Cerri. *Chem Synth* 2023;3:18
DOI: 10.20517/cs.2022.37

Chemical Synthesis

Opinion

Open Access

Check for updates

Information, knowledge, and human learning for chemistry: the visionary contribution of Professor Alain Krief

Stefano A. Cerri[1,2]

[1]LIRMM, Laboratory of Informatics, Robotics and Microelectronics of Montpellier, University of Montpellier, Montpellier F-34095, France.
[2]Digital Knowledge Technologies Services (DKTS) SRL, Via Ampère 61/a, Milan 20131, Italy.

Correspondence to: Prof. Stefano A. Cerri. LIRMM, Laboratory of Informatics, Robotics and Microelectronics of Montpellier, University of Montpellier, 161 Rue Ada, Montpellier F-34095, France. E-mail: cerri@lirmm.fr; VP. Digital Knowledge Technologies Services (DKTS) SRL, Via Ampère 61/a, Milan 20131, Italy. E-mail: sacerri@didaelkts.it

How to cite this article: Cerri SA. Information, knowledge, and human learning for chemistry: the visionary contribution of Professor Alain Krief. *Chem Synth* 2023;3:18. https://dx.doi.org/10.20517/cs.2022.37

Received: 29 Nov 2022 **First Decision:** 9 Mar 2023 **Revised:** 21 Mar 2023 **Accepted:** 22 Mar 2023 **Published:** 10 Apr 2023

Academic Editor: Bao-Lian Su **Copy Editor:** Ke-Cui Yang **Production Editor:** Ke-Cui Yang

Abstract

This is not a paper about the current state of chemistry but rather the author's perspective on a strategy for the future of chemistry and other scientific and technical disciplines. This future vision is based on synergies of natural sciences and technical disciplines on one side and the science of information, i.e., informatics, on the other. It relates the author's interactions during 2003-2023 in professional collaborations with Alain Krief, now Emeritus Professor at the University of Namur. It is argued that the scientific value of Prof. Krief's work - particularly related to his well-known interest in "chemical knowledge transfer" - extends beyond the scientific-technical domain of chemistry to include a much broader area of natural sciences and associated technologies. Prof. Krief's signal contribution has resulted from his profound interest in information, knowledge, and human learning, which has been at the core of our exchanges during the last 20 years.

Keywords: Information, knowledge, human learning, interaction, agents, collective intelligence

THE EMERGENCE OF INFORMATICS AS A NATURAL SCIENCE

Natural sciences can and should seek to answer questions about the essential elements (components,

Cerri. *Chem Synth* 2023;3:18 | https://dx.doi.org/10.20517/cs.2022.37

properties, functions) of natural phenomena, to understand them and to forecast their evolutions as much as possible. Each of the traditional disciplines (e.g., physics, chemistry, biology) answers questions about one of the components/properties of natural phenomena. The primary concern of physics is energy; for chemistry, the goal is to understand matter; and for biology, it is to understand life.

In ancient times, these areas were regarded as part of philosophy. Aristoteles, for example, wrote about physics (*περί φύσεως* means *about nature)*, with philosophy considered to be a general reflection on nature. These disciplines were originally part of philosophy and emerged as distinctive subjects; furthermore, each continues to split with increasing speed and acceleration today. For example, we may now speak about nanotechnologies as a technical offshoot of modern physics and chemistry or genetics as an offshoot of contemporary biology.

Disciplines also fuse with one another: biochemistry and bioinformatics are examples. Their interactions are increasingly perceived, measured, and controlled, so no scientist wishing to be taken seriously may admit to ignoring these interactions if they want to be regarded as credible in understanding nature. This evolving picture illustrates a recognition of the extreme complexity of nature itself and an acknowledgment of the diverse and sometimes paradoxically contradictory qualities of real scientists, including ambition and modesty. Scientists should be ambitious to explore the complexity of natural phenomena but also modest because no single competence may comprehend the subtleties of interactions among all significant factors. This limitation may justify a strong interest in multidisciplinary studies supported by collective intelligence - but the reality is far from theory, as we all know.

The emergence of sciences over many centuries was marked, in the 19^{th} century, by significant progress in understanding “what is energy” (physics) and “what is matter” (chemistry), and in the 20^{th} century by progress in the “what is life” (biology). Another crucial property, steering all natural phenomena (and enabling our ability to study them), was only gradually emerging to be identified, isolated, and studied in this later period: information.

The invention of the computer had a similar triggering effect on the construction of knowledge as had been the case for the printing press centuries earlier. Indeed, essential computer scientists from 1940 to 1960 (e.g., Turing, Wiener, Shannon) paved the way for a foundational new discipline. A well-known aphorism reminds us that computer science is not more about computers than astronomy is about telescopes. In the second half of the 20^{th} century, informatics became recognized as the science of information, a property of nature of the same importance as energy, matter, and life, the four being in perpetual interaction in all natural phenomena. Examples of extreme complexity and significance in the human body are the brain and the immune system (see, for example, debates on the meaning of informatics: https://en.wikipedia.org/wiki/Informatics). Being the computer programs (i.e., the control) stored in the same memory as data, programs are able to treat program code as data. Therefore the invented artefact - the computer - is much more than a machine, it is rather a meta machine, i.e.: a machine capable of learning, i.e.: able to produce totally different machines as a function of the course of interactions with the environment. This phenomenon is very clear and popular today, with the emergence of machine learning in deep neural nets with big data and large language models and consequently the emergence of interactive, generative artificial intelligence systems such as ChatGPT.

As colleagues at the University of Edinburgh (https://www.ed.ac.uk/informatics/about/what-is-informatics) have synthesized the subject, informatics, the science of information, consists of computer science (the components of modern computers, including robots), cognitive science (how animals, individually and collectively, perceive, understand, reason, decide and act) and artificial intelligence, which links the two complementary disciplines. This vision of informatics includes

1. Technical/mathematical components, such as logic, complexity theory, algorithms, software, databases, hardware, and networks typical of computer science; the anatomy of artificial systems.

2. Functional processes, such as perception, reasoning, communication in natural language, and action: the physiology of natural systems typical of cognitive science.

3. The synergy between artificial and natural systems typical of artificial intelligence, for instance, in the approach called "multi-agent systems" pioneered by Norbert Wiener in his book on Cybernetics: (https://en.wikipedia.org/wiki/Cybernetics:_Or_Control_and_Communication_in_the_Animal_and_the_Machine).

4. And, in recent years, emotions and social studies.

In other words, information and knowledge are essential elements of "agents," whether the latter are living or artificial, engaged collectively in collaboration and competition (see, e.g., evidence of these processes in the development of species produced by Darwin), as has long been recognized as the case for animals, from single cells to complex heterogeneous living societies.

In this vein, the author will revisit his interactions with Prof. Krief over the last 20 years to demonstrate that the latter's vision merits profound interest regarding the future of chemistry and natural sciences more broadly and, consequently, of related technologies.

THE ENCORE PROJECT: KNOWLEDGE CONSTRUCTION AND USE BY INTERACTIONS

Around 2003, the author met Prof. Krief, who was interested in our work and considered it somewhat different from other computer scientists he had encountered previously. At the time, Prof. Krief wanted to devote his experience, competence, and reputation to an enterprise that was attractive and simultaneously daunting - the EnCOrE project.

EnCOrE is an acronym for Encyclopedie Chimie Organique Electronique. The idea was to build a repository (an encyclopedia) in the domain of organic chemistry that could help scientists and students "understand and forecast" natural phenomena in organic chemistry (e.g., reactions).

As an informatician, the author was enthusiastic about the project, which could be a very concrete, compelling, and valuable testbed for constructing a series of interactive, knowledge-based systems but recognizing that critical bottlenecks would include using such systems once they existed and the construction itself. By 2005, we could include EnCOrE in the European Union (EU) 6th Framework Programme Project E-LeGI (European Learning Grid Infrastructure https://cordis.europa.eu/project/id/002205). The proposal convinced the EU that the interactions necessary for constructing and using EnCOrE were significant for human learning. As one of the critical application testbeds (or SEES: Service Elicitation Exploitation Scenarios), EnCOrE won part of a significant support gained by E-LeGI (€10 million for 23 partners) from the EU for a period of four years.

 Cerri. *Chem Synth* 2023;3:18 | https://dx.doi.org/10.20517/cs.2022.37

To express the needs adequately, the following is a paragraph quoting remarks by Prof. Krief, founder of EnCOrE, which he made when addressing the E-LeGI consortium during its meeting in Brazil in 2004[1,2]. The author highlights a few issues and phrases that he finds particularly valuable:

"*Currently, the information is only delivered flat according to a* ***single point of view*** *dictated by tradition of 'book organization' following the 'Johannes Gutenberg age' (ca.1400-1468). In fact, chemists' brains work differently, and the usual delivery message is* ***context-oriented****. There is a huge number of different contexts which are covered, and it is impossible using a book or even a lecture to describe them all (experimental-oriented, starting material-oriented, product-oriented, mechanism-oriented, stereochemical-oriented, calculation-oriented...)... Not only methods and tools needed for each context are different (flasks, molecular models, heavy calculation), but* ***even the words used in each of these contexts are not properly defined***."

"*The construction of the EnCOrE Dictionary is extremely important for our project. It will fix the language and the related ideas and will play an important role in questioning EnCOrE. Its production is an act of power. If this power is not well understood, the chemists will ignore it. For that purpose, it is extremely important that chemists accept and use it. For that purpose, it should be elaborated through a* ***collaborative work*** *implying discussions, contextualization, and consensus between the chemist's community. We want to* ***archive the discussions*** *in order to keep the dictionary always alive by* ***reactivating the discussions*** *on a single word from time to time according to new needs. We believe that the times where confirmed chemists, sent by their respective governments, were gathering in palaces sponsored by IUPAC (Union of Pure and Applied Chemistry, https://iupac.qmul.ac.uk/) to build the compendium of chemical terms in a non contextual manner, is over.*"

The two fundamental messages from Prof. Krief were 1. the importance of words, language, and agreements about the semantics and 2. the strong dependence of these on the context. Both concerns remain at the core of modern informatics, thus encouraging a joint project. Prof. Krief's vision of the future of chemistry was one of a pioneer, without any doubt.

A crucial point concerns the difference between "information" and "knowledge." We were perfectly aware that it would be possible to build a repository of information, a kind of electronic library of concepts, properties of concepts, and relations among concepts and their properties, often called "ontologies." There are many databases, particularly in chemistry; however, the goal of EnCOrE was much more ambitious. Prof. Krief explained that he wanted to identify and implement a process of construction and use of chemical *knowledge*, not merely chemical *information*.

For years, we discussed the difference between information and knowledge. We proposed a simple definition: knowledge is information necessary and sufficient in the context of a decision. Prof. Krief agreed, but simultaneously, problems emerged: the "context" where knowledge is exploited depends heavily on the individual and their previous knowledge, goals, strategies, and tactics, in the case of an expert chemist constructing the encyclopedia and when the knowledge is being exploited, whether by a student or an expert.

After months of work on a small subset of chemical concepts (e.g., chemical equation, chemical structure, element, functional group, named reaction, pure substance, retro-synthesis scheme, segment, reaction vessel), a small encyclopedia was built; nevertheless, many different points of view remained unresolved among the contributing senior chemists. However, the participants all progressed significantly in "learning," as described below. In retrospect, this failure led us to revise our plans: Prof. Krief and the chemists were encouraged to investigate more deeply the tools, methods, and initiatives dedicated to human learning and dissemination of scientific knowledge, particularly in chemistry, while the informaticians started another project (ViewpointS) which continues (see, e.g., Refs.[2-6]). This latter project is the best testimony of the influence of the experience we gathered previously working with Prof. Krief in EnCOrE.

A synthesis view (2) of the interactive process is presented in Figure 1, in which the agents reach a consensus even in the presence of different points of view about the world.

HUMAN LEARNING AS A SIDE EFFECT OF INTERACTIONS

The E-LeGI project in which Prof. Krief participated with EnCOrE was meant to build a European infrastructure - based on GRID services[7] - that would support human learning. The technical progress of the infrastructure was based on two primary aspects:

1. GRID infrastructures, i.e., network-based predecessors of the cloud, distributed/centralized repositories of information accessible via middleware/web; and

2. GRID services (unlike web services) are "stateful" ones, i.e., they retain a memory of the history of interactions, an essential element for the personalization of teaching-learning processes as well as any other significant interaction. Thanks to these stateful services, one may realize the goal of Prof. Krief when he wanted to store and retrieve the history of conversations that justify any decision about the Encyclopedia.

Both these critical technical aspects were discussed with Prof. Krief, whose enthusiasm and engagement were unlimited. He adopted and reinforced the author's own vision of the requirements for successful learning and for successful progress in science - in fact, we agreed that learning, discovering, and inventing are human activities strongly related to each other (see, for instance, the subsequent work in Refs.[4-6]).

Below are a few snapshots of the work undertaken with Prof. Krief, fully described in the papers and deliverables for the EU.

The Grid Shared Desktop was the interface of a system called AGORA, which enabled the definition and exploitation of Agents (human and eventually software agents), Groups (sets of Agents), Organizations (the structure of rights and duties of groups), Resources (computational and communication resources), and activities (the processes activated within the system). AGORA interactions run on Grid middleware (i.e., worldwide on the web) services (processes activated by software) were stateful, i.e., with memory, enabling the storage and retrieval of histories of the interactions.

In Figure 2, from a 2012 presentation in Brazil by the author, we see six persons connected to the system. In addition to Prof. Krief, who was in Namur, Belgium, participants included a junior chemist (Catherine Colaux-Castillo: at the time, a student of Prof. Krief) from Paris, a senior chemist (Claude Laurenco, Ecole de Chimie, Montpellier) from Montpellier, and two senior informaticians (Philippe Lemoisson and Pascal Dugénie) from different labs in Montpellier. The "conductor" of the interaction was Monica Crubezy, the author of the interface of Protégé[8] from Stanford University, who actively supported the project. Protégé is (even now) one of the most successful editors for ontologies worldwide. The session aimed to define the concepts and relations (ontology) concerning the synthesis of the natural product carpanone. The session involved the use of different shared applications such as Flashmeeting (http://flashmeeting.open.ac.uk/) for communication (courtesy of the Open University, Knowledge Media Institute, UK: Prof. Marc Eisenstadt, one of our partners), ChemDraw (a standard software for drawing chemical elements), Resyn Assistant (a

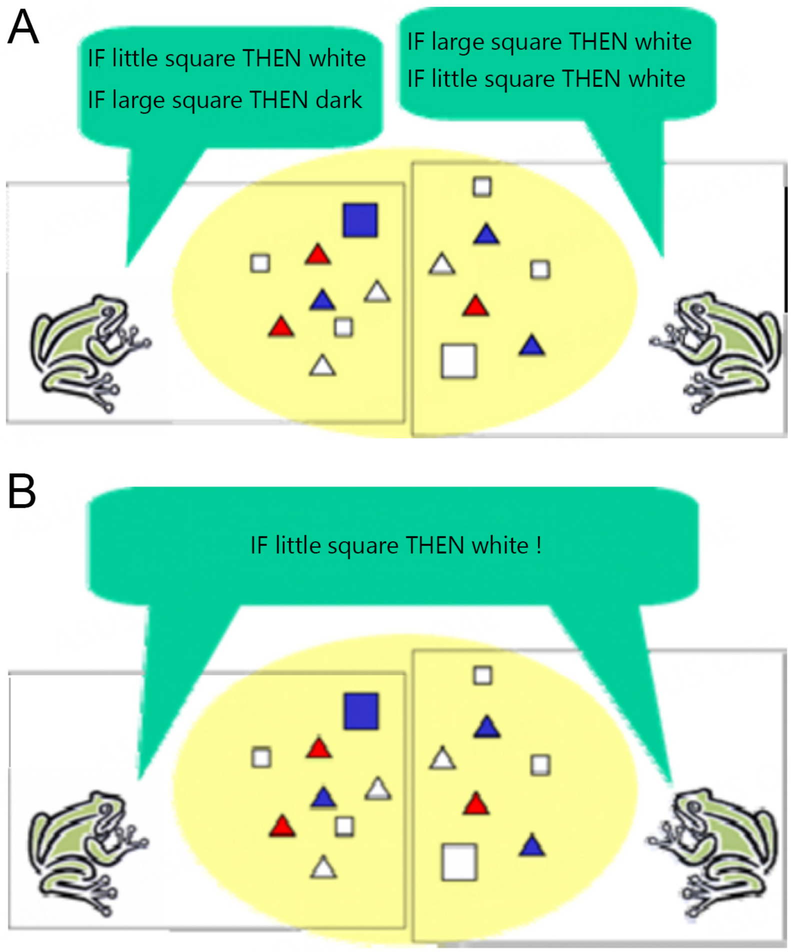

Figure 1. Building a consensus in presence of different viewpoints. (A) Each agent makes his own abduction, according to his local horizon; (B) each agent takes into account the other's abduction and proceeds to a revision of one's theory.

research tool developed by Dr. Laurenco and seldom shown outside his laboratory), and Microsoft PowerPoint (operated by Prof. Krief during a "classical" lecture at a distance).

The duties and rights of the five components of AGORA, the core software behind its interface, the Grid Shared Desktop, were set previously to coordinate the reasoned exploitation of the collaborative processes during the interactions at a distance. To the author, no other software tool currently offers the same robust control and communication functionalities.

Human learning in these experiments occurred in many ways because of the interactions. Each participant "learned" continuously from the others and the whole experience. Two examples highlight how this experiment was an exceptional source of human learning at a distance.

The first was linked to Prof. Krief, a senior scientist learning to use complex software. Previously, he had never been in a situation where he could build an ontology, i.e., a written, formal symbolic representation of concepts, properties, and relations, that software could use to infer new concepts/properties/relations. His profound knowledge of chemistry was indispensable for the success of the enterprise. No other (younger) scientist - probably better acquainted with the tool - could replace him. Therefore, he needed to

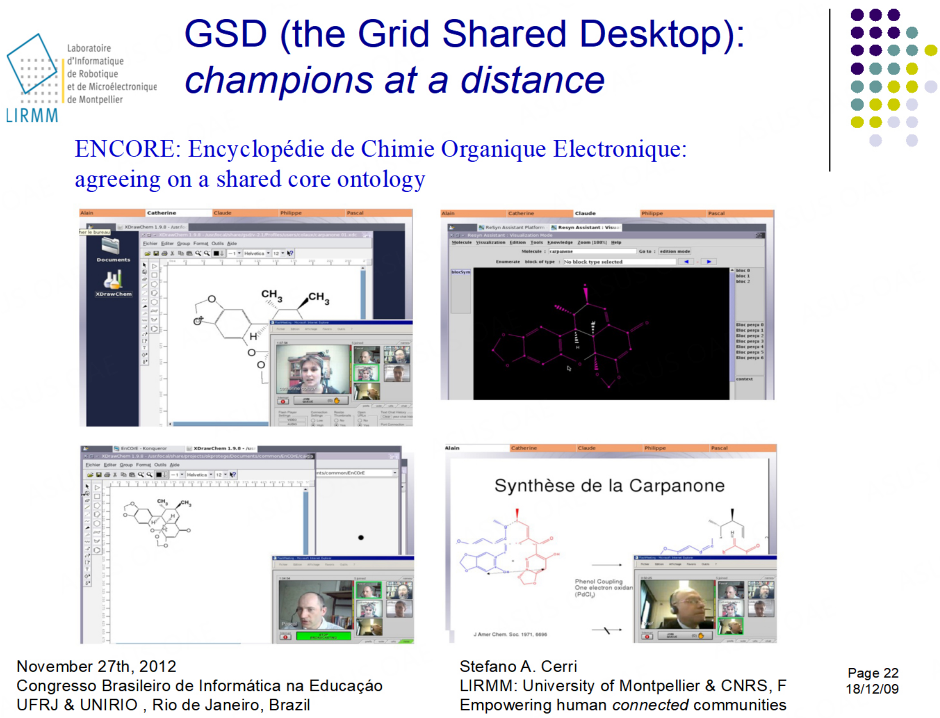

Figure 2. Working and learning at a distance.

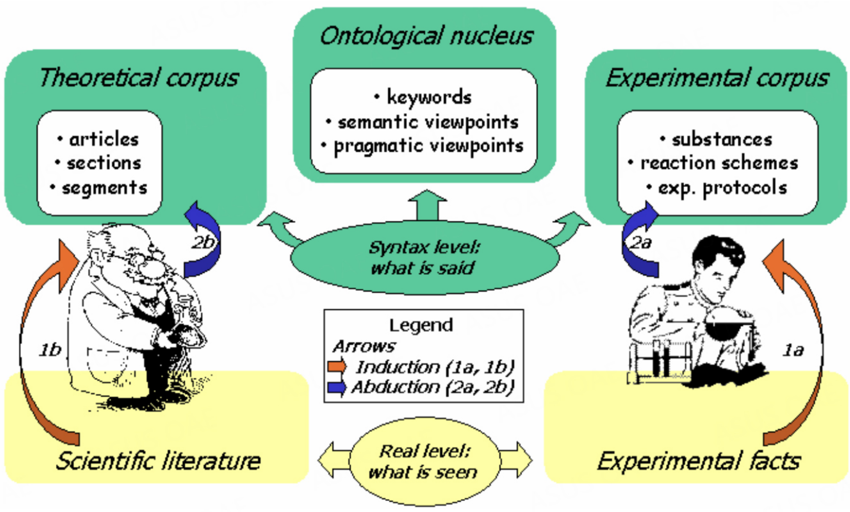

Figure 3. Information, knowledge, and human learning in the EnCOrE project.

"understand and exploit" the powerful, complex Protégé tool. Monica Crubezy, the author of the interface, helped him (at a distance) to become familiar with Protégé in the context of knowledge construction.

The second was related to Claude Laurenco, a senior chemist learning how to teach others to use complex software. Having never previously shown his Resyn Assistant tool, which he had developed over the years, outside his laboratory, he had to learn how to communicate its functionalities to other participants and disseminate its operation and value at a distance.

In experiences from 2005 to 2007, the power of technology then being put into action suggested what today appears normal for people working and learning at a distance. At the time, however, it was not yet widely exploited in its complete spectrum of capacities to the extent true within EnCOrE. We were convinced of the importance of our progress and began to propose to several agencies to advance the work. Unfortunately, none understood the interest of working on the "collective brain"[4,6] put into action at a distance, and the projects proposed were not considered research in either chemistry or informatics (computer science or artificial intelligence) that merited long-term support. Inter- and multi-disciplinarity and the power of information and communication technologies for science construction did not attract those evaluating the submitted research proposals.

Later, information technology companies (e.g., Zoom or Microsoft and many others) presented tools that were apparently similar but realistically were incomplete concerning the one we had demonstrated years ago. Despite what is often declared, the lesson learned is that interdisciplinary progress is rarely sponsored because advances are all too often expected as a side effect of pure disciplinary research plans. Of course, this policy can be understood but does not fit with the observation - supported by many historical examples - that most new discoveries and inventions occurred serendipitously (see, e.g., Ref.[5]) within projects involving high risk, not planned in detail according to the current state of the art. This reality was embedded in the construction of EnCOrE as an encyclopedia supporting information, knowledge, and human learning (Figure 3, from Ref.[1]).

CONCLUSION

One of Prof. Krief's essential attributes is his invaluable curiosity for innovative concepts, methods, and tools, even when they seem far from his previous experience. This ambition to master advances from other disciplines (concerning our collaboration: computer science, cognitive science, artificial intelligence) by interacting with experts has always been related to a profound modesty, witnessing his character as a genuinely modern scientist.

The author, who has known several colleagues with different expertise who have been interested in his approach to informatics, considers Prof. Krief one of the few able to establish a true partnership between our two very different disciplines. If modern scientists are expected to be ambitious and modest, he is such a person. It has always been a pleasure to accompany Prof. Krief in explorations regarding the future of chemistry based on information, knowledge, and human learning, the three components we have long identified as necessary, if not sufficient, for the progress of the discipline and, in general, for the advancement of science. The author is confident that the future will acknowledge the pioneering work of Prof. Krief for his various achievements in organic chemistry, documented by several contributions in this journal issue, and especially for his broad and telescopic vision of the conditions fostering the advancement of science by collective knowledge, such as those he has demonstrated to the author in the last 20 years.

DECLARATIONS

Authors' contributions

The author contributed solely to the article.

Availability of data and materials

Not applicable.

Financial support and sponsorship
None.

Conflicts of interest
The author declared that there are no conflicts of interest.

Ethical approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

REFERENCES

1. Lemoisson P, Untersteller E, Netto Nunes MAS, Cerri SA, Krief A, Paraguaçu F. Interactive construction of EnCOrE (learning by building and using an encyclopedia). Available from: https://hal-lirmm.ccsd.cnrs.fr/lirmm-00108797 [Last accessed on 3 Apr 2023].
2. Lemoisson P, Cerri S. Interactive knowledge construction in the collaborative building of an encyclopedia. *Appl Artif Intell* 2005;19:933-66. DOI
3. Lemoisson P, Cerri SA, Douzal V, Dugénie P, Tonneau JP. Collective and informal learning in the ViewpointS interactive medium. *Information* 2021;12:183. DOI
4. Lemoisson P, Cerri SA. ViewpointS: a collective brain. In: Frasson C, Bamidis P, Vlamos P, editors. Brain function assessment in learning. Cham: Springer International Publishing; 2020. pp. 34-44. DOI
5. Cerri SA, Lemoisson P. Serendipitous learning fostered by brain state assessment and collective wisdom. In: Frasson C, Bamidis P, Vlamos P, editors. Brain function assessment in learning. Cham: Springer International Publishing; 2020. pp. 125-36. DOI
6. Cerri SA, Lemoisson P. Sovereignty by personalization of information search: a collective wisdom may influence my knowledge. In: Cristea AI, Troussas C, editors. Intelligent Tutoring Systems. Cham: Springer International Publishing; 2021. pp. 376-83. DOI
7. Foster I, Kesselman C, Tuecke S. The anatomy of the grid: enabling scalable virtual organizations. *Int J High Perform Comput Appl* 2001;15:200-22. DOI
8. Noy NF, Crubézy M, Fergerson RW, et al. Protégé-2000: an open-source ontology-development and knowledge-acquisition environment. *AMIA Annu Symp Proc* 2003;2003:953. PubMed PMC