

Ascus - Collaboration Finder

Saad Arif

Fourth Year Dissertation - First Deliverable

Supervisor: Ruth Aylett

Second Reader: Patricia Vargas

November 21, 2013

Contents

1	Aims and Objectives	1
2	Literature Review	2
2.1	Organisational need for Expert locator	2
2.2	Why do individuals want to seek experts?	3
2.2.1	Why people do not want to seek experts?	3
2.2.2	Stages of finding Expertise	4
2.2.3	Traditional Approach	5
2.2.4	Contemporary Approach	6
3	Requirements	7
3.0.5	High Priority	7
3.0.6	Medium Priority	8
3.0.7	Low Priority	9
4	Methodology	10
5	Technology Analysis	11
5.1	Database	11
5.1.1	MySQL	11
5.1.2	SQLite	11
5.1.3	Conclusion	11
5.2	Server Side	12
5.2.1	Ruby on Rails (RoR)	12
5.2.2	Active Server Pages (ASP)	12
5.2.3	PHP Hypertext Processor (PHP)	12

5.2.4	Conclusion	12
5.3	Testing	13
5.3.1	PHPUnit	13
5.3.2	SimpleTest	13
5.3.3	Conclusion	13
6	Risks Assessment	14
7	Initial Testing Strategy and Evaluation	15
7.1	System Evaluation	15
7.2	System Testing	16
8	Project Plan	17

Chapter 1

Aims and Objectives

The aim of this project is to develop a web application in order to allow artist and scientist to find collaboration opportunities.

The main objectives of the new system are:

- Develop a web application that will reduce work load on artist and scientist when looking for collaboration opportunities.
- Develop facilities that allow artist and scientists to discuss ideas for collaborations.
- The design of the application must allow both artist and scientists to make effective use of it.

Chapter 2

Literature Review

2.1 Organisational need for Expert locator

Many organisations have identified the need to locate knowledgeable individuals within there organisations. It is important for organisations to effectively use there knowledge in order to enable organisational learning, providing better technical assistance and creating teams to deal with critical situations among other goals(Ackerman, Pipek, and Wulf 2003). Furthermore an organisation may end up "reinventing the wheel" even though a solution had already been made for a similar problem before. Thus it becomes necessary to catalogue skills and expertise of individuals, who knows what, in way that it can later be queried(Becerra-Fernandez 2000).

Examples of organisations developing expert finding systems:

Hewlett Packard (HP), a company in the computers and electronic equipment market developed an Expert-Finder. The goal of the project was to build a network of experts which consisted of a database user profiles. The user profiles gave a summary of the users knowledge and skill(Becerra-Fernandez 2000).

The National Security Agency (NSA) has also attempted to build a system to locate experts with in the organisations. The goal of the his project was similar to HP, identification and cataloguing of knowledge and skills with in the organisation(Becerra-Fernandez 2000).

ASCUS has similar needs to the previous companies named. The organisation requires a tool that allows its members to find collaborators. Thus it needs to catalogue its members expertise, so they can be mined by members who are looking for collaborators.

2.2 Why do individuals want to seek experts?

Yimam-Seid and Kobsa(2003) offer a few reasons as to why individuals may seek experts They state there are two major reasons why individuals seek experts. (a) They need specific information from the expert and (b) They need the expert who to perform some function. The people seeking experts for the first reasons are usually looking to replace or complement other sources of information such as documents. Some scenarios for this reason include seeking information that is not documented, using experts to minimize ones own effort or individuals may prefer interacting with humans rather than documents or computers.

People seeking experts for the second reasons need experts for a continued period of time where the expert will be working for them or with them. Usually the search for this type of reason is performed more carefully than for obtaining information from experts.

ASCUS members looking for collaboration fall into both categories. While members may be looking for active collaborators to participate with in their project, they may also be seeking advice or information into a field they have no expertise in. Thus it is important for the system to facilitate both types of members.

2.2.1 Why people do not want to seek experts?

Allen (1977) noted some reasons as to why information seekers may not want to use their colleagues as a source of information but rather would rather use other information channels. In his study of 19 engineers, he found a higher correlation between frequency of use accessibility than quality of the source of information. He further found that information seekers, found the transaction of information seeking as a costly one. The the cost was perceived in the chance of a response that maybe "ego threatening", a loss in status and seeming incompetent. For these reasons engineers would first look at documentation as a source of information. T

While it has been shown there are issues when seeking experts, they are not of primary concern. ASCUS is organisation for collaboration thus it is assumed that members are looking for interaction. Further it can be assumed that many members have already been part of collaborations and thus the previously stated issues should not apply to these members.

2.2.2 Stages of finding Expertise

McDonald and Ackerman(1998) identify three stages individuals went through to find expertise with in an organisation, Expertise Identification, Expertise Selection and Escalation.

Expertise Identification

Expert identification is defined as “the problem of knowing what information or special skills other individuals have.” It is further noted that expertise identification is difficult problem to solve.It contains many varying factors such as what is expertise , how will it be used within the given context and the problem of handling the change of individuals skills and expertise as time goes on. One solution to such expertise identification is to “consider the types of historical artifacts that are employed by local users as resources and then incorporate use of those within the system.”(McDonald and Ackerman 1998)

Expertise Selection

After determining who has what expertise it is intuitive to then select the most appropriate individual(s) that will solve the problem McDonald and Ackerman. (1998) define expertise selection as “appropriately choosing among people with the required expertise.” Furthermore they observed that expert seekers usually used three expertise selection criteria, “organizational criteria, load on the source and performance.” Expert seekers tried to find experts that were local and when that failed they went to different departments within the organisations. Expert seekers, further more took into account how busy experts were, approaching the least busy first. Finally they firstly approached experts that were better at explaining solutions or had better “attitudes”.

Expertise Escalation

Escalation is the process by which people resolve the failure of the expertise identification or selection mechanism. The expertise seeker may try to identify other experts or pursue other experts that maybe able to solve the problem. This does not necessary involve asking members higher up in the organisation hierarchy, it may involve asking help from less desirable experts or even searching for experts in a different department within the organisation(McDonald and Ackerman 1998).

The system being built should emulate the first two stages and try to avoid the third stage as much as possible. The system ,with regards to expertise identification, should try to match with as much accuracy as possible if a person is an expert in a particular area. In the second stage the system should order the found candidates in an ranked list with the top candidate the most useful to expert seeker. The ideal solution would be to never reach the expertise escalation stage at all but their is no known system that is 100% accurate in expertise identification and selection. Furthermore it is unclear from the literature as to how a system can help the user in the expertise escalation stage.

2.2.3 Traditional Approach

Many organisations contain roles which act as expertise or information locators. In Allen's(1977) discussion he presented a highly linked role within an organisation , which served to bring relevant information to informations seekers. Other researchers have have found similar roles within different types of organisations. Ehrlich and Cash(1994) found a what they called an “information mediator”, who because of his breadth of knowledge and interpretation skills was the go to person in case of any problems. They also noted that the information mediator was a critical part of the organisation. McDonald and Ackerman(1998) found role that they called “expertise concierge”. The expertise concierge has the knowledge of who within the organisations knows what. When a person who is looking for expertise , they ask the expertise concierge about people who maybe be able to help then the concierge will use their knowledge to suggest individuals that match the query.

One way in which we can emulate and automate such roles is by building an expert database. Such an idea works by manually entering expertise data into the database, which can then be queried. The expert database will return a set of individuals that may be of help. Furthermore the expert database may return the individuals in a ranked list such that higher up the list individuals are, more likely they are to be of help. However such a system do have limitations(Yimam-Seid and Kobsa 2003).

1. Developing the databases is a labour intensive and expensive.
2. For the such a system to work , it relies on the experts willingness to spend time initially providing information about their expertise.
3. Due to a continuous change in peoples expertise it is hard to keep the databases up to

date.

4. There is usually a disconnect between expertise description entered into the database and the expert related query. The expertise description are usually general and incomplete while the expert related queries are very specific.

2.2.4 Contemporary Approach

In order to combat the limitations of the traditional approach the expert finding process is further automated. This is done by automatically obtaining information on experts from many different sources and not relying solely on human sources(Yimam-Seid and Kobsa 2003). Using this method the experts do not need to update their profile manually and thus it is always up to date.

However this approach has many difficulties. Firstly it is can be difficult to find sources on which to judge the skill level of an expert in a particular area. An example source for scientist could be articles published but for artists this much more difficult as their work is more difficult to analyse. A common interdisciplinary source that can be used to within an organisation, are email communications. Emails can demonstrate expertise, as queries can be answered by the expert and communication patterns can help determine who has what knowledge with in the organisation(Campbell et al. 2003). However using such a source requires handling privacy concerns. Even with a source of information it still difficult to determine the weight of the source. for example if a scientist co authors an article does that make the scientist an expert in that area? did both authors equally contribute to the article?. Such questions are difficult to answer without further information and if sources are not properly weighted then it is more likely for incorrect experts to be identified.

Chapter 3

Requirements

The requirements of the project are prioritised in order to manage the project effectively. The most important requirements are given the tag High. These requirements must be completed before the end of the project. The next priority tag is Medium, any priorities with the Medium tag will be started only after the High tag requirements have been finished. While the Medium tag priorities are not essential, it is recommended that at least some of them be part of the system if not all. Finally the least important requirements will be given the tag Low. These requirements are not needed for the main functionality and only provide extra features for convenience .

3.0.5 High Priority

Search for Collaborators

This is has been identified as the most important requirement. The users shall be able to search for collaborators in the database. The criteria for search shall be:

Area of work The user must be able to search for collaborators in the database who are Artist or Scientists.

Location Search for collaborators in a particular city.

Distance from user Search for collaborators with in a certain distance from the city the user is in. The user will not need to register to use this functionality.

Register and Log in

User shall be able to register as member and customize their profile. To register the user must provide a first and last name as well choose a user name and password. Once the user has registered they can then customize their profile. The profile shall contain the following items:

- first and last name (Required)
- age (optional)
- email (optional)
- Phone/Mobile number (optional)
- Biography (optional)
- Artistic and/or Scientific interest (optional)
- Profile Picture (optional)
- Extra pictures - in order to demonstrate some previous work (optional)
- External Links - in order to show previous work done, such as blogs and journal articles (optional)

The username and password provided will be used to login, which will allow the members to edit their profile. Members can only edit their own profiles. While only administrators can remove members, which in turn will also remove the profile.

High usability

This has been given a high priority as an application with high usability and consistent design will attract more users. Even with a functional application in order for the application to be successful and help collaboration, the user must find it intuitive and easy to use.

3.0.6 Medium Priority

Map visualization

This feature allows users to visually see the location of members in a map. The map will show user dots representing members on map. The dots will become bigger as more members are in

a particular area. There should be a zoom functionality, as the user zooms in the more detailed density of the members becomes clear. This is likely to be the most difficult requirement and is likely to be scoped down in order to be manageable within the project deadlines.

Projects

This feature will allow members to create projects, that users can search for. The project should have a Title, description, images to demonstrate ideas, goals and what skills are required. The original member to start the project can edit the project and delete the project. There will be a comment section for each project allowing members to give feedback and ask questions. Administrator can only delete the project.

3.0.7 Low Priority

Forum

User will have the standard forum functionality available. Users will be able to create threads and post comments.

Chapter 4

Methodology

Due the user centric nature of the of project, it is best to use iterative development and avoid the traditional waterfall methodology. The waterfall methodology is not suited for this project as it does not involve users as much as required for high user interaction projects.

Using iterative development, this project will focus on cycles of prototyping small parts of the project, implementing the prototype and finally getting feedback. The original prototype will be continually improved as more feedback is obtained.

By using iterative development , the unpredictable nature of working users can be managed. Due to continually prototyping and implementation small parts, the user will still have some functionality even if the project cannot be finished. If the waterfall methodology was used and significant difficulties were encountered, it is likely the project would not finish and have no functionality implemented at all.

Any issues regarding requirements or user interface design can be solved early on in the development process. Using the waterfall method it would too late to go back and make changes to the design, if there were any misunderstandings in the requirements. In conclusion many of the problems in encountered when developing interactive software can be dealt with using iterative development.

Chapter 5

Technology Analysis

5.1 Database

The database will be used to store user details such as log in and password as well as the users information displayed on the users profile.

5.1.1 MySQL

MySQL server is open source relational database with extensive customisation options for performance tuning. It also offers capabilities to set user accounts and permissions. The author has experience in working with MYSQL.

5.1.2 SQLite

SQLite is a relational database in the public domain. Unlike the other databases servers mentioned, this software does not operate in a client-server model, the database is contained within a file. Thus the SQLite databases can be embedded into an application while other databases works by having a separate database server with which to communicate.

5.1.3 Conclusion

Both database have similar features and both perform slightly different tasks. MySQL offers more control, while SQLite offers easy of set up. Due to the short time in which this project will take place, the MySQL database is chosen as the author has previous experience in it, allowing for immediate productivity.

5.2 Server Side

5.2.1 Ruby on Rails (RoR)

Ruby on Rails is a framework built on ruby programming languages and uses a MVC architecture. The RoR philosophy is convention over configuration. By following the set of conventions the developer can be more productive and the framework "just works". Thus it is possible to be very productive but it does have a steep learning curve. Further it is relatively new, being created in 2005, it lacks the mature documentation of '.

5.2.2 Active Server Pages (ASP)

Active Server Pages (ASP) is a server-side environment from Microsoft, that you can use to create and run dynamic web applications. In order to use ASP a the server must run a Microsoft I.I.S. (Internet Information Server) operating system. Further more the database must be Microsoft SQL Server(MS-SQL). Though it is only limited to the windows platform, without considering third party tools allow for cross platform functionality, it does integrate well with other windows technologies.

5.2.3 PHP Hypertext Processor (PHP)

PHP is a general purpose language suited for web development. It provides many features suited for web development and can be used as scripting language. PHP is one of the most widely used programming languages for the web, thus has mature documentation as well as plethora of tools to support its development. It also considered to the have lowest barrier to entry as it is very easy to set up. Furthermore the author has previous experience in PHP.

5.2.4 Conclusion

PHP will be used to perform server side scripting for the website system, as it is mature and well documented. Furthermore the previous experience allows the author to be productive very quickly in comparison to the other technologies.

5.3 Testing

In order to perform thorough testing it is important to automate the testing by using a testing framework. By automating it becomes easy to perform tests after every change to the application and any issues will be immediately reported. From this, it saves the developer from wasting time manually running algorithms on data. Furthermore Automated tests can easily execute many of different types of test cases during every test run providing a wide coverage which is difficult to do manual tests.

For PHP two popular unit testing frameworks will be considered, PHPUnit and SimpleTest. Both testing frameworks provide similar features such as creating unit tests, running them automatically and showing failed tests and passed test. Therefore evaluating the frameworks will focus on other aspects such as documentations, integration with other tools.

5.3.1 PHPUnit

PHPUnit is the standard unit testing framework that is included in many PHP frameworks such as Zend Framework, Cake PHP. PHPUnit is well maintained as updates have been done regularly. PHPUnit seems to have larger user base as well as more online tutorials. PHPUnit is integrated into many PHP IDEs such as Eclipse, Netbeans, Zend Studio, PhpStorm. PHPUnit requires a terminal in order to run tests.

5.3.2 SimpleTest

SimpleTest is another unit testing framework with the focus on simplicity. SimpleTest is not well maintained with last update over a year ago. SimpleTest does offer the ability to run tests in a web browser thus not needing a terminal, which can be useful when working on a remote web server. Furthermore adding SimpleTest to a project is easy as simply using the include command in a PHP script. While many tutorials are available for SimpleTest, many of them are over a year old. Eclipse IDE offers a SimpleTest plugin.

5.3.3 Conclusion

While SimpleTest and PHPUnit offer similar features in terms of creating tests and automatically running them, PHPUnit is better maintained thus easier to find documentations and help with issues should any arise.

Chapter 6

Risks Assessment

This section will give an an overview of the potential risks to project development.

Risk	Probability	Severity	Mitigation Strategy
Browser compatibility issues	High	Moderate	Constantly test browser compatibility through out the development cycle.Leverage automatic test frameworks in order to supplement manual testing.
Fall behind on plan	Moderate	Moderate	Give small leeway time for each part of the application. If significantly behind schedule then only complete vital parts of the application.
Person of interest is unavailable	Moderate	Moderate	Another client is kept up to date with the main development of the application.
Users find the system unusable	Moderate	High	Follow user centred design approach. Maintain regular meetings with the client. User testing must be carried out extensively and feedback should be integrated into the next iteration.
Low quality feedback	Low	Moderate	Pilot all methods used to gain user feedback. If low quality feedback is given reassess previous methods for weaknesses and integrated any findings into a new- questionnaire.

Table 6.1: Risk Assessment

Chapter 7

Initial Testing Strategy and Evaluation

This chapter describes the plan for testing and evaluating the system.

7.1 System Evaluation

In order to meet the high usability requirement, it is important to involve users when developing the system and user interface. If the system is not usable then the users will not use the system, rendering the project pointless.

In order to evaluate the system, prototypes will be used. The system evaluation will be conducted by allowing expected system users to interact with the prototypes. Afterwards the user will be asked to fill out an questionnaire, which will include open and closed questions. The closed questions will be used focus on a particular aspect of the design, while open questions will be used to get in depth feedback.

In order to gain even more feedback a semi structured interview will be used. It will be decided in advance which themes and areas are needed to be explored during the interview. Using a semi structured interview allows for more freedom to discuss various areas and explore areas based on the users answers. The interviews will be recorded, as writing the users answers will likely not capture exactly what the user said. It also frees the interviewer to focus on the users answers rather than with writing them down.

Before any questionnaire and interviews are conducted the user will be given an consent form detailing what will the questionnaire/interview involve, all information gathered will be

anonymised, the user can withdraw their participation as they see fit and any recordings will only be kept as long as necessary for the project.

7.2 System Testing

Chapter 8

Project Plan

A high level plan below describes the key tasks to be achieved.

Due Date	Version	Task
15/10/2013	R1	Gather initial requirements.
01/11/2013	N/A	Conduct Literature review and research into relevant technologies
22/11/2013	Final	Hand in final version of the first deliverable
03/12/2013	R2	Create Questionnaire to gather further requirements. Pilot the questionnaire and improve it if needed.
31/12/2013	WD1	Create Website design prototype.
31/12/2013	DD1	Create initial database design and implementation.
31/12/2013	IMP1	Create an initial implementation of high priority requirements
10/01/2014	N/A	Gather feedback on initial prototype and implementations. Improve the current system according to feedback.
17/01/2014	WD2	Create website design prototype for high and medium priority requirements.
27/01/2014	DD2	Update database design in order to facilitate high and medium priorities.
07/02/2014	IMP2	Implement medium priority requirements and integrate with previous requirements.
17/02/2014	N/A	Gather feedback and improve system based on feedback.
27/02/2014	Final	Create final website design prototype for high,medium and low priority requirements.
07/03/2014	Final	Final update to database design in order to facilitate high,medium and low priorities.
17/03/2014	Final	Implement low priority requirements and integrate with previous requirements.
01/04/2014	Final	Test the systems, fix any bugs and release the system.

Table 8.1: Project Plan

Bibliography

- [1] Mark S Ackerman, Volkmar Pipek, and Volker Wulf. *Sharing expertise: Beyond knowledge management*. MIT press, 2003.
- [2] Thomas J Allen. “Managing the flow of technology: Technology transfer and the dissemination of technological information within the R & D organization(Book)”. In: *Research supported by the National Science Foundation*. Cambridge, Mass., MIT Press, 1977. 329 p (1977).
- [3] Becerra-Fernandez. “The role of artificial intelligence technologies in the implementation of People-Finder knowledge management systems”. In: *Knowledge-Based Systems* 13.5 (2000), pp. 315–320. ISSN: 0950-7051. DOI: [http://dx.doi.org/10.1016/S0950-7051\(00\)00091-5](http://dx.doi.org/10.1016/S0950-7051(00)00091-5). URL: <http://www.sciencedirect.com/science/article/pii/S0950705100000915>.
- [4] Christopher S. Campbell et al. “Expertise Identification Using Email Communications”. In: *Proceedings of the Twelfth International Conference on Information and Knowledge Management*. (New York, NY, USA). CIKM '03. New Orleans, LA, USA: ACM, 2003, pp. 528–531. ISBN: 1-58113-723-0. DOI: 10.1145/956863.956965. URL: <http://doi.acm.org/10.1145/956863.956965>.
- [5] Kate Ehrlich and Debra Cash. “Turning information into knowledge: Information finding as a collaborative activity”. In: *Proceedings of digital libraries*. Vol. 94. 1994, pp. 119–125.
- [6] *Just talk to me: a field study of expertise location*. (New York, NY, USA). CSCW '98. Seattle, Washington, USA: ACM, 1998, pp. 315–324. ISBN: 1-58113-009-0. DOI: 10.1145/289444.289506. URL: <http://doi.acm.org/10.1145/289444.289506>.
- [7] Dawit Yimam-Seid and Alfred Kobsa. “Expert-Finding Systems for Organizations: Problem and Domain Analysis and the DEMOIR Approach”. In: *Journal of Organizational Computing and Electronic Commerce* 13.1 (2003), pp. 1–24. DOI: 10.1207/S15327744JOCE1301_

1. eprint: http://www.tandfonline.com/doi/pdf/10.1207/S15327744JOCE1301_1. URL:
http://www.tandfonline.com/doi/abs/10.1207/S15327744JOCE1301_1.