Visualizing Expert Network via User-Interface Markup Language and ActionScript

Abstract—**As with many other countries, Thailand has a large number of experts in various domains working in many different public and private organizations. Searching for experts in a specific domain of knowledge is difficult, even by using a search engine. The reason is due to the unstructured characteristics of Web pages and also the difference of terms used to represent the area of expertise. In this paper, we describe our ongoing project to build an expert finder system for identifying experts in different research areas in Thailand. The social network and information visualization techniques are adopted to provide the solution to the problem. We designed a framework for visualizing online expert network and applied visualization techniques to support exploration of different areas of expertise. The proposed framework is based on the User-Interface Markup Language (MXML) and ActionScript to provide powerful interaction between users and the system. The system which is implemented under the proposed framework can support more efficient and effective expertise information searching and browsing.**

Keywords- Expert finder, social network, information visualization, ActionScript, MXML

I. INTRODUCTION

Among many practical and domain-specific tasks, expertise retrieval (ER) has recently gained increasing attention in the information retrieval and knowledge management communities. ER can be broadly classified into two tasks: expert finding and expert profiling. The expert finding task aims to identify a list of people who carry some certain knowledge specified by the input query [2, 7]. The expert profiling, on the other hand, focuses on identifying the area of expertise associated with a given person [3]. To construct an expert profile, two types of information which can used to describe an expert are *topical* and *social* information. The topical information represents domain and degree of knowledge in which an expert possesses. The social information measures an association aspect among experts such as research project collaboration, publication co-authoring and program committee assignment.

Due to the necessity of Thailand in S&T development, governments in the past and at present have been supporting and providing a lot of findings in scholarships and research projects. However, there are many departments and organizations such as the Office of Civil Service Commission, Ministry of Science and Technology, various educational institutions, who have authorities in granting these scholarships and research findings. As a result, searching for research related information in Thailand has been a difficult task. Therefore, we proposed a project is to design and develop Thailand's Research Information Portal and Search Engine (ThaiReSearch) [12]. The main goal is to integrate research-related information in Thailand into one-stop information service. One of the ongoing tasks is to design and develop a framework for building an expert finder system. The expert finder analyzes expertise of both individual researcher and single research organization.

To make the system easy to use and understand, the concepts and techniques of social networking and information visualization are adopted. Social network analysis is the construction of social relationships in terms of nodes and links. Nodes are the individual persons within the networks, and links are the relationships between persons. Social networks can be realized in many levels, from families up to the level of nations, and play a critical role in determining the way problems are solved, organizations are run, and the degree to which individuals succeed in achieving their goals.

Using a graph representation, a node can be used to display entities such as researchers and organizations, and a link can represent the similarity in expertise between researchers and organizations. To provide the users with user-friendly features, we applied the social network and information he proposed For example, R&D managers can use the system to search for a group of experts in each research area. Funding agency can use such system to allocate project funding more effectively by not giving the funding to any overlapping research projects. Finally, anyone can search for expert who can help answer some particular questions.

We are interested in using the visualization tools without having to pay any licensing fee. Therefore, our system is designed and implemented based on the open source software concept. We first performed some thorough survey of the open source information visualization libraries and animation tools. We found Flare [10], a visualization toolkit, to be a suitable choice for implementing our visualization framework. Flare makes it easier to build interactive data visualizations. Flare provides interactive visualization and animation tools on the web for ActionScript and the Adobe Flash Player. Flare is one of the most widely used information visualization library. To extend the library to support other languages, developers must

add language specific analysis package into the library. Integrating and extending Flare library is very simple and easy, since its creator has designed the framework via the object-oriented programming concept of ActionScript. This object-oriented feature is another reason that makes flare a very attractive choice for implementing our system framework.

As a case study, we construct a Thailand's researcher network based on the publication database of Science Citation Index (SCI). From the SCI database, we prepare the publication collection by querying the records which has the word “Thailand” as the affiliation. The resulting records of approximately 23,000 are used to build the network of researchers in Thailand. We consider two types of relationships: topical and social. Topical relationship is based on the similarity of research topics represented by the controlled keywords among researchers. Social relationship is based on the co-authoring degree among the researchers. The proposed visualization framework is then applied to construct the researcher network from this extracted database.

The remainder of this paper is organized as follows. In the next section, we review some related work in visualization technology. In Section 3, we give a detailed discussion on the software tools used for implementing the system. In Section 4, we proposed the framework design of visualizing expert. Section 5 gives the conclusion and future works.

II. Related Works

In recent years, there are many related research work in the field of visualization on the web for constructing social networks. For example, Takama et. al. [6] proposed a method for visualizing and examining news distribution in Blog space by using resources which are extracted from a Blog community. This Blog space constantly reacts to news articles about specific topic. The combinations of both resources are online news sites and Blog. Three kinds of objects to be visualized are news articles, Blog entries, and Blog sites which shows relationship between the objects. The social network is visualized on a 2D space with interactive information visualization system. This allows users to generate maps from multiple viewpoints.

Heer and Boyd [5] designed and implemented a visualization system for playful end user exploration and navigation of large-scale online social network. Their proposed framework was applied to map the web of Friendster relationships called Vizster. They are applying the results of ethnographic finding to build visualization. Viszter provides a social network as images for each person link via connectors to show relationship in large graph structures. It support connectivity highlighting, search features, X-ray mode. For the visualization, the social network is built by using Prefuse (http://prefuse.org/), a user interface toolkit for interactive information visualization written in Java using the Java2D graphics library.

Our work differs from the above paper in that: we extract researcher’s expertise/experience from abstracts; we use Flare toolkit that provides visualization and animation; we use ActionScript and MXML language to provide the user interface which is more attractive than the Java2D graphics library. Furthermore, our proposed framework supports Rich Internet Applications (RIAs) which is a web application running inside a web browser. However, an RIA looks and feels more like a desktop application and improves user satisfaction.

III. Background and Related Technology

In this section, we describe several software tools which are related to information visualization and also introduce the process of constructing a Flex application.

A. Flex

Flex SDK [11] is a free open source framework for building and maintaining expressive web applications, highly interactive that deploy consistently on all major browsers, desktops and operating systems. Then there is the basic framework, which is a set of predefined class libraries, and applications services that include layout and display systems to manage the user interface (UI), data binding, and style parameters. Flex helps to create buttons, check boxes and radio buttons, data grids graph/chart rich text editors and component supporting drag and drop. For example, a data visualization application built in Flex can pull data from multiple back-end sources and display it visually. Flex applications can be developed either by using freely available command-line tools that are distributed as part of the Flex SDK or by using Adobe’s commercially licensed Flex Builder an IDE built upon the open source Eclipse platform.

B. ActionScript

ActionScript [8] is the object-oriented programming (OOP) language based on ECMAScript. ActionScript is used primarily for the development of websites and software using the Adobe Flash Player platform (in the form of SWF files embedded into Web pages). But ActionScript is also used by Flex. ActionScript enables efficient programming of Flash applications for everything from simple animations to complex, data-rich, interactive application interfaces.

C. MXML

MXML is an XML-based user interface markup language first introduced by Macromedia in March 2004. Adobe, which acquired Macromedia in December 2005, kept the name, which some developers believe stands for “Multimedia eXtensible Markup Language.” Application developers use MXML in combination with ActionScript to develop rich Internet applications. MXML is used mainly to declaratively lay out the interface of applications, and can be used to implement business logic and internet application behaviors. MXML is often used with Flex Server, which dynamically compiles it into standard binary SWF files. However, Adobe's Flex Builder IDE and free Flex SDK can also generate SWF files from MXML for using without Flex Server.

D. Flare

Flare is an ActionScript library for creating interactive visualizations for the Web that runs into the Adobe Flash player. For example, flare can be used to build basic charts,

complex animations, network diagrams, treemaps and more. Flare is written in the ActionScript 3 programming language. Flare applications can be built using the free Adobe Flex SDK or Adobe's Flex Builder IDE. Flare is based on Prefuse, a full-featured visualization toolkit written in Java. Flare is open source software licensed under the terms of the BSD license, and is freely used for both of commercial and non-commercial purposes. Even better, flare features a modular design that lets developers create customized visualization techniques without having to reinvent the wheel.

E. The process of constructing a Flex application

Flex is an XML-based language that is compiled into Flash applications. The process to build a Flex application can be split up into four steps.

- *Design and configuration:* The first step in designing an application is to set the scope and functionalities. This step includes the selection of various tools and programming languages including platforms.
- *Build:* Building a Flex application consists of compiling and debugging. Might be building the application by without using the debuggers.
- *Compiling:* The compiling step creates the common binary SWF file that the flash player will use to execute application. The binary file is created from source files including MXML, ActionScript, SWC (Shockwave Component) and media files such as images, audio, video files and XML data files. The process is similar to the way Java files are compiled into a binary *bytecode* (the equivalent to the SWF file) that is then executed by the java Runtime Environment (instead of the Flash player)
- *Deploy:* To make the application accessible to end users. We must publish the SWF file on a web server. The SWF file is integrated into a web page. The web page may be simple HTML or a dynamically generated web page like a PHP, JSP, ASP or ColdFusion page. The ended users will access the SWF file by having web browse request the SWF file and launch the Flash player to process.

The SWF file is the end product of the compilation process and once deployed will be available for end users to request from a web server and execute within their Flash Player. [4] You can see the process graphically in Figure 1.

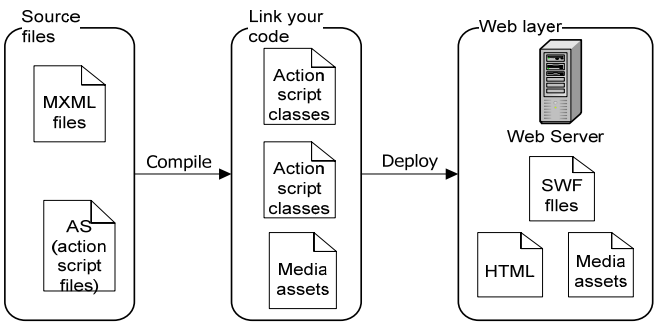

Figure 1. The Flex project creating a single SWF

IV. SYETEM FRAMEWORK DESIGN

We propose a design framework of visualizing expert network. In Figure 2 illustrates the overall system framework design. The system consists of two main functions: *AMF Gateway* and *visualizing researcher network on the web*. For the first function, there are three modules which support to send data back and forth between server and client.

- *AMFPHP:* This module contains *database class* for connection with the databases and returns the results to Flex application and provide Remote Procedure Call (RPC) functionality allowing Flex application to call server-side functions. AMFPHP [1, 9] was designed to allow data communication between Flash Player and PHP classes.
- *Serializer:* In this module the AMF gateway will convert text into a binary format that process takes place on the server side.
- *Deserializer:* In this module the AMF gateway will convert binary data to text that process takes place on the server side.

The outputs from the first task are binary format of Action Script (AS2, AS3) native types and data objects. The second function consists of four files which support the visualizing researcher network.

- *MXML file:* This file is used mainly to declaratively lay out the user interface of applications.
- *AS3 file:* This file is used primarily for the development of websites and help for creating interactive data visualizations.
- *Flare library:* This file is an ActionScript library for creating visualizations expert network that run in the Adobe Flash Player.
- *AS3FlexDB:* This file is an open source library that allows Adobe Flex applications to connect to a MySQL.

The source code in this framework is written in PHP, MXML and ActionScript 3.0 including with Flare library based on object-oriented programming. End users can enter keyword to search, is access to the SWF file on a web server by having web browser. The SWF file including ActionScript bytecode, MXML file, and embedded assets (graphics, sound, video, and fonts). Then SWF file sending binary ActionScript (AMF3) by using RemoteObject class via Remote Procedure Call (RPC) that functionality allowing SWF applications to call server-side functions. Deserializer will transforms the data types between binary ActionScript into PHP data types. The generated data object is then passed to be executed by the SQL. The results are returned from database to AMF gateway. Serializer will transform the data object into SWF file, then process on the server and return the search results back to the web browser. After web browser will present the expert network with Flash Player.

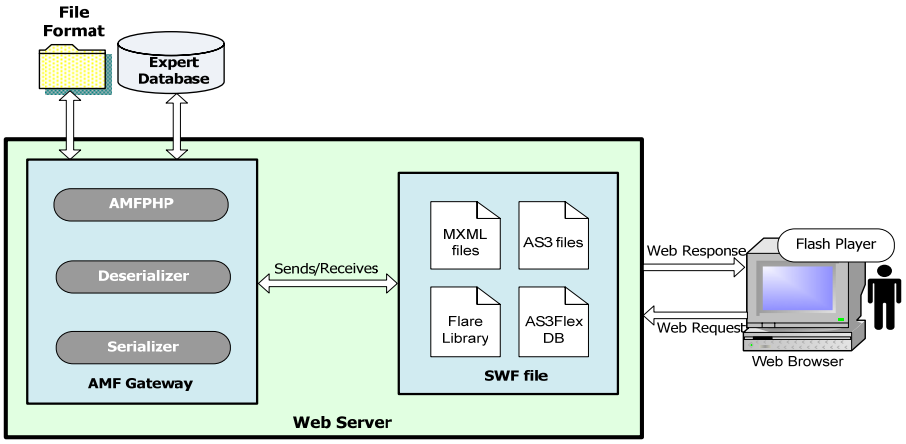

Figure 2. Framework of visualizing expert network

V. A CASE STUDY: CONSTRUCTING A THAILAND'S RESEARCHER NETWORK FROM BIBLIOGRAPHICAL DATABASE

We developed a prototype for visualizing online expert network. The data set we collected the publications/documents was results from Science Citation Index [13] by manual on covered almost all the researchers in all majors in Thailand such as computer, biology, and chemical. We believe this data set contains high quality and comprehensive information to reveal the expertise of Thailand. The data set contained about 2000 researchers which have publications between years 2001 to 2008. From this collection of approximately 23,000 publications, we extracted four related fields including authors (denoted by AU), controlled terms (denoted by ID), keywords (denoted by DE) and subject category (denoted by SC).

To build a researcher network, we consider two types of relationships: direct and indirect. The direct (or social) relationship is defined as the co-authoring degree between one researcher to others. The co-authoring degree between two researchers, co-authoring (A, B), can be calculated based on the co-occurrence frequency between A and B found in the field AU of 23,000 retrieved records. The indirect (or topical) relationship is defined when two researchers have publications under the same topics. The topical degree between two researchers, topical (A, B), can be calculated based on the similarity measure between two sets of extracted keywords, keyword(A) and keyword(B), representing researcher A and B, respectively. In this paper, we use the Jaccard coefficient for the similarity calculation.

Once the extracted information is preprocessed and relationship scores are calculated, we store the information including relationship values, keywords, and profiles data in a backing MySQL database by called expert database. We adopt visualization technique to drawing graph of expert and expertise field based on expert database. We design the expert database structure that consists of following four tables.

Expert_information Table consists of 4 different fields: ID, Name, Picture, and Affiliation. Example of our expert_information table structure is shown in Table I.

TABLE I. EXPERT_INFORMATION

Field	*Type*	*Comments*
ID	Integer	The unique identifying, auto increment to author.
Name	Varchar	The full name's author.
Picture	Varchar	The photo's author.
Affiliation	Varchar	The office's author.

Keyword_information Table consists of 3 fields: ID, Keyword, and Score. The keyword field represents keywords associated with each Author. The score shows the frequency of that associated keyword for each Author. Example of our keyword_information table structure is shown in Table II.

TABLE II. KEYWORD_INFORMATION

Field	*Type*	*Comments*
ID	Integer	The unique identifying of author.
Keyword	Varchar	The first name provided by the user.
Score	Decimal	The second name provided by the user.

Co-author_relation Table consists of the author pairwise (ID1 and ID2) and their associated score based on co-authoring similarity calculation. Example of our co-author_relation table structure is shown in Table III.

TABLE III. CO-AUTHOR_RELATION

Field	*Type*	*Comments*
ID1	Integer	The identifying of author one.
ID2	Integer	The unique identifying of author second.
Score	Decimal	The relation value from similarity calculation.

Keyword_relation Table consists of the author pairwise (ID1 and ID2) and their associated score based on Jaccard similarity calculation. Example of our keyword_relation table structure is shown in Table IV.

TABLE IV. KEYWORD_RELATION

Field	*Type*	*Comments*
ID1	Integer	The identifying of author one.
ID2	Varchar	The unique identifying of author second.
Score	Varchar	The relation value from similarity calculation.

The system load data from database or external files such as file formats are tab-delimited text or GraphML file formats [12] that have the relation value in each pair. End user will enter query to search for web browser and access the SWF file on a web server by having web browser request the SWF file. The returned visualizing expert network search results presented on the web browser with the highlighted person node. The center person being placed in the center of the display with their friends placed around. Individual nodes include the name of the person. The right side consists of a panel displaying a selected member's profile information. User can click on the center node to read full information and see relationship with other person. You can see the process graphically in Figure 3 and Figure 4.

Figure 3. Visualized social network for the query "Virach" by see from co-authoring.

Figure 4. Visualized social network for the query "Virach" by see from same research topic.

VI. CONCLUSION AND FUTURE WORKS

In this paper, we proposed a system framework of visualizing expert network from expert database for be helpful end user exploration an online expert network. The system is developed by using ActionScript, MXML, and Flare library is an open-source provide data visualization on the web and easy to create interactive data visualization. The MySQL database is used to store expert and expertise. AMFPHP provided database connectivity. Our system is run on an Apache Web Servers. To provide an easy to use interface we adopt the MXML integrated with web browser. The system support keyword search, user type search queries in the box then press bottom of the display matching nodes immediately highlight in center node and show expert profile data, and visualization of expert network. We aimed to apply information visualization techniques developed in our system framework. In future work, we additionally other features such as photo, tool tip in corner of the display, drag control interaction, panning and zooming to further explore the displaying that includes the advance search for further improving search effectiveness of visualized expert system in database. We are interested in applying other visualization techniques to more complex system framework.

ACKNOWLEDGMENT

This work is part of the Thai Web Archive Search Engine and Text Mining System project, which is supported by the Information and Mobile Applications (IMA) program under the National Science and Technology Development Agency (NSTDA), Thailand.

REFERENCES

[1] C. Allen, W. Arnold, A. Balkan, N. Cannasse, J. Grden, M. Gunesch, M. Hughes, R. Jon MacDonald, and A. Zupko, The Essential Guide to Open Source Flash Development, FriendSoft, 2008.

[2] K. Balog, L. Azzopardi and M. de Rijke, "Formal models for expert finding in enterprise corpora," In Proc. of SIGIR 2006, pp. 43-50, 2006.

[3] K. Balog and M. de Rijke, "Determining expert profiles (with an application to expert finding)," In Proc. of IJCAI 2007, pp. 2657-2662, 2007.

[4] M. E. Davis and J. A. Phillips, Flex3: A Beginner's Guide, Mc Graw Hill, 2008.

[5] J. Heer and D. Boyd, "Vizter: Visualizing online social networks," IEEE Symposium on Information Visualization (InfoVis 2005), 2005.

[6] Y. Takama, A. Matsumura and T. Kajinami, "Visualization of News Distribution in Blog Space," In Proc. of the 2006 IEEE/WIC/ACM international conference on Web Intelligence and Intelligent Agent Technology, p.413-416, 2006.

[7] J. Zhang, M. S. Ackerman and L. Adamic, "Expertise networks in online communities: structure and algorithms," In Proc. of WWW 2007, pp. 221-230, 2007.

[8] ActionScript and MXML, *http://www.actionscript.org/*

[9] AMFPHP, *www.amfphp.org/*

[10] Flare library, *http://flare.prefuse.org/*

[11] Flex SDK, *http://www.adobe.com/products/flex/*

[12] GraphML, *http://graphml.graphdrawing.org/*

[13] ScienceDirect, *http://www.sciencedirect.com/*

[14] Thailand's Research Information Portal (ThaiReSearch), available at *http://www.thairesearch.in.th*