about the learner, his characteristics, profile, background, preferences, level, achievements and evaluations. Portfolios or e-portfolios are being used by teachers and learners themselves to store all these data. Learning Analytics on the other hand stands for collecting data about learner behavior and performance within a system or environment with a view to improve its design, and indirectly, learner performance. Finally, *Research Information* constitutes a more comprehensive approach in collecting data, often triangulating quantitative and qualitative approaches. Research data are being made more accessible and are being published more and more together with research articles.

But data can also become content when implemented in the learning and teaching process with a view to have a significant effect on learning. This content fans out into a wide range of possibilities:

- Traditional textbooks by publishers (Decoo, 2010).
- Self-authored materials produced by the teacher for use with his/her own students only.

- Authentic documents found on the Web. The Semantic Web (Web 3.0) promotes common data formats and exchange protocols for authors to add meaning to content, to describe the structure of the knowledge about that content, and in so doing to offer promising possibilities for retrieving relevant and meaningful materials for teaching.
- Open Educational Resources are materials that are being shared, reused, improved and re-shared again. They are supposed to reduce workload for teachers and to increase learning effect considerably.
- Interactive materials, also called interactive language courseware or tutorial CALL, have been around since the early eighties (Colpaert & Decoo, 1999).
- Massive Online Open Courses are the most remarkable phenomenon since 2010. Platforms such as *Moodle* and *OpenLearning* allow teachers to create and use language courses accessible worldwide.
- Virtual Worlds (such as *Second Life*) and Serious Games (Cornillie et al., 2012).
- Ambient Intelligence (the adaptation of an electronic device to the presence of the learner) and Augmented Reality (the view of a real-world environment whose elements are augmented by sensory input such as sound, video, graphics or GPS data).
- The Internet of Things stands for real-world objects and artifacts which carry readable data that can be used as content in tasks.

The availability of huge amounts of data that can turn into Information or Content leads to anxiety and choice stress for teachers, next to the pressure they are already experiencing. And all the efforts made to help teachers with this choice, like this article, lead to even more data. The question is: how to make a justifiable choice ?

4. Educational engineering and distributed design

Neither technology nor pedagogy is starting points for designing learning environments (Colpaert, 2015). Nor is content. The specification of the required technology, pedagogical model and content should be the result of a methodological design process (Colpaert & Stockwell, 2016). We will briefly try to explain our approach in the following paragraphs.

Our starting point is that education will never be perfect. Education has always been *l'art du possible*, and this for four reasons. First, by its very nature, education can and will never be perfect. Because we are humans. Secondly, lack of time and resources often prevent us from duly implementing the required changes. Thirdly, any change, even the most justifiable one, entails some kind of resistance, often from stakeholders we misjudge. Last but not least, there is not enough knowledge available in terms of substantiated findings which would enable us to improve education, solve problems or design solutions in a systematic, methodological and justifiable way.

Engineering is 'the strategy for causing the best change in a poorly understood situation within the available resources' (Koen, 2003) or, in other words, the strategy to be used for devising the best possible real-world solutions when not enough knowledge is available for doing so. It is a way of thinking in the first place, and does not necessarily imply technology.

Engineering is about building knowledge through real-world implementations, in a systematic and verifiable way, using working hypotheses that ate based on theory and practice and that should be empirically and theoretically validated. Hypothesis testing analyzes the effect of modified parameters, taking into account the specificity of the context. Engineering is about formulating and validating working hypotheses regarding the role, order, weight and intensity of these parameters.

Engineering is cyclic, iterative and probabilistic: Engineering seldom leads to proven facts in one project, but it often requires several iterations due to resistance, financial limitations, technological challenges or practical constraints in order to observe significant changes in the effect of the parameters in play (Bayesian epistemology).

In the same vein, *Educational Engineering* is about building the best possible educational artifacts. These educational artifacts can be documents, tools, content, concepts, models and solutions such as textbooks, syllabi, lesson plans, curricula, graded readers, exercises, tests, applications or electronic learning platforms.

Our research in educational engineering focuses on the theoretical and empirical validation of the following hypotheses, which we have grouped under the term *Distributed Design*, referring to the idea that the design process should take into account as many actors and factors as possible.

The four paradigm shifts stand for a radically new way of thinking about ICT in education:

- The *Ecological Paradigm Shift*: No technology has an inherent, measurable and generalizable effect on learning. Only the entire learning environment, seen as an ecology of interacting components, can have this effect.
- The *Process-Oriented Paradigm Shift*: The targeted effect of a learning environment does not depend on product features, but is proportional to its *designedness*. Designedness stands for the extent to which the learning environment has been designed in a methodological and justifiable way. This methodological approach is universally applicable, but leads to polymorphous results.
- The *Psychological Paradigm Shift*: In cases of problematic or lesser motivation, we tend to insist more on our pedagogical goals. This appears to be counterproductive: it is better to focus on personal goals first. Personal goals are defined here as subconscious volitions which hinder or stimulate acceptance and willingness to engage in the learning process. The problem with personal goals is that they are difficult to elicit (Colpaert, 2010).
- The *Demand-driven Paradigm Shift*: Neither technology nor pedagogy is appropriate starting points for design (Colpaert, 2015). The methodological design process creates a need, a strong demand for theoretical knowledge, content and technology.

Distributed Design, our Educational Engineering model based on these four paradigm shifts, can be considered an Instructional Design model of the ADDIE type for guiding the Analysis, Design, Development, Implementation and Evaluation of educational artifacts for learning, testing and teaching.

The Analysis stage is all about understanding the problem in its context. Its output is an accurate description of which aspects can, cannot and/or should change. The Design stage has three substages: Conceptualization, Specification and Prototyping. The goal of the Conceptualization stage is to detect and elicit subconscious goals, to identify the points where these personal goals conflict with the set pedagogical goals and to formulate a hypothesis about the best possible way to find a compromise between these conflicting goals. The output of Conceptualization is the formulation of the expected outcome, which will be compared with the actual outcome during the Evaluation stage.

The Specification stage is nothing more than a detailed (ontological) description of what is needed to realize the formulated construct in terms of pedagogical models (for teaching, learning and evaluation), content, technology and infrastructure.

The purpose of the Distributed Design model is to enable teachers to decide for themselves which pedagogical approach, content and technology to use, when, where, how and why.

5. Big Data in an educational engineering approach

The specification of content in the projects where we have applied the model, has led to a series of interesting observations.

Regarding Information, learners need Information on what to do, on their degrees of freedom, on how to bridge the gap in a TBLT approach, on how to reach the next level in a constructivist approach. They need Just-in-Time Information and Supportive Information in a 4C/ID approach. They want to know where they stand and how well they are doing.

Teachers rely on Learner Analytics and Portfolios in order to keep track of their learners' progress. Learning Analytics can reveal useful Information for diagnostics, remediation, and redesign of the learning environment, which was the main theme of the European project VITAL (http://www.project-vital.eu/en/). In an educational engineering approach, Learning Analytics should in the first place be geared towards validating the design hypothesis by comparing the actual outcome with the expected outcome. Data from Learning Analytics can become Information on how our design process can be improved.

Content is an important part in a Task-based Language Teaching (TBLT) approach. However, in order to make tasks really effective ('optimal' or 'activating'), our research has pointed out that the following aspects should not be neglected:

- Task design is a process (Colpaert et al., 2015). Teachers should follow steps in analyzing the context and specify the best possible tasks in order to realize the set pedagogical goals in the best possible way. Now, tasks 'fall from the sky' too frequently without any methodological approach.
- Any tasks are a hypothesis and should be formulated as such. Task validation consists in comparing the expected outcome with the actual outcome ("I expect my students to invest more time and energy in this paper if their paper will be reused later by other students"). Students even seem to become more interested in a task when teachers actually let them know that it is a hypothesis.
- There is no such thing as a good task or a bad task. Task effectiveness depends on the context. A simple drill-and-practice focus on form exercise may be useful in one context, and totally not in another. It is important that in task design, the specification phase makes the teacher choose from a wide range of skill types (21st century skills, Higher-order Thinking Skills, SAMR model, Digital Bloom Taxonomy).
- Tasks should create acceptance and willingness in the learners' mind. Self-Determination Theory (Deci & Ryan, 2002), Dörnyei's L2 SELF model (Dörnyei & Ushioda, 2009) or the author's Personal Goal Theory (Colpaert, 2010) help in explaining why tasks should be meaningful (what is in there for me?) and useful (what does it mean for others?).

What seems to work particularly well in many contexts is co-construction of knowledge. Our students write the content of our Instructional Design course themselves, adding every year another layer and another focus. Also five-minute knowledge clips as tasks are quite effective in this respect, both on the level of secondary education ("Explain the battle of Hastings", "Who was Confucius?", "Explain uniformly accelerated motion") and on the level of higher education and teacher training (as part of teaching materials they can upload in any digital learning environment).

The already mentioned Open Educational Resources looked very promising, but their success seems to be hampered (Colpaert, 2012) by psychological ('what will others say about my content?'), technological ('what should I use to share my content?'), epistemological ('what does Open exactly mean?') and juridical ('are Creative Commons enough?') challenges.

MOOCs do not seem to break through either, at least not to the extent we had expected. As explained in Colpaert (2014), there is a conceptual problem: MOOCs are not really massive, not really open and they cannot be considered real courses. Their highest potential lies in reaching and bringing together smaller groups of isolated learners on a specific specialized topic.

Learning from all this, the main challenges for the future seem to be personalization (adaptation of difficulty level, task type, … to the learner) and contextualization (adaption of content to the geotemporal location of the learner). In this respect, the role of interactive language courseware becomes more interesting again as long as it is situated in a wider array of content types. But what exactly is the problem with interactive language courseware?

6. Towards a generic structure for interactive content

Since the 1997 CALICO conference with its theme 'Content! Content! Content!', the issue of sustainability, exchangeability and reusability of content has not been out of the public discussion. The emergence of interactive language courseware (Colpaert, 2004), also called Tutorial CALL, made the problem become apparent: due to the complex functionalities needed for the required interactivity, content became less reusable and got lost. The authoring of this content for interactive language courseware was very labor-intensive, hence expensive. This is why, as explained in more detail in Colpaert 2013, we focused our research more and more on how to make content more generic in a first phase.

By using the term learning content, we do not only refer to traditional textbooks, but also to materials such as syllabi and handouts, interactive exercises in applications like *Hot Potatoes* or *QuestionMark*, course content in electronic learning environments like *Blackboard*, video and sound clips (podcasts), presentation slides in *PowerPoint* or *Prezi*, materials for Interactive Whiteboards, web pages, wikis and e-reader content. Writing learning content does not happen in a linear way, but it involves an arduous cyclic process of creating, editing, combining, structuring and formatting materials in several layers. Content should comply with many pedagogical, linguistic and cultural requirements such as to be linguistically correct, adapted to a specific level and context, engaging and attractive, as interactive and relevant as possible and to be politically correct by avoiding any statements or images which could insult or irritate individuals or minorities.

On the other hand, it is not easy to retrieve, select, evaluate and integrate materials developed by others, due to the fact that they are protected by copyright, not accessible or difficult to copy-paste. These materials contain text, images, sound and video, all with or without some levels of tagging, metadata or interactive functionality. Especially in the case of language learning, this functionality can become very complex (Colpaert, 2006). The more 'enriched' or interactive these materials, the higher the cost.

Moreover, language learning content should continuously be updated, adapted, rearranged and rechecked at every change in the learning context. These changes can be due to a curriculum change, a new pedagogical approach such as the 4CD/ID model (Van Merriënboer & Kirschner, 2013) or the Dynamic Systems Approach (Ellis & Larsen-Freeman, 2009), the integration of a new technology such as tablets or Interactive Whiteboards (Van Laer, Beauchamp, & Colpaert, 2012), and to changing learning styles, attitudes and preferences. But existing learning content is not easy to change. This is mainly due to the fact that most learning content has been created in a dedicated format: it is determined by the medium or the technology of the educational artifact as product.

Learning content gets lost far too quickly due to this inability to adapt, to be reused, exported, transferred. In other words, due to its lack of sustainability. In order to remedy this problem, learning content should become more *sustainable*. We define sustainable in this context as the sum of four properties: generic, reusable, interactive and open.

- *Generic*: Content should be authored, structured and accessed independently from any concrete device or medium and should be stored in a separate database. Its structure should not be influenced by any product as possible output.
- *Reusable*: Learning content should be made as transferable or exportable as possible to a wide variety of media, technologies and carriers, such as traditional hard copy textbooks, digital customized printed material on demand, mobile app exercises and materials for Interactive Whiteboard use.
- *Interactive*: Learning content can be 'flat' text, audio or video. There are however several possibilities for offering more information (e.g. enriched materials by semantic tagging afford more accurate selection of suitable learning materials) or more functionality (e.g. interactive exercises containing exercise types, answer possibilities, feedback scenarios, error analysis and remediation, reporting and logging).
- *Open*: Learning content should be as accessible, open and authorable as possible for allowing easier co-construction, updating and adaption.

We have been developing since 1986 a long series of applications and project tools (Colpaert & Decoo, 1999), representing a total of more than 150 man-years in projects for universities, governments, institutions, industrial companies, publishers and Europe. Initially the content of these developed programs was stored in a specific dedicated database structure (every application had its own structure). In 1997 we developed a new platform in Windows that focused on two requirements: generate a wide variety of applications with the same source code, and have the content in a separate database. In fact, there were two databases: the first contained all information for the application to run (identity and appearance, menu systems, behavior and interaction) while the second contained the learning content. Both databases were based on different object models, but they were both open, readable and updatable, at least for authorized people. The learning content was stored and structured in a relational Access database. The advantage of this approach was not only the strong integration with Visual Basic programming, but also the fact that authors could easily make their own interfaces as forms, based on queries, and reports. Some authors even succeeded in writing their own error-checking routines in VBA (Visual Basic for Applications).

In the *Eventail/Arcades Interactive Textbooks* project (with Wilfried DECOO), a longer-term project with publisher Van In, we initially converted language textbooks into interactive applications. Gradually, we started first structuring a database of learning content so that CD-ROM and textbook could be generated at more or less the same time as different output products. We gradually applied this approach in projects where possible: the BIS Online project for the Flemish department of Education, SELOR language tests for the Belgian Civil Service Commission, and a series of European projects. Finally, we ended up working more with the same object models behind the database structure of several different programs. These object models were not technical, but reflected a reasoning for opening, reading, editing and updating a specific database.

In 2004, we developed an object model that complied with all possible requirements and defended this research as doctoral dissertation (Colpaert, 2004; Colpaert, 2006).

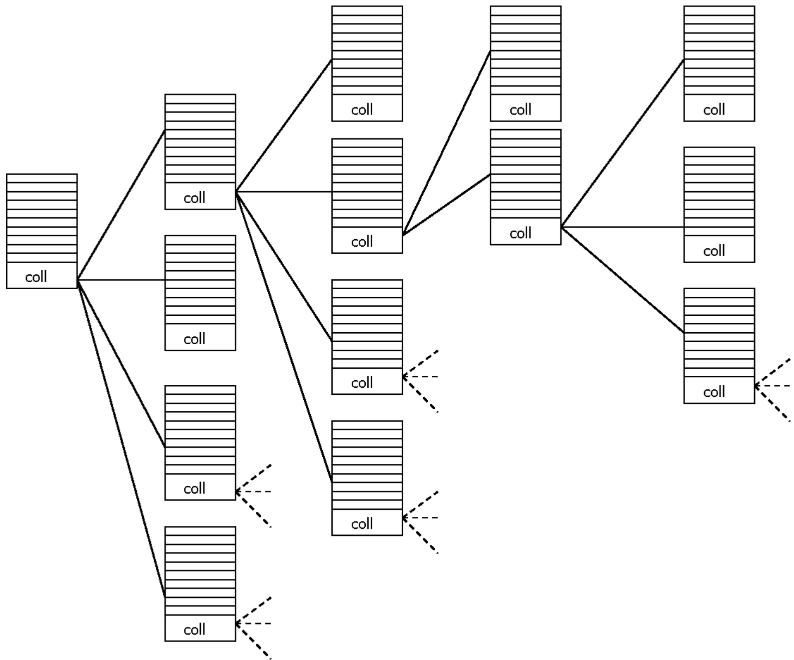

Figure 1. Generic Object Model

While generic structuring (for instance if carried out in a relational database) may appear fairly readable and authorable, surrounding factors in a normal working environment can make things quite complex: large data, many co-authors, complex functionality, author support (queries, forms and reports for content analysis), error checking and prevention (quality control), integration and reuse of existing materials, and generation of content into a wide range of products and services. In order for the object model to be implemented and to lead to significant and sustainable results, we wanted to develop an authoring interface, or at least define an ontology. The adopted methodology was a mixed-method approach consisting of a theoretical, an empirical and an engineering component.

Our research (Colpaert & Cornillie, 2008) has shown that language learning content can and should be structured and stored in a sustainable way, defined as generic, reusable, authorable (open) and allowing interactivity of scalable complexity. Language learning content can be structured as a collection of a collection of a collection of items, with as many collection levels as deemed necessary: from a plain text (only one level with one collection) to complex task-based scenarios. In this collection structure, it is possible to use the same, simple, object model, consisting of properties and methods which govern appearance and behavior of the object content. The proposed structure appears to work well in client/service environments such as web-based applications or mobile apps, but also for generating a wide range of products and services, also non-digital.

But our research has also shown that it remains quite a challenge, if not impossible from a practical point of view, to have this sustainable content structuring implemented in the real life of teachers, authors and publishers. The transition from the actual publishing tools to a generic authoring interface is an enormous undertaking for publishers worldwide, which in the best case would take five to ten years to realize. Teachers and authors need to be informed about the best way to structure their data in a sustainable way. The technological side of the interface was not really the problem, but the psychological side. It will take several other research projects to convince authors of an interface perceived as useful and easy to use at the same time.

7. A special case: Open Data

Open Data is a recent phenomenon which deserves our attention. The adjective 'Open' in Open Data has a slightly different connotation than 'Open' in Open Educational Resources, Open Access (for research articles), Massive Open Online Courses and Open Source. Open Data, according to the open definition (http://opendefinition.org), is any kind of data in any kind of format, that can be accessed, used, modified and shared, by anyone for any purpose. An open license on the container of the data is the only requirement to comply with this definition. More interesting is the goal behind Open Data, which could be interpreted as maximizing the reuse of a dataset. It may be in the best interest of a company that its datasets are being used within other app(lication)s as much as possible. In the same vein, one could consider that it is in a publisher's interest to have learning content reused in as many ways as possible.

While many data sources are already openly licensed (Bizer et al., 2009), as already explained in this article, there are still many hurdles to overcome when trying to publish data for maximized reuse (Colpaert, 2014). Enhancing interoperability with all other published datasets would increase their reuse. Interoperability denotes the ability of diverse systems and organizations to work together (inter-operate). In this case, it is the ability to interoperate - or intermix - different datasets. Interoperability is important because it allows for different components to work together. Measuring this interoperability is difficult (Colpaert et al., 2014), yet dividing the concept into properties may help: technical interoperability, syntactic interoperability, semantic interoperability and querying interoperability. Many of these interoperability problems are being tackled by standardization, yet in a constantly evolving world, standards often lack behind, especially in education.

Open Data, in a simplified view, provides an interface for reading the structure and content of the dataset before actually accessing it. Linked Open Data maximizes the effect of Open Data by linking together large series of datasets. The connection of these datasets to the Semantic Web creates an even bigger enrichment.

The advantages are obvious in science and for government, but the potential for education is huge. Datasets can be made accessible without having to restructure them first. On the level of Information, as defined earlier in this article, datasets about research, education, schools, students, teachers and even parents. On the level of Content: any existing material could be made reusable again. Provided we add a layer (API or other) which explains how to read the data. In theory, all these interfaces can be different, but the evolution towards some kind of standardization in terms of learning objects or student records may seem attractive and daunting at the same time.

We are currently analyzing to what extent the previously developed Generic Object Model can be reused as an interface object model for any kind of learning content. We are aware that new issues and problems will arise, but the fact that so much lost content may be reactivated again gives us a lot of courage to continue in this endeavor. Open Data can lead to Big Learning Content. There is a promising compatibility with recent paradigms in TBLT (Colpaert et al., 2015), Complex Dynamic Systems (Ellis & Larsen-Freeman, 2009), personalization and contextualization of the learning process, skills (SAMR, Bloom digital taxonomy, 21st century skills, Higher Order Thinking skills…).

8. Conclusions and recommendations

Teachers do not only find themselves under increasing pressure to use technology – and to adopt a new pedagogy – digital or not – , they are also confronted with a tsunami of data, which we have called Big Data. These data belong to two important categories: Information and Content. Teachers need Information to guide and support their learners and to organize their learning environment in the best possible way. Learners need information to know where they stand, where they can go and what their degrees of freedom are. Learning Content on the other hand is fanning out in a wide variety of types, so that teachers are – again – confronted with choice stress. How to decide which content to use, where, how, why, when?

Content can only become useful if it is needed based on a specification of an optimal learning environment. This specification remains a hypothesis within what we have called an educational engineering approach. This angle of attack makes us look in a different way at content. In a more critical way to OERS and MOOCs, but also in a more worried way to interactive language courseware: it is labor-intensive due to the required linguistic-didactic functionalities, and it gets lost all the time due to its non-generic structure.

Open Data, in the same spirit as Open Source, Open Access, Open Educational Resources, but with a slightly different acceptation for ‘Open’, seems like a promising phenomenon to solve this problem. Publishers and authors can continue to produce learning content the way they were, but they will gradually add interfaces to the extent that they want their content to be shared and reused again.

Researchers and developers in CALL will find in Open Data a promising direction which is very compatible with current approaches such as TBLT, personalization and contextualization of the language learning process, Complex Dynamic Systems and digital skills. Teacher and learner roles will change considerably, especially regarding collaboration in activities such as telecollaboration, course co-construction or knowledge clips.

Again, the most appropriate pedagogical model, technology and content are not starting points for design, but they are the result of good design, and they will always depend on the context. And in that context, the needs, goals and volitions of learners and teachers are important. Like in neuromarketing, the subconscious volitions are more and more considered as decisive starting points for good design. So how can we 'Open' up this source of 'Information'?

References

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked data-the story so far. *Semantic Services, Interoperability and Web Applications: Emerging Concepts*, 205-227.

Colpaert, J. (2004). *Design of online interactive language courseware: Conceptualization, specification and prototyping. Research into the impact of linguistic-didactic functionality on software architecture.* (Doctoral dissertation). Antwerp, Belgium: University of Antwerp.

Colpaert, J. (2006). Toward an ontological approach in goal-oriented language courseware design and its implications for technology-independent content structuring. *Computer Assisted Language Learning*, *19*, 109-127.

Colpaert, J. (2010). Elicitation of language learners' personal goals as design concepts. *Innovation in Language Learning and Teaching*, *4*(3), 259-274.

Colpaert, J. (2012). Open educational resources for language teachers: A goal-oriented approach. In *Proceedings of the joint CMC and teacher education EuroCall SIGS workshop Bologna*, 29-30 March 2012. Retrieved from http://eurocallsigsbologna.weebly.com/papers.html

Colpaert, J. (2013). Ontological specification of an authoring interface for creating sustainable language learning content. In L. Wong, C. Liu, T. Hirashima, P. Sumedi, & M. Lukman (Eds.), *Proceedings of the 21st International Conference on Computers in Education* (pp 644-653). Indonesia: Asia-Pacific Society for Computers in Education.

Colpaert, J. (2014). Reflections on Present and Future: towards an Ontological Approach to LMOOCs. In E. Martín-Monje, & E. Bárcena (Eds.), *Language MOOCs: Providing learning, transcending boundaries.* Mouton de Gruyter (Open Series), 162-173.

Colpaert, J. (2015, April 8). Neither technology nor pedagogy should come first in design. *The Educationist*. Retrieved from https://www.openlearning.com/educationist/NorTechnologyNorPedagogyShouldComeFirstInDesign0.

Colpaert, J., & Cornillie, F. (2008). Between dedicated and generic requirements: The Entreposage Universel Model as a grammar of interactive content. In A. Gimeno, & R. Seiz Ortiz (Eds.), *Aprendizaje de lenguas asistido por ordenador: herramientas de autor para el desarrollo de cursos a través de la Web* (pp.57-92). Valencia, Spain: Editorial Universidad Politécnica de Valencia.

Colpaert, J., & Stockwell, G. (in press). *Technology and the Language Learning Environment*.

Colpaert, J., & Decoo, W. (1999). The role of didactic functions in CALL design. In K. C. Cameron (Ed.), *CALL & the learning community* (pp. 65-74). Exeter, UK: Elm Bank Publications.

Colpaert, J., Aerts, A., Oberhofer, M., Gutiérrez-Colón Plana, M. (Eds.). (2015). *Task design and CALL. Proceedings of the Seventeenth International CALL Research Conference*. Antwerp, Belgium: University of Antwerp.

Colpaert, J., Aerts, A., Wu, W. and Chao, Y. (Eds.). (2012). *The Medium matters: Proceedings of the Fifteenth International CALL Conference*. Antwerp, Belgium: University of Antwerp.

Colpaert, P. (2014). Publishing data for maximized reuse. In *11th Extended Semantic Web Conference*. Retrieved from http://ceur-ws.org/Vol-1165/salad2014-1.pdf

Colpaert, P., Van Compernolle, M., De Vocht, L., Dimou, A., Vander Sande, M., Verborgh, R., & Mannens, E. (2014). Quantifying the interoperability of open government datasets. *Computer*, *10*, 50-56.

Cornillie, F., Clarebout, G., & Desmet, P. (2012). Between learning and playing? Exploring learners' perceptions of corrective feedback in an immersive game for English pragmatics. ReCALL, 24(3), 257–278. doi:10.1017/S0958344012000146.

Deci, E., & Ryan, R. (Eds.). (2002). *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.

Decoo, W. (2010). *Systemization in foreign language teaching: Monitoring content progression.* New York: Routledge Research in Education.

Dörnyei, Z., & Ushioda, E. (Eds.). (2009). *Motivation, language identity and the L2 self*, Bristol, UK: Multilingual Matters.

Ellis, N.C., & Larsen-Freeman, D. (Eds.). (2009). *Language as a complex adaptive system*. Oxford, UK: Wiley-Blackwell.

Koen, B. V. (2003). Discussion of the method: Conducting the engineer's approach to problem solving. NY: Oxford University Press.

Van Laer, S., Beauchamp, G., & Colpaert, J. (2012). Teacher use of the interactive whiteboards in Flemish secondary education—mapping against a transition framework. *Education and Information Technologies*, 1-15.

Van Merrienboer, J. J. G., & Kirschner, P. A. (2013). *Ten steps to complex learning (Second Revised Edition).* New York, US: Routledge.

Weaver, W., & Shannon, C. (1963). *The mathematical theory of communication*. Champaign, IL: University of Illinois Press.