

# Interrogation of Electronic Health Record Data by Biomedical Researcher

Chunhua Weng\*

Department of Biomedical Informatics,  
University of Michigan Medical School,  
USA

**Corresponding author:** Chunhua Weng

 ChunhuaWeng678@gmail.com

Department of Biomedical Informatics,  
University of Michigan Medical School, USA

**Citation:** Weng C (2022) Interrogation of  
Electronic Health Record Data by Biomedical  
Researcher. *Transl Biomed*, Vol. 13 No. 7: 242.

## Abstract

For research purposes such as cohort identification, Phenotyping, Pharmacovigilance, and public health surveillance, electronic health records are an essential data source. Enabling effective and independent EHR data interrogation by end users, such as biomedical researchers, is essential to achieving the promise of EHR data for advancing clinical research. This essay examines contemporary strategies and significant methodological issues for achieving this goal. With the goal of making it easier for end users to query EHR data, we modified a previously published conceptual framework for interactive information retrieval, which identifies three entities: user, channel, and source. We demonstrate that the current development in biomedical informatics is predominantly driven by support for query execution and information modelling, mostly as a result of emphasis on the creation of data infrastructure. Access through self-service query tools, but has overlooked the necessity for user help during laborious and erroneous repetitive query formulation processes. On the other hand, complex theories and techniques for user modelling and question formulation support have been provided in the information science literature.

**Keywords:** Phenotyping; Pharmacovigilance; Biomedical informatics

**Received:** 01-July-2022, Manuscript No IPTB-22- 13011; **Editor assigned:** 04-July-2022, PreQC No. IPTB-22-13011; **Reviewed:** 18-July-2022, QC No. IPTB-22- 13011; **Revised:** 23-July-2022, Manuscript No. IPTB-22- 13011 (R); **Published:** 29-July-2022, DOI: 10.21767/ 2172-0479.100242

## Introduction

The complementarity between the two bodies of literature suggests chances for interdisciplinary idea sharing. Based on this, we propose the routes for future informatics research to enhance our comprehension of user demands and specifications for enabling autonomous EHR data analysis by biological researchers [1]. We propose that our research in supporting effective data access in the life sciences can benefit from cross-disciplinary translational research between biomedical informatics and information science. Patient health records have long been a useful and affordable data source for biomedical research. For instance, analysis of patient medical information was used to create the Goldman multifactorial index of cardiac risks and the Apgar Scale [2]. Many have recognised the rich clinical data made increasingly available by EHRs as a promising data resource for expediting medical knowledge creation and for enabling comparative effectiveness research as a result of the widespread deployment of electronic health records around the world [3]. As a result, the need for reusing EHR data for research has been quickly increasing

among biomedical researchers [4]. The biomedical informatics research community has made it a priority to help biomedical researchers analyse EHR data. However, there are substantial technological and human impediments to this undertaking [5]. Task encounters tremendous technological and human hurdles. Many institutions use intermediary data analysts to retrieve EHR data for biomedical researchers, with varying degrees of help from self-service query tools, because the data currently captured by EHRs is not optimised for secondary uses beyond clinical care or administration centered documentation practises [6]. In massive data networks like the clinical data research networks included in the cornet created by the Patient centred Outcomes Research Institute, the usage of intermediates might not scale [7]. Users of the data in cornet face significant challenges due to factors including the variability of data representations among institutions and the intricate, individualised local data gathering methods that frequently function as black boxes middlemen. Many institutions have to charge clinician scientists for reusing such data obtained during patient care for research in order to manage costs for implicated pricey operations [8]. Self-service

query support, however, is still in its early stages of development and might not be able to support complex data queries. We intend to inform the design of next-generation EHR data interrogation aids that directly assist biomedical researchers to autonomously acquire and reuse this data for clinical and translational research by identifying and assessing existing theories and best practises for EHR data interrogation [9]. This essay provides a survey of the related literature in support of that objective. We outlined current strategies, found research gaps, and suggested research objectives [10]. Despite the fact that this evaluation focuses on EHR data, the learned information may apply to interactive end-user data. The learned skills may be applied to interactive end-user data interrogation for additional reusable health data sources. Utilizing a series of tasks, an information retrieval process responds to information needs [11].

## Discussion

The information retriever's a priori understanding of the information demand, the information retrieval procedure required by data owners, and the complexity of each task required to complete the information retrieval process all influence how complicated the task sequence will be [12]. For describing the information retrieval process or for examining how information technologies empower users during this activity, numerous models have been developed. For instance, the sense-making model and the berry-picking model both centre on how the user iteratively refines his or her information demands based on those conceptualizations. Among all models currently in use, only concentrate on the user's iterative refinement of information requirements based on concepts in the information space. Only one of the current models, created by Bystrom and Jarvelin, clearly specified the user, channel, and source as the three entities that determine the complexity of an information retrieval process. The user entity is concerned with the user's profiles, communication preferences, and data literacy. The topic of the source is data representations for the best data retrieval effectiveness. The channel transforms user information requirements into data representations while hiding the intricacies of the source. As a result, the source is the place where information is stored, and the channel directs the source's efficient navigation. This conceptual framework was used by us to arrange the research on interactive EHR data retrieval [13]. In this essay in this study, we review relevant theories and approaches in the context of EHR data retrieval for secondary use by end users, such as biomedical researchers and clinical scientists, who are not familiar with the data. Since the goal of this work is to improve query formulation for end users, we prioritise activities that support the user and the channel while briefly describing current efforts on the source [14]. In order to further translate this query from contextualised data elements into executable database queries made up of various data types and represented by local terminologies, we co-opted the constructs of user, channel, and source and combined them with the concepts of query formulation and query execution, as shown in the step of query execution. This allowed us to consider all experience as information. Following this strategy, we examined the informatics and computer and information science literature in addition

to the fields of biological and clinical research informatics that were obviously relevant. Additionally, we divided up all included citations according to their emphasis on user, source, and channel to provide quantitative information that helped us see the broad picture and identify evidence gaps [15]. The creation of the search query was based on these 74 articles. We iteratively searched and reviewed the identified articles using our initial search string, incorporating additional search terms as they appeared, and modifying our search string and article inclusion/exclusion criteria as needed based on the relevance of the results. Via manual inspection. We looked at both the literature in information science and biomedical informatics. We felt that the primary journal citation databases for the domains of information science and biomedical sciences, the ACM Digital Library, and Medline, would provide a representative sample for our topic, therefore we restricted our search to these sources. The inclusion and exclusion criteria used to choose the papers for this study were based on the PubMed and ACM databases. The articles that met any of the exclusion criteria were deleted from the pool. The first author created the final search phrase and checked the titles and abstracts of the retrieved articles.

## Conclusion

The included articles were then iteratively examined and annotated by the first author using the conceptual framework created in Section. Each annotation had a summary written by the first author. According to our conceptual framework arranges the articles. As can be seen, in the field of information science rather than in the field of biomedical informatics, more focus was placed on user modelling, human intermediates, and reference interviews. The key ideas from each discipline will be summarised in the parts that follow, along with a comparison and contrast of their ideas from other sources. There are two types of obstacles to accessing task-based data in the life sciences. Human elements, such as a user who does not correctly understand the complexity of the work. System variables traditional systems' technological constraints, such as data fragmentation and heterogeneity. We provided examples of well-known obstacles and the accompanying suggested fixes. Human factors concern the user, whereas system factors concern the source's metadata, or in this case, its absence in her theoretical underpinning of the information seeking process, kuhlthau gives a great aggregation of these user traits, which was well validated by review. The process of finding information has also been studied in the biomedical literature. Mendonça has presented models for the process of finding information in biomedical literature that aim to support users' search strategies by utilising well-structured clinical queries and, accordingly, the expertise of human search experts. Information-seeking behaviours are influenced by user cognitive types. The two orthogonal axes of analysis and description are frequently used to describe these styles. The descriptive cognitive style represents a passive approach to information seeking, whereas the analytic cognitive style captures an active approach that uses conceptual level questioning to address information needs. Self-service inquiry systems fall short in meeting user needs, making human mediators essential in many institutions. These middlemen work with the user to negotiate their demands. The

user and intermediary's relative lack of medical and technological understanding are obstacles to this process. Communication may be more effective if these knowledge gaps between the intermediary and the user are closed. Users frequently give a hazy knowledge and description of their information needs. A uniform format for organising requests could be advantageous for intermediaries since it could lessen the ambiguity of the request and free the intermediary to concentrate on other activities, including query execution.

## Acknowledgement

None

## Conflict of Interest

None

## References

- 1 Garies S, Birtwhistle R, Drummond (2017) Data Resource Profile: National electronic medical record data from the Canadian Primary Care Sentinel Surveillance Network (CPCSSN). *Int J Epidemiol* 46: 1091-1092.
- 2 O Connor BD, Merriman B, Nelson SF (2010) SeqWare query engine: storing and searching sequence data in the cloud. *BMC Bioinformatics* 11: S2.
- 3 Richesson RL, Hammond WE, Nahm M (2013) Electronic health records based phenotyping in next-generation clinical trials: a perspective from the NIH Health Care Systems Collaboratory. *J Am Med Inform Assoc* 20: 226-231.
- 4 Son M, Young R, Harper B, Machion B, Getz K, et al. (2018) Baseline Assessment of the Evolving 2017 clinical Landscape. *Ther Innov Regul Sci* 1-10.
- 5 Ireson CL, Slavova S, Steltenkamp CL, Scutchfield FD (2009) Bridging the care continuum: patient information needs for specialist referrals. *BMC Health Serv Res* 9: 163.
- 6 Gennings C, Olmo JL, Ventura S, Cios KJ, Clore JN, et al. (2014) Impact of HbA1c measurement on hospital readmission rates: analysis of 70,000 clinical database patient records. *BioMed Res Int*.
- 7 Hosseini V (2015) Algorithm and related application for smart wearable devices to reduce the risk of death and brain damage in diabetic coma. *J Diabetes Sci Technol* 10: 802-803.
- 8 Gerritsen MG, Sartorius OE, Vd Veen FM, Meester GT (1993) Data management in multi-center clinical trials and the role of a nationwide computer network. *Proc Annu Symp Compute Appl Med Care* 659-662.
- 9 Fleurence RL, Curtis LH, Califf RM, Platt R, Selby JV, et al. (2014) Launching PCORnet, a national patient-centered clinical research network. *J Am Med Inform Assoc* 21: 578-582.
- 10 Anderson S, Allen P, Peckham S, Goodwin N (2008) asking the right questions: scoping studies in the commissioning of research on the organisation and delivery of health services. *Health Res Policy Syst* 6: 1-12.
- 11 Plaisant C, Lam S, Shneiderman B, Smith MS, Roseman D, et al. (2008) Searching electronic health records for temporal patterns in patient histories: a case study with microsoft amalga. *AMIA Annu Symp Proc* 601-605.
- 12 Jing L, Tian Y (2020) Self-supervised visual feature learning with deep neural networks: A survey. *IEEE Trans Pattern Anal Mach Intell*.
- 13 Alvarez Madrazo S, McTaggart S, Nangle C (2016) Data Resource Profile: The Scottish National Prescribing Information System (PIS). *Int J Epidemiol* 45: 714-715.
- 14 Burn E, You SC, Sena AG (2020) Deep phenotyping of 34,128 adult patients hospitalised with COVID-19 in an international network study. *Nat Commun* 11: 5009.
- 15 Daudt HM, van Mossel C, Scott SJ (2013) Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. *BMC Med Res Methodol* 13: 48.