MUResearch: Bring Big Data to Effective Communication

Student Leader:

Qiuming Yao, PhD, Department of Computer Science

Faculty Leader/Sponsor:

Dong Xu, PhD, Director, Information Technology Program, Investigator, Bond Life Sciences Center

0. Summary

Effective scientific communication has drawn high attention among research communities these years. We have successfully built a centralized search system bridging research topics and scientists based on various data resources from internal repositories to public databases. The system is able to help MU students learn more knowledge from their interested research topics and at the same time communicate with the right researcher. In this document, we state that although we have finished the basic functionalities of the whole system, we would like to request an extension for further search optimization, large-scale testing and outreach, which are very critical steps for real-world production and achieving potential higher impact.

1. Project Accomplishment

(1) The backend data repository is done.

With the help of MU library and DoIT, we use either remote APIs or local storage to fetch different types of data. The backend data sets are all indexed properly for effective searching with some challenging techniques solved. We have accomplished this data layer as we originally designed.

(2) The searching portal was done.

Key word search is supported from specific data resources or whole data sets (Figure 1). The search procedure is implemented as we originally designed based on the backend indexing and multiple databases. The results are sorted by the initial ranking of the search system we build. We are further optimizing the ranking algorithm. The corresponding MU researchers or contacts for certain research topic are then recommended by our system. The student as our valued user can initialize a chatting with any of the suggested contacts (Figure 2).

Figure 1. MU Research search portal

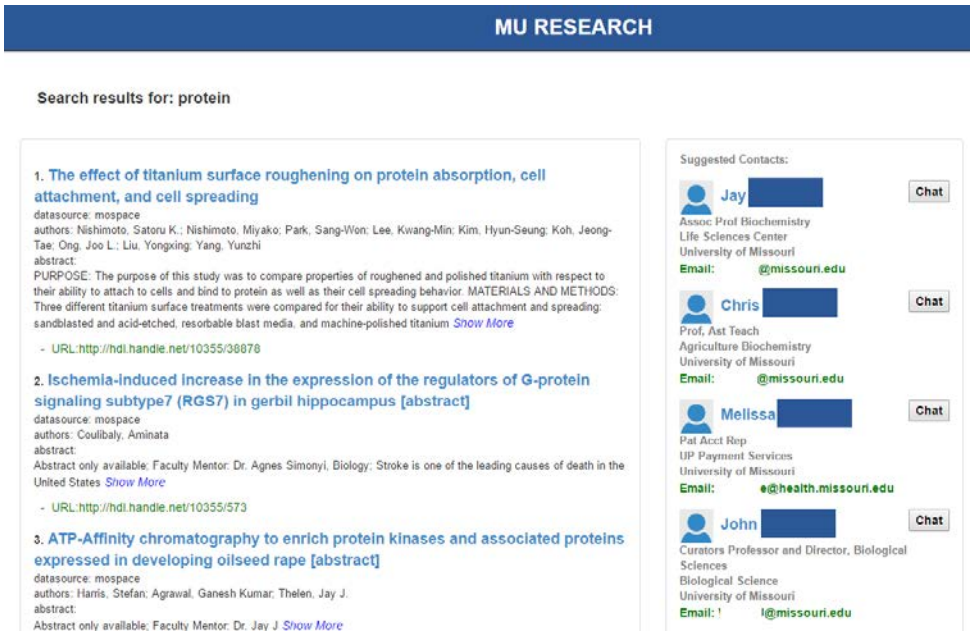

Figure 2. The search results page (some personal information is masked in this figure)

(3) Researcher's profile module is done.

We integrated basic contact information and research publications for our MU researchers (Figure 3). A collaboration network is also built for each person. The black and gold color marked people are all MU contacts. The student can therefore have a general overview before communication about what this person is doing, and whom he is doing with. This will potentially increase the effectiveness for communication. This part is also implemented as we originally designed.

(4) The chat room / offline message board is done.

Student can contact any person (if registered) by initiating a chatting in our system. If the contact is not online, this chat room will also serve as an offline message board. The user can also maintain a contact list in our system. (Figure 4).

We understand that the student can initiate communication using any of the offline methods such as making appointment or calling through phone. However, our system helps them to increase the convenience of communication in this chat room. Besides, the student can keep an effective contact list. This module is viewed as part of our personalized service so the user has to register and signed in when he wants to use this functionality. This module is also implemented as we originally designed.

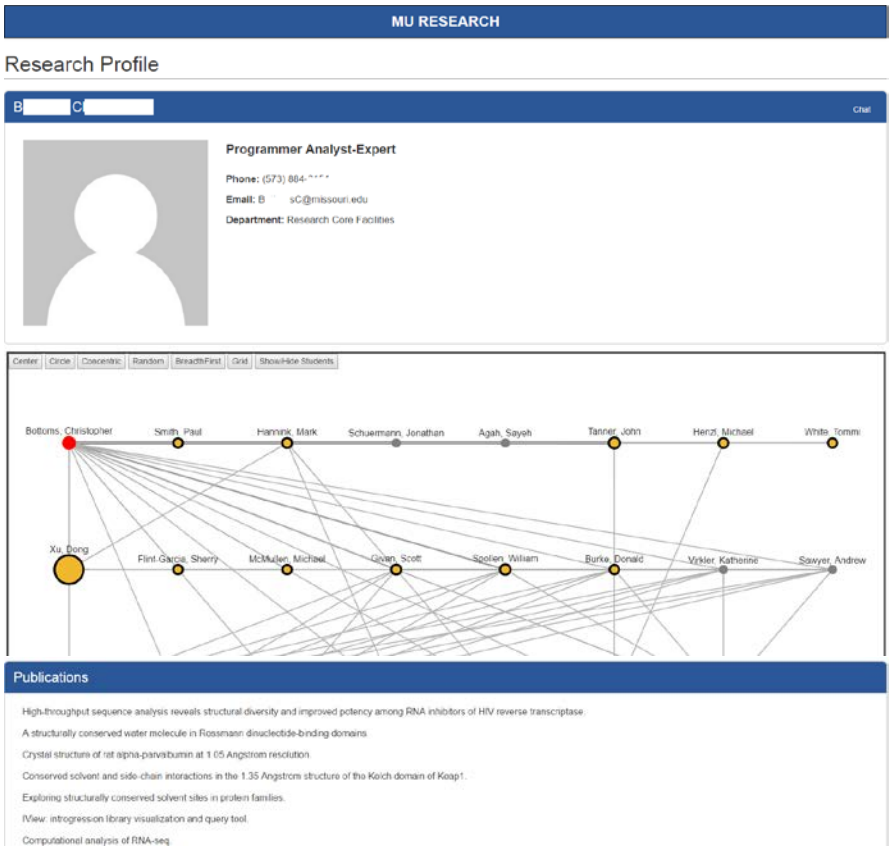

Figure 3. Profile page (some information is masked)

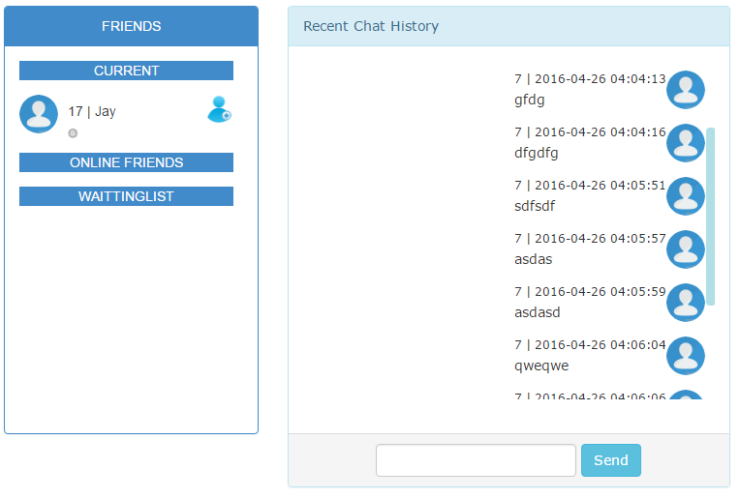

Figure 4. Message / chatting board

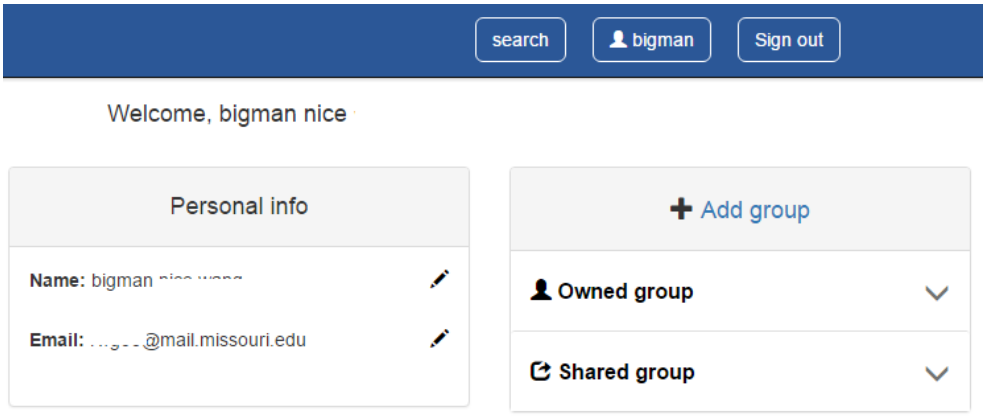

Figure 5. User account center

(5) The user account center is done.

The registered users can change the personal settings or other information in our user center (Figure 5). Personalized services to enhance the searching optimization can be also set here.

(6) Network structures to optimize the searching

We implemented four different types of key words based networks to enhance the user experience in searching. We can re-order the results based on user's preference and habit. The system is able to identify the user's personal interest.

(7) Lab wise group page

We added lab-wise information for professors and principal investigators to advertise their groups. These also helps the students to identify more information for potential research opportunities.

(8) Statistics in searching

We added some statistics for key words based searching. These statistics help us to filter the key words more reasonably in searching processes. We also suggest more correlated key words based on these statistical features.

2. Evaluation Points

(1) The project management is effective.

The project is very well managed to have weekly meetings, and the meeting minutes are updated on Google document so that progresses or problems can be tracked back.

(2) Students involved in this project are from inter-disciplines and are successful.

From Jan 2015 to Jan 2017, eleven students from electrical engineering, computer science, information technology participated in our project. We also help two students to accomplish their master degrees by using part of this project in their final defense. This project has been serving as an important training tool for our students.

(3) We have finished the basic functionalities of the system as we originally designed, mentioned in above sections 1 (1)-(8).

The web service is implemented at: muresearch.missouri.edu

3. Current Expenditure

Four Dell laptops (tax included) :	$5864
Macbook Pro (tax included):	$1469
Nexus 9 tablet (tax included):	$349
iPhone 6:	$749
Seagate Backup Plus Slim Portable Drive, model number STDR2000103 (* 2)	$198
Regular Undergraduate Research (*4*2 semesters):	$8000
Amazon EC2 backup system	$200
Summer Undergraduate Research (2016 Summer)	$4000
Miscellaneous expenditures	$4021

Total	$24,850

4. Further Plan

We have several reasons to continue this project without any cost.

(1) Students are still interested in working with us on this project to gain more experience in data science.

(2) We plan some student activities to promote our product on campus. We will invite people to test. We are interested in helping serve MU better for science.

(3) We are also planning to outreach and pursue more opportunities to serve the BIG Missouri Area such as Kansas. This will definitely increase the impact of the whole project.

(4) This project should be further considered to fulfill business potentials, which also meet the IIF initiative purpose.