J. Health Inform. 2023 Abril-Setembro; 15(2): 53-59

Artigo de Revisão

DOI: https://doi.org/10.59681/2175-4411.v15.i2.2023.970

JHI JOURNAL OF HEALTH INFORMATICS

Methodology for developing OpenEHR archetypes: a narrative literature review

Metodologias para desenvolvimento de arquétipos OpenEHR: uma revisão narrativa de literatura

Metodologías para desarrollar arquetipos OpenEHR: una revisión narrativa de la literatura

Daiane Evangelista Ferreira[1,2], Jano Moreira de Souza[3]

Keywords: Electronic Health Record, Semantics, Methodology

ABSTRACT

Objective: This article aims to present a literature review in narrative form in order to identify, analyze, and characterize the state of the art methodologies for developing openEHR archetypes. **Method:** An exhaustive literature search in the computer science field was carried out, based on the following databases: IEEE Digital Library, ACM Digital Library, Science Direct, Scopus and Springer Link. The screening process involved applying suitable selection criteria to 361 publications to define the scope for selecting the appropriate papers. **Results:** The nine selected papers were grouped into five categories, in which we identified some connection points between the papers, and we realized that any gaps in one paper are complemented by the other papers. **Conclusion:** The research contributed to the construction of a theoretical reference on methodologies for developing openEHR archetypes, as well as showing that it is a growing research topic and there are some aspects that require further study.

Descritores: Registros Eletrônicos de Saúde, Semântica, Metodologia

RESUMO

Objetivo: Apresentar uma revisão narrativa de literatura para identificar, analisar e caracterizar o estado da arte sobre metodologias para o desenvolvimento de arquétipos openEHR. **Método:** Pesquisa exaustiva na literatura da área de ciência da computação. Utilizaram-se as bases de dados IEEE Digital Library, ACM Digital Library, Science Direct, Scopus e Springer Link. O processo de revisão envolveu a aplicação de critérios de seleção aos 361 textos encontrados, de modo a selecionar os artigos que se adequassem ao escopo. **Resultados:** Os 9 artigos selecionados foram agrupados em cinco categorias, onde identificamos algumas conexões e notamos que as lacunas de alguns artigos eram complementadas pelas lacunas de outros. **Conclusão:** A pesquisa contribuiu para a construção de um referencial teórico sobre metodologias para desenvolvimentos de arquétipos openEHR, mostrando que é um tópico de pesquisa em crescimento e que alguns aspectos ainda necessitam de mais estudos.

Descriptores: Registros Electrónicos de Salud, Semántica, Metodología

RESUMEN

Objetivo: Presentar una revisión narrativa de la literatura para identificar, analizar y caracterizar el estado del arte de las metodologías para el desarrollo de arquetipos openEHR. **Método:** Búsqueda exhaustiva en la literatura en el área de informática. Utilizamos las bases de datos IEEE Digital Library, ACM Digital Library, Science Direct, Scopus y Springer Link. El proceso de revisión implicó la aplicación de criterios de selección a los 361 textos encontrados, con el fin de seleccionar los artículos que se ajusten al alcance. **Resultados:** Los 9 artículos seleccionados se agruparon en cinco categorías, donde identificamos algunas conexiones y notamos que los vacíos de algunos artículos se complementaban con los vacíos de otros. **Conclusión:** La investigación contribuyó a la construcción de un marco teórico sobre metodologías para el desarrollo de arquetipos de openEHR, mostrando que es un tema de investigación en crecimiento y que algunos aspectos aún necesitan más estudios.

[1] *Business Analyst, ProntLife, Rio de Janeiro (RJ), Brazil.*
[2] *Doctoral Student, Systems Engineering and Computer Science Program, Federal University of Rio de Janeiro (PESC/COPPE/UFRJ), Rio de Janeiro (RJ), Brazil.*
[3] *PhD, Full Professor, Systems Engineering and Computer Science Program, Federal University of Rio de Janeiro (PESC/COPPE/UFRJ), Rio de Janeiro (RJ), Brazil.*

Autor Correspondente: **Daiane Evangelista Ferreira**
e-mail: **daiane@cos.ufrj.br**

Artigo recebido: 14/06/2022
Aprovado: 02/10/2023

1. INTRODUCTION

Over the past two decades, there has been a growing need to develop electronic health records (EHR), which have been introduced into hospitals and clinics to manage patient data. However, organization and standardization are required to enable transfer of the data and information between systems[1]. The openEHR approach emerged as a way of standardizing these health data, using an archetype paradigm (with a two-level model) as a new technique to transfer patient data between EHR systems[1]. OpenEHR is the name of one particular technology for the healthcare area, more specifically for e-health, and it consists of open access specifications, clinical models and software that can be used to create standards, information, and interoperability solutions for the health area[2].

Archetypes are structures that define how clinical data should be stored. Each has an identifier, and each data point can be accessed through a path within the archetype. These identifiers and paths are unique and independent of the context in which the archetype is used[3]. Thus, archetype development is an important part of the openEHR standard[4]. And seeking to understand the state of the art about the development of openEHR archetypes, we conducted a literature review in narrative format of the computer science field in order to find methodologies for the development of openEHR archetypes for health.

2. METHOD

To understand the state of the art in computing and healthcare in terms of the use and development of openEHR archetypes, we did a narrative literature review using some guidelines proposed by Kitchenham and Charters[5,6]. The objective of this narrative literature review was to identify scientific publications on methodologies for the development of openEHR archetypes, with a focus on publications related to the area of computing and health, and those which present the steps for the development of openEHR archetypes in their respective methodologies.

For this, the first step of the narrative literature review was to define the research questions:

1) Which articles involve research about developing openEHR archetypes using specific methodologies?

2) How do these articles describe these methodologies for developing openEHR archetypes?

2.1. Search strategy

From the definition of the research questions, we selected the following keywords: archetype, methodology and openEHR. From this, we then constructed the following string for the search:

("openEHR archetype" AND methodology (OR "openEHR archetype methodology"))

It is important to highlight that we defined the term "openEHR archetype" instead of the individualized terms "openEHR" and "archetype" because, in both scenarios, the search result would be very comprehensive and may even have no relation to the theme (in the case of the term archetype, which is used in several areas).

The databases chosen for the research were: IEEE Digital Library, ACM Digital Library, Science Direct, Scopus and Springer Link. These databases were chosen because they are internationally recognized and are the main vehicles for publication in the field of computer science. To optimize the results of the publications in these digital libraries, we applied the search string only in metadata (i.e., title, abstract, and keywords), thus refining the results found.

2.2. Selection criteria

According to the guidelines proposed by Kitchenham and Charters[5,6], the process of selecting the criteria ensures that only studies aligned with the objective of the narrative literature review are selected.

Thus, in the selection stage, the focus was on analyzing the papers found in the databases, and defining which ones were within the scope of the narrative literature review. In order to include papers that were aligned with the objective of the research, the inclusion criteria established were: (1) papers written in English, (2) papers published between 2002 and 2020, (3) papers that presented some theoretical reference related to methodologies for developing openEHR archetypes, (4) papers that presented some theoretical reference related to methodologies for developing openEHR archetypes with examples of archetypes, and (5) papers that clearly presented some approach that uses methodologies to develop openEHR archetypes.

Regarding the discarding of papers that were not in accordance with the objective of the narrative literature review, the exclusion criteria established were: (1) type of publication — texts like reviews, reports, posters, brief reports, books, textbooks, theses and dissertations, editorial letters, brief communications, commentaries, and unpublished working papers were discarded; (2) papers that were not written in English; (3) papers with text not available in full; (4) duplicated papers; and (5) papers that did not directly address methodologies for developing openEHR archetypes.

3. STUDY CONDUCTION

The narrative literature review was conducted using the online tool Parsifal[1], which is a tool for organizing literature reviews — it assists in the process of selecting and conducting research. The first step was the inclusion of

[1] https://parsif.al

the following narrative literature review planning items: objectives, research questions, keywords, search string, sources (databases), and selection criteria. Using the defined search string, we then did searches of the IEEE Digital Library, ACM Digital Library, Science Direct, Scopus and Springer Link databases. The search results were imported into the Parsifal tool through bibtex files. After importing the bibtex files, the studies were loaded into the tool, and we started the study selection process, which involved the following three stages: duplicate verification, selection, and refinement. In the duplicate verification stage, the Parsifal tool performed the automatic verification, but the researcher responsible also performed manual verification to ensure that the process was correct. In the selection stage, the researchers analyzed the abstracts of the studies, considering the inclusion and exclusion criteria. This resulted in 54 studies being accepted and 279 being excluded. Finally, in the refinement stage, the researchers read the full texts, paying attention to the following characteristics: (1) if the paper presented a process of creation and/or reuse of openEHR archetypes, (2) if the paper described the steps for the process of creating and/or reusing openEHR archetypes, and (3) if the paper described the actors involved in these steps for the process of creating and/or reusing openEHR archetypes. After this analysis, 9 papers were accepted. These numbers can be seen in detail in Table 1.

Table 1: Number of selected papers

Databases	Preliminary result	Duplicated	Selection stage		Refinement stage	
			Included	Excluded	Included	Excluded
IEEE	17	0	12	5	2	10
ACM	17	0	5	12	0	5
Science Direct	94	0	23	71	4	19
Scopus	104	28	5	71	2	3
Springer Link	129	0	9	120	1	8
Total	**361**	**28**	**54**	**279**	**9**	**45**

4. RESULTS

As a result of the refinement step, 9 papers were selected for analysis regarding methodologies for developing openEHR archetypes. After reading and analyzing them, we identified some common themes among them and performed a step called categorization. We identified five categories, which are summarized in Table 2.

Table 2: Categories for analyzed papers

Categories	Selected studies
Methodology design	7. and 8.
Methodology applied to specific clinical scenarios	9., 10., 11. and 12.
Methodology applied to exams	13.
Archetype modeling for EHR system	14.
Archetype modeling and interoperability	15.

4.1. METHODOLOGY DESIGN

The methodology design category includes papers that deal with methodologies for the development of openEHR archetypes in general, as they are a kind of guideline for other studies. According to Eguzkiza et al.[7], the openEHR archetype development process has three steps, in which each step has its own sub-steps, summarized in Table 3:

Table 3: Three steps for developing OpenEHR archetypes[7]

Definition of the project	Design of the clinical process	Building the electronic model
Detection of a need within a healthcare process	Definition of the clinical process	Creation and update of archetypes
Establishment of a work team	Study of clinical concepts	Definition of semantic links to clinical terminologies
Analysis of resources	Hierarchical organization of knowledge artefacts	Building templates
		Modeling guideline rules and workflow
		Modeling UI templates

Eguzkiza et al.[7] did the following in their study: explained in detail each step of the process; proposed a methodology that would allow experts in the clinical domain to be actively involved in the definition of EHR content without having to resort to software developers; and provided a proof of concept in the clinical setting of ophthalmology. It is worth noting that their paper highlights the need for training a multidisciplinary team to apply the methodology, so that consistent results can be achieved for the desired clinical case. Furthermore, it is a paper that highlights that the relationship between archetypes and clinical terminologies is important but difficult to implement, which causes delays in the implementation of the methodology.

According to Moner, Maldonado and Robles[8], the process of developing openEHR archetypes involves the steps summarized in Table 4.

Table 4: Five steps for developing openEHR archetypes[8]

Analysis	Design	Development	Validation	Publication
Scope definition and selection of work group	Information structuration	Archetype reuse (new archetype; specialize or modify archetype; existing archetype without changes)	Archetype review	Archetype and template publication
Clinical concept discovery	Constraint definition	Archetype structure development	Template review	
Information elements gathering		Archetype terminology binding		
		Template structure development		
		Template terminology refinement		

The methodologies of the two papers[7,8] complement each other, because the second one presents two more processes: the validation and the publication of the archetypes. Furthermore, both highlight the difficulty of managing the relationship with medical terminologies, as the same information can be represented as part of the archetype structure or as terminology concepts[7,8].

4.2. Methodology applied to specific clinical scenarios

Regarding this category, the analyzed papers involve case studies under specific clinical scenarios, in which the openEHR archetypes are developed according to these scenarios.

4.2.1. Eating disorder

The paper by Maranhão et. al.[9] presents the process for developing openEHR archetypes for eating disorders, as well as describing the process and challenges involved in creating openEHR archetypes associated with screening tools for eating disorders. The authors believe that the archetypes created can help improve data storage for eating disorders and patient care. As part of the archetype modeling process, the authors did research on the Clinical Knowledge Manager (CKM)[16] — an openEHR archetype repository — in order to verify the existence of archetypes on the topic of eating disorders, and they used the Archetype Editor tool to develop the three archetypes. Validation through submission to the CKM was not part of the scope of the paper but was identified as a future work[9].

4.2.2. Nutrition and childhood obesity

Maranhão et al.[10] described the process of developing a framework, based on openEHR archetypes, for the representation of clinical practice focused on the treatment of childhood nutrition and obesity. Additionally, they described the use of a four-step methodology: identification of a clinical practice guideline, analysis in the CKM, modeling of openEHR archetypes, and submission of archetypes for review by the CKM community. The actors involved were: one nutritionist, one physician, and three IT professionals. The authors[10] developed a guideline involving the clinical terms of interest: data, diagnosis, risk factors, comorbidities, laboratory tests, body fat composition, family story, nutrient intake, and dietary assessment. According to the authors, the process of validating through submission to the CKM was not within the scope of the paper, but it was identified as a future work[10].

4.2.3. Demographic and obstetric data

The paper by Pahl et al.[11] had three objectives: to model demographic and obstetric data used in Brazilian healthcare units, by reusing, editing, or creating new openEHR archetypes; describe the process and issues related to openEHR content development at different moments in the obstetric setting; and validate the use of openEHR in relation to obstetric data in the Brazilian scenario. Thus, the process of developing the openEHR archetypes included extracting, classifying, and reorganizing data into demographic information and clinical information[11].

4.2.4. Diabetic retinopathy

The paper by Eguzkiza et al.[7] deals with a methodology for modeling patient-centered clinical processes, which seeks interoperability and knowledge reuse for continuity of patient care. The authors emphasize that it is a methodology that allows healthcare specialists to be involved in defining EHR content without resorting to IT professionals. As a proof of concept, the methodology presents a standardized clinical process, modeled on an e-ophthalmology-based service for handling the diabetic retinopathy screening scenario[7].

4.2.5. Multiple sclerosis functional composite

The paper by Braun et al.[12] deals with the archetype review process, with a focus on the team's role in

improving the semantic quality of archetypes, beyond what is defined by the openEHR reference model. The basis of the article is the Multiple Sclerosis Functional Composite, which is a performance scale based on three neurological tests for evaluating patients with multiple sclerosis. Four openEHR archetypes were developed in the archetype modeling process: one for each neurological test and one for the total score. These archetypes were developed without correlations to terminologies, but they were validated by the CKM community[12].

4.3. Methodology applied to exams

The paper by Papež and Mouček[14], in this category, deals with the feasibility of applying openEHR to model the data stored in EEGBase, which is a portal for managing experimental data from electroencephalograms and related potential events. The authors evaluated the reuse of existing openEHR archetypes and proposed a set of new archetypes together with openEHR templates from the same domain. Their paper also shows the linking step with the terminology, where the odML electrophysiology terminology was chosen[14].

4.4. Archetype modeling for EHR system

In this category is the paper by Santos, Bax and Kalra[13], which discusses the modeling process used by the Health Department of Minas Gerais (Brazil) to support the development of its regional EHR system, as well as the lessons learnt during the process. In their openEHR archetype development process 20 archetypes were developed in the ADL format using LinkEHR tool, but they were not submitted to CKM for validation[13].

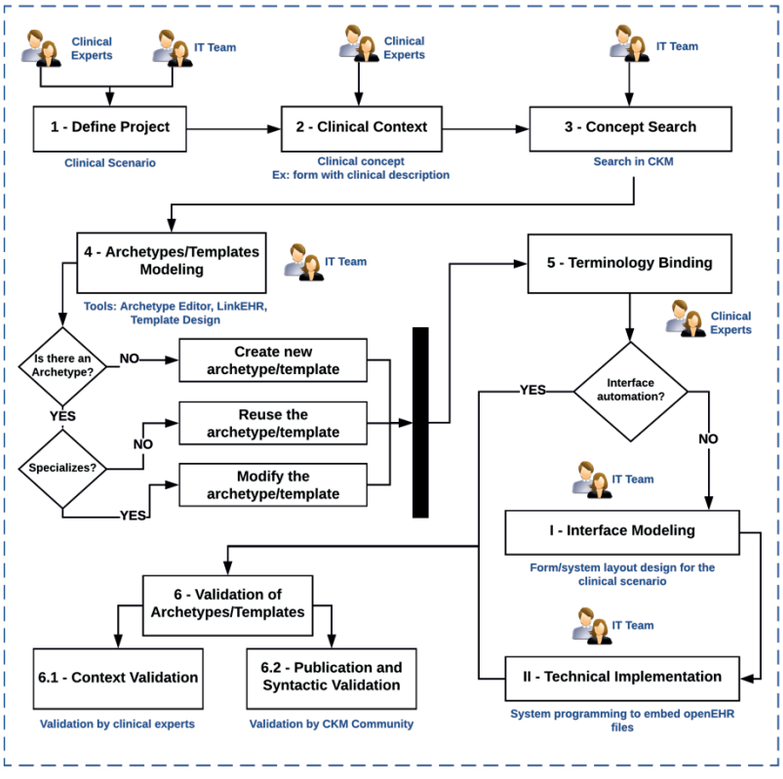
4.5. Archetype modeling and interoperability

In this category is the paper by Marcos et. al.[15], which addresses the problem of interoperability between clinical decision support systems and EHR systems, as well as describing the implementation of a prototype that demonstrates the feasibility of the interoperability between them, based on clinical guidelines through openEHR archetypes. The authors described the process for specialization of openEHR archetypes obtained from the CKM, stating that they had the support of an oncologist in the process, and they briefly addressed the specialization process and broadly described the concepts needed for each stage of specialization. Additionally, they demonstrated the integration process between the clinical decision support system and the EHR system (which is the most detailed process of the study), with the definition of archetypes being the first step in the process[15].

5. DISCUSSION

To understand the methodologies for developing the openEHR archetypes described in the papers analyzed, we drew a flowchart with a systematization of these methodologies (Figure 1) to get an overview of what the different papers cover.

Figure 1: Systematization of the OpenEHR archetype modeling process from the analyzed papers



We then identified points of connection between the papers, and we realized that any gaps in one paper are complemented by the other papers. In Table 5 we present a summary of the characteristics of the analyzed papers.

Table 5: Summary of the characteristics of the analyzed papers

Characteristics versus Papers	7.	8.	9.	10.	11.	12.	13.	14.	15.
Archetype modeling process	x	x	x	-	x	x	x	x	x
Develops archetypes (creates, specializes)	x	-	x	x	x	x	x	x	x
Reuses archetypes	x	-	-	x	x	-	x	-	x
Searches archetypes in CKM	x	-	x	x	x	x	x	x	x
Shows the stages of development	x	-	x	x	x	x	x	x	x
Describes the actors in the process	x	x	x	x	-	x	-	x	x
Involves health specialists	x	x	x	x	x	x	-	x	x
Uses clinical terminologies	-	-	-	-	-	-	x	x	-
Validates the archetypes in CKM	-	-	-	-	-	x	-	-	-

When viewing the summary of Table 5, we can see that the paper by Moner et al.[8] does not have most of the characteristics, because it is a text without practical application, which only describes a methodology. Another highlight is that the papers mention the need to use clinical terminology, but only two papers perform this step in their respective scenarios. In the case of Braun et al.[12], it is worth noting that it is the only article that presents the archetype validation process via the CKM community.

Most papers followed an openEHR archetype modeling flow that involves healthcare experts as well as IT professionals, and almost all the articles are based on archetype research on CKM. Thus, we can highlight some points in the literature in relation to development of openEHR archetypes, which deserve greater emphasis: the classification of relationships with medical terminology — only[13,14] performed this step; and validation of archetypes by the CKM community.

Furthermore, in the analysis of the results, we noticed that the selected papers fall within the period 2013–2018, which shows that the discussion about methodologies for developing openEHR archetypes is still recent and has been growing over the last few years.

6. CONCLUSIONS

The aim of this article was to conduct a narrative literature review in the field of computer science and health on methodologies for developing openEHR archetypes. After conducting the narrative literature review, a total of nine papers were selected to be analyzed, and from this it was possible to understand the steps of the openEHR archetype development process, as well as draw a flowchart with these steps and list some characteristics.

When analyzing the steps of the openEHR archetype development process, we noticed that the steps for defining projects and concepts and modeling the archetypes are common among many papers. However, there are few examples in the literature addressing the "terminology binding" and "validation" steps, which makes them important research points for future work.

Some of the limitations of this work were the language choice — only papers in English were selected for analysis. Additionally, the focus of the narrative literature review was papers that presented methodologies for developing openEHR archetypes in the context of the computer science and health field — studies addressing other fields were discarded.

We concluded that conducting a narrative literature review contributed to the construction of a theoretical reference on methodologies for developing openEHR archetypes, as well as showing that it is a research topic that is still growing.

ACKNOWLEDGMENTS

We would like to take this opportunity to thank the National Council for Scientific and Technological Development of Brazil for funding this research.

REFERENCES

1. Conde AM. *Towards best practice in the Archetype Development Process* [Dissertation]. Dublin: The University of Dublin - Trinity College Dublin, Center for Health Informatics, Department of Computer Science; 2010. [Internet]; [cited 2020 Oct 29]. [Accessed 2022 Jun 13]. Available from: <https://www.scss.tcd.ie/publications/theses/diss/2010/TCD-SCSS-DISSERTATION-2010-016.pdf>.
2. OpenEHR Foundation. *What is openEHR?*. c2022. [Internet]; [cited 2020 Oct 29]. [Accessed 2022 Jun 13]. Available from: <https://www.openehr.org/>
3. Cruz-Correia R, Ferreira D, Bacelar G, Marques P, Maranhão P. Personalised medicine challenges: quality of data. *International Journal of Data Science and Analytics* 2018; 6:251-259.

4. Braga RD, Lucena FN, Leitão-Junior PS. Registro Eletrônico de Saúde em Saúde Interoperável: os desafios do padrão OpenEHR. *J. Health Inform.* 2016; 8(3):103-9.
5. Kitchenham B, Charters S. Guidelines for performing systematic literature reviews in software engineering, *Tech. Rep. EBSE* 2007–001, Keele University and University of Durham, 2007.
6. Kitchenham B. Procedures for performing systematic reviews, Keele University, Department of Computer Science, Keele University, UK, *Tech. rep.,* 2004.
7. Eguzkiza, A., Trigo, J.D., Martínez-Espronceda, M., Serrano, L., Andonegui, J. Formalize clinical processes into electronic health information systems: Modelling a screening service for diabetic retinopathy. *Journal of Biomedical Informatics* 2015; 56: 112-126. Available from: <https://doi.org/10.1016/j.jbi.2015.05.017>.
8. Moner D, Maldonado JA, Robles M. Archetype modeling methodology. *Journal of Biomedical Informatics* 2018; 79:71-81. Available from: <https://doi.org/10.1016/j.jbi.2018.02.003>.
9. Maranhão PA, Bacelar-Silva GM, Gonçalves-Ferreira DN, Calhau C, Vieira-Marques P, Alvarenga M, et al. OpenEHR modeling applied to eating disorders in clinical practice. In: 2018 *IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)*, 2018, 36-41. [Internet]; [cited 2020 Oct 29]. [Accessed 2022 Jun 13]. Available from: <https://ieeexplore.ieee.org/document/8417209>.
10. Maranhão PA, Bacelar-Silva GM, Gonçalves-Ferreira DN, Marques PV, Cruz-Correia RJ. Relevant lifelong nutrition information for the prevention and treatment of childhood obesity - Design and creation of new openEHR archetype set. In: 2017 *IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, 2017, 236-241. [Internet]; [cited 2020 Oct 29]. [Accessed 2022 Jun 13]. Available from: <https://doi.org/10.1109/CBMS.2017.96>.
11. Pahl C, Zare M, Nilashi M, Borges MAF, Weingaertner D, Detschew V, et al. Role of OpenEHR as an open source solution for the regional modelling of patient data in obstetrics. *Journal of Biomedical Informatics* 2015; 55:174-187. Available from: <https://doi.org/10.1016/j.jbi.2015.04.004>.
12. Braun M, Brandt AU, Schulz S, Boeker M. Validating archetypes for the Multiple Sclerosis Functional Composite. *Medical Informatics & Decision Making* 2014; 14: 64. Available from: <https://doi.org/10.1186/1472-6947-14-64>.
13. Santos MR, Bax MP, Kalra D. Dealing with the Archetypes Development Process for a Regional EHR System. *Appl Clin Inf.* 2012; 3:258-275. Available from: <http://dx.doi.org/10.4338/ACI-2011-12-RA-0074>.
14. Papež V, Mouček R. Applying an Archetype-Based Approach to Electroencephalography/Event-Related Potential Experiments in the EEGBase Resource. *Front. Neuroinform.* 2017; 11:24. Available from: <https://doi.org/10.3389/fninf.2017.00024>.
15. Marcos M, Maldonado JA, Martínez-Salvador B, Boscá D, Robles M. Interoperability of clinical decision-support systems and electronic health records using archetypes: A case study in clinical trial eligibility. *Journal of Biomedical Informatics* 2012; 46:676-689. Available from: <https://doi.org/10.1016/j.jbi.2013.05.004>.
16. CKM. *Clinical Knowledge Manager.* c2022 [Internet]; [cited 2020 Oct 29]. [Accessed 2022 Jun 13]. Available from: <https://www.openehr.org/ckm>.