Proceedings of TENCON 2018 - 2018 IEEE Region 10 Conference (Jeju, Korea, 28-31 October 2018)

Redesign of Clinical Decision Systems to Support Precision Medicine

Muhammad Afzal
Department of Software
Sejong University
Seoul, South Korea
mafzal@sejong.ac.kr

Maqbool Hussain
Department of Software
Sejong University
Seoul, South Korea,
maqbool.hussain@sejong.ac.kr

Sungyoung Lee *
Department of CS and Engineering
Kyung Hee University
Seoul, South Korea
sylee@oslab.ac.kr

Hasan Ali Khattak
Department of Computer Science
Comsats University
Islamabad, Pakistan,
hasan.alikhattak@comsats.edu.pk

***Abstract*—Clinical decision systems (CDSs) showed promising results in different areas including reminder systems, diagnostic, drug dose, prescription, and other pharma related domains. However, many aspects are subject to rethink and redesign in the era of precision medicine. Current clinical decision systems are more focusing to serve one or the other area of clinical care, imparting individualistic characteristics of a single and probably an isolated domain or sub-domain. While precision medicine requires a comprehensive data and knowledge for making precise decisions. The comprehensive knowledge shall contain information about disease sub-types, disease risk, diagnosis, therapy, and prognosis. Building such a comprehensive knowledge and executing queries to get the results from the decision support system may require federation at query level and integration at data level that is accumulated from different databases situated in one or more than one setup. In this paper, we research to provide an initial idea and guide of redesigning the CDS architecture in a way to serve the very need of precision medicine. We describe the limitation of existing CDSs, challenges to address them and propose a solution to address those challenges. This work may lay down a foundation for the architectures of futuristic CDSs in the era of precision medicine.**

Keywords—clinical decision support, precision medicine, big data, knowledge base, integration, federated queries.

I. INTRODUCTION

The purpose of clinical decision support (CDS) is to keep the knowledge and other person-specific information, updated and intelligently inferred for the betterment of health and healthcare [1]. Different stakeholders are the beneficiaries of CDSs including clinicians, staff, patients, and other individuals. Taking a gander at the history, CDSs demonstrated promising outcomes in various areas including reminder systems, diagnostic, drug dose, prescription, and other pharma related domains [2]–[4]. Despite the significance of CDS in medicine, do we have to rethink about the design and the development of CDS's sub-components such as knowledge base and inference engine in the era of precision medicine? The answer to this inquiry is **yes**, because many of the researchers have emphasized in their work [4]–[6] that a CDS with a comprehensive knowledge base is required to be designed to fulfill the very need of precision medicine. Current CDSs are not adequate as they either can serve one or the other area of clinical care where in precision medicine a collective decision will require to collect data from diverse resources. Researchers in their work on informatics research agenda to support precision medicine [6] emphasized on the very need of designing a comprehensive knowledge base (KB) to contain information about disease sub-types, disease risk, diagnosis, therapy, and prognosis. As indicated by the authors, the present KBs and databases are separated from each other, subsequently unfit to offer help for executing and dealing with the federated inquiries. Consequently, they should be patched up to help not just the federated queries rather also stretched out reasoning abilities notwithstanding adaptability and versatility. In the review study [4], the authors featured the isolation issue from a different perspective that is the scientific and clinical data sets are normally located in different databases worldwide as information silos. They need to be linked to bring them in one place in the form to be easily reviewed by physicians and researchers. This raises the need to apply Intelligent algorithms with a use of standardized languages/phenotypes for the decision of choosing the right piece of information to satisfy the need of a specific question. The Clinical Pharmacogenetics Implementation Consortium (CPIC) Informatics Working Group is actively working to develop a set of standard set of CDS functions and EHR-agnostic implementation resources into its guidelines for the successful adoption of Pharmacogenetics into routine clinical care [7]. This group likewise demonstrated the confinement of contemporary CDS issue of tending to single gene by depending on local forms of national guidelines and stressed to move from local to national usage, therefore persuaded them to take a shot at planning and actualizing of modern resources.

II. LIMITATION AND CHALLENGES

Contemporary CDSs are developed in isolation with knowledge bases not sufficient to serve the federated queries composed of an array of information accumulated from different sub-domains. Current architectures are not designed to support functions of accumulating data from different sub-domains and to manage federated queries. Also the structure of knowledge bases is not scalable to accommodate diverse knowledge of different disease sub-types as well as the clinical, genomic, and other wellness information of users.

Moreover, there is an integration issue of clinical data with scientific research data. Making a precise decision requires both aspects to be integrated thus raises to a big challenge of designing methods to support the acquisition of clinical as well as scientific research data and their integration. The knowledge base is needed to be inspired from the integrated data scope rather than only one of them.

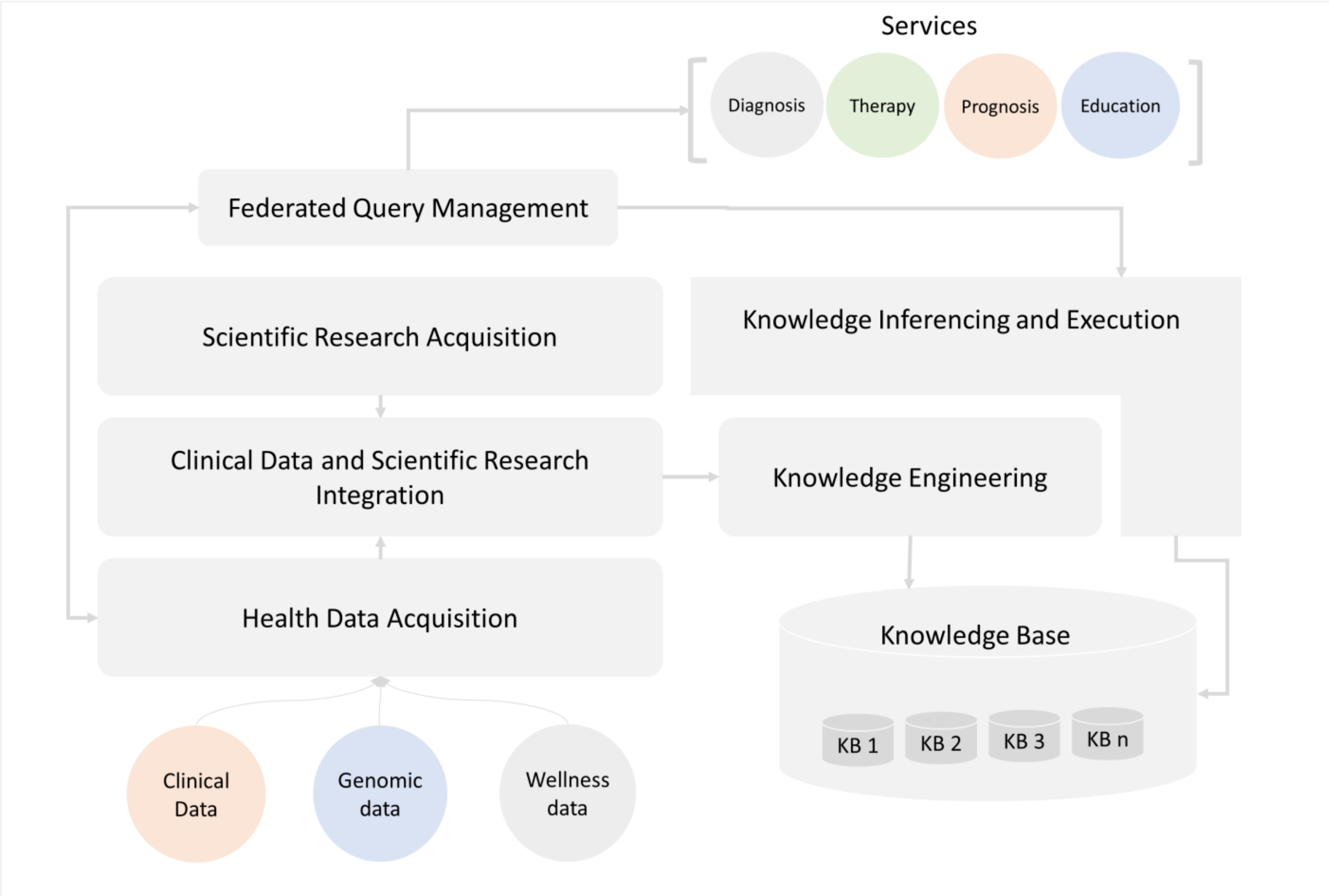

Figure 1: Proposed architecture of clinical decision system

III. PROPOSED SOLUTION

Based on the initial investigation, we come up with an architecture as depicted in Figure 1 to combine different components required by the next generation CDS. The architecture abstractly describes the modules containing the required functions of envisioned CDS's architecture for making precise decisions. We here introduce and explain the details of each of these modules for a better understanding of the proposed idea.

A. Health Data Acquisition

The core requirement of precision medicine is to use diverse information of a user for a precise decision, including clinical data, genomic data, and lifestyle related wellness data. Currently, the electronic health record (EHR) systems are dealing with only clinical data. Although, work has been started to link clinical data with bio-specimen resources for genomic research.

- Clinical data: It is the set of data that is used in current patient care, such as demographics, medical history, physical examinations, risk factors, diagnosis, treatments, and different dates [8], [9].
- Genomic data: It refers to the Genome and DNA data of an organism. The aim of genomic data is to analyze it for determining the function of a specific gene. In the past, it was not feasible to include genomic data in patient care because of two main factors: (i) it was expensive activity and (ii) it was computationally expensive and time taking activity. Today, however, with the advent of advanced technology, genomic data could be used alongside clinical data to precisely diagnose and treat different sub-types of diseases such as cancer and others. This data itself renders a set of challenges to acquire, such as designing of a database to hold voluminous data and controlling the variability of the data and so forth [10].
- Wellness data: it refers to the information about a person's lifestyle data, e.g. physical activity, dietary information, and sleeping patterns. An array of services and tools such as Fitbit [11], Samsung Health [12], Google Fit [13], and Noom Coach [14], are geared towards monitoring user health status of activities, diet, and to some extent sleep and the effects of environmental factors. These tools and services could be used to facilitate the acquisition of user lifestyle data to be utilized together with clinical and genomic data.

The acquisition of clinical data, genomic data, and wellness data involves processing at two levels: one at individual data level, for instance methods to collect genomic data and preprocessing, and second, the integration of all these three types of data at one place.

B. Acquisition and Integration of scientific research data

The health data acquired from local may or may not be sufficient to answer the questions of every kind. There is a need of looking towards scientific literature that includes a list of primary and secondary studies and reviews. Acquiring data from the literature is itself a challenging task. Currently, different databases like MEDLINE, ClinVar [15], and MedGen [16] are accessible through PubMed services. The challenge is the algorithm accuracy and efficiency of identifying relevant and quality studies applied in the context of patient care. In other words, linking the clinical data with scientific data need contextual interpretation and processing of the sub-elements of both the data types in order to answer the clinical questions.

C. Knowledge Base Design

The current knowledge bases of CDSs are not designed to uphold divers kind of data. They need to be redesigned to in a way to accommodate individual knowledge bases for each type of disease as well as the combination. Also, the current top-down and model-driven approaches are not sufficient in the era of precision medicine. A bottom-up and data-driven approach is required to be incorporated alongside the expert-driven approach. The knowledge bases should be indexed and configured for answering the question in due clinical time. An ontological approach associated with data-driven models is one of the potential approach to be consider for designing the overall knowledge base structure.

D. Knowledge Inference and Execution

The capabilities of reasoning algorithms that can work on a model-driven approach to be enhanced to cover up the results of data-driven approaches. Our proposition for such algorithms is to design them in a hierarchical fashion so that the results of one algorithm to be used as an input for another algorithm. It means the overall inference algorithm shall be a combination of sub-algorithms, including both expert-driven and data-driven. Moreover, the execution environment is enriched with methods to accept federated query and run on the knowledge base. Also, the execution environment shall provide the results with logical and semantic interpretations in response to the federated query input in order to fulfill the ultimate user requirements.

E. Federated queries

Federated query is the ability to provide solutions based on information from many different sources. In other words, it connects the end of one query to another query to form one wholesome query. The main purpose of federated queries to meet the need of searching multiple disparate content sources with one query. This allows a user to search multiple data sources at once in real time.

IV. RESULTS

We have worked in the area of designing and developing clinical decision support systems for the last couple of years. All together, we developed Smart CDSS [17] system for facilitating physicians in decision for head and neck cancer treatment. Similarly, we also worked on the acquisition and appraisal of scientific literature using our proposed technique of KnowledgeButton [18]. Using KnowledgeButton approach, we were able to acquire quality articles from PubMed database at the first stage and check for the quality using a machine learning approach at the second stage. The model of quality appraisal was trained on a subset of about 50000 annotated documents acquired from a team of expert physicians. Moreover, we also developed a framework of Mining Minds [19] to curate lifestyle data, manage, and reason over it and discriminate the results to the user. Briefly the outcomes of the existing work are described in Table 1.

Table 1: Authors' existing work outcomes

System/Service	Characteristics	Outcomes
Smart CDSS	• Core CDS system for Head and Neck Caner disease treatment • Expert-driven rule-based system • Interactive knowledge authoring environment • Knowledge Base encoded in HL7 Arden Syntax	• System is demonstrated at multiple venues. • Limited version is implemented in one of the collaborating hospitals. • About 15 times physician's performance enhanced in rule creation. • Publications published [17], [20]–[23].
KnowledgeButton	• A service for scientific research acquisition and appraisal. • Knowledge-based query automated construction • Quality of research contents with machine learning approach.	• Tested on PubMed service • Quality model is trained on a subset of 50000 annotated documents. • Overall, the system has showed 4% better performance than the existing service. • Publications list [18], [24], [25]
Mining Minds	• A health and wellness framework • Big data curation framework • Activity and diet recognitions • Data-driven and expert-driven knowledge acquisition methods • Personalized recommendation services	• Capable of writing 2.2 requests or packets per second in average • Activity recognition with 0.95 F-Score • Overall the system gets 7.5 out of 10 user satisfaction level in the initial evaluations. • Publication list [19], [26]–[28]

A. Experimentation roadmap of the proposed research

In light of our existing work, we plan to perform experiments on real data collected from hospital and the research publication from PubMed database. As noted above, we will integrate the data from local databases with the data extracted from the research publications. For collecting data from local databases, we will follow the data acquisition strategy of Mining Minds [19] with additional functions. To retrieve data from PubMed database, we will use the strategy followed by KnowledgeButton [24], [25] with additional functions. For data preparation, training, and testing, we will utilize the services of data science tools such as Rapid Miner [29]. We will utilize the design of the Smart CDSS [21] knowledge base for the proposed architecture with extension of a comprehensive inference engine to support not only the reasoning over the rules but also the federated queries.

V. Conclusion and Future Work

Precision medicine is the future of medicine and we need to prepare for it. This paper highlighted the limitations of contemporary clinical decision systems, associated challenges and guidelines to revamp and redesign the architecture of the era of precise decision. We proposed a comprehensive architecture of a futuristic decision support system that contains modules for acquisition of diverse data, their integration, management of federated queries, comprehensive knowledge base and the execution environment of reasoning and inferencing. This work provided an initial roadmap of the future journey towards designing a big system for precision medicine. In the future, we will expand our previous work to build new components and enhance the capabilities of existing components to construct a clinical decision system for the era of precision medicine.

Acknowledgment

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2017-0-01629) supervised by the IITP(Institute for Information & communications Technology Promotion)".

References

[1] J. A. Osheroff, J. M. Teich, and B. Middleton, “JAMIA Perspectives on Informatics A Roadmap for National Action on Clinical Decision Support,” *III J Am Med Inf. Assoc*, vol. 14, pp. 141–145, 2007.

[2] D. L. Hunt, R. B. Haynes, S. E. Hanna, and K. Smith, “Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review.,” *JAMA*, vol. 280, no. 15, pp. 1339–46, Oct. 1998.

[3] D. Blum *et al.*, “Computer-Based Clinical Decision Support Systems and Patient-Reported Outcomes: A Systematic Review,” *Patient - Patient-Centered Outcomes Res.*, vol. 8, no. 5, pp. 397–409, Oct. 2015.

[4] C. Castaneda *et al.*, “Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine,” *J. Clin. Bioinforma.*, vol. 5, no. 1, p. 4, Dec. 2015.

[5] I. Maglogiannis Theodosios Goudas, A. Billiris, H. S. Karanikas Datamed, and I. Valavanis Olga Papadodima Georgia Kontogianni Aristotelis Chatziioannou, “Redesigning EHRs and Clinical Decision Support Systems for the Precision Medicine Era.”

[6] J. D. Tenenbaum *et al.*, “An informatics research agenda to support precision medicine: seven key areas,” *J. Am. Med. INFORMATICS Assoc.*, vol. 23, no. 4, pp. 791–795, Jul. 2016.

[7] J. M. Hoffman *et al.*, “Developing knowledge resources to support precision medicine: Principles from the Clinical Pharmacogenetics Implementation Consortium (CPIC),” *J. Am. Med. Informatics Assoc.*, vol. 23, no. 4, pp. 796–801, Jul. 2016.

[8] R. B. Conley *et al.*, “Core Clinical Data Elements for Cancer Genomic Repositories: A Multi-stakeholder Consensus,” *Cell*, vol. 171, no. 5, pp. 982–986, Nov. 2017.

[9] NIH, “Clinical Data Elements | NCI Genomic Data Commons.” [Online]. Available: https://gdc.cancer.gov/clinical-data-elements. [Accessed: 08-Jun-2018].

[10] E. Birney *et al.*, “Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project,” *Nature*, vol. 447, no. 7146, pp. 799–816, Jul. 2007.

[11] “Fitbit,” 2017. [Online]. Available: https://www.fitbit.com/kr/home. [Accessed: 19-May-2017].

[12] Samsung, “Samsung Health,” 2016. [Online]. Available: http://www.samsung.com/global/galaxy/apps/samsung-health/. [Accessed: 09-Mar-2018].

[13] Google Inc., “Google Fit,” 2016. [Online]. Available: https://developers.google.com/fit/. [Accessed: 11-Sep-2018].

[14] Noom Inc., “Noom Coach.” [Online]. Available: https://www.noom.com/. [Accessed: 09-Mar-2018].

[15] M. J. Landrum *et al.*, “ClinVar: public archive of interpretations of clinically relevant variants,” *Nucleic Acids Res.*, vol. 44, no. D1, pp. D862–D868, Jan. 2016.

[16] N. R. NCBI Resource Coordinators, “Database resources of the National Center for Biotechnology Information.,” *Nucleic Acids Res.*, vol. 44, no. D1, pp. D7-19, Jan. 2016.

[17] M. Hussain *et al.*, “Cloud-based Smart CDSS for chronic diseases,” *Health Technol. (Berl).*, vol. 3, no. 2, pp. 153–175, Jun. 2013.

[18] M. Afzal, M. Hussain, W. A. Khan, T. Ali, S. Lee, and B. H. Kang, “KnowledgeButton: An evidence adaptive tool for CDSS and clinical research,” in *2014 IEEE International Symposium on Innovations*

in Intelligent Systems and Applications (INISTA) Proceedings, 2014, pp. 273–280.

[19] O. Banos *et al.*, "The Mining Minds digital health and wellness framework," *Biomed. Eng. Online*, vol. 15, no. S1, p. 76, Jul. 2016.

[20] M. Hussain, W. A. Khan, M. Afzal, and S. Lee, "Smart CDSS for Smart Homes," Springer, Berlin, Heidelberg, 2012, pp. 266–269.

[21] M. Hussain *et al.*, "Data-driven knowledge acquisition, validation, and transformation into HL7 Arden Syntax," *Artif. Intell. Med.*, Oct. 2015.

[22] T. Ali *et al.*, "Multi-model-based interactive authoring environment for creating shareable medical knowledge," *Comput. Methods Programs Biomed.*, vol. 150, pp. 41–72, Oct. 2017.

[23] T. Ali, M. Hussain, W. Ali Khan, M. Afzal, and Sungyoung Lee, "Authoring tool: Acquiring sharable knowledge for Smart CDSS," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 1278–1281.

[24] M. Afzal *et al.*, "Knowledge-Based Query Construction Using the CDSS Knowledge Base for Efficient Evidence Retrieval," *Sensors*, vol. 15, no. 9, pp. 21294–21314, Aug. 2015.

[25] M. Afzal, M. Hussain, R. B. Haynes, and S. Lee, "Context-aware grading of quality evidences for evidence-based decision-making," *Health Informatics J.*, p. 146045821771956, Aug. 2017.

[26] O. Banos *et al.*, "Mining Minds: an innovative framework for personalized health and wellness support," in *Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare*, 2015.

[27] M. Afzal *et al.*, "Personalization of wellness recommendations using contextual interpretation," *Expert Syst. Appl.*, vol. 96, pp. 506–521, Apr. 2018.

[28] M. Amin *et al.*, "On Curating Multimodal Sensory Data for Health and Wellness Platforms," *Sensors*, vol. 16, no. 7, p. 980, Jun. 2016.

[29] RapidMiner, "Lightning Fast Data Science for Teams | RapidMiner©." [Online]. Available: https://rapidminer.com/. [Accessed: 11-Sep-2018].