

A preliminary proposal of a conceptual Educational Data Mining framework for Science Education

Scientific competences development and self-regulated learning

Rita Tavares*
CIDTFF, DEP, UA
Aveiro, Portugal
ritaveigatavares@ua.pt

Rui Vieira
CIDTFF, DEP, UA
Aveiro, Portugal
rvieira@ua.pt

Luís Pedro
CIC.DIGITAL, DECA, UA
Aveiro, Portugal
lpedro@ua.pt

Abstract—The present paper is part of a wider study, focussed on the development of a digital educational resource for Science Education in primary school, integrating an Educational Data Mining framework. The proposed conceptual framework aims to infer the impact of the adopted learning approach for the development of scientific competences and students' self-regulated learning. Thus, students' exploration of learning sequences and students' behaviour towards available help, formative feedback and recommendations will be analysed. The framework derives from the proposed learning approach, as well as from the literature review. Before introducing it, the authors present an overview of the digital educational resource learning approach and the adopted Educational Data Mining methods. Finally, we present the proposed conceptual Educational Data Mining framework for Science Education, focussing its relevance on the development of students' scientific competences and self-regulated learning.

Keywords—Educational Data Mining; Latent Knowledge Estimation; Causal Data Mining; Domain Structure Discovery; Digital Educational Resources; Scientific Competences; Self-regulated Learning

I. INTRODUCTION

In recent years, as a result of technological advances in network systems and intelligent tutoring systems, new methods based on human-computer interaction have emerged in Educational research, namely Educational Data Mining (EDM) and Learning Analytics (LA) [1]–[3]. According to John Behrens at the LAK 2012 conference, mentioned by Baker & Inventado [1], EDM has a greater focus on learning, aiming to design/develop educational solutions that automatically adapt to the user [2]. Essentially, LA seeks to find ways to report user performance in a certain system – analysis and report of educational data [4]. EDM is an interdisciplinary and emerging research area that results from the application of data mining methods to educational technological systems, showing great potential in various educational subjects and for different stakeholders [3]. When focussed on the students, EDM methods can be used to customise their learning paths, and help them according to their needs/difficulties. When focussed on the teachers, these methods can be used to identify which students need more educational support, and help teachers analyse and reflect about adopted teaching and learning approaches. When used by course designers and Educational researchers (Cf. [5]) they can be used to evaluate the

effectiveness of learning using different environments, and to infer about EDM methods' potential for different research objectives. These methods can also be used by schools/universities and training entities, to suggest new training courses/units, according to students' profiles, and to find class patterns and thus design educational strategies accordingly. From the exposed, EDM methods can be used in diverse systems, including (i) Learning (Content) Management Systems (e.g., online courses participants' monitoring, supporting, and real-time evaluation); (ii) intelligent tutoring systems (e.g., provide recommendations and feedback through system modelling of students' performance/behaviour); (iii) adaptive and intelligent hypermedia systems (e.g., build students' goals, preferences, and knowledge models based on their interaction); and (iv) test and quiz systems (e.g., assessing students' knowledge in a particular concept/subject, using a sequence of items and storing data related to students' scores and statistics) [3]. The present study is focussed on the use of EDM methods for students, teachers and Educational researchers, aiming at the development of a Digital Educational Resource (DER) for Science Education, integrating a framework to infer the impact of the proposed learning approach on students' scientific competences development and self-regulated learning. Thus, students' exploration of learning sequences – correlated (interactive) digital educational contents – and students' behaviour towards available help, formative feedback, and recommendations will be analysed.

Once the wider study predicts the conception and development of components grounded in theory and data collection (e.g., primary school teachers' questionnaire to define DER target), the authors adopted the Educational Design Research (EDR) methodological approach [6]. EDR is focussed on "real world" educational problems, aiming to solve them through scientific knowledge deepening and the development of educational solutions. This approach predicts interactive and iterative phases (Preliminary research phase, Development or Prototyping phase, and Assessment phase), developed according to the ADDIE model [ibid.]. This paper is part of the Preliminary research, that predicts seven moments related to analysis and design processes. The design of the conceptual EDM framework for Science Education is the last moment of this phase, based on previous moments, namely Moment 6 related to the design of DER learning approach. In this regard, the present paper is structured to first present the

This article reports research developed within the Research Fellowship #SFRH/BD/107808/2015, under the *Human Capital Operational Program*, supported by the *European Social Fund* and national funds of the *Ministry of Science, Technology and Higher Education*, funded by *Fundação para a Ciência e a Tecnologia – FCT I.P. – Portugal*; and within the *PhD Program Technology Enhanced Learning and Societal Challenges*, funded by *Fundação para a Ciência e a Tecnologia – FCT I.P. – Portugal*, under contract #PD/00173/2014. This article is, also, financially supported by National Funds through *Fundação para a Ciência e a Tecnologia – FCT I.P. – Portugal*, under the project UID/CED/00194/2013.

DER learning approach (section ii), then clarify the adoption of EDM methods (section iv) and, finally, present the proposed conceptual framework (section v).

II. PROPOSED DER LEARNING APPROACH

The proposed DER learning approach crosses the Inquiry-Based Science Education methodology with the BSCS 5Es instructional model [7], [8]. Inquiry-Based Science Education methodology (IBSE) proposes five phases: *Orientation*, *Conceptualization*, *Investigation*, *Conclusion*, and *Discussion* [7]. The *Orientation* phase aims to stimulate students' curiosity about a certain scientific concept/subject. The *Conceptualization* phase aims to confront and/or inquire students' (pre-)concepts and promote new ideas generation and/or assumptions related to a presented problem/challenge. The *Investigation* phase aims that students plan and apply exploration and investigation processes, collecting, analysing and interpreting data to test the assumptions of the previous phase (e.g., experimental activities). The *Conclusion* phase proposes that the students draw conclusions about the previous phase, comparing/confronting their (pre-)concepts with the collected evidences. Finally, the *Discussion* phase is transversal to the previous phases, aiming at students' ideas and/or results confrontation, promoting students' reflection and (self-)evaluation of the learning process. In the last decade, several authors have approached the IBSE methodology according to the BSCS 5Es instructional model (5Es) [9]–[11]. This model also proposes five phases: *Engage*, *Explore*, *Explain*, *Elaborate*, and *Evaluate* [8]. The *Engage* phase aims to stimulate students' interest and promote their personal and active involvement in learning. This phase should be of short duration, relating and/or confronting students' previous knowledge. The *Explore* phase proposes that the students, once involved in the concept/subject, build their own understanding about it, by confronting and experimenting with scientific phenomena (e.g., experimental activities). This phase should foresee moments for the students to inquire, collect and analyse data, and reflect about the processes and results. The *Explain* phase aims to promote the opportunity for the students to communicate their own findings and establish a theoretical framework about their meaning, predicting students' prior knowledge confrontation. The *Elaborate* phase aims at students' new knowledge application, in order to deepen scientific concepts/subjects and/or proceed towards new learning paths. Finally, the *Evaluate* phase is transversal to the previous phases, and aims to help students realize how much they have learned and how their conceptual frameworks have evolved according to expectations. As much as possible, in this phase formative feedback about the students' learning performance and results should be provided.

Besides the intrinsic relationship between these two approaches, by crossing them we aim to underline and theorize on opportunities to promote primary school students' (a) interest, personal and active involvement in learning (objects); (b) ideas communication, confrontation and inquiry; and (c) real-time reflection and (self-)evaluation about their learning paths. Thus, the proposed DER is designed to first provide the possibility for students to contact with (new) scientific concepts/subjects and/or confronting themselves with their

previous knowledge, and then promote students' involvement in active, reflective, exploratory and (self-)evaluative activities [7]. For that, DER provides a set of correlated (interactive) digital educational contents, aiming at scientific concepts/subjects' contextualization, exploration, application and deepening, allowing students to go through the five phases of the adopted approaches (the IBSE and the 5Es). Regarding the development of scientific competences and the promotion of self-regulation, it is important to briefly clarify both constructs. In the last years, literature has underlined the holistic character of competences development, predicting not only knowledge, but also skills and attitudes as part of human comprehensive development [10], [12]–[14]. Thus, scientific knowledge is the ability to understand and establish relationships, meanings, appreciations and interrelations when confronted with (new) information (factual, conceptual and procedural knowledge of (inter)disciplinary nature) [ibid.]. Scientific skills are the cognitive, social, emotional, physical and practical abilities in a certain scientific subject [ibid.]. That is, the ability to establish complex and organized schemes of thought and/or action to reach a (personal) goal (e.g., to be able to analyse and critically evaluate (new) information and meanings). Finally, scientific attitudes are the dispositions to use scientific knowledge, to understand and reflect about scientific subjects, and adopt competent, critical and reflexive behaviours regarding Science (e.g., to adopt responsible behaviour towards a problem) [ibid.]. Thus, scientific competences are knowledge in action, and so, the three components should be looked at from a holistic point of view.

In line with the exposed, the proposed DER learning approach was designed to promote the development of scientific competences (scientific knowledge, skills and attitudes), highlighting self-regulation as a responsible, critical and reflexive attitude regarding the learning process [15], [16]. Self-regulation is a mindful process in which is used a variety of strategies, resulting in the students' ability to think and act in an active, organized, articulated, critical, reflexive and motivated way, regarding learning [16], [17]. Some self-regulation abilities include: (a) to identify personal interests and learning needs; (b) to set learning objectives and pathways according to personal interests and needs; and (c) to search for personal skills consolidation and deepening opportunities [12]–[14]. In the present study, the emphasis on self-regulation is related to the authors' willingness to promote opportunities for students' self-awareness learning and autonomy, as well as to improve students' ability to adopt informed decision-making, develop self-confidence and remain motivated to learn. To clarify how the intersection of the adopted approaches (IBSE and 5Es) facilitates these aspects, in the following sections we summarize the DER learning approach proposed, exemplifying its operationalization, as well as highlighting its potential for students' scientific competences development and self-regulated learning.

A. Orientation and Engage

For the *Orientation* and *Engage* phases, one proposes that the students watch and explore *interactive animations* (e.g., answering questions about fluctuation so the animation can proceed). This type of animations aims to (i) stimulate

students' curiosity about a particular concept/subject, addressing a problem/challenge (*Orientation* – IBSE). *Interactive animations* also represent an opportunity for (ii) students' self-evaluation about previous knowledge (e.g., establish relationships between previous learning; contact with new stimuli about a concept/subject). The proposed animations will have a short duration, aiming to (iii) draw students' attention/interest; (iv) involve them in a personal way; and (iv) stimulate them to predict, relate and evaluate their previous knowledge (*Engage* – 5Es). In terms of students' scientific competences development, it is expected that *interactive animations* will help the students develop factual scientific knowledge (e.g., concept/subject-specific details); scientific skills (e.g., identify or formulate criteria to draw possible answers); and attitudes (e.g., access available help to solve a problem) [12], [14].

B. Conceptualization and Explore

For the *Conceptualization* and *Explore* phases, we propose that the students explore *games* (e.g., catch falling objects that do not float and prevent them from sinking). *Games* aim to lead the students to form assumptions related to the presented problem/challenge and to test them according to the established dynamics through inquiring (*Conceptualization* – IBSE). *Games* also aim at students' knowledge mobilization, representing an opportunity for them to (i) actively learn; (ii) stimulate them to analyse information, observe and compare phenomena, variables and concepts; (iii) identify requirements and variables that influence outcomes; (iv) interpret results; and (v) draw and confront conclusions (*Explore* – 5Es). In terms of students' scientific competences development, it is expected that *games* will help students develop conceptual scientific knowledge (e.g., classes, categories, principles, systems and scientific phenomena); scientific skills (e.g., decide (by attempting) about the best action/procedure); and attitudes (e.g., follow recommendations of learning reinforcement and/or deepening) [12], [14].

C. Investigation and Explain

For the *Investigation* and *Explain* phases, one proposes that the students explore *simulations* (e.g., perform experimental activities related to fluctuation controlling variables). Starting from a research question, *simulations* aim to (i) lead students to form assumptions; (ii) plan processes; (iii) test assumptions; and (iv) collect, analyse and interpret data (*Investigation* – IBSE). *Simulations* also aim to (v) stimulate students' reflection about how they structure their conceptual framework and the designed research path; (vi) help students draw conclusions and structure their knowledge; (vii) confront their initial ideas with the results of the experimental activity; (viii) establish a theoretical framework about their meaning; and (ix) establish relationships between their choices and the initial research question (*Explain* – 5Es). In terms of students' scientific competences development, it is expected that *simulations* will help students develop procedural scientific knowledge (e.g., define and/or interpret experimental procedures); scientific skills (e.g., observe scientific systems and/or phenomenon variations); and attitudes (e.g., find alternatives to validate the set criteria) [12], [14].

D. Conclusion and Elaborate

For the *Conclusion* and *Elaborate* phases, we propose that students answer *knowledge tests* without the possibility to access help. Before proceeding with a *knowledge test*, the DER recommends that the students access *information areas* to reinforce and/or deepen knowledge (e.g., access an information area related to the application of the "Archimedes' principle" in ships and submarines design, and then answer a knowledge test addressing fluctuation, predicting questions relating the principle to everyday situations). *Information areas* aim to (i) lead the students to deepen and (ii) expand their knowledge, as well as (iii) help them clarify doubts (*Conclusion* – IBSE). *Information areas*, in addition to these phases, can be accessed at any time during the DER's exploration (a "Help" icon is always available in content screens so that the students can dissipate doubts and/or deepen knowledge) (*Elaborate* – 5Es). In terms of students' scientific competences development, it is expected that *information areas* will help students develop conceptual scientific knowledge (e.g., deepen scientific phenomena); scientific skills (e.g., identify necessary assumptions to understand scientific concepts/subjects); and attitudes (e.g., find ways to be well informed about scientific concepts/subjects) [12], [14]. *Knowledge tests*, according to these phases, aim to lead the students to (iv) draw conclusions and (v) reflect about how they construct their knowledge in a particular scientific concept/subject (*Conclusion* – IBSE). *Knowledge tests* also aim at the (vi) students' knowledge mobilization; (vii) to help them discover and understand the implications of the phenomena explored; and (viii) to establish relationships with other concepts/subjects (*Elaborate* – 5Es). In terms of students' scientific competences development, it is expected that *knowledge tests* will help students develop conceptual scientific knowledge in order to deepen their knowledge (e.g., deepen scientific concepts and/or specific details related to the concept/subject addressed); scientific skills (e.g., identify or formulate criteria for possible answers); and attitudes (e.g., analyse statements and (ir)relevant information) [ibid.].

E. Discussion and Evaluate

For the *Discussion* and *Evaluate* phases, one proposes the integration of *formative feedback* about students' results and learning paths and, simultaneously, the availability of *recommendations* to (i) reinforce or deepen students' knowledge, helping them to (ii) self-regulate their learning (e.g., what content to (re-)explore). *Formative feedback* and *recommendations* also aim at (iii) students' reflection on knowledge construction (e.g., decide to access an information area to learn more about a particular concept/subject and, thus, improve their performance); and (iv) self-awareness of their learning (e.g., performance level) [10]. In these phases, *knowledge tests* are also proposed as a knowledge assessment strategy. In this regard, *knowledge tests* are aimed at (v) the evaluation of the students' understanding of a particular scientific concept/subject. *Knowledge tests* also aim at leading students to (vi) apply their new knowledge, and (vii) deepen their conceptual framework or advance towards new research paths (*Evaluate* – 5Es). In terms of students' scientific competences development, it is expected that *knowledge tests* will help students develop conceptual scientific knowledge in

order to assess knowledge (e.g., verify the domain of scientific concepts); scientific skills (e.g., interpret statements and answer questions); and attitudes (e.g., use their knowledge to analyse statements, relevant information, and answer correctly and critically) [12], [14]. *Formative feedback and recommendations* aim to lead the students (viii) to constantly and continuously be aware about how much they have learned and how their conceptual framework evolved; (ix) to a greater understanding of the scientific competences developed; and (x) to find ways of self-correction and readjustment according to what is expected (*Evaluate* – 5Es). It is, therefore, a formative and immediate assessment, provided under a simulated student-teacher and peer communication approach, as well as teacher, self- and peer assessment environment (*Discussion* – IBSE).

Aiming to infer about the potential of the proposed DER learning approach on students' scientific competences development and self-regulated learning, the integration of an EDM framework for Science Education in the DER is proposed and presented in the following sections.

III. EDM METHODS

EDM methods application, as a possibility of knowledge discovery, require the establishment of clear objectives, so data collection, processing, analysis and interpretation can result in relevant inferences [1]–[3], [18]. In the last years, several authors have guided their research in order to (a) predict students' learning behaviours (knowledge, motivation and attitudes); (b) study the effects of different types of pedagogical support to improve students' learning; and (c) infer about the optimal (sequences of) contents/subjects for each student, based on their difficulties, gains and preferences (Cf. [19]). According to the intended objective, there are several methods (1, 2,...) and techniques (a, b,...) for data mining in Educational research, as briefly presented below [1], [3], [20]:

- 1) **Prediction:** the goal is to infer about a single aspect of the data collected (predicted variable – like the dependent variable in statistical analysis), by combining other aspects of the data (predictor variables – similar to the independent variable in statistical analysis). As the name indicates, it is a method that predicts what will happen in the future, and it can be used to predict students' educational success rates, and the students' behaviour according to certain stimuli.
 - a) **Classification:** a technique used when the predicted variable is categorical. It can be applied in training or test situations, mapping data to predefined classes. This technique can be used to analyse students' educational performance, and to monitor students' learning paths in digital platforms.
 - b) **Regression:** a technique used when the predicted variable is continuous. It can be used in intelligent systems that support learning agents/tutors, in which the tutor learns to predict the probability of the users' following action, answer, and/or behaviour.
 - c) **Latent Knowledge Estimation:** a technique used to estimate students' knowledge through correctness patterns (mapping correct and incorrect answers)

according to defined objectives and, occasionally, mapping other relevant information that influence knowledge development. “The word ‘latent’ refers to the idea that knowledge is not directly measurable, it must be inferred from a student's performance” [1, p. 64] and actions.

- 2) **Relationship Mining:** the goal is to identify relationships between variables and to codify them into rules for later use, trying to find out which variables are most strongly correlated to a particular variable of interest, or what is the correlation between two variables of interest. It can be used to identify students' behaviour patterns and difficulties or learning mistakes that frequently occur at the same time.
 - a) **Association Rule Mining:** a technique used to find any relationship between variables, aiming to find “if-then” rules. It can be used to find relationships such as “if the students intend to improve their performance, then they will frequently use the available help”.
 - b) **Sequential Pattern Mining:** a technique used to find temporal associations between variables or events. It can be used to find students' requests for help patterns over time in software exploration.
 - c) **Correlation Mining:** a technique used to find linear correlations between variables (positive or negative). It can be used to find relationships between students' attitudes towards an activity (*positive* – they try to finish, or *negative* – they leave the activity) and help request frequency).
 - d) **Causal Data Mining:** a technique used to find relationship causes between variables, i.e., to find out if an event is caused/originated by another. It can be used to predict which factors influence students' performance in an activity, such as acceptance of software recommendations.
- 3) **Structure Discovery:** the goal is to find data structure (relationships) without any predefined idea/premise about what should be found. This method is, therefore, opposed to predicting methods, since it does not provide a previous definition of variables correlations before data mining method application.
 - a) **Clustering:** a technique used to group similar data into clusters, in order to discover data groups. It can be used to map students' preferences in the exploration of different types of educational contents, and to find interaction-learning patterns.
 - b) **Factor Analysis:** a technique used to find correlated variables, dividing each set of variables into a set of latent facts (i.e., not directly observable). It can be used to determine correlated contents in an online course, and to find which events result in other events.
 - c) **Domain Structure Discovery:** a technique used to discover which factors influence students' specific competences development. It can be used to map students' performance and interactions during the exploration of an intelligent tutoring system.

Attending to the study goal, to infer about the potential of the proposed DER learning approach (i) to promote students' scientific competences development through the exploration of learning sequences and the available recommendations to reinforce or deepen students' knowledge; and (ii) to promote students' self-regulated learning through recommendations, formative feedback and available help; *Prediction*, *Relationship Mining*, and *Structure Discovery* methods will be adopted in the proposed EDM framework for Science Education. The following section presents the aspects that consubstantiated the authors' options and the proposed conceptual framework.

IV. CONCEPTUAL EDM FRAMEWORK FOR SCIENCE EDUCATION

As previously mentioned, in order to infer about the potential of the proposed DER learning approach in students' scientific competences development and self-regulated learning, the integration of an EDM framework for Science Education in the DER is proposed. The choice of EDM methods and techniques emerged from the set of questions presented in Fig. 1, as well as the need to collect, analyse and draw inferences about the data resulting from the represented events.

Regarding Q1) *What is the impact of correlated (interactive) digital educational contents sequences in students' scientific knowledge and skills development?* we intend (i) to infer about the increase of students' scientific knowledge and skills levels, through correctness patterns (mapping students' correct and incorrect answers), according to the defined objectives; and (ii) to find events patterns that influence knowledge and skills development. In other words, we intend to infer about the positive impact of correlated (interactive) digital educational contents in students' educational performance in the learning sequences, using tests to verify knowledge construction. To infer about the data collected and analysed from Q1) we propose the use of *Prediction – Latent Knowledge Estimation*, to estimate students' scientific knowledge and skills levels, through correctness patterns, and *Relationship Mining – Causal Data Mining* to find relationship causes between the “complete learning sequences” and “educational performance improvement” events.

Regarding Q2) *What is the impact of formative feedback and recommendations in students' self-regulated learning?* and Q2a) *In what situation do the students accept DER recommendations?*, we intend (i) to infer about the increase of students' self-regulation through the awareness of learning path and the availability of recommendations (Q2 – proceed according to the recommendation / do not proceed according to the recommendation); and (ii) to infer about the situations in which the students accept the recommendations (Q2a – reinforcement / deepening). Simultaneously, we intend (iii) to infer about events caused by another event, that is, to infer about the impact of formative feedback and recommendations on students' scientific knowledge and skills development (Q3). In this regard, we also intend (iv) to infer if the fact that the students accept the recommendations (Q2a) promotes students' educational performance improvement in the learning sequences and in the tests (Q3). To infer about the data

collected and analysed from Q2) and Q2a) we propose the use of *Relationship Mining – Causal Data Mining* to find relationship causes between the “proceed according to the recommendation / do not proceed according to the recommendation”, “learning reinforcement / no learning reinforcement”, and “learning deepening / no learning deepening” events. To infer about the data collected and analysed from Q2a) and Q3) we also proposed the use of *Relationship Mining – Causal Data Mining* to find relationship causes between the “learning reinforcement / deepening” and “educational performance improvement” events.

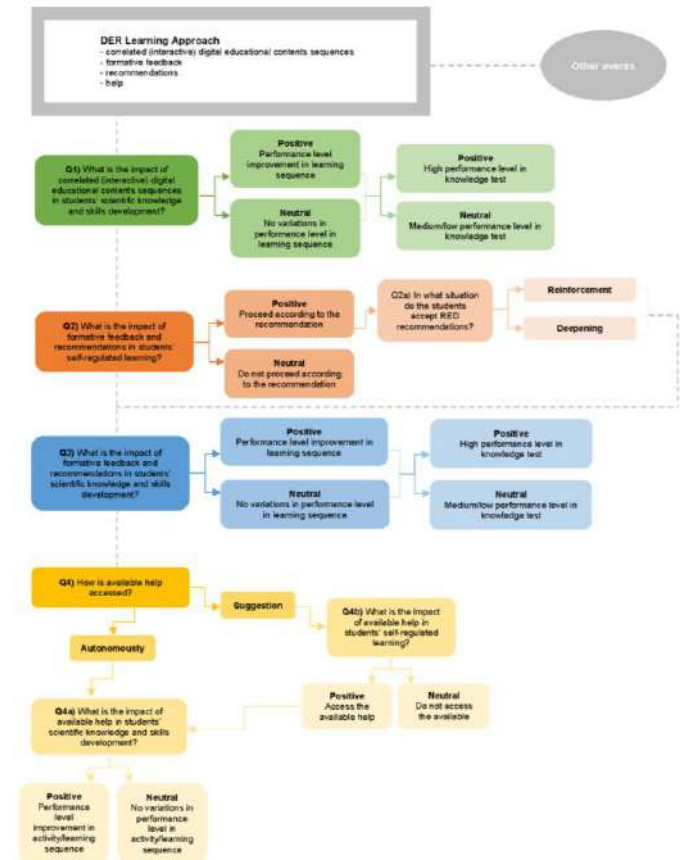


Fig. 1 Relational structure: questions and events that result in the conceptual EDM framework for Science Education

Finally, regarding Q4) *How is available help accessed?*, Q4a) *What is the impact of available help in students' scientific knowledge and skills development?*, and Q4b) *What is the impact of available help in students' self-regulated learning?*, we intend to infer (i) if the students access available help autonomously or by suggestion (Q4); (ii) if students accept DER help suggestions (Q4b); and (iii) if the available help has impact in students' scientific knowledge and skills development (Q4a), that is, to infer about events caused by another event. To infer about the data collected and analysed from Q4), Q4a) and Q4b) we propose the use of *Relationship Mining – Causal Data Mining* to find relationship causes between the “access the available help autonomously or by suggestion” and “self-regulated learning levels” events; and the “access the available help autonomously or by suggestion” and

“educational performance improvement in activity/learning sequence” events.

In addition to the exposed, and attending to EDM methods potential, we also propose to explore “Other events” that will result in a deeper understanding about the potential of the proposed DER learning approach on students’ scientific competences development and self-regulated learning, among others (i) students’ most accessed content type; (ii) students’ educational performance in each content type; (iii) students’ educational performance in a learning sequence each time they explore it; (iv) students’ global educational performance; (v) students’ time spent in contents/sequences exploration each time they repeat them; (vi) students’ most accessed scientific concepts/contents/subjects; (vii) students’ autonomous and suggested total accesses to available help; (viii) students’ total acceptances of DER recommendations; and (ix) students’ total times they complete and abandon a content/learning sequence. To infer about the data collected and analysed from “Other events” we propose the use of *Structure Discovery – Domain Structure Discovery*, to unveil which unpredicted correlated events influence educational performance improvement, and therefore, students’ scientific competences development and self-regulated learning.

V. CONSIDERATIONS

The proposed conceptual EDM framework for Science Education presents a holistic approach, attending to its application potential, as well as to the knowledge that may emerge from it. Once it allows to infer about students’ scientific competences development and find events’ patterns that influence scientific knowledge and skills development, this framework presents potential benefits for students (e.g., learning personalization); teachers (e.g., identify learning needs/gaps); and Educational researchers (e.g., evaluate the effectiveness of students’ exploration of learning sequences to reinforce/deepen Science learning). Regarding the possibility to find causal relationship between events, inferring about the impact of formative feedback, help and recommendations as to students’ self-regulated learning and scientific competences development, the framework presents potential increments for students (e.g., recommend students’ most appropriate contents to improve their educational performance); teachers (e.g., identify which students need more educational support); and Educational researchers (e.g., investigate new approaches to improve students’ Science learning). Not less important, deriving from “Other events” data, the framework offers great potential to infer about several aspects that influence students’ learning and pedagogical approaches (e.g., students’ most accessed content type; and students’ most accessed scientific concepts/contents/subjects). From the exposed, the proposed framework will allow us to (1) infer about the DER learning approach potential on students’ scientific competences development and self-regulated learning; (2) validate the proposed DER for Science Education in primary school; (3) improve future developments (e.g., improvement of the (interactive) digital educational contents); and (4) conduct new studies based on the data collected and analysed (e.g., extended implementation of the proposed DER learning approach).

REFERENCES

- [1] R. Baker and P. Inventado, “Educational Data Mining and Learning Analytics,” in *Learning Analytics: From Research to Practice*, J. Larusson and B. White, Eds. New York: Springer, 2014, pp. 61–75.
- [2] R. Baker and G. Siemens, “Educational Data Mining and Learning Analytics,” *Columbia Univ.*, 2013.
- [3] C. Romero and S. Ventura, “Data mining in education,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 3, no. 1, pp. 12–27, 2013.
- [4] M. Bienkowski, M. Feng, and B. Means, “Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief,” 2012.
- [5] S. ElAtia, D. Ipperciel, and O. Zaiane, *Data mining and learning analytics: applications in educational research*. New Jersey: John Wiley & Sons, Inc., 2016.
- [6] S. McKenney and T. Reeves, *Conducting Educational Design Research*. New York: Routledge, 2012.
- [7] M. Pedaste, M. Mäeots, L. Siiman, T. de Jong, S. van Riesen, E. Kamp, C. Manoli, Z. Zacharia, and E. Tsourlidaki, “Phases of inquiry-based learning: Definitions and the inquiry cycle,” *Educ. Res. Rev.*, no. 14, pp. 47–61, 2015.
- [8] R. Bybee, “The BSCS 5E instructional model and 21st century skills,” in *Paper prepared for the Workshop on Exploring the Intersection of Science Education and the Development of 21st Century Skills*, National Research Council, 2009, pp. 2–24.
- [9] L. Chitman-Booker and K. Kopp, *The 5Es of Inquiry-Based Science*. Huntington Beach: Shell Education, 2013.
- [10] W. Harlen, *Assessment & Inquiry-Based Science Education: Issues in Policy and Practice*. Trieste: Global Network of Science Academies (IAP) Science Education Programme (SEP), 2013.
- [11] C. Wilson, J. Taylor, S. Kowalski, and J. Carlson, “The Relative Effects and Equity of Inquiry-Based and Commonplace Science Teaching on Students’ Knowledge, Reasoning, and Argumentation,” *J. Res. Sci. Teach.*, vol. 47, no. 3, pp. 276–301, 2010.
- [12] C. Gomes, J. Brocardo, J. Pedroso, J. Carrillo, L. Ucha, M. Encarnação, M. Horta, M. Calçada, R. Nery, and S. Rodrigues, *Perfil dos Alunos à Saída da Escolaridade Obrigatória*. Lisboa: Ministério da Educação, 2017.
- [13] OECD, *Global competency for an inclusive world*. Paris: Secretary-General of the OECD, 2016.
- [14] C. Tenreiro-Vieira and R. Vieira, *Construindo práticas didático-pedagógicas promotoras da literacia científica e do pensamento crítico*. Madrid: OEI & IBERCENCIA, 2014.
- [15] R. Carneiro, P. Lefrere, K. Steffens, and J. Underwood, *Technology Enhanced Learning - Self-Regulated Learning in Technology Enhanced Learning Environments: A European Perspective*, 5th ed. Rotterdam: Sense Publishers, 2011.
- [16] P. Winne and A. Hadwin, “Studying as Self-Regulated Learning,” in *Metacognition in Educational Theory and Practice*, D. Hacker, J. Dunlosky, and A. Graesser, Eds. Oxon: Routledge, 2009, pp. 277–304.
- [17] B. Zimmerman, “Self-Regulated Learning and Academic Achievement: An Overview,” *Educ. Psychol.*, vol. 25, no. 1, pp. 3–17, 2010.
- [18] L. Prabha and M. Shanavas, “Educational Data Mining applications,” *Oper. Res. Appl. An Int. J.*, vol. 1, no. 1, pp. 23–29, 2014.
- [19] A. Peña-Ayala, “Educational data mining: A survey and a data mining-based analysis of recent works,” *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1432–1462, 2014.
- [20] K. Sukhija, M. Jindal, and N. Aggarwal, “The Recent State of Educational Data Mining: A Survey and Future Visions,” in *2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE)*, 2015, pp. 354–359.