Name:	Hojun Park
Affiliation:	School of Medicine, Ajou University
Email:	bacojun@ajou.ac.kr
Presentation type (select one):	Poster

AURORA: analytic code managing and distributing tool between researchers and data partners

Hojun Park,BE.[1], JungHyun Byun, B.E. [1], MinSeok Jeon, BE. [1], Sungjae Jung, BE. [1] ,Dukyong Yoon, MD, Ph.D[1], Rae Woong Park, M.D, Ph.D[1]

[1] Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Korea

Abstract

We need to distribute an analytic code to the data partners to perform a study using the distributed research network like Observational Health Data Sciences and Informatics (OHDSI). Although we have successfully performed many analyses by using email to communicate and distribute the analytic codes, however we need a more sophisticated managing tool for them. We develop a managing and distributing tool for the analysis codes between the researcher and data partners and named AUtomated inteR-ORganization Analytics (AURORA). With the AURORA, a researcher can easily send their study protocol and analytic codes to multiple data partners and check the status of the study progress.

Introduction

Distributed research network (DRN) has been emerging as an alternative mean for maintaining confidentiality among analysis using information from multiple institutions. To perform an analysis with many data partner, we need to send the analysis codes to many data partners. Although we have successfully performed many analyses by using email to communicate and distribute the analytic codes, however we need a more sophisticated managing tool for them. We tried to develop a managing and distributing tool for the analysis codes between the researcher and data partners.

Method

Design of the system

We named the system as AUtomated inteR-ORganization Analytics (AURORA). Through the system, a researcher can send a research protocol and analytic codes to multiple data partners via SMTP. Each administrator of the data partner can determine to approve study or not. For the approved study, the analysis is performed in the organization. The decision, comments from the administrator, or the results of the analysis are automatically provided to the researcher. Delivery of messages between researcher and administrator of each data partners is performed by e-mail (e.g., sending code, notifying the change of status) (**Figure 1**).

Construction of the server and the environment

We built a web-based system using JSP, HTML, Apache Tomcat Server. In order to save the information of the researchers as well as the organizations, and record the details and progress log of the study, we constructed a log database with MS-SQL.

Distribution and production of the web interface

Using Java Database Connectivity (JDBC**)** API, we let the system be able to access the database and record the log. The system send all of the events that occur during the process of the study by e-mail to the researchers and the administrators of the organizations, to enable real-time monitoring of the research.

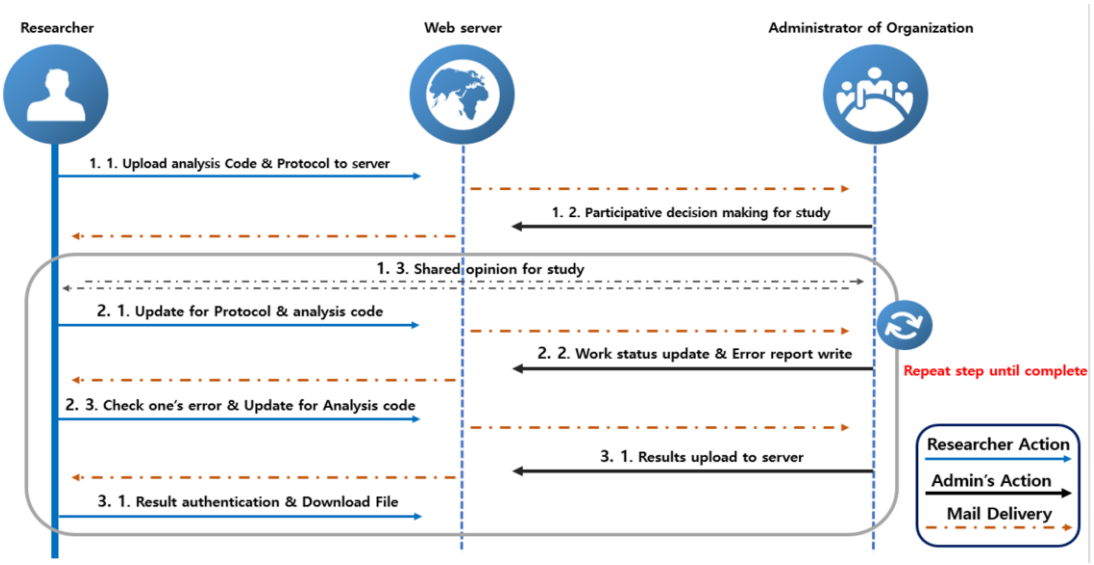

Figure 1. Sequence diagram of AURORA

Result

Figure 2 shows the AURORA user interfaces.

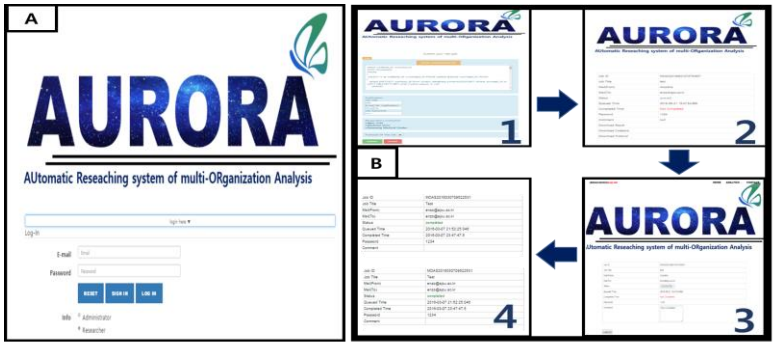

Figure 2. A: Main page of AURORA. B: Overview of research process in AURORA, 1. AURORA send email containing the analysis code generated from Code generator and research protocol written by researcher to administrator 2. After email sending, researcher can check out the status of jobs. 3. Administrator can update the status. (e.g., 'Not completed' to 'Completed') 4. Researcher can confirm updated job status.

Only authorized user can access the system. The system has the following functions: uploading and downloading for protocol, analytic code and result files; mailing for alarm and link; status update for job and error; member administration.

Conclusion

With the AURORA, a researcher can easily send their study protocol and analytic codes to multiple data partners even without information who is in charge of the system. Each institution can confirm the study protocol and perform the analysis by the received analytic code. Eventually, each administrator of data partner can send the results from the analysis to the original researcher through the AURORA. Since only the result of the analysis is requested and transmitted, the personal information is thoroughly protected, and the anonymity is maintained. Although the AURORA is limited in their function as it is in early phase of development, we believe the system will be upgraded soon.

References

1. Forum of Observational Health Data Sciences and Informatics (OHDSI) http://forums.ohdsi.org/
2. Lee at el (2014), Proteome-wide remodeling of protein location and function by stress (ConditiOnal Function Predictor ,CoFP;) Proceedings of the National Academy of Sciences of the United States of America,E3157–E3166, doi: 10.1073/pnas.1318881